Economic evaluation of supplementary screening using imaging modalities for women with dense breast tissue: systematic literature review

External review against programme appraisal criteria for the UK National Screening Committee

Version: 1

Author: Rodolfo Hernández, Mohammad Azharuddin, Sinéad N Duggan, Clare Robertson, Paul Manson, Gianni Virgili, Mike Clarke, Shaun Treweek, Miriam Brazzelli

NIHR Evidence Synthesis Group: Aberdeen-Belfast Evidence Collaboration (ABEC)

Date: February 2025

The UK National Screening Committee secretariat is hosted by the Department of Health and Social Care

Contents

Plain English Summary

In the UK, women aged 50-70 are offered mammograms every three years to help detect breast cancer early. This national screening programme has successfully reduced breast cancer deaths by 20-40%. However, mammograms are less effective for women with dense breasts, who also have a higher risk of cancer. About half of screened women have dense breasts, which can lead to missed cancers.

A 2019 review by the National Screening Committee found that there was no evidence that adding ultrasound to mammograms reduced cancers found between screenings, lowered mortality, or reduced NHS costs in participants with dense breasts who had negative screening results.

Recently, other screening methods have been considered. Clinical guidelines in the US and Europe now recommend that women be informed about their breast density and, in some cases, receive additional screening like Magnetic Resonance Imaging (MRI).

The UK National Screening Committee (NSC) commissioned a review looking at existing studies looking at the cost-effectiveness (value for money) of additional imaging for women with dense breasts. We found 12 economic studies comparing different screening approaches. Results varied:

- MRI was found to be cost-effective in some studies, particularly for younger women with the densest breasts, but this evidence was from non-UK studies.
- Results for digital breast tomosynthesis (DBT) and ultrasound were mixed, with some studies finding it cost-effective and others not.
- Contrast-enhanced mammography (CEM) was found to be cost-effective compared to mammography alone but was outperformed by other imaging methods.

Overall, findings were inconsistent, especially for UK-based studies. MRI alone (rather than as an additional test after a negative mammogram) showed potential benefits for younger women with very dense breasts. However, UK-specific research is needed to determine the screening strategy with the best value for money, considering factors like age, risk level, imaging accuracy, and screening frequency.

Executive summary

Background

The UK National Screening Committee (NSC) plays a vital role in early cancer detection, inviting women aged 50-70 years for digital mammography every three years. This national programme has led to a significant 20-40% reduction in breast cancer mortality risk. However, underdiagnosis remains a critical concern, particularly for women with dense breast tissue – a factor that not only increases cancer risk but also diminishes the effectiveness of mammograms. Nearly half of all women in screening programmes are classified as having heterogeneously or extremely dense breasts – categorised as BI-RADS (Breast Imaging Reporting and Data System) grades C and D, respectively - placing them at a heightened risk of missed breast cancer diagnoses. Given this substantial gap in detection, the question of whether additional imaging methods could improve outcomes has been a subject of ongoing debate.

In 2019, the NSC commissioned an evidence review to evaluate the potential benefits of supplemental ultrasound screening for women with dense breasts following a negative mammogram. This review concluded that existing evidence did not demonstrate that ultrasound (U/S) could reduce interval cancers and mortality, nor did it prove to be a cost-effective option for the NHS, leading to its rejection as an additional screening modality.

Meanwhile, the landscape of breast imaging is rapidly evolving. In 2022, the European Society of Breast Imaging (EUSOBI) issued guidelines recommending that women be informed of their breast density and offered supplemental magnetic resonance imaging (MRI) screening for those with extremely dense breasts. In the US, the Food and Drug Administration (FDA) has mandated that, from September 2024, mammogram reports must disclose breast density information and advice should be provided to patients that additional screening could aid cancer detection. In the UK, the recently completed multicentre BRAID (Breast Screening - Risk Adapted Imaging for Density) study investigated whether abbreviated MRI (Ab-MRI), contrast-enhanced mammography (CEM), and automated breast ultrasound (ABUS) could enhance cancer detection in women with dense breasts.

To inform future policy, the UK NSC has commissioned a suite of evidence reviews to assess supplemental imaging modalities to detect breast cancer in women with dense breasts and the role of these modalities within the national breast screening programme.

Focus of the review

- Objective 1: To determine the agreement (concordance) between automated and manual measurement of mammographic breast density

- Objective 2: To determine the effect of an additional imaging modality to supplement standard mammography compared with standard mammography alone for identifying breast cancer in women with dense breasts.

- Objective 3: To review evidence on existing economic models assessing the costs and consequences of enhanced mammographic screening for women with dense breasts.

This document addresses **Objective 3** and complies with **NSC criterion 14**, which indicates that i) the opportunity cost of the screening programme (including testing, diagnosis and treatment, administration, training and quality assurance) should be economically balanced in relation to expenditure on medical care as a whole (i.e., value for money) and ii) evidence from cost-benefit and/or cost-effectiveness analyses as well as the effective use of available resources should be considered.

Findings and gaps in the evidence of this review

Twelve economic evaluation studies were identified with three of these including more than one imaging modality. All studies used decision analytic models and compared new strategies with strategies using mammography only. Overall, 7 studies evaluated MRI, 4 DBT, 5 U/S, and 2 CEM

Study quality

The reporting standard for the 12 modelling economic evaluations was guided by the Philips checklist for good practice in decision-analytic modelling in health technology assessment. The reporting quality was variable with the proportion of positive answers ranging from 40% to 89%.

As expected, all models incorporated lower mammography sensitivity for individual with dense breast. However, 2 studies made questionable assumptions for the quality of life for women with no cancer and those recovering from treatment that could bias results in favour of the additional imaging modality. Another study obtained suspiciously high quality-adjusted life year (QALYs) differences between screening programmes that did not correspond with the life years gained. Half of the studies used validated pre-existing cancer models.

Key study findings

The seven studies assessing MRI show mixed results with five studies concluding that MRI was cost-effective. However, among the studies where MRI was cost-effective, 2 studies made questionable quality-of-life assumptions that could bias their results, and 1 study was from the USA and used a very high cost-effectiveness threshold. Two studies from the Netherlands, using pre-existing validated models, found that MRI and Ab-MRI were cost-effective, against biennial mammography:

- MRI alone every 4 years for women with BI-RADS D (extremely dense breasts)
- Biennial Ab-MRI for women with BI-RADS C and D (heterogeneous and extremely dense breasts) aged 50 to 65 years and mammography thereafter.

Two studies (one from Canada, and one from the UK) found very low QALY gains from MRI and Ab-MRI, concluding that neither were cost-effective when added to mammography.

Three of the four studies evaluating DBT found this modality to be cost-effective. However, from those studies where DBT was cost-effective, one reported implausible QALY gain and another from the USA used a very high cost-effectiveness threshold. One cost-effectiveness study from the Netherlands found that a sensitivity above 75% was needed for DBT to be cost-effective for biennial screening women with BI-RADS C&D. A more recent Canadian study, found that DBT was not cost-effective.

Five high quality studies evaluated the use of supplementary U/S. Two studies, one from the USA and one from the UK, found U/S to be cost-effective. The remaining three studies conducted in Canada, USA and the UK, found U/S was not cost-effective.

CEM was evaluated in two studies; one assessing its use as a supplementary tool after negative mammography and the other examining its addition between triennial mammography screenings. While both studies found CEM to be cost-effective compared with mammography alone, CEM strategies were dominated by those including other imaging modalities such as U/S, Ab-MRI, and ABUS.

In summary, findings were mixed or unfavourable for most modalities when used as an adjunct to mammography alone. MRI alone, as opposed to supplemental after a negative mammography, is potentially cost-effective, particularly among the youngest women with the highest breast density, who have the greatest scope to benefit. However, it is worth noting that

this evidence comes from non-UK settings. DBT results are also mixed with the latest cost-utility analysis conducted for Canada finding this modality not being cost-effective. CEM was not cost-effective compared with strategies using alternative supplementary imaging modalities. In contrast, a recent UK study found automated U/S to be cost-effective, contradicting the findings from an earlier UK study.

Considerations and uncertainties

The evidence for the cost-effectiveness of supplemental imaging modalities for breast cancer screening in women with dense breasts has been explored using good quality economic models, but with variable or unfavourable findings. MRI alone, as opposed to supplemental after negative mammography, is potentially cost-effective, particularly in the youngest women with the highest breast density, who have the greatest scope to benefit. However, this evidence comes from non-UK studies.

We identified two studies assessing the cost-effectiveness of alternative screening strategies from the UK NHS perspective. Both studies used good quality economic models, one of which – the MANC-RISK-SCREEN model - has undergone both internal and external validation. Notably, these models yielded very different cost-effectiveness results for similar risk-stratified screening strategies compared with triennial mammography. While an in-depth analysis of these models was beyond the scope of this review, a clearer understanding of their differences is essential before they can be considered for informing policy decisions.

Limitations and strengths

This systematic review of economic evaluations is based on studies with diverse methodologies and healthcare settings, limiting direct applicability to the UK. Furthermore, variations in economic modelling parameters introduce uncertainties in the observed cost-effectiveness estimates.

All the included studies used economic evaluation decision analytic models, many of which included complex mathematical sub-models based on a series of assumptions. One example is the choice of tumour growth model and the estimated timeframe during which a tumour can be detected through screening or clinical diagnosis - factors that ultimately will determine the

maximum possible benefit of early screening. Without access to the original models, it is challenging to pinpoint the key drivers of clinical and economic results or to fully assess the suitability of these economic models for UK decision-making.

Overall conclusions

The cost-effectiveness of supplemental breast cancer screening for women with dense breasts has been explored using high-quality economic models, yet findings remain variable or unfavourable. MRI alone - rather than as a supplemental modality after negative mammography - appears potentially cost-effective, particularly for younger women with the highest breast density, who have the greatest scope to benefit. However, no study comprehensively assessed, the full spectrum of viable alternatives from the UK NHS perspective, including varying age ranges, imaging modalities (alone or in combination), and screening frequency. There is a clear need for a UK-based cost-effectiveness study that evaluates a range of feasible approaches to enhance screening for women with dense breasts, taking into account individual risk profiles, possible adjustments to the frequency of screening, age, and the accuracy of available imaging modalities when used alongside or as a replacement for mammography.

Introduction and approach

Background

Breast cancer is the most common type of cancer among women in the UK, accounting for 15% of all new cancer cases. Each year around 55,900 are diagnosed, more than 150 daily.[1] Whilst breast cancer can occur at any age, it most commonly affects postmenopausal women over 50 years of age.

The UK national screening programme invites women aged 50-70 years for mammograms (digital mammography) every three years, significantly reducing breast cancer mortality risk by 20-40%.[2-4] However, underdiagnosis remains a concern, particularly for women with dense breast tissue, which not only increases cancer risk but also makes mammograms less effective.[5]

Breast density is determined by the proportion fibroglandular tissue visible on a mammogram. It is classified into four categories (A,B,C,D) according to the American College of Radiology Breast Imaging Reporting and Data System (ACR BI-RADS) system[6] with categories C and D indicating moderate to extremely dense breasts. Nearly half of women in screening programmes fall into these two groups, making them more prone to missed diagnoses.[7] In clinical practice, breast density has traditionally been assessed through a subjective manual evaluation, where radiologists visually inspect mammograms to categorise breast density. There are also automated systems to measure density such as Volpara software that measures volumetric breast density providing a density score that corresponds with the BI-RADS density classification.

Ethnicity also plays a role. For example, women in Japan have denser breasts. Unlike Western countries, where screening has reduced mortality, Japan has not observed a similar reduction in mortality rates[8], possibly due to underdiagnosis with standard mammography.[8]

Risk-adapted screening: current UK position

In 2019, the National Screening Committee (NSC) commissioned an evidence review to evaluate the benefits of additional ultrasound (U/S) screening for women with dense breasts after a negative mammogram.[9]

The review found that increased breast density was associated with a reduced mammography sensitivity and a higher risk of interval cancers. However, it also highlighted challenges in

validating the breast density measurement methods and reported high false positive rates with U/S. Notably, no evidence supported U/S in reducing interval cancers and mortality, nor demonstrating its cost-effectiveness, leading to its rejection as a supplemental screening modality.[9]

The field of breast imaging is rapidly advancing. In 2022, the European Society of Breast Imaging (EUSOBI) guidelines recommended informing women about their breast density and offering supplemental screening for those with extremely dense breasts.[10] In the US, the Food and Drug Administration (FDA) has mandated that, from September 2024, mammogram reports must disclose breast density and inform patients that additional screening may aid cancer detection.[11]

In the UK, the multicentre Breast Screening - Risk Adapted Imaging for Density (BRAID) trial, which investigates whether abbreviated (Ab-MRI), contrast-enhanced mammography (CEM), and automated breast ultrasound (ABUS) improve cancer detection in women with dense breasts, has recently been completed.[12, 13]

To inform future policy, the UK NSC has commissioned a suite of evidence reviews to assess supplemental imaging modalities to detect breast cancer in women with dense breasts and the role of these modalities within the national breast screening programme.

Objectives

Specifically, these reviews address the following objectives:

- Objective 1: To determine the agreement (concordance) between automated and manual measurement of mammographic breast density
- Objective 2: To determine the breast cancer screening performance of supplemental imaging modalities for women with dense breasts at risk of breast cancer.
- Objective 3: To review evidence on existing economic models assessing the costs and consequences of enhanced mammographic screening for women with dense breasts.

This document addresses Objective 3 and complies with NSC criterion 14, which indicates that i) the opportunity cost of the screening programme (including testing, diagnosis and treatment, administration, training and quality assurance) should be economically balanced in relation to expenditure on medical care as a whole (i.e., value for money) and ii) evidence from cost-benefit

and/or cost-effectiveness analyses as well as the effective use of available resources should be considered.

This report is structured in a classical way with the next section describing the methods, followed by the results, discussion and conclusion.

Methods

General

This systematic review was commissioned by the NSC and was conducted in accordance with the recommendations of the Cochrane Handbook for Systematic Reviews of Interventions and in adherence with the Preferred Reporting Items for Systematic Reviews (PRISMA) guidelines. The methods were pre-specified in a protocol and registered with the PROSPERO International Prospective Register of Systematic Reviews, available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42024550250.

Patient and public involvement (PPI)

Two PPI partners were part of the study Advisory Group, which also included academic and clinical experts. One PPI partner has lived experience of undergoing mammography for routine breast screening and the other has lived experience of breast cancer. PPI partners participated in regular Advisory Group meetings, where they contributed to discussions and made recommendations at each stage of the project.

Language and inclusivity statement

Most, but not all, people who use the UK's breast screening programme identify as women. While exclusively gender-neutral language can be more inclusive, it can also make writing harder to understand. None of the studies identified in our reviews presented non-binary genders. We have chosen to use both 'women' and, where we can, gender-neutral language. We acknowledge that this is a compromise, but when readers see 'women', we encourage readers to interpret this as meaning all users of the breast screening service, not only those who identify as women.

Role of the funding source

The NIHR Aberdeen-Belfast Evidence Collaboration (ABEC) was funded by the NIHR Evidence Synthesis Programme to conduct this review (project no. NIHR164221). The funder of the study and the NSC contributed to the conceptualisation of the research question and study design, but had no role in data collection, data analysis, data interpretation, or writing of the report.

Methods for reviewing economic evaluation studies

Search strategy for identification of studies

An Information Specialist developed a comprehensive literature search strategy to identify relevant published peer-reviewed economic evaluation studies. Major electronic databases were searched, including MEDLINE, Embase, NHS Economic Evaluations Database, International HTA Database (INAHTA), Cost-Effectiveness Analysis (CEA) Registry, International Society for Pharmacoeconomics and Outcomes Research (ISPOR), EconPapers, EconLit, the Canadian Agency for Drugs and Technologies in Health (CADTH, now Canada's Drug Agency (CDA), and the Institute for Clinical and Economic Review (ICER).

The search focused on primary economic evaluations published in English since 2014 that assess the costs and consequences of supplemental imaging modalities for the detection of breast cancer in women with dense breasts.

All references were exported to EndNote X9 for recording and deduplication. A MEDLINE search is detailed in Appendix 1. The MEDLINE search was adapted for use in other electronic databases.

Eligibility/Inclusion Criteria

The eligibility criteria for the review of economic evaluation studies are summarised in Table 1.

The included population were individuals aged 40 to 70 years who were undergoing breast cancer screening and had been stratified by breast density categories using either visual or automated methods. The intervention focused on supplemental imaging modalities for breast cancer detection in women with dense breast tissue. These included modalities such as digital breast tomosynthesis (DBT), U/S techniques (i.e., hand-held ultrasound -HHUS- or ABUS), magnetic resonance imaging (MRI), either using full (Fp-MRI) or abbreviated (Ab-MRI) protocols,

and CEM. The comparator intervention was standard mammography, which included digital mammography (DM) and x-ray mammography (XM).

Relevant outcome measures included medical resource use and associated costs, life years (LYs) or life years gained (LYG), quality adjusted life years (QALYs), the incremental cost-effectiveness ratio (ICER) per life year or QALY gained, net health benefit, and cost per accurate diagnosis or cost per true positive detection.

We focus on primary economic evaluations assessing the costs and consequences of supplemental imaging modalities for breast cancer detection in women with dense breasts, published in English within the last 10 years. Studies comparing the costs and consequences of two or more alternative courses of action were eligible. These included cost-effectiveness, cost-utility, cost-benefit, and cost-consequence analyses.

Table 1. Eligibility criteria for the review of economic evaluations

Population	**Women aged 40-70 undergoing screening, stratified by breast density using visual or automated methods**
Intervention	Supplemental imaging modalities for breast cancer detection in women with dense breasts, including: • Magnetic resonance imaging (MRI) (full/abbreviated), magnetic resonance mammography (MRM) • Contrast-enhanced mammography (CEM) • Ultrasound (U/S) (hand-held HHUS/automated ABUS) • Digital breast tomography (DBT)
Comparator intervention	Digital mammography (DM), X-ray based mammography (XM), 2-dimension Digital Mammography (2D DM)
Outcomes	Relevant outcome measures included: • Health care resource use and costs • Life years (LYs)/life years gained (LYG) • Quality adjusted life years (QALYs)/QALYs gained • Incremental cost-effectiveness ratio (ICER) per life year or QALY • Net health benefit (NHB) • Cost/accurate diagnosis • Cost/true positives
Study design	Included full primary economic evaluations (e.g., cost-effectiveness, cost-utility, cost-benefit, and cost-consequence analyses) assessing costs and outcomes of supplemental imaging for breast cancer detection. Restricted to studies published in English in the last 10 years.
Healthcare setting	Not restricted to a specific healthcare setting but apply to any context where breast cancer screening has been conducted.

Study selection

One reviewer (MA) screened the citations identified by the search strategies, while a second reviewer (RH) independently screened a random 20% sample of these citations. Potentially relevant articles were then retrieved in full and assessed by one reviewer (MA) according to pre-specified inclusion criteria, with 20% of studies assessed independently by a second reviewer

(RH). The number of excluded studies and the main reasons for their exclusion were documented. The study selection process is illustrated in a PRISMA flow diagram.[14]

Data extraction and synthesis

Data were extracted by one reviewer (MA) using a standardised data extraction template in Microsoft Excel®.

The following information was recorded from each study:

1. Characteristics of studies: first author, year of publication, country, aim/objective
2. Characteristics of study participants: population/screening strategy (age, population)
3. Characteristics for the analysis: perspective, time horizon, discounting, currency, and cost year
4. Characteristics of the intervention and comparators
5. Characteristics of the models: model type, health states, cycle length
6. Relevant outcomes: benefit measures/outcome measures, threshold for measuring cost-effectiveness, total costs/LYs/QALYs for intervention and comparator, incremental costs/LYs/QALYs, base case ICER and NHB

The findings of each included study were tabulated and summarised narratively for each outcome of interest.

Critical appraisal

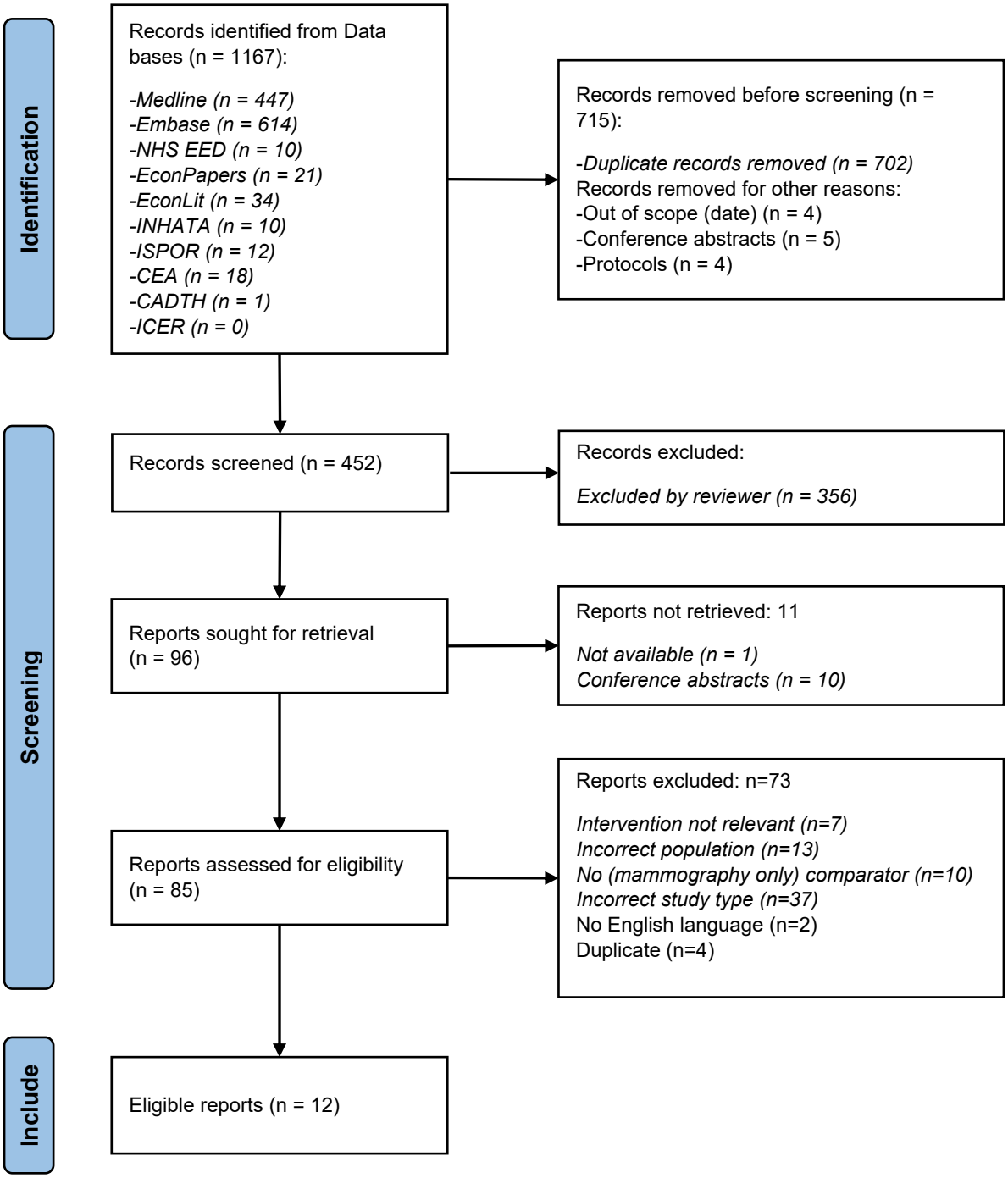
The reporting standard of any identified trial-based economic evaluations was assessed using the Consolidated Health Economic Evaluation Reporting (CHEERS) checklist.[15] The quality of decision models was critically appraised by assessing the appropriateness of the model structure and data inputs such as treatment effects, utility weights, resources, and costs, as well as the way parameters were incorporated in the decision model. The quality check was guided by the criteria used in the Philips checklist for good practice in decision-analytic modelling in health technology assessment.[16]

Results

Number of studies retrieved from the searches

The literature searches identified a total of 452 titles and abstracts. Out of these, 96 publications were selected for full-text screening, and 12 met our inclusion criteria. Figure 1 provides a PRISMA flow diagram that outlines the study selection process. Reasons for exclusion at full text screening stage are provided in Appendix 2.

Figure 1: PRISMA flow diagram



Description of included studies

A total of 12 publications were included in this review. Five studies were conducted for the USA,[17-19, 23, 26] three for the Netherlands,[20-22] two in the UK,[27, 28] and one study each for Brazil,[24] and Canada,[25] (see Table 2).

Table 2: Summary characteristics of included studies

Study (first author, year) (Country)	Aim/Objective	Population/screening strategy	Intervention/ Comparator	Model type	Health states/ other related information
Cost-utility analysis					
Couto, 2024[24] (Brazil)	To estimate the cost effectiveness of switching from DM to DBT + s2D vs. DM used in a biennial breast screening of women aged 40-69 years with scattered areas of fibroglandular breast density and heterogeneous dense breasts.	40-69 years of age, scattered areas of fibroglandular breast density & heterogeneous dense breasts, Biennial screening	DBT + s2D mammograms Comparator: DM	Decision tree and Markov model	Decision: Positive (BI-RADS 4 and 5*), suspect (BI-RADS 0), or negative/benign (BI-RADS 1, 2, and 3) Markov: DCIS, TNM 1, TNM 2, TNM 3 or TNM 4, distant metastasis, death
Hill 2023,[28] UK	To assess the cost-effectiveness if introducing new risk-stratified breast cancer screening programmes into the UK NHS Breast Screening Programme.	50-70 years of age, Triennial mammogram BI-RADS A&B and Additional 18-monthly imaging for BI-RADS C&D	ABUS, CESM, Ab-MRI Comparator: Triennial mammogram	Individual-level microsimulation model	DCIS or invasive. If invasive categorised as less than 10mm, ≥10mm and <20mm, ≥ 20mm and < 50, or ≥50mm. TNM classification used to estimate survival.
Ontario Health, 2023[25] (Canada)	To assess the cost-effectiveness of supplemental screening with U/S, MRI, or DBT as an adjunct to mammography.	50 to 74 years of age, heterogeneously & extremely dense breasts, Biennial screening, & annually for extremely dense breasts	Mammography screening + supplemental MRI, U/S (HHUS), or DBT Comparator: Mammography alone	Individual-level microsimulation model	Absence of BC; stage-specific BC DCIS, stages I to IV) accounting for varied treatment phases; and death.
Blankenburg, 2022[17] (USA)	To evaluate the cost-effectiveness of supplemental breast imaging modalities for women with heterogeneously and extremely dense breasts	40-74 years of age, heterogeneously & extremely dense breasts	Supplemental imaging modalities including (Fp-MRI, Ab-MRI),	Decision tree linked to a Markov chain	No BC; undetected BC; detected invasive BC tumour size: ≤ 1 mm, 1-5 mm, 5-10 mm, 10-20 mm, 20-50 mm, >50 mm; detected-DCIS; post-treatment A; post-treatment B; tumour-related death; non-tumour-related death

Study (first author, year) (Country)	Aim/Objective	Population/screening strategy	Intervention/ Comparator	Model type	Health states/ other related information
	and average or intermediate risk of breast cancer in the USA and analyse capacity requirements for supplemental MRI and CEM.	Annual screening	CEM, and U/S (HHUS) as add-on to XM or DBT Comparator: XM		
Geuzinge, 2021[20] (The Netherlands)	Estimate the cost-effectiveness of MRI screening compared with mammography in women with extremely dense breast tissue by using the results of the DENSE trial and microsimulation modelling. Quantify the effects and costs of several different screening scenarios by varying the screening interval between MRIs and mammograms offered for women aged 50-75 years.	50-75 years of age, extremely dense breasts Biennial screening	MRI Comparator: Mammography	MISCAN: Individual-level microsimulation model	No BC, after a (false) positive screening result, after undergoing screening, DCIS/localized BC, regional BC, metastasis, death
Kaiser, 2021[26] (USA)	Assess the cost-effectiveness of MRM in comparison to XM in screening patients of intermediate risk due to elevated breast density.	50 years at simulation start extremely dense breast, Biennial screening	MRM Comparator: XM	Decision and Markov Model	No breast malignancy, Undetected breast malignancy, detected breast malignancy >1cm, detected breast malignancy <1cm, After extended therapy, After therapy of local disease, tumour related death
Tollens, 2021[23] (USA)	To evaluate the cost-effectiveness of MRM in comparison to XM in screening women of intermediate risk for breast cancer due to their elevated	55 years at simulation start, extremely dense breast, Biennial screening	MRM Comparator: XM	Decision and Markov Model	Decision: breast neoplasm, no breast neoplasm, true positive, false negative, true negative and false positive Markov: absence of malignancy, undetected breast malignancy, detected breast malignancy > 1 cm,

Study (first author, year) (Country)	Aim/Objective	Population/screening strategy	Intervention/ Comparator	Model type	Health states/ other related information
	breast density, considering the changes in specificity and false positives in a follow-up situation.				detected breast malignancy < 1 cm, post-treatment status and death
Gray, 2017[27] (UK)	CUA to identify the potential impact of introducing stratified National Breast Screening Programme in the UK and key drivers of the relative cost-effectiveness of different types of stratified NBSPs.	50-75 years of age heterogeneously & extremely dense breasts (altogether defined as high dense), Screening every 3 years	Supplemental U/S (HHUS or ABUS) for high breast density), Supplemental MRI (both high breast density and high risk of breast cancer) Comparator: Current UK NBSP (Mammography)	DES	The model was conceptualized to include three components: stratification, breast cancer natural history with screening, and diagnosis/treatment after detection. A DES model was used to represent these elements. The economic model includes a continuous time tumour size growth mathematical model.
Lee, 2015[18] (USA)	To evaluate the effectiveness of combined biennial DM and tomosynthesis screening, compared with biennial DM screening alone, among women with dense breasts.	50-74 years of age, heterogeneously or extremely dense breasts, Biennial screening	DM + tomosynthesis Comparator: DM	DES (Wisconsin model; CISNET)	Healthy, stage I, in-situ cancer, stage II, local cancer, stage III, regional cancer, stage IV, distant metastases
Sprague, 2015[19] (USA)	To evaluate the benefits, harms, and cost-effectiveness of supplemental U/S screening for women with dense breasts.	40 years at simulation start, heterogeneous & extremely dense breasts, Biennial or annual screening	Supplemental U/S (HHUS) screening Comparator: Mammography	CISNET: Individual-level microsimulation model	<u>Used three micro simulation models</u> Model E: preclinical in situ disease; subset does not progress from in situ to invasive Model W: preclinical in situ disease; subset does not progress from early invasive and may regress if undetected Model G-E: preclinical in situ disease; subset does not progress from in situ to invasive

Study (first author, year) (Country)	Aim/Objective	Population/screening strategy	Intervention/ Comparator	Model type	Health states/ other related information
Cost-effectiveness analysis					
Wang, 2022[21] (The Netherlands)	To investigate the cost effectiveness of Ab-MRI in women with dense breasts (heterogeneously/extremely dense) in a population-based screening program.	50-74 years of age, heterogeneously & extremely dense breast, Biennial screening	Ab-MRI Comparator: Mammography	SiMRiSc: Micro simulation model	Women's lifetimes were simulated by considering their life expectancy, the chance of developing cancer, tumour growth, tumour self-detection probability and survival probability.
Wang, 2020[22] (The Netherlands)	To evaluate at which sensitivity DBT would become cost-effective compared to DM in a population breast cancer screening program, given a constant estimate of specificity.	50-75 years of age, heterogeneously & extremely dense breast, Biennial screening	DBT Comparator: DM	SiMRiSc: Micro-simulation model	Women's lifetimes were simulated based on life expectancy, cancer risk, tumour growth, self-detection probability, and breast cancer survival.

Abbreviations: DM: digital mammography, DBT: digital breast tomosynthesis, s2D: two-dimensional mammograms, BI-RADS: Breast imaging-reporting and data system, DCIS: ductal carcinoma in situ, TNM: Tumour Nodes and Metastasis, U/S: ultrasound, MRI: magnetic resonance imaging, HHUS: Hand-held ultrasound, CESM: Contrast-enhanced spectral mammography, , Fp-MRI: full--protocol magnetic resonance imaging, Ab-MRI: abbreviated-protocol magnetic resonance imaging, CEM: contrast-enhanced mammography, XM: x-ray mammography, MISCAN: Microsimulation Screening Analysis, BC: breast cancer, MRM: MR-Mammography, DES: discrete event, NBSPs: national breast screening programs simulation. Current UK NBSP: Women between 50 and 70 years with screening every 3 years using mammography.

* The BI-RADS assessment categories differ from the BI-RADS density classification. Assessment categories evaluate the risk of malignancy, emphasizing imaging findings and cancer suspicion. Mammogram results into six numbered categories, from 0 to 6. Categories: 0) results are incomplete, additional imaging needed, 1) negative, no abnormalities detected, 2) benign findings, no cancer, 3) probably benign, follow-up in a short time frame needed, 4) suspicious abnormality, biopsy recommended. Subcategories 4A (low probability, 2%-10%), 4B (moderate probability, 10%-50%), 4C (high probability, 50%-95%), with biopsy recommended, 5) highly suggestive of malignancy (>95%), immediate diagnostic action or treatment required, 6) biopsy-proven malignancy, immediate and appropriate action based on the confirmed diagnosis (*Guide to Mammography Reports: BI-RADS Terminology REBECCA B. BITTNER, MD George Washington University, Washington, DC, https://pubmed.ncbi.nlm.nih.gov/29083600/*).

Among the 12 studies, 10 were cost-utility analyses (CUA),[17-20, 23-28] and two were cost-effectiveness analyses (CEA)[21, 22] (Table 2). Whilst CEA and CUA measure costs in monetary units (e.g., British Pounds Sterling), CEA measures health consequences using a clinical measure such as the number of cancer cases detected or LYs, and CUA uses a measure of quality adjusted life expectancy (most commonly the QALYs).

All 12 studies were economic modelling studies (Table 2). The most commonly used model type was microsimulation model, employed in six studies.[19-22, 25, 28] Discrete event simulation (DES) was used in two studies,[18, 27] while a combination of decision tree and Markov models were used in four studies*.[17, 23, 24, 26]

Some of the modelling studies adapted previously published models rather than developing de novo economic evaluation decision models to address the question:

- Geuzinge et al. (2021) used an updated version of the Microsimulation Screening Analysis (MISCAN) model to extrapolate findings of the DENSE trial (dwell times and sensitivities of mammography).[20] The MISCAN computer simulation package was developed to estimate the effects of implementing a nationwide breast cancer screening programme in the Netherlands.[29, 30] The model incorporates the natural history of breast cancer, epidemiological data, the design of the screening programme, and the performance of screening strategies[31, 32] simulating individual life histories for women and tracking the progression of breast cancer over time.
- The studies by Wang et al. (2020)[22] and Wang et al. (2022)[21] employed the validated SiMRiSc model, updating relevant parameters for their specific population of interest.

* In a micro-simulation model, individuals are passed through the model one at a time, with each individual's results stored separately. The experience of a cohort is then derived by aggregating the outcomes of all individuals. DES models the individual experiences of patients over time, tracking and summarising the events that occur and their consequences; however, the transitions between a patient's health states are driven by events (e.g., diagnosis) that occur at varying times. A decision tree model uses distinct branches to represent potential outcomes for an individual patient or a patient cohort. The tree consists of a series of "nodes", that branch out from left to right, representing the options or chance events that emanate from the nodes. Each node can represent either a "choice" (a decision regarding which intervention to apply) or a "probability" (the likelihood of an event occurring or not, based on chance). Costs and outcomes are assigned to each segment of every branch, including the end ('leaf') of each branch. Finally, Markov models use distinct disease states to represent all possible outcomes of an intervention such as progression-free, post-progression, and dead for a cancer intervention. Time is divided into discrete periods, called "cycles" and people can transition between these states at the end of each cycle, as their condition changes over time. Cost and health outcomes are attached to the health states and are aggregated for a cohort of patients to provide a summary of the cohort experience. (https://yhec.co.uk/resources/glossary/).

SiMRiSc is a micro-simulation model[33-35] that simulates the lifetime of women, taking into account their life expectancy, the chance of developing a tumour, tumour growth, and survival from breast cancer by stage at diagnosis. The model adjusts mammographic sensitivity according to breast density.

- Lee et al. (2014)[18] and Sprague et al. (2015)[19] used models within the Cancer Intervention and Surveillance Modelling Network (CISNET). The CISNET Breast Cancer programme, sponsored by the National Cancer Institute, USA, is a collaborative research effort in which various modelling groups have studied the impact of screening and adjuvant treatment on trends in breast cancer incidence and mortality from 1975 to 2000. Lee et al. (2014)[18] simulated the model population based on the US breast cancer epidemiological model (part of CISNET) using the University of Wisconsin model, and Sprague et al. (2015)[19] used three established CISNET models (Models E -Erasmus University Medical Centre-, Model G-E -Georgetown University Medical Centre & Albert Einstein College of Medicine model-, and Model W -University of Wisconsin & Harvard Medical School model) to evaluate the benefits, harms, and cost-effectiveness of supplemental screening with U/S.

Population, intervention and comparator across studies

The population of interest was women aged 40 to 70 years stratified by breast density categories using either visual or automated methods.

Seven modelling studies considered women with heterogeneously and extremely dense breasts (BI-RADS C & D).[17-19, 21, 22, 25, 27] Three studies included women with extremely dense breasts only (BI-RADS D),[20, 23, 26] and one study women with scattered areas of fibroglandular breast density & heterogeneous dense breasts (BI-RADS B & C).[24]

Four studies assessed the use of MRI in women in high and/or very high breast density,[20, 23, 26] with Wang et al. (2022) assessing Ab-MRI.[21] Three studies evaluated DBT, either alone[22] or in combination with DM[18] or 2D mammography[24] (Table 2). Two studies assessed supplemental U/S with mammography;[19, 27] Gray et al. (2017),[27] also included MRI in their intervention strategy for individuals with high breast density and high risk of breast cancer. The studies by Blankenburg et al. (2022)[17] and Ontario Health (2023),[25] and Hill et al. (2023)[28] evaluated multiple individual modalities as supplemental to mammography screening. Blankenburg et al. (2022)[17] included Fp-MRI, Ab-MRI, CEM and U/S (HHUS) while Ontario Health (2023)[25] included MRI, U/S (HHUS), and DBT. Hill et al. (2023) considered automatic U/S, CEM, and Ab-MRI.[28]

The studies exhibited inconsistency and heterogeneity with respect to the type of "mammography" used as the comparator. Six studies referred to mammography alone,[19-21, 25, 27, 28] three used DM,[18, 22, 24] two used XM,[17, 23] and one used the abbreviation XM but described it as 'conventional mammography'.[26] Due to the varying use of terms regarding comparators, we have gathered and derived the definition from multiple sources. Mammography is an x-ray imaging technique used to examine the breast for early detection of cancer and other breast diseases serving as both a diagnostic and screening tool. DM is a modern form of mammography that utilises solid-state detectors instead of traditional film to record the X-ray pattern passing through the breast. These detectors convert the X-rays into electronic signals, which are then processed by a computer to create images displayed on a monitor and stored for future use. Compared to film mammography, DM offers several advantages, such as the ability to manipulate image contrast for enhanced clarity, use computer-aided detection for identifying abnormalities, and easily transmit digital images for second opinions. Additionally, the digital nature of DM can reduce the need for re-takes, which are more common with film mammography due to issues like incorrect exposure or film development problems, leading to lower X-ray exposure.

Descriptions of the individual studies

This section provides a brief description of the population and screening interventions for each individual study. These are grouped by the imaging modality assessed.

Magnetic Resonance Imaging (MRI)

Geuzinge et al. (2021)[20] evaluated supplemental MRI screening compared to mammography alone, offered to women with extremely dense breasts (BI-RADS D) and aged 50-75 years, simulating several screening strategies with varying intervals for MRI and mammography (from every two to every six years). The model (MISCAN-breast microsimulation) was calibrated using results for the DENSE trial and adjusted to incorporate decreasing density with increasing age. The DENSE trial is a study embedded within the Dutch biennial mammography screening programme, in which women with extremely dense breasts and a normal mammography result were randomly assigned to two groups: an invitation for additional MRI screening or no additional screening (control group). Within the defined screening strategies, Geuzinge et al. (2021)[20] included those corresponding to the study groups for the DENSE trial: mammography plus MRI every two years (2Mx_2MRI) for the intervention group, and mammography every two years (2Mx) for the control group.

The studies by Kaiser et al. (2021)[26] and Tollens et al. (2021)[23] evaluated the cost-effectiveness of biennial screening with magnetic resonance mammography (MRM) compared to XM. The decision models used a starting age of 55 years and were run for 30-year[26] or 20-year time horizons.[23] These studies were conducted by the same academic group. Kaiser et al. (2021)[26] used data for the first screening round from the DENSE study while Tollens et al. (2021)[23] updated the analysis using data for the first and second screening rounds. Of particular interest was the data on specificity for MRI observed between screening rounds (i.e., 92% for first and 97% for second) that could result in lower false positive result rates.

Wang et al. (2022)[21] explored the cost-effectiveness of Ab-MRI as an alternative to mammography for biennial screening. Six base-case intervention biennial screening strategies were examined: (1) Women with heterogeneously dense breasts (BI-RADS C) receive mammography from ages 50-74, while women with extremely dense breasts (BI-RADS D) receive Ab-MRI from 50-65 and mammography from 66-74; (2) Women with BI-RADS C receive mammography from 50-74, and women with BI-RADS D receive Ab-MRI from 50-69 and mammography from 70-74; (3) Women with BI-RADS C receive mammography from 50-74, and women with BI-RADS D receive Ab-MRI from 50-74; (4) All women with BI-RADS C&D receive Ab-MRI from 50-65 and mammography from 66-74; (5) All women with BI-RADS C&D receive Ab-MRI from 50-69 and mammography from 70-74; and (6) All women with BI-RADS C&D receive Ab-MRI from 50-74. The comparator was biennial mammography screening for women 50 to 74 years old.

Digital breast tomosynthesis (DBT)

DBT, also known as a 3D mammography, creates a three-dimensional image of the breast using multiple X-ray images taken from different angles.[36] Five studies assessed DBT: Couto et al. (2024),[24] Lee et al. (2015),[18] Wang et al. (2020),[22] Blankenburg, et al. (2021)[17] and Ontario Health (2023).[25] The last two studies assessed more than one technology and are reported in a separate section (see *Modelling studies assessing more than one imaging technique*).

The study by Couto et al. (2024)[24] estimated the cost-effectiveness of switching from DM to DBT + synthesized two-dimensional (s2D) mammography in biennial breast screening following the guidelines of the Brazilian supplementary health system. Participants underwent biennial breast screening and received DBT combined with s2D mammography in the intervention arm, compared to DM in the control arm. s2D, generated through slab reconstruction from the

tomosynthesis acquisition, was approved by the U.S. Food and Drug Administration in May 2013 as an alternative to DM when using DBT. This approach maintains screening effectiveness while reducing radiation exposure, with the radiation dose derived solely from the DBT exam, which is approximately 45% lower than the combined dose of digital mammography and DBT.[37]

Similarly, Lee et al. (2015)[18] studied combined biennial DM and DBT screening, comparing it with biennial DM screening alone. The model of this study is based on a U.S. breast cancer epidemiological model (CISNET). This study used a hypothetical cohort of women followed from age 50 to 74 years through 12 screening rounds. The model was modified to study only the U.S. subpopulation with heterogeneously dense or extremely dense breasts as seen on mammography. In the screening scenarios, all women underwent mammography at age 50 (since breast density is determined from a baseline mammogram) and were then screened with either mammography alone or combined with tomosynthesis starting with their first follow-up visit at age 52 and continuing through to age 74.

The cost-effectiveness analysis by Wang et al. (2020)[22] assessed biennial breast cancer screening simulated for a population aged 50-75 in The Netherlands, evaluating the impact of replacing DM with DBT for all women or for those with high breast density, compared to the current practice of screening with DM alone. Three scenarios were explored: in the first scenario, DBT was used for women with high-density breasts (BI-RADS 4th edition density scores 3 and 4), while DM was used for women with non-dense breasts; the second scenario involved using DBT for all women aged 50-75 years; and a third scenario that served as the reference, simulated the current practice of biennial DM screening for all women in this age group. The first and third scenarios are relevant for our review.

Supplemental Ultrasound (U/S)

U/S imaging can be differentiated according to the mode of application between HHUS and ABUS. HHUS appears to have significant limitations that have restricted its widespread integration into the screening environment, such as the lack of standardisation of the technique, the need for high skill and experience, time consumption, and its small field of view (FOV). ABUS seems to address these challenges of HHUS.[38] Our review revealed that the reporting of supplemental U/S varies across studies that included U/S alone or as one of the multiple imaging technologies, with most using HHUS,[17, 19, 25] and one study employing a combination of HHUS and ABUS,[27] and one using ABUS.[28] The studies (Blankenburg, et al. (2021)[17], Ontario Health (2023)[25], and Hill 2023) assessed more than one technology and are

reported in a separate section (see *Modelling studies assessing more than one imaging technique*).

Gray et al. (2017)[27] included women eligible for the UK National Breast Screening Programme (NBSP), with a mean age of 49 years. The current UK NBSP offers screening every three years using mammography for women aged 50 to 70 years; however, the mean age for this early technology assessment model was 49 years. Gray et al. (2017)[27] used the NBSP as a comparator and evaluated four potential approaches to stratify screening within the NBSP using a 10-year cancer risk algorithm enhanced by density:

- Risk 1, three strata with associated screening intervals defined by 10-year risk of breast cancer: a) <3.5% (3-yearly), b) 3.5%–8% (2-yearly), and 3) >8% (annually)
- Risk 2, screening population divided into thirds (tertiles): a) the lowest risk tertile (3-yearly), b) the middle tertile (2-yearly), and c) the highest risk tertile (annually)
- Masking, current screening approach with supplemental U/S offered to women with high breast density (heterogeneous or extremely dense) and MRI offered to women with high breast density and high 10-year risk of cancer defined >8%.
- Risk 1 & masking, used the stratification for Risk 1 together with the strategy described for the Masking approach.

The *Masking* and *Risk 1 & masking* strategies in Gray (2017)[27] are the only ones that met our inclusion criteria.

Sprague et al. (2015)[19] simulated cohorts of 40-year-old women with initial breast density assigned based on BI-RADS categories from the Breast Cancer Surveillance Consortium (BCSC) and adjusted at age 50 according to the same source to match the observed prevalence for postmenopausal women. The study compared three strategies for annual screening: 1) mammography alone; 2) mammography combined with screening U/S (HHUS) after a negative mammogram for women with extremely dense breasts; and 3) mammography combined with HHUS after a negative mammogram for women with heterogeneously or extremely dense breasts (base-case). In addition, a strategy of no screening was used but this is not considered in this review as this is not relevant for UK decision making where a screening programme is in place.

Modelling studies assessing and comparing more than one supplemental imaging technique

Hill et al. (2023)[28] evaluated risk stratified breast cancer screening (RSBCS) strategies that differed in frequency and imaging modality by level of risk. The authors included eight RSBCS developed by three independent research groups (Breast Screening Risk Adaptive Imaging for Density -BRAID;[39] Adapting Breast Cancer Screening Strategy Using Personalised Risk Estimation -ASSURE;[40] and Predicting Risk of Cancer at Screening -PROCAS.[41] Three of the RSBCS strategies only considered alternative frequency for mammography (ASSURE 1, ASSURE 2, PROCAS), and another one incorporated various frequency for mammography for women with non-dense breast together with supplementary imaging for women BI-RADS C&D (ASSURE 4), and are therefore excluded for this review. The four strategies considered in this review are:

- BRAID 1: triennial mammography for women with BI-RADS A&B and additional 18-monthly imaging with ABUS for women with BI-RADS C&D
- BRAID 2: triennial mammography for women with BI-RADS A&B and additional 18-monthly imaging with CEM for women with BI-RADS C&D
- BRAID 3: triennial mammography for women with BI-RADS A&B and additional 18-monthly imaging with MRI for women with BI-RADS C&D
- ASSURE 3: triennial mammography for women with Volumetric Breast Density (VBD) 1&2, triennial mammography with HHUS for women with VBD 3&4, and MRI for women with VBD 3&4 and 10-year breast cancer risk >8%

Screening for all these strategies started at 50 years and continued until age 70 with the last invitation for screening.

Ontario Health (2023)[25] assessed the supplemental screening with HHUS, MRI, or DBT as an adjunct to mammography, compared to mammography alone. The authors simulate a cohort that represented age and sex distributions and all-cause mortality for the Ontario population, Canada, born between 1949 and 1973 (aged 50 to 74 years in 2023). Screening strategies for each modality (U/S, MRI, or DBT) were applied to two groups: individuals with dense breasts (BI-RADS C&D) and those with only extremely dense breasts (BI-RADS D). For individuals with BI-RADS D, the screening frequency was adjusted to annual screening, in line with Ontario Breast Screening Programme (OBSP) recommendations for those with breast density of 75% or higher. The remaining screen-eligible population continued with biennial mammography as per OBSP guidelines for average-risk individuals. The model also incorporated modified participation and

retention rates (64.81%) from the Cancer System Quality Index (CSQI) 2020 Ontario Cancer System Performance report.[25]

The study by Blankenburg et al. 2022,[17] included asymptomatic women aged 40-74 years with BI-RADS C&D who were invited for annual screening. Participants were categorised into two subpopulations based on the American College of Radiology (ACR) breast cancer risk classification:[42] average risk (women with unknown personal or family history of breast cancer or a lifetime risk of less than 15%) and intermediate risk (women with a personal history, first-degree family history, or a lifetime risk of 15-20%).[17] The authors evaluated multiple supplemental imaging modalities including full-protocol MRI (Fp-MRI), Ab-MRI, CEM, and HHUS after a negative XM or DBT (XM being relevant for this review). The study also examined the capacity for additional scans required for Fp-MRI and CEM, suggesting that one additional Fp-MRI scan per day per existing general scanner could accommodate the demand in intermediate-risk subpopulations.

Characteristics of the economic analyses

Table 3 reports further characteristics of the economic analyses. Eight modelling studies reported the perspectives for the analysis:[17-20, 23-25, 27] four studies included costs from the health-care perspective of their respective countries,[17, 20, 23, 24] three from the federal payer perspective,[18, 19, 25] and one from the UK National Health Service perspective (Table 3).[27] Four studies did not clearly state the perspective for the analysis;[21, 22, 26, 28] however, these are inferred to be from a healthcare provider or payer perspective based on the cost categories included and the source of unit costs (i.e., national costing source for The Netherlands;[21, 22] UK NHS,[28] and USA Medicare.)[26]

Ten studies used a lifetime horizon (Table 3).[17-22, 24, 25, 27, 28] One study each utilised a 20-year,[23] and one a 30-year time horizon.[26] The four studies that used Markov Models specified using an annual Markov cycle length. [17, 23, 24, 26] All the modelling studies discounted future costs and QALYs or life years, albeit using different discount rates. Discounting is the common practice of weighting future costs and outcomes less heavily compared to those that occur in the present; the higher the discount rate the less weight given to future costs and outcomes (e.g., costs and QALYs).

Table 3 Characteristics of the economic analyses

Study (first author and year)	Perspective	Time horizon	Cycle length	Discounting	Benefit measures/ Outcome measures	Currency, year	Threshold
Cost-utility analysis							
Couto, 2024[24]	Brazilian supplementary health system perspective	Lifetime	Annual	5%	Mean cost, incremental costs, LYG, QALYs, ICER/LYG, ICER/QALYs	Euro, 2023	€72,00
Hill, 2023[28]	UK National Health Service	Lifetime	N/A	3.5%	Mean total costs, mean LYs, QALYs, Net Health Benefit, mean number of screens, number diagnosed (DCIS, Invasive cancers & stage)	GBP, 2019/20	£20,000
Ontario Health, 2023[25]	The Ontario Ministry of Health	Lifetime	N/A	1.5%	LYs, QALYS, total and disaggregated costs (e.g., screening, diagnostic imaging and assessment, and breast cancer management)	Canadian dollars, 2022	$50,000
Blankenburg, 2022[17]	Health-care perspective	Lifetime	Annual	3%	False negative/positives cost, total costs, QALYs, LYs, ICER/QALYs	USD, 2021	$100,000
Geuzinge, 2021[20]	Health-care perspective	Lifetime	N/A	3%	LYs, QALYs, total costs, ICER/QALY	Euro, 2018	€22,000 (£20,000)
Kaiser, 2021[26]	Healthcare system	30 years	Annual	3%	QALYs, total costs, ICER/QALY	USD, 2015	$100,000
Tollens, 2021[23]	U.S. healthcare system	20 years	Annual	3%	Cumulative costs and QALYs, incremental costs and QALYs, ICER/QALY	USD, NR	$100,000
Gray, 2017[27]	National health care service	Lifetime	N/A	3.50%	Total costs, LYs, QALYs, ICER/QALY	GBP, 2014	£20,000
Lee, 2015[18]	Federal payer perspective	Lifetime	N/A	3%	Total costs, QALYs, ICER/LYG, ICER/QALY	USD, 2013	$100,000
Sprague, 2015[19]	Federal payer perspective	Lifetime	N/A	3%	LYs, LYG, QALYs, QALY gained, incremental costs, cost/QALY, ICER/QALY	USD, 2013	$100,000
Cost-effectiveness analysis							
Wang, 2022[21]	Healthcare provider or payer	Lifetime	N/A	3%	LYG, ACER/LYG, ICER/LYG	Euro, 2019	€ 20,000
Wang, 2020[22]	Healthcare provider or payer	Lifetime	N/A	3%	LYG, ICER/LYG	Euro, NR	€ 20,000

Abbreviations: LYs: life years, QALYs: quality adjusted life years, ICERs: incremental cost-effectiveness ratios, LYG: life year gain, NR: not reported, N/A: not applicable, USD: United States dollar, NOK: Norwegian Kroner, GBP: Pound sterling, ACERs: average cost-effectiveness ratios

Quality of the included studies

The reporting standard for the 12 modelling economic evaluations was assessed using the questions from the Philips checklist for good practice in decision-analytic modelling in health technology assessment.[16] The checklist questions were answer as 'Yes', 'No', 'Partially, Unclear', or 'Not Applicable' (Appendix 3). The quality of reporting is variable for the 12 studies with the proportion of positive answers to questions ranging from 40% to 89%. Gray 2017[27] and Couto 2024[24] obtained the higher proportion of questions answered as 'Yes'. To note, only four studies reported their models being calibrated against independent data[17-20] with three of these using pre-existing models.[18-20]

Further considerations of the quality and appropriateness of the economic evaluation models

The 12 included studies used a variety of models with substantial differences, which influence the interpretation of the results. Particularly, it is important to consider whether adjustments were made for:

- mammography sensitivity due to density
- density level due to age
- utility decrements due to positive screening test results
- utility weight used for individuals with no cancer

All models recognise that the sensitivity of mammography is reduced when used to test women with dense breasts. Whilst some models assumed a unique value for sensitivity throughout the model run,[17, 23, 26] others have modelled sensitivity as a function of density, age and screening interval,[19] or as a function of density and tumour size.[21, 27, 28] Individuals in the microsimulation by Geuzinge et al. (2021)[20] may have their level of density reduced at age 55 and 65, and in Hill et al. (2023)[28] density declines annually based on previous year density and age, but cancer risk increases based on age. Of note, the mammography sensitivity estimates used in Kaiser et al. (2021)[26] and Tollens et al. (2021)[23] (41.2%) and in Blankenburg et al. (2023)[17] (30%) are substantially lower than those used in other studies (e.g., 61% in Ontario Health 2023).[25]

Geuzinge et al. (2021),[20] Lee et al. (2015)[18] and Couto (2024)[24] applied utility decrements for individuals obtaining a positive test result. However, the decrement used varied substantially (i.e., 0.105 for 5 weeks[18, 20] or 0.25 for 1 year).[24] Couto et al. (2024) applied 0.25 quality of life reduction across false positives, false negatives and true positive results.[24]

Most studies used the general population utility weights for individuals with no cancer. However, Kaiser et al. (2021)[26] and Tollens et al. (2021)[23] assumed utility weight of 1, 0.99 and 0.95 for individuals with no cancer, after successful non-invasive and invasive cancer treatments, respectively. Similar assumptions were used in Blankenburg et al. (2023) using 0.99 utility weight for no cancer, and 0.98 for undetected breast cancer.[17] The common practice in economic evaluation studies is to use the general population utility weights for individuals without disease and adjust these utilities according to the data obtained from the literature for the event of interest. Assumptions of perfect health for individuals with no disease, and nearly perfect health for those recovering from treatment can substantially bias cost effectiveness results.

Further differences include the screening participation rate (e.g., 100% assumed Blankenburg et al. (2023)[17] and Sprague (2015)[19] but 65% in Ontario Health (2023)[25] and 80% in Wang et al. (2020)[22] and Wang et al. (2022)[21] and the adverse effect of radiation (induced tumours) that was considered only by Wang et al. (2020)[22] and Wang et al. (2022).[21]

Finally, the models used by Geuzinge et al. (2021),[20] Ontario Health (2023),[25] Wang et al. (2020),[22] Wang et al. (2022),[21] Lee et al. (2015)[18] and Sprague et al. (2015)[19] have been extensively validated. Hill et al. (2023)[28] reports validation against UK NBCP data. Gray et al. (2017) reports that no external validation was conducted of their early economic model‡.[27]

Economic evaluation results

For all included studies, we present the total costs, QALYs, LYs, incremental costs and QALYs, LYG, and ICERs, when these were reported. Costs and ICERs are reported in the same currency and year as in the original studies. We have not converted data to British pounds to avoid making assumptions about the comparability of healthcare systems across different countries. Where an ICER threshold is provided, it is shown in Table 4. In the absence of specific thresholds, we assumed standards that are country specific.

Due to inconsistencies in the reporting of results across the included studies, we have summarised the base-case results in Table 4 to Table 8 and grouped them by intervention whenever possible. For studies reporting results for several strategies, results relevant to this

‡ The model reported in Gray et al. (2017) was further developed and validate in Wright et al. (2024)[44]

review (e.g., supplemental imaging for women with dense breast tissue) and/or closer to the UK NBSP (e.g., three-year screening) were selected.

Three studies considered more than one imaging technology.[17, 25, 28] Ontario Health (2023) produced three separate analyses for MRI, DBT and U/S, and results are reported within the corresponding imaging subsection.[25] Blankenburg et al. (2021) produced a comparative analysis between Fp-MRI, Ab-MRI, DBT, U/S and CEM[17] and Hill et al. (2023) compared Ab-MRI, automatic U/S, and CEM.[28] Both studies results are reported in the section, *Studies assessing multiple screening imaging technologies*.

MRI

Geuzinge et al. (2021), reported the results for several screening strategies that varied the frequency of mammography and/or MRI offered to women with extremely dense breast tissue (Table 4).[20] The strategy representing the current screening in The Netherlands is biennial mammography (2Mx) and the results are reported for 1,000 screened individuals. Biennial screening with mammography resulted in the lowest total costs (€10,681,842) but also the lowest QALYs (49,520) for all the screening strategies. Adding MRI after a negative mammography every 2 years (2Mx_2MRI) resulted in the highest cost (€11,943,649) but did not produce the highest number of QALYs. Moreover, the strategies containing mammography together with MRI were either dominated (a less costly strategy generated more QALYs) or weakly dominated (a combination of two alternative strategies was less costly and generated more QALYs). All the strategies considering MRI as a replacement to mammography were on the cost-effectiveness frontier. However, the authors state that screening women with extremely dense breasts with MRI every four years (4MRI) had the highest acceptable ICER, when applying the National Institute for Health and Care Excellence (NICE) cost-effectiveness threshold (i.e., £20,000, circa €24,000).

Kaiser et al. (2021) assessed the cost-effectiveness of MR-Mammography against mammography (XM) for biennial screening in women at intermediate risk for breast cancer due to elevated breast density (extremely dense breast tissue).[26] Similar to Geuzinge et al. (2021),[20] Kaiser et al. (2021) used MRI test accuracy data from the DENSE study conducted in The Netherlands. The study results show screening women with extremely dense breast tissue with MRM generated more QALYs (i.e., 18.92 for MRM and 18.87 for XM) but at higher costs (i.e., $5,877 for MRM and $5,493 for XM), and an ICER of $8,798 per QALY gained (Table 4). This ICER is well below the $100,000 cost-effectiveness threshold used in the USA. Tollens et al. (2021)

updated the analysis conducted by Kaiser et al. (2021) using data for the first and second rounds for the DENSE study.[23] Similarly to Kaiser et al. (2021), Tollens et al. (2021) found that MRM resulted in more QALYs than XM but also incurred higher costs.[23] Over a time horizon of 20-years (10-year shorter than Kaiser et al. (2021)), the MRM strategy produced average cumulative costs of $6,081 per woman and 15.12 QALYs, compared to XM, which costs $5,810 and resulted in 15.10 QALYs. This led to an ICER of $13,493 per QALY gained also below the USA threshold of $100,000 per QALY gained. The specificity of MRM was set to 92% for the first screening round, and when the specificity for subsequent rounds was varied from 92% to 99%, the ICER ranged from $38,849 to $5,062 per QALY gained (data not shown). Despite the higher costs, the study concluded that biennial supplemental MRI screening was cost-effective.

In the study by Ontario Health (2023), supplemental MRI screening as an adjunct to mammography was compared with mammography alone.[25] For women with heterogeneously or extremely dense breasts, the total healthcare cost for supplemental MRI screening was € 3,620 per screened person compared with €3,367 for mammography alone (Table 4). The total discounted QALYs and LYs per person for supplemental MRI screening were 16.8766 and 22.341, respectively, compared with 16.8758 and 22.334 for mammography alone. This resulted in a slight increase of 0.0008 QALYs per screened person (0.007LYGs) and an ICER of $314,470 per QALY gained. For individuals with extremely dense breasts, the average total cost for supplemental MRI screening was €3,435 per screened person and €3,367 for mammography alone. The total discounted QALYs and LYs per person for supplemental MRI screening were 16.8764 and 22.336, respectively, compared to 16.8758 and 22.334 for mammography alone. For this group, the average increase in LYs was 0.002 years per person, and the average increase in QALYs was 0.0007 per screened person, with an ICER of $101,813 per QALY gained. The ICERs for screening women with heterogeneously and/or extremely dense breast with supplementary MRI are above the authors' stated cost-effectiveness threshold of $50,000 Canadian dollars and therefore are not cost-effective.

Table 4. Summary results of the included economic evaluation studies: Magnetic Resonance Imaging (MRI).

Study (first author, year)	Total costs		Total QALYs	Total LYs	Incremental costs	Incremental QALYs/Lys	Base case ICER/QALYs/LYs
Geuzinge, 2021[20]	2Mx	€10681842	49520	57870	NR	NR	**ICER/QALYs**: Weakly dominated
	5MRI	€11110699	49560	57913			€12410
	4MRI	€11246367	49569	57921			€15620
	2Mx/MRI	€11330895	49566	57919			Strongly dominated
	3MRI	€11411722	49573	57926			€37181
	2Mx_4MRI	€11431163	49565	57918			Strongly dominated
	6Mx_2MRI	€11763234	49577	57929			Weakly dominated
	2MRI	€11805633	49581	57933			€46971
	4Mx_2MRI	€11903811	49581	57933			Strongly dominated
	2Mx_2MRI	€11943649	49576	57929			Strongly dominated
	Results for 1000 women from age 25 until death (MISCAN model). Screening from 50 to 75 years. Strategy notation (examples): 2Mx = mammography every 2 years; 5MRI = MRI every 5 years; 2Mx/MRI = alternate Mammography or MRI every 2 years; 2Mx_4MRI = mammography every 2 years and MRI every 4 years (for the years with both modalities, the model assumes mammography first and MRI conducted 1 month after, allowing for the cancelation of MRI due to drop in density).						
Kaiser, 2021[26]	MRM	$5,877	18.92	NR	NR	NR	**ICER/QALYs**: $8,798
	XM	$5,493	18.87				
Ontario Health, 2023[25]	***People with dense breast; total cohort (per person screened)***						
	Mammography + supplemental MRI	$ 3,620	16.8766	22.341	$252	**QALYs (LYs)** 0.0008 (0.007)	**ICER/QALYs:** $314,170
	Mammography alone:	$3,367	16.8758	22.334			
	People with extremely dense breast (per person screened)						
	Mammography + supplemental MRI	$3,435	16.8764	22.336	$67	**QALYs (LYs)** 0.0007 (0.002)	**ICER/QALYs:** $101,813
	Mammography alone:	$3,367	16.8758	22.334			
Tollens,2021[23]	MRM	$6,081	15.120	NR	$271	0.020	**ICER/QALYs:** $13,493
	XM	$5,810	15.099				

Study (first author, year)	Total costs	Total QALYs	Total LYs	Incremental costs	Incremental QALYs/Lys	Base case ICER/QALYs/LYs
Wang, 2022[21]	NR	NR	*Biennial screening*			
			LYG: A: 132	NR	NR	**ICER/LYG:** €18,201
			B: 145			Extended dominance
			C: 149			Extended dominance
			D: 501			€24,700
			E: 554			Extended dominance
			F: 562			€58,700
			Alternative scenarios			
			Triennial screening			
			A: 91	NR	NR	€15,500
			B: 95			Extended dominance
			C: 98			Extended dominance
			D: 334			€23,000
			E: 353			Extended dominance
			F: 362			Extended dominance
			Quadrennial screening			
			A: 52	NR	NR	€14,700
			B: 59			Extended dominance
			C: 60			Extended dominance
			D: 158			Extended dominance
			E: 204			Extended dominance
			F: 205			Extended dominance
	Screening for women 50y-74y. Ab-MRI for: A) BI-RADS D 50y-65y; B) BI-RADS D 50y-69y; C) BI-RADS D 50y-74y; D) BI-RADS C/D 50y-65y; E) BI-RADS C/D50y-69y; F) BI-RADS C/D 50-74; in all strategies women go to mammography if not in the group going to Ab-MRI. LY per 10,000 women screened. Comparator: Mammography for women 50y-74y.					

Abbreviations: QALYs: quality-adjusted life years, LYs: Life years, LYG: life year gain ICER: incremental cost-effectiveness ratio, NR: not reported, MRM: MR-Mammography, XM: x-ray mammography

The cost-effectiveness analysis by Wang et al. (2022) explored replacing biennial mammography with Ab-MRI for the Dutch breast screening programme.[21] The six evaluated strategies use Ab-MRI for three alternative age groups (50 to 65; 50 to 69; or 50 to 74) for women with either heterogeneously and extremely dense breast or extremely dense breast only. Those women not eligible for biennial Ab-MRI continue biennial mammography and the reference strategy was biennial mammography for women between 50 and 74 years. Strategies A, B and C in Table 4 are those allowing Ab-MRI for women with extremely dense breast (BI-RADS D) and differ by the upper age limit for screening with Ab-MRI (i.e., A=65, B=69 or C=74). Similarly, strategies D, E, and F apply Ab-MRI for both heterogeneously and extremely dense breasts, differing on the upper age limit for screening with Ab-MRI (i.e., D=65, E=69 or F=74). The reported ICERs show Ab-MRI biennial screening strategies B, C and E as extended/weekly dominated. That is, a combination of two alternative strategies will be less costly and generate more LYs. The ICER for strategy A (Ab-MRI for 50-65 years; extremely dense) was €18,201 per LYG, and €24,700 and €58,700 for strategies D (Ab-MRI for 50-65; heterogeneous & extremely dense) and F (Ab-MRI for 50-74; heterogeneous & extremely dense), respectively. The results for triennial screening scenarios are also reported in Table 4 showing strategies A and D as the only non-dominated strategies with ICERs of €15,500 and €23,000, respectively. The authors also ran quadrennial screening scenarios where strategy A was the only non-dominated strategy (ICER €14,700). The authors concluded that at a threshold of €20,000 per LYG, the optimal strategy was identified as biennial Ab-MRI from ages 50 to 65, followed by mammography only for ages 66 to 74 years (strategy A) for women with extremely dense breasts. The authors' probabilistic analysis showed this strategy having 79% probability of being cost-effective at the €20,000 per LYG threshold.

DBT

Results for Couto et al. (2024),[24] Lee et al. (2015),[18] Ontario Health (2023)[25] and Wang et al. (2020)[22] for the evaluation of DBT are reported in Table 5 and summarised below.

Couto et al. (2024) found that biennial breast screening with DBT combined with synthesised 2D mammography (DBT + s2D) for women with scattered areas of fibroglandular breast density and heterogeneous dense breasts (BI-RADS B & C) was both more effective and less costly compared to DM alone for the Brazilian health care system.[24] DBT + s2D resulted in 18.92 QALYs (16.49 LYs), compared to 13.72 QALYs (16.47 LYs) for DM, offering an additional 5.2 QALYs (0.016 LYs). The total costs were €2,095 for DBT + s2D and €3,049 for DM alone, saving €954. These results are explained by the higher cancer detection rate (relative risks -RR- 1.35

and 1.48 for invasive cancers), the lower recall rate (R 0.81), and the lower biopsy rate (RR 0.89) for DBT + s2D with respect to DM alone, that the authors obtained from their own systematic review and meta-analysis of 18 studies. The authors probabilistic analysis shows DBT + s2D as the dominant strategy with lower costs and higher QALYs for all 1000 iterations.

Lee et al. (2015) evaluated the addition of DBT to DM for biennial screening for women with dense breast tissue (heterogeneously and extremely dense), compared with DM alone.[18] The discounted cost for DBT+DM was $4,440 and $4,091 for DM alone (Table 5). DBT+DM produced 16.814 QALYs (20.652 LYs) and DM alone 16.807 QALYs (20.647 LYs); and an ICER of $53,893 per QALY gained ($70,500 per LY gained). The authors did not produce a probabilistic sensitivity analysis but ran a series of one-way sensitivity analyses showing ICERs above the commonly used $100,000 threshold value used for USA for extreme cases such those where limited test accuracy improvement was used for DBT+DM. The authors concluded that biennial combined screening for women aged 50-74 years with dense breasts is cost-effective from the federal payer perspective.

A study from the Ontario Health Technology Assessment Programme compared supplemental DBT screening as an adjunct to mammography with mammography alone, evaluating two different populations: women with dense breasts (which included both heterogeneously dense breasts and extremely dense breasts) and only those with extremely dense breasts.[25] For women with dense breasts, the total discounted cost per person for supplemental DBT screening was €3,462, compared with €3,367 for mammography alone (Table 5). The total discounted QALYs and LYs per person for supplemental DBT screening were 16.8762 and 22.336, respectively, compared to 16.8758 and 22.334 for mammography alone. This led to a slight increase in QALYs (0.0004) and LYs (0.002), with a high ICER of $212,707 per QALY gained. For individuals with extremely dense breasts, the total cost per person for supplemental DBT screening €3,393, and €3,367 for mammography alone. The total discounted QALYs and LYs per person for supplemental DBT screening were 16.8760 and 22.3344, respectively, compared to 16.8758 and 22.334 mammography alone; and increment of 0.0002 QALYs (0.0006 LYs) per person, resulting on an ICER of $142,730 per QALY gained. These ICERs are well above the $50,000 cost-effectiveness threshold the authors state is commonly used for decision making in Canada.

Table 5. Summary results of the included economic evaluation studies: Digital Breast Tomosynthesis (DBT)

Study (first author, year)	Total costs		Total QALYs	Total LYs	Incremental costs	Incremental QALYs/LYs	Base case ICER/QALYs/LYs
Couto, 2024[24]	DBT + s2D	€2094.54	18.9185	16.4855	€- 954.02	**QALYs (LYG)** 5.1989 (0.0160)	**ICER/QALYs (LYG):** Dominant (Dominant)
	DM alone	€3048.57	13.7196	16.4695			
Lee, 2015[18]	DM + DBT	$4,440	16.814	20.652	NR	NR	**ICER/QALYs (LYG):** $53,893 ($70,500)
	DM alone	$4,091	16.807	20.647			
Ontario Health, 2023[25]	***People with dense breast; total cohort (per person screened)***						
	Mammography + supplemental DBT	$3,462	16.8762	22.336	$95	**QALYs (LYG)** 0.0004 (0.002)	**ICER/QALYs:** $212,707
	Mammography alone	$3,367	16.8758	22.334			
	People with extremely dense breast (per person screened)						
	Mammography + supplemental DBT	$3,393	16.8760	22.3344	$25	**QALYs (LY)** 0.0002 (0.0006)	**ICER/QALYs:** $142,730
	Mammography alone	$3,367	16.8758	22.334			
Wang, 2020[22]	NR		NR	**LYG** At DBT sensitivity	NR	NR	**ICER/LYG** At unit cost for DBT €96 / €80
				75% 5			€180,265 / €91,076
				80% 19			€41,021 / €20,768
				85% 30			€24,407 / €12,390
				90% 42			€17,254 / €8,779
				95% 54			€13,228 / €6,749
				100% 65			€11,034 / €5,639
	Results for LYG per 10,000 women screened. ICERs for study *Scenario 1 vs. Reference*. *Scenario 1:* DBT screening only for women with dense breasts; Reference: maintaining DM screening. Discounted at 3% cost and Lys.						

Abbreviations: QALYs: quality-adjusted life years, LYs: Life years, LYG: life year gain ICER: incremental cost-effectiveness ratio, DBT: digital breast tomosynthesis, s2D: two-dimensional mammograms, DM: digital mammography, NR: not reported

Wang et al. (2020) evaluated biennial screening in different breast density populations of 10,000 through three scenarios, two of these are relevant for this review: scenario 1) DBT for women with dense breasts and DM for those with fatty breasts, and a reference scenario maintaining DM screening.[22] The authors report their result using two unit costs for DBT of €96 and €80 per screen (1.5 and 1.25 times the cost for DM, respectively). The authors stated that these estimates account for the higher equipment costs, additional digital storage capacity, more expensive reading stations, and the longer reading time required for DBT. The cost of mammography in the Dutch national screening programme was calculated by dividing the whole programme cost by the number of participants, resulting in €64 per screen. Moreover, the results were presented for various sensitivity levels of DBT, ranging from 65% to 100%; however, DBT was dominated (more costly and less effective) for DBT sensitivities below 72%. The authors found that DBT offered increased LYGs claiming DBT was cost-effective when its sensitivity was at least 75%. Using a cost-effectiveness threshold of €20,000 per LY gained, the authors found that DBT was cost-effective at sensitivities of 86% and 80% for unit cost for DBT of €96 and €80, respectively (Table 5).

Whilst the analyses show improved clinical outcomes such as a higher cancer detection rate from DBT, these studies show mixed cost-effectiveness results: from DBT dominating DM for Couto et al. (2024),[24] being cost-effective for Lee et al. (2015)[18] and Wang et al. (2020)[22] from US and Dutch payer perspectives, to ICERs that are well above the threshold used in Canada ($50,000 Canadian dollars) for Ontario Health (2023).[25]

Supplemental U/S

The UK based study included four strategies (Table 6), two of these involving supplemental screening offered to women with high breast density (heterogeneous or extremely dense), and compared these with the current UK National Breast Screening Programme, which uses mammography every three years for women between 50 and 70 years old.[27] The "*Masking"* strategy involved triennial screening with supplemental U/S screening for women with high breast density and supplemental MRI for women with both high breast density and 10-year risk of breast cancer >8%. *Risk 1 & masking* strategy involved the same supplemental screening modalities but with screening frequency dependent on the 10-year risk of breast cancer: a) triennial for <3.5%, b) biennial for risk between 3.5%–8%, and 3) annually for risk >8%. Gray et al. (2017) analysis showed that the *Masking* strategy resulted in higher mean costs (£809 versus £654 for mammography) and more QALYs (17.7102 versus 17.7095 for mammography) with an ICER of £212,947 against the current UK NBSP. The *Risk 1 & masking* strategy also resulted in

higher cost (£870) and QALYs (17.124) with an ICER of £75,254. The analysis used the usual 3.5% discount rate for costs and QALYs. When a lower discount rate of 1.5% was used for QALYs, the ICERs for reduced to £105,412 for *Masking* and £33,199 for *Risk & masking*. All these ICERs are above the usual cost-effectiveness threshold used for decision making in the UK (e.g., £20,000 per QALY gained).

In the study by Ontario Health (2023), supplemental HHUS screening as an adjunct to mammography was compared with mammography alone.[25] For women with dense breasts, the total average cost for supplemental HHUS screening was €3,451 compared with €3,368 for mammography alone. The total discounted QALYs and LYs per person for supplemental HHUS screening were 16.877 and 22.338, respectively, compared with 16.876 and 22.334 for mammography alone. This resulted in a slight increase of 0.0007 QALYs (0.004 LYGs), with a high ICER of $119,943 per QALY gained. For individuals with extremely dense breasts, the total average cost for supplemental HHUS screening was €3,390, compared with €3,368 for mammography alone. The total discounted average QALYs and LYs for supplemental HHUS screening were 16.8761 and 22.335, respectively, compared with 16.8759 and 22.334 for mammography alone. For this group, the increase in QALYs was 0.0003 (0.001 LYG), with an ICER of $83,529 per QALY gained. The authors used a cost-effectiveness threshold of Canadian $50,000 at which HHUS was not cost-effective.

Sprague et al. (2015) assessed the use of annual supplemental U/S after a negative screening mammogram for women with dense breasts (heterogeneously and extremely dense breasts) and extremely dense breasts only.[19] The study used three validated simulation models (models E, W, and G-E), and presented results per 1,000 women. The authors also report median results and range for the three models, and these are the results reported in Table 6.

The study estimated the total costs and QALYs for mammography plus U/S: $3.08 million and 19,059.9 for women with extremely dense breasts, and $3.39 million and 19,060.8 for women with heterogeneously or extremely dense breasts, compared with $3.02 million and 19,059.8 for mammography alone. The study found high ICERs of $246,000 per QALY gained (range $74,000 - $535,000) for women with extremely dense breasts and $325,000 per QALY gained (range $112,000 - $766,000) for women with heterogeneously or extremely dense breasts. These ICERs exceed the commonly accepted threshold of $100,000 per QALY in the USA, suggesting that supplemental U/S is not cost-effective in these populations. The study concluded that while supplemental U/S screening for women with dense breasts provides limited health benefits, it significantly increases costs.

Table 6. Summary results of the included economic evaluation studies. Supplementary U/S vs. Mammography.

Study (first author, year)	Total costs		Total QALYs	Total LYs	Incremental costs	Incremental QALYs/LYs	Base case ICER/QALYs	
Gray, 2017[27]	Current UK NBSP	£654	17.7095	NR	NR	NR	vs. UK NBSP (3.5% DR)	vs. UK NBSP (1.5% health, 3.5% costs discount)
	Risk 1	£694	17.7119				£16,689	£11,565
	Risk 2	£858	17.7181				£23,924	£11,592
	Masking	£809	17.7102				£212,947	105,412
	Risk 1 and masking	£870	17.7124				£75,254	£33,199
	Screening frequency defined by risk of cancer and/or breast density. *Risk 1*, three strata defined by 10-year risk of breast cancer: a) <3.5% (3-yearly), b) 3.5%–8% (2-yearly), and 3) >8% (annually); *Risk 2*, screening population divided into thirds (tertiles): a) the lowest risk tertile (3-yearly), b) the middle tertile (2-yearly), and c) the highest risk tertile (annually); *Masking*, current screening approach (triennial) with supplemental U/S offered to women with high dense breast (heterogeneous or extremely dense) and MRI offered to women with high dense breast and 10-year risk of cancer >8%.; *Risk 1 & masking* used the stratification for Risk 1 together with the strategy described for the Masking approach.							
Ontario Health, 2023[25]	***People with dense breast; total cohort***							
	Mammography + supplemental HHUS	$3,451	16.877	22.338	$83	QALYs (LYs) 0.0007 (0.004)	$119,943	
	Mammography alone	$3,368	16.876	22.334				
	People with extremely dense breast							
	Mammography + supplemental HHUS	$3,368	16.8761	22.335	$22	QALYs (LYs) 0.0003 (0.001)	$83,529	
	Mammography alone	$3,390	16.8759	22.334				

Study (first author, year)	Total costs		Total QALYs	Total LYs	Incremental costs	Incremental QALYs/LYs	Base case ICER/QALYs	
Sprague, 2015[19]	*Median (range) of outcomes per 1000 women across the three simulation models. Biennial Screening 50-74. (All outcomes computed from age 40 until death. Life years, QALYs and total costs ($ millions)*							
	Mammography	3.02 (2.87-3.05)	19,059.8 (18,970.4-19,405.4)	23,108.5 (22,981.0-23,548.7)	-	-	(ICER/QALYs with respect to Mammography)	(ICER/QALYs for Mammography + U/S heterogeneously or extremely DB versus Mammography + U/S for extremely DB)
	Mammography + U/S for extremely DB	3.08 (2.91-3.08)	19,059.9 (18,970.9-19,405.5)	23,108.7 (22,981.6-23,548.9	$287,000 ($271k - $411k)	1.1 (0.8, 3.9)	$246,000 ($74k-$535k)	$338,000 ($121k-$562k)
	Mammography + U/S for heterogeneously or extremely DB	3.39 (3.20-3.42)	19,060.8 (18,873.3-19,405.9)	23,109.8 (22,984.4-23,549.4)	$560,000 ($529k-$652k)	1.7 (0.9, 4.7)	$325,000 ($112-$766k)	

Abbreviations: QALYs: quality-adjusted life years, LYs: Life years, LYG: life year gain ICER: incremental cost-effectiveness ratio, UK NBSP: National breast screening programme, NR: not reported, HHUS: Hand-held ultrasound, U/S: ultrasound, DB: dense breasts

Studies assessing multiple screening imaging technologies

The study by Blankenburg et al. (2022) (Table 7) assessed multiple supplemental imaging modalities added after a negative result to annual mammography for women with dense breasts (heterogeneously or extremely dense) for two subpopulations: average risk (women with unknown personal or family history of breast cancer, or less than 15% lifetime risk of breast cancer) and intermediate risk (women with personal history or first-degree family history, or who have 15-20% lifetime risk of breast cancer).[17] The included imaging technologies were U/S, Ab-MRI, Fp-MRI and CEM and results are reported for 1,000 women screened (Table 7). For both population risk subgroups all supplemental imaging modalities produce more QALYs but cost more than mammography only. ICERs with respect to mammography only for all modalities in both risk subgroups are below the $100,000 threshold used for the USA. However, the ICERs with respect to the next less costly strategy (calculated from the reported data) are lower than the $100,000 threshold for U/S and Ab-MRI but show CEM and Fp-MRI being dominated by Ab-MRI.

Fp-MRI produces the same LYs and QALYs as Ab-MRI but is more costly, and CEM is more costly and produces fewer QALYs than Ab-MRI.

The UK study by Hill et al. (2023) evaluated additional imaging to triennial screening with mammography in four risk-stratified strategies that are relevant for this review: BRAID 1 to 3[**] considered additional 18-monthly ABUS, CEM and Ab-MRI, respectively, for women with heterogeneously or extremely dense breasts (VBD3&4), and ASSURE 3 triennial mammography with HHUS for women with VBD3&4, and MRI for women with VBD3&4 and 10-year breast cancer risk >8%.[28] All risk-stratified strategies produce more QALYs than triennial mammography (Table 8) but are also more costly. The total mean cost for mammography only was £1,537, and the total mean QALYs was 16.641. The next costly strategy is ASSURE 3 with a total mean cost of £1,743 and 16.648 QALYs, resulting in an ICER of £29,429. This strategy is extendedly dominated by BRAID 2 (CEM); that is, a combination of mammography only and BRAID 2 strategies can be less costly and produce more QALYs. BRAID 1 (ABUS) is the next costly strategy after BRAID 2, with an ICER of £14,400, below the usual cost-effectiveness threshold used in the UK (i.e., £20,000 per QALY gained). Finally, BRAID3 (Ab-MRI) is more

[**] Hill et al. (2023) Results Table 4 labelled BRAID strategies as BRAID 2, BRAID 3 and BRAID 4, corresponding to the authors' Table 1 BRAID 1, 2, and 3, respectively. (personal communication Dr. Harry Hill, February 2025)

costly than BRAID 1 (ABUS) but produces equal mean QALYs and therefore is dominated by BRAID 1.

Table 7. Summary results of the included economic evaluation studies. Various supplemental modalities (Blankenburg et al. 2022)

Study (first author, year)	Total costs		Total QALYs	Total LYs	Incremental costs	Incremental QALYs/LYs	Base case ICER/QALYs/LYs	
Blankenburg, 2022[17]	***Supplemental imaging modality-add-on to annual XM for the DB and average breast cancer risk subpopulation/1,000 screenings***							
	Mammography only	$12,536.702	21.916	22.213	NR	NR	With respect to Mammography only	With respect to less costly non-dominated
	U/S	$15,407.411	22.039	22.322			$23,394	$23,339
	Ab-MRI	$19,445.162	22.096	22.375			$38,423	$70,838
	CEM	$21,181.214	22.090	22.369			$49,824	Dominated
	Fp-MRI	£21,599.596	22.096	22.375			$50,476	Dominated
	Supplemental imaging modality-add-on to annual XM for the dense breast population and intermediate breast cancer risk subpopulation/1,000 screenings							
	Mammography only	$16,176.899	21.786	22.097	NR	NR	-	-
	U/S	$18,556.882	21.966	22.256			$13,241	$13,222
	Ab-MRI	$22,438.861	22.050	22.333			$23,772	$46,214
	CEM	$24,058.983	22.041	22.325			$31,009	Dominated
	Fp-MRI	$24,583.844	22.050	22.333			$31,960	Dominated

Abbreviations: QALYs: quality-adjusted life years, LYs: Life years, LYG: life year gain ICER: incremental cost-effectiveness ratio, XM: x-ray mammography, DB: dense breasts, U/S: ultrasound, Ab-MRI: abbreviated-protocol magnetic resonance imaging, CEM: contrast-enhanced mammography, Fp-MRI: full-protocol magnetic resonance imaging, NR: not reported

Table 8. Summary results of the included economic evaluation studies. Various supplemental modalities (Hill et al. 2023)

Study (first author, year)		Total costs	Total QALYs	Total LYs (undiscounted)	NHB	Incremental costs	Incremental QALYs/LYs	Base case ICER/QALYs (calculated from reported data)	
Hill, 2023[28]	***Supplemental imaging modality-add-on to DM***								
	Mammography only:	1537	16.641	35.963	16.564	NR	NR	With respect to Mammography only	With respect to less costly non-dominated
	ASSURE 3	1743	16.648	35.981	16.560			£29,429	£29,429
	BRAID 2	1804	16.671	36.022	16.581			£8,900	£2,652
	BRAID 1	1876	16.676	36.019	16.582			£9,686	£14,400
	BRAID 3	1972	16.676	36.024	16.577			£12,429	Dominated
	ASSURE 3: triennial mammography for women with VBD1&2, triennial mammography with HHUS for women with VBD3&4, and MRI for women with VBD3&4 and 10-year breast cancer risk >8%; BRAID 1: Triennial mammogram for BI-RADS A&B and additional ABUS 18-monthly for BI-RADS C&D; BRAID 2: Triennial mammogram for BI-RADS A&B and or additional CEM 18-monthly for BI-RADS C&D; BRAID 3: Triennial mammogram for BI-RADS A&B and additional MRI 18-monthly for BI-RADS C&D								

Abbreviations: QALYs: quality-adjusted life years, LYs: Life years, NHB: Net health benefit, ICER: incremental cost-effectiveness ratio, DM: digital mammography, NR: not reported

Discussion

Summary of the findings

This systematic review identified 12 economic evaluation studies with three studies assessing multiple imaging technologies. Overall, seven studies evaluated MRI, four DBT, five U/S, and two CEM.

The seven studies assessing MRI show mixed results with five studies concluding MRI was cost-effective. For those studies where MRI was cost-effective, two reported very low ICERs that could be partially explained by the high utility weights assumed for people without cancer and those recovering from treatment, as well as the relatively low sensitivity for mammography used in the model.[23, 26] These assumptions could overestimate the QALY difference between the MRI and mammography strategies. Two Dutch studies using preexisting validated models found that MRI and Ab-MRI were cost effective. One study concluded that MRI when substituting mammography was cost-effective every 4 years for women with BI-RADS D.[20] The other found that Ab-MRI was cost-effective every two years for women with BI-RADS C&D aged 50 to 65 followed by mammography thereafter.[21] One study found that Ab-MRI after a negative mammography was cost-effective for the USA;[17] however, the Dutch study by Geuzingue et al. (2021) found that all the strategies using MRI added to mammography were dominated by MRI only strategies.[20] Finally, one Canadian and one UK study showed very low QALY gains from MRI[25] and Ab-MRI,[28] respectively, concluding that this imaging technology is not cost-effective when added to mammography.

Three of the four studies evaluating DBT found this modality to be cost-effective. However, the Brazilian study by Couto et al. (2024) reports an average QALY difference of 5.2 between DBT and DM that is not consistent with the study reported LYG (0.016), or with the small QALY differences reported in any of the other studies assessing DBT.[24] The other two studies found biennial DBT screening for women with BI-RADS C&D to be cost-effective for the USA[18] and The Netherlands.[22] The Canadian study found that mammography screening supplemented by DBT every two years for women with BI-RADS C&D and annually for women with BI-RADS D was not cost-effective.[25]

Five studies considered supplemental U/S with only two finding U/S cost-effective: Blankenburg et al. (2022)[17] assessed supplementary U/S after negative mammography for the USA, and Hill et al. (2023)[28] evaluated automatic U/S between triennial mammography for the

UK. However, biennial or annual supplemental U/S screening after negative mammography was not cost-effective for the USA in the analysis conducted by Sprague et al. (2015) using three pre-existing and validated economic models.[19] In addition, Gray et al. (2017)[27] and Ontario Health (2023)[25] found that adding U/S to mammography was not cost-effective from the UK and Canadian healthcare system perspectives, respectively.

CEM was evaluated in two studies: as supplementary after a negative mammography[17] and in addition to mammography between triennial mammography screening.[28] Both studies found CEM to be cost-effective compared with mammography only, but the CEM strategy was dominated by strategies including other imaging technologies (i.e., U/S or Ab-MRI,[17] and automated U/S).[28]

In summary, findings are mixed or unfavourable for most modalities when used as an adjunct to mammography alone. MRI alone, as opposed to supplemental after a negative mammography result, is potentially cost-effective, particularly in the youngest women with the highest breast density, who have the greatest scope to benefit. However, this evidence comes from non-UK settings (e.g., USA). DBT results are also mixed, with the latest cost-utility analysis for Canada finding this technology not being cost-effective from the Canadian NHS perspective. Finally, automated U/S appears cost-effective in a recent UK study, and CEM was not cost-effective compared with strategies using alternative supplementary imaging modalities.

Quality of the economic models assessing supplementary imaging modalities

As part of this review, we were asked to answer the following question: *has the cost-effectiveness of supplemental breast cancer screening in women with dense breasts been explored in high-quality modelling studies?* Good quality pre-existing validated models have been used by Ontario Health (2023),[25] Geuzinge et al. (2021),[20] Lee et al. (2015),[18] Sprague et al. (2015) (three models),[19] Wang et al. (2020),[22] and Wang et al. (2022).[21] Two other studies developed de novo models stating that external[17] or internal[28] validation has been conducted.

Two good quality economic evaluation models assessing screening from the UK NHS perspective were identified: Hill et al. (2023)[28] and Gray et al. (2017).[27] The early model used in Gray et al. (2017)[27] has been updated to assess alternative risk stratification strategies using mammography only at various frequencies that are defined upon the level of risk.[43] The MANC-RISK-SCREEN model has been internally and externally validated.[44] Hill et al. (2023) also

assessed the cost-effectiveness of risk-stratified breast cancer screening.[28] However, these models' cost-effectiveness results differ substantially when assessing similar screening strategies against the current UK screening practice (mammography every three years). Both models assessed a strategy that included triennial mammography for women with BI-RADS A&B, triennial mammography with supplemental U/S for women with BI-RADS C&D, and triennial mammography with supplemental MRI for women with BI-RADS C&D and 10-year cancer risk >8%. Whilst Hill et al. (2023)[28] found this strategy cost-effective, the opposite result was obtained by Gray et al. (2017).[27] While the difference in total mean cost between the models can help to explain the opposite cost-effectiveness results, it is the difference in total QALYs that is of more concern. Notably, the total average QALYs for the current NHS NBCSP are 16.641 for Hill et al. (2023)[28] versus 17.7095 for Gray et al. (2017),[27] and the average QALYs generated by the evaluated risk stratified strategy are 16.648 and 17.7102 for Hill et al. (2023)[28] and Gray et al. (2017),[27] respectively. These differences result in a much higher QALY difference for Hill et al. (2023) compared with Gray et al. (2017) (0.007[28] versus 0.0007[27]), explaining the 7-fold difference in the ICER (£29,429[28] and £212,947[27]). A thorough inspection and understanding of these models and their differences is needed before they are used to inform policy decision making.

Resource capacity

Only one study estimated the capacity requirements based on the economic model results. The USA study concluded that while the existing MRI capacity could accommodate additional imaging needed to screen women with extremely dense breasts, a substantial investment would be needed to provide MRI screening for women with heterogeneously and extremely dense breasts.[17] A capacity analysis to assess potential investment needs is also required for informed decision making.

Conclusions

Evidence for the cost-effectiveness of supplemental breast cancer screening in women with dense breasts has been explored using good quality economic models but with variable or unfavourable findings. No study assessed the full range of plausible alternatives from the UK NHS perspective, varying age range, imaging modalities -alone or combined-, and frequency. There is scope for a UK-based cost-effectiveness study considering the range of viable approaches to enhance screening for women with dense breasts, taking into account the risk

status, the frequency of screening, age, and the accuracy of available modalities when used in addition to or as a replacement for mammography.

References

1. Cancer Research UK. Cancer statistics for the UK. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics-for-the-uk (Accessed January 2025).

2. Lauby-Secretan B, Loomis D, Straif K. Breast cancer screening: viewpoint of the IARC Working Group. N Engl J Med. 2015;**373**(15):1479.

3. Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. Lancet. 2012;**380**(9855):1778-86.

4. Duffy SW, Tabár L, Yen AM, et al. Mammography screening reduces rates of advanced and fatal breast cancers: results in 549,091 women. Cancer. 2020;**126**(13):2971-9.

5. Edmonds CE, O'Brien SR, Conant EF. Mammographic breast density: current assessment methods, clinical implications, and future directions. Semin Ultrasound CT MR. 2023;**44**(1):35-45.

6. American College of Radiology. ACR BI-RADS Atlas, Breast Imaging Reporting and Data System 2013. 2013. Available from: https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/Bi-Rads (Accessed January 2025).

7. Sprague BL, Kerlikowske K, Bowles EJA, et al. Trends in clinical breast density assessment from the Breast Cancer Surveillance Consortium. J Natl Cancer Inst. 2019;**111**(6):629-32.

8. Katanoda K, Ito Y, Sobue T. International comparison of trends in cancer mortality: Japan has fallen behind in screening-related cancers. Jpn J Clin Oncol. 2021;**51**(11):1680-6.

9. Patterson J, Stinton C, Alkhudairy L, et al. UK NSC Evidence Review. Additional Breast Screening with Ultrasound in Women with Dense Breasts Final 27 02 2019 2019. 2019. Available from: https://view-health-screening-recommendations.service.gov.uk/document/465/download. (Accessed January 2025)

10. Mann RM, Athanasiou A, Baltzer PAT, et al. Breast cancer screening in women with extremely dense breasts recommendations of the European Society of Breast Imaging (EUSOBI). Eur Radiol. 2022;**32**(6):4036-45.

11. Federal Register. Mammography Quality Standards. 2023. Available from: https://www.federalregister.gov/documents/2023/03/10/2023-04550/mammography-quality-standards-act (Accessed January 2025).

12. Clinical Trials.Gov. Breast Screening - Risk Adaptive Imaging for Density (BRAID) Protocol ID NCT04097366. 2019. Available from:

https://clinicaltrials.gov/study/NCT04097366?term=BRAID&checkSpell=false&rank=2#study-record-dates (Accessed January 2025).

13. Gilbert FJ. Breast Screening - Risk Adaptive Imaging for Density. BRAID trial results (personal communication: Brazzelli, M. 2024). 2024.

14. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;**372**:n71.

15. Husereau D, Drummond M, Augustovski F, et al. Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) Statement: updated reporting guidance for health economic evaluations. Clin Ther. 2022;**44**(2):158-68.

16. Philips Z, Bojke L, Sculpher M, Claxton K, Golder S. Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. Pharmacoeconomics. 2006;**24**(4):355-71.

17. Blankenburg M, Sanchez-Collado I, Soyemi BO, et al. Economic evaluation of supplemental breast cancer screening modalities to mammography or digital breast tomosynthesis in women with heterogeneously and extremely dense breasts and average or intermediate breast cancer risk in US healthcare. J Med Econ. 2023;**26**(1):850-61.

18. Lee CI, Cevik M, Alagoz O, et al. Comparative effectiveness of combined digital mammography and tomosynthesis screening for women with dense breasts. Radiology. 2015;**274**(3):772-80.

19. Sprague BL, Stout NK, Schechter C, et al. Benefits, harms, and cost-effectiveness of supplemental ultrasonography screening for women with dense breasts. Ann Intern Med. 2015;**162**(3):157-66.

20. Geuzinge H, Bakker MF, Heijnsdijk EAM, et al. Cost-Effectiveness of Magnetic Resonance Imaging Screening for Women With Extremely Dense Breast Tissue. J Natl Cancer Inst. 2021;**113**(11):1476-83.

21. Wang J, Greuter MJW, Vermeulen KM, et al. Cost-effectiveness of abbreviated-protocol MRI screening for women with mammographically dense breasts in a national breast cancer screening program. Breast. 2022;**61**:58-65.

22. Wang J, Phi XA, Greuter MJW, et al. The cost-effectiveness of digital breast tomosynthesis in a population breast cancer screening program. Eur Radiol. 2020;**30**(10):5437-45.

23. Tollens F, Baltzer PAT, Dietzel M, et al. Cost-Effectiveness of MR-Mammography in Breast Cancer Screening of Women With Extremely Dense Breasts After Two Rounds of Screening. Front Oncol. 2021;**11**:724543.

24. Couto HL, Gargano LP, de Oliveira VM, et al. Cost-Effectiveness Analysis of Digital Breast Tomosynthesis Added to Synthetic Mammography in Breast Cancer Screening in Brazil. PharmacoEconomics - Open. 2024.

25. Ontario Health. Supplemental Screening as an Adjunct to Mammography for Breast Cancer Screening in People With Dense Breasts. Canada: Ontario Health; 2023. Available from: https://www.hqontario.ca/evidence-to-improve-care/health-technology-assessment/reviews-and-recommendations/supplemental-screening-as-an-adjunct-to-mammography-for-breast-cancer-screening-in-people-with-dense-breasts.

26. Kaiser CG, Dietzel M, Vag T, Froelich MF. Cost-effectiveness of MR-mammography vs. conventional mammography in screening patients at intermediate risk of breast cancer - A model-based economic evaluation. Eur J Radiol. 2021;**136**:109355.

27. Gray E, Donten A, Karssemeijer N, et al. Evaluation of a Stratified National Breast Screening Program in the United Kingdom: An Early Model-Based Cost-Effectiveness Analysis. Value Health. 2017;**20**(8):1100-9.

28. Hill H, Kearns B, Pashayan N, et al. The cost-effectiveness of risk-stratified breast cancer screening in the UK. Br J Cancer. 2023;**129**(11):1801-9.

29. de Koning HJ, Boer R, Warmerdam PG, Beemsterboer PM, van der Maas PJ. Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials. J Natl Cancer Inst. 1995;**87**(16):1217-23.

30. Groenewoud JH, Pijnappel RM, van den Akker-van Marle ME, et al. Cost-effectiveness of stereotactic large-core needle biopsy for nonpalpable breast lesions compared to open-breast biopsy. Br J Cancer. 2004;**90**(2):383-92.

31. van Oortmarssen GJ, Habbema JD, van der Maas PJ, et al. A model for breast cancer screening. Cancer. 1990;**66**(7):1601-12.

32. Habbema JD, van Oortmarssen GJ, van Putten DJ, Lubbe JT, van der Maas PJ. Age-specific reduction in breast cancer mortality by screening: an analysis of the results of the Health Insurance Plan of Greater New York study. J Natl Cancer Inst. 1986;**77**(2):317-20.

33. Jacobi CE, Nagelkerke NJ, van Houwelingen JH, de Bock GH. Breast cancer screening, outside the population-screening program, of women from breast cancer families without proven BRCA1/BRCA2 mutations: a simulation study. Cancer Epidemiol Biomarkers Prev. 2006;**15**(3):429-36.

34. Greuter MJ, Jansen-van der Weide MC, Jacobi CE, et al. The validation of a simulation model incorporating radiation risk for mammography breast cancer screening in women with a hereditary-increased breast cancer risk. Eur J Cancer. 2010;**46**(3):495-504.

35. Lu W, Greuter MJ, Schaapveld M, Vermeulen KM, Wiggers T, de Bock GH. Safety and cost-effectiveness of shortening hospital follow-up after breast cancer treatment. Br J Surg. 2012;**99**(9):1227-33.

36. National Institute of Biomedical Imaging and Bioengineering. 2022.

37. Zuckerman SP, Conant EF, Keller BM, et al. Implementation of synthesized two-dimensional ammography in a population-based digital breast tomosynthesis screening program. Radiology. 2016;**281**(3):730-6.

38. Vourtsis A, Kachulis A. The performance of 3D ABUS versus HHUS in the visualisation and BI-RADS characterisation of breast lesions in a large cohort of 1,886 women. Eur Radiol. 2018;**28**(2):592-601.

39. Vinnicombe S, Harvey H, Healy NA, et al. Introduction of an abbreviated breast MRI service in the UK as part of the BRAID trial: practicalities, challenges, and future directions. Clin Radiol. 2021;**76**(6):427-33.

40. European Commission. Adapting Breast Cancer Screening Strategy Using Personalised Risk Estimation. 2015. Available from: https://cordis.europa.eu/project/id/306088/reporting (Accessed 3 February 2025).

41. Evans DG, Astley S, Stavrinos P, et al. Improvement in risk prediction, early detection and prevention of breast cancer in the NHS Breast Screening Programme and family history clinics: a dual cohort study. Programme Grants for Applied Research. 2016;**4**(11).

42. Mainiero MB, Moy L, Baron P, et al. ACR Appropriateness Criteria Breast Cancer Screening. AJR Journal of the American College of Radiology. 2017;**14**(11 Supplement):S383-S90.

43. Wright SJ, Rogers G, Gray E, et al. Evaluation of a Risk-Stratified National Breast Screening Programme in the United Kingdom: An updated cost-effectiveness analysis. medRxiv. 2024.

44. Wright SJ, Gray E, Rogers G, Donten A, Payne K. A structured process for the validation of a decision-analytic model: application to a cost-effectiveness model for risk-stratified national breast screening. Appl Health Econ Health Policy. 2024;**22**(4):527-42.

Appendices

Appendix 1: search strategy

Ovid MEDLINE(R) and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations, Daily and Versions <1946 to May 28, 2024>

1 Mass Screening/ or Early Detection of Cancer/ 148056
2 breast/ 46402
3 1 and 2 1402
4 exp Breast Neoplasms/ 354119
5 ((breast or mammogr* or mammary) adj3 (screen* or cancer or tumo?r or neoplasm?)).tw. 386950
6 3 or 4 or 5 479694
7 Mammography/ or mammogra*.tw,kf. 47520
8 Magnetic Resonance Imaging/ or ("magnetic resonance imaging" or MRI).tw,kf. 717621
9 Ultrasonography, Mammary/ or (sonogra* or ultrasound* or ultrasonogra* or echomammogra* or ABUS or HHUS).tw,kf. 470164
10 (("contrast-enhanced" adj3 mammogra*) or CEM).tw,kf. 7592
11 (tomosynthesis or "3D mammogra*" or "3-D mammogra*" or "digital breast tomogra*" or DBT).tw,kf. 5370
12 ((supplement* or enhance* or adjunct* or addit* or "risk-adapted" or "risk adapted") adj5 (screen* or imag*)).tw,kf. 94175
13 or/7-12 1238053
14 *economics/ 10814
15 exp *"costs and cost analysis"/ 81123
16 (economic adj2 model*).mp. 15584
17 (cost minimi* or cost-utilit* or health utilit* or economic evaluation* or economic review* or cost outcome or cost analys?s or economic analys?s or budget* impact analys?s).ti,ab,kf,kw. 44834
18 (cost-effective* or pharmacoeconomic* or pharmaco-economic* or cost-benefit or costs).ti,kf,kw. 92696
19 (life year or life years or qaly* or cost-benefit analys?s or cost-effectiveness analys?s).ab,kf,kw. 42691
20 (cost or economic*).ti,kf,kw. and (costs or cost-effectiveness or markov).ab. 77357
21 or/14-20 223418
22 6 and 13 and 21 792
23 limit 22 to yr="2014 -Current" 361 (finds 18/20 Azad's examples)

Embase <1974 to 2024 Week 21>

1 Mass Screening/ or early cancer diagnosis/ 76339
2 breast/ 95214
3 1 and 2 436
4 exp breast cancer/ or breast tumor/ 673808
5 ((breast or mammogr* or mammary) adj3 (screen* or cancer or tumo?r or neoplasm?)).tw. 556586

6 3 or 4 or 5 755162
7 Mammography/ or mammogra*.tw,kf. 71599
8 nuclear magnetic resonance imaging/ or breast magnetic resonance imaging/ or ("magnetic resonance imaging" or MRI).tw,kf. 1246229
9 echomammography/ or (sonogra* or ultrasound* or ultrasonogra* or echomammogra* or ABUS or HHUS).tw,kf. 717394
10 contrast enhanced mammography/ or (("contrast-enhanced" adj3 mammogra*) or CEM).tw,kf. 9010
11 digital breast tomosynthesis/ or (tomosynthesis or "3D mammogra*" or "3-D mammogra*" or "digital breast tomogra*" or DBT).tw,kf. 7770
12 ((supplement* or enhance* or adjunct* or addit* or "risk-adapted" or "risk adapted") adj5 (screen* or imag*)).tw,kf. 128454
13 or/7-12 2006682
14 *economics/ 27824
15 exp *economic evaluation/ 79393
16 *health economics/ 16762
17 (economic adj2 model*).mp. 10557
18 (cost minimi* or cost-utilit* or health utilit* or economic evaluation* or economic review* or cost outcome or cost analys?s or economic analys?s or budget* impact analys?s).ti,ab,kf,kw. 68250
19 (cost-effective* or pharmacoeconomic* or pharmaco-economic* or cost-benefit or costs).ti,kf,kw. 135718
20 (life year or life years or qaly* or cost-benefit analys?s or cost-effectiveness analys?s).ab,kf,kw. 65263
21 (cost or economic*).ti,kf,kw. and (costs or cost-effectiveness or markov).ab. 120027
22 or/14-21 295056
23 6 and 13 and 22 1197
24 conference abstract.pt. 5165989
25 23 not 24 987
26 limit 25 to yr="2014 -Current" 431 (finds 16/16 Azad's examples)

EBM Reviews - NHS Economic Evaluation Database <1st Quarter 2016>

1 ((breast or mammogr* or mammary) adj3 (screen* or cancer or tumo?r or neoplasm?)).tw. 604
2 (mammogra* or "magnetic resonance imaging" or MRI or sonogra* or ultrasound* or ultrasonogra* or echomammogra* or ABUS or HHUS or (("contrast-enhanced" adj3 mammogra*) or CEM or tomosynthesis or "3D mammogra*" or "3-D mammogra*" or "digital breast tomogra*" or DBT).tw. 975
3 ((supplement* or enhance* or adjunct* or addit* or "risk-adapted" or "risk adapted") adj5 (screen* or imag*)).tw. 254
4 2 or 3 1142
5 1 and 4 200
6 limit 5 to yr="2014 - 2015" 9

EconLit (Proquest)

noft((breast OR mammogr* OR mammary) NEAR/3 (screen* OR cancer OR tumo?r OR neoplasm?)) AND noft((mammogra* OR "magnetic resonance imaging" OR MRI OR sonogra*

OR ultrasound* OR ultrasonogra* OR echomammogra* OR ABUS OR HHUS OR ("contrast-enhanced" NEAR/3 mammogra*) OR CEM OR tomosynthesis OR "3D mammogra*" OR "3-D mammogra*" OR "digital breast tomogra*" OR DBT)) AND stype.exact("Scholarly Journals") AND pd(20140101-20241231)

Appendix 2: Excluded studies

Ineligible study type

Anonymous. Studies weigh cost, effectiveness of mammography. Cancer Discov. 2014;**4**(5):OF5.

Anonymous. 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. 2016;**16**(20):1-30.

Anonymous. Ontario health technology assessment series ultrasound as an adjunct to mammography for breast cancer screening: A health technology assessment. 2016;**16**(15):1-71.

Autier P. Efficient treatments reduce the cost-efficiency of breast cancer screening. Ann Intern Med. 2016;**164**(4):297-8.

Autier P, Boniol M. Mammography screening: a major issue in medicine. Eur J Cancer. 2018;**90**:34-62.

Baltzer PAT. Supplemental screening using breast MRI in women with mammographically dense breasts. Eur J Radiol. 2021;**136**:109513.

Berg WA. Breast MRI for "the Masses". Eur Radiol. 2022;**32**(6):4034-5.

Bulliard J-L. Estimating the cost-effectiveness of modern screening mammography programmes. Evid Based Med. 2014;**19**(2):80.

Clift AK, Dodwell D, Lord S, et al. The current status of risk-stratified breast screening. Br J Cancer. 2022;**126**(4):533-50.

Cong Z, Goldsmith-Martin G, Phalguni A, Brown A. Economic Evaluations of Cancer Screening Tests in the US: A Systematic Literature Review (SLR). Value Health. 2024;**27**(6).

Cong Z, Goldsmith-Martin G, Phalguni A, Brown A. An Overview of Methods Employed in Economic Models of Cancer Screening Tests: A Systematic Literature Review (SLR). Value Health. 2024;**27**(6).

Harkness EF, Astley SM, Evans DG. Risk-based breast cancer screening strategies in women. Best Practice and Research: Clinical Obstetrics and Gynaecology. 2020;**65**:3-17.

Harvey JA. Our new normal. Journal of Breast Imaging. 2020;**2**(5):415.

Heller SL, Moy L. Breast Cancer Screening and Health Care Costs. JAMA internal medicine. 2020;**180**(11):1552-3.

Kennard K, Wang O, Kjelstrom S, et al. Outcomes of Abbreviated MRI (Ab-MRI) for Women of any Breast Cancer Risk and Breast Density in a Community Academic Setting. Ann Surg Oncol. 2022;**29**(10):6215-21.

Khan SA, Hernandez-Villafuerte KV, Muchadeyi MT, Schlander M. Cost-effectiveness of risk-based breast cancer screening: A systematic review. Int J Cancer. 2021;**149**(4):790-810.

Ko ES, Morris EA. Abbreviated magnetic resonance imaging for breast cancer screening: Concept, early results, and considerations. Korean Journal of Radiology. 2019;**20**(4):533-41.

Kuhl CK, Baltzer P. You Get What You Pay For: Breast MRI Screening of Women With Dense Breasts Is Cost-effective. J Natl Cancer Inst. 2021;**113**(11):1439-41.

Kunst N, Long JB, Xu X, et al. Understanding Regional Variation in the Cost of Breast Cancer Screening Among Privately Insured Women in the United States. Med Care. 2021;**59**(5):437-43.

Kunst N, Long JB, Xu X, et al. Use and Costs of Breast Cancer Screening for Women in Their 40s in a US Population With Private Insurance. JAMA internal medicine. 2020;**180**(5):799-801.

Lee CI, Lee JM, Tosteson ANA. Annual combined mammography and tomosynthesis screening: is it really cost-effective? AJR American journal of roentgenology. 2016;**207**(5):1156-8.

Lee CS, Moy L. Lessons Learned from the Randomized Controlled TOmosynthesis plus SYnthesized MAmmography (TOSYMA) Trial. Radiology. 2023;**306**(2):e222178.

Lee SJ, Li X, Huang H, Zelen M. The Dana-Farber CISNET Model for Breast Cancer Screening Strategies: An Update. Med Decis Making. 2018;**38**(1_suppl):44S-53S.

Lessard J, Blancquaert I. Avis: tomosynthèse mammaire et dépistage du cancer du sein. Canada: Institut national d'excellence en sante et en services sociaux (INESSS); 2019. Available from: https://www.inesss.qc.ca/publications/repertoire-des-publications/publication/tomosynthese-mammaire-et-depistage-du-cancer-du-sein.html.

Lessard J, Provost V. Tomosynthesis and Breast Cancer: Diagnostic Evaluation and Other Uses [French]. Canada: Institut national d'excellence en sante et en services sociaux (INESSS); 2020. Available from: https://www.inesss.qc.ca/publications/repertoire-des-publications/publication/tomosynthese-et-cancer-du-sein-evaluation-diagnostique-et-autres-utilisations.html.

Li J, Jia Z, Zhang M, et al. Cost-Effectiveness Analysis of Imaging Modalities for Breast Cancer Surveillance Among BRCA1/2 Mutation Carriers: A Systematic Review. Front Oncol. 2022;**11**:763161.

Luczynska E, Pawlak M, Popiela T, Rudnicki W. The role of ABUS in the diagnosis of breast cancer. Journal of Ultrasonography. 2022;**22**(89):e76-e85.

Mango VL, Goel A, Mema E, Kwak E, Ha R. Breast MRI screening for average-risk women: A monte carlo simulation cost-benefit analysis. Journal of magnetic resonance imaging : JMRI. 2019;**49**(7):e216-e21.

Mann RM, Kuhl CK, Moy L. Contrast-enhanced MRI for breast cancer screening. J Magn Reson Imaging. 2019;**50**(2):377-90.

Mayo RC, Leung JWT. Impact of artificial intelligence on women's imaging: Cost-benefit analysis. American Journal of Roentgenology. 2019;**212**(5):1172-3.

Mokhtar O, Mahmoud S. Does dynamic post contrast MRI of the breast have a role in screening of high risk patients? Egyptian Journal of Radiology and Nuclear Medicine. 2014;**45**(3):1029-41.

Movik E, Dalsbo TK, Fagelund BC, Friberg EG, Haheim LL, Skar A. Digital Breast Tomosynthesis with Hologic 3D Mammography Selenia Dimensions System for Use in Breast Cancer Screening: A Single Technology Assessment. 2017.

Patel BK, Gray RJ, Pockaj BA. Potential Cost Savings of Contrast-Enhanced Digital Mammography. AJR American journal of roentgenology. 2017;**208**(6):W231-W7.

Thornton J. Targeted breast screening could reduce overdiagnosis and be more cost effective, study finds. BMJ (Clinical research ed). 2018;**362**:k3009.

Tollens F, Baltzer PAT, Froelich MF, Kaiser CG. Comment on: Cost-effectiveness of magnetic resonance imaging screening for women with extremely DENSE breast tissue. Eur J Radiol. 2022;**149**:110186.

Uematsu T. The need for supplemental breast cancer screening modalities: a perspective of population-based breast cancer screening programs in Japan. Breast Cancer. 2017;**24**(1):26-31.

van den Broek JJ, van Ravesteyn NT, Heijnsdijk EA, de Koning HJ. Simulating the Impact of Risk-Based Screening and Treatment on Breast Cancer Outcomes with MISCAN-Fadia. Med Decis Making. 2018;**38**(1_suppl):54S-65S.

Ineligible population

Barfar E, Rashidian A, Hosseini H, Nosratnejad S, Barooti E, Zendehdel K. Cost-effectiveness of mammography screening for breast cancer in a low socioeconomic group of Iranian women. Arch Iran Med. 2014;**17**(4):241-5.

Bonafede MM, Kalra VB, Miller JD, Fajardo LL. Value analysis of digital breast tomosynthesis for breast cancer screening in a commercially-insured US population. ClinicoEconomics and outcomes research : CEOR. 2015;**7**:53-63.

Chung W-S, Wan TTH, Shiu YT, Shi H-Y. Cost-Effectiveness Analysis of Digital Breast Tomosynthesis and Mammography in Breast Cancer Screening: A Markov Modeling Study. Journal of epidemiology and global health. 2024.

Froelich MF, Kaiser CG. Cost-effectiveness of MR-mammography as a solitary imaging technique in women with dense breasts: an economic evaluation of the prospective TK-Study. Eur Radiol. 2021;**31**(2):967-74.

Kalra VB, Wu X, Haas BM, Forman HP, Philpotts LE. Cost-Effectiveness of Tomosynthesis in Annual Screening Mammography. AJR American journal of roentgenology. 2016;**207**(5):1152-5.

Lowry KP, Trentham-Dietz A, Schechter CB, et al. Long-Term Outcomes and Cost-Effectiveness of Breast Cancer Screening With Digital Breast Tomosynthesis in the United States. J Natl Cancer Inst. 2020;**112**(6):582-9.

Malek Pascha VA, Sun L, Gilardino R, Legood R. Telemammography for breast cancer screening: a cost-effective approach in Argentina. BMJ health & care informatics. 2021;**28**(1).

Miller JD, Bonafede MM, Herschorn SD, Pohlman SK, Troeger KA, Fajardo LL. Value Analysis of Digital Breast Tomosynthesis for Breast Cancer Screening in a US Medicaid Population. Journal of the American College of Radiology : JACR. 2017;**14**(4):467-74.e5.

Moger TA, Holen A, Hanestad B, Hofvind S. Costs and Effects of Implementing Digital Tomosynthesis in a Population-Based Breast Cancer Screening Program: Predictions Using Results from the To-Be Trial in Norway. PharmacoEconomics - open. 2022;**6**(4):495-507.

Morton R, Sayma M, Sura MS. Economic analysis of the breast cancer screening program used by the UK NHS: should the program be maintained? Breast cancer (Dove Medical Press). 2017;**9**:217-25.

Phi XA, Greuter MJW, Obdeijn IM, et al. Should women with a BRCA1/2 mutation aged 60 and older be offered intensified breast cancer screening? - A cost-effectiveness analysis. Breast. 2019;**45**:82-8.

Sun L, Legood R, Sadique Z, Dos-Santos-Silva I, Yang L. Cost-effectiveness of risk-based breast cancer screening programme, China. Bull World Health Organ. 2018;**96**(8):568-77.

Ulloa-Perez E, Mohar-Betancourt A, Reynoso-Noveron N. Estimation of the Cost-Effectiveness of Breast Cancer Screening Using Mammography in Mexico Through a Simulation. Revista de investigacion clinica; organo del Hospital de Enfermedades de la Nutricion. 2016;**68**(4):184-91.

Ineligible comparator

Cevik M, Ayer T, Alagoz O, Sprague BL. Analysis of mammography screening policies under resource constraints. Production and Operations Management. 2018;**27**(5):949-72.

Davidovic M, Zielonke N, Lansdorp-Vogelaar I, Segnan N, de Koning HJ, Heijnsdijk EA. Disability-Adjusted Life Years Averted Versus Quality-Adjusted Life Years Gained: A Model Analysis for Breast Cancer Screening. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2021;**24**(3):353-60.

Pataky R, Ismail Z, Coldman AJ, et al. Cost-effectiveness of annual versus biennial screening mammography for women with high mammographic breast density. J Med Screen. 2014;**21**(4):180-8.

Shih Y-CT, Dong W, Xu Y, Etzioni R, Shen Y. Incorporating Baseline Breast Density When Screening Women at Average Risk for Breast Cancer : A Cost-Effectiveness Analysis. Ann Intern Med. 2021;**174**(5):602-12.

Stout NK, Lee SJ, Schechter CB, et al. Benefits, harms, and costs for breast cancer screening after US implementation of digital mammography. J Natl Cancer Inst. 2014;**106**(6):dju092.

Tina Shih Y-C, Dong W, Xu Y, Shen Y. Assessing the cost-effectiveness of updated breast cancer screening guidelines for average-risk women. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2019;**22**(2):185-93.

Tollens F, Baltzer PAT, Dietzel M, Rubenthaler J, Froelich MF, Kaiser CG. Cost-effectiveness of digital breast tomosynthesis vs. Abbreviated breast mri for screening women with intermediate risk of breast cancer-how low-cost must mri be? Cancers (Basel). 2021;**13**(6):1-14.

Trentham-Dietz A, Chapman CH, Jayasekera J, et al. Breast Cancer Screening With Mammography: An Updated Decision Analysis for the U.S. Preventive Services Task Force. 2024.

van Luijt PA, Heijnsdijk EAM, de Koning HJ. Cost-effectiveness of the Norwegian breast cancer screening program. Int J Cancer. 2017;**140**(4):833-40.

Vilaprinyo E, Forné C, Carles M, et al. Cost-Effectiveness and Harm-Benefit Analyses of Risk-Based Screening Strategies for Breast Cancer. PLoS One. 2014;**9**(2):1-10.

Ineligible intervention

Christensen MK, Niclasen BV, Iburg KM. Societal costs and effects of implementing population-based mammography screening in Greenland. Int J Circumpolar Health. 2017;**76**(1):1373580.

Evans DG, Astley S, Stavrinos P, et al. Improvement in risk prediction, early detection and prevention of breast cancer in the NHS Breast Screening Programme and family history clinics: a dual cohort study. 2016.

Fenton JJ, Lee CI, Xing G, Baldwin L-M, Elmore JG. Computer-aided detection in mammography: downstream effect on diagnostic testing, ductal carcinoma in situ treatment, and costs. JAMA internal medicine. 2014;**174**(12):2032-4.

Hruska CB, Conners AL, Jones KN, et al. Diagnostic workup and costs of a single supplemental molecular breast imaging screen of mammographically dense breasts. AJR American journal of roentgenology. 2015;204(6):1345-53.

Kregting LM, Sankatsing VDV, Heijnsdijk EAM, de Koning HJ, van Ravesteyn NT. Finding the optimal mammography screening strategy: A cost-effectiveness analysis of 920 modelled strategies. Int J Cancer. 2022;**151**(2):287-96.

Tollens F, Baltzer PAT, Dietzel M, et al. Economic potential of abbreviated breast MRI for screening women with dense breast tissue for breast cancer. Eur Radiol. 2022;**32**(11):7409-19.

Vargas-Palacios A, Sharma N, Sagoo GS. Cost-effectiveness requirements for implementing artificial intelligence technology in the Women's UK Breast Cancer Screening service. Nature communications. 2023;**14**(1):6110.

Duplicate

Anonymous. Ontario health technology assessment series; 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. 2016;**16**(20):1-30.

Movik E, Dalsbo TK, Fagelund BC, Friberg EG, Haheim LL, Skar A. Digital Breast Tomosynthesis with Hologic 3D Mammography Selenia Dimensions System for Use in Breast Cancer Screening: A Single Technology Assessment. 2017.

Wang J, Greuter MJW, Vermeulen KM, et al. Cost-effectiveness of abbreviated-protocol MRI screening for women with mammographically dense breasts in a national breast cancer screening program. Breast (Edinburgh, Scotland). 2022;**61**:58-65.

Wang J, Phi X-A, Greuter MJW, et al. The cost-effectiveness of digital breast tomosynthesis in a population breast cancer screening program. Eur Radiol. 2020;**30**(10):5437-45.

Non-English language

Gazieva TV, Mikhailov IA, Malkov PG, Suleimanov EA, Frank GA. Comparison of cost-effectiveness of breast cancer screening by mammographic and ultrasonic methods in the healthcare system of the Chechen Republic. Profilakticheskaya Meditsina. 2022;**25**(5):52-60.

Gazieva TV, Mikhailov IA, Malkov PG, Suleymanov EA, Frank GA. Cancer screening cost-effectiveness analysis of the Chechen Republic population for erly detection of breast cancer and colorectal cancer. Profilakticheskaya Meditsina. 2021;**24**(3):22-9.

Appendix 3: Quality assessment for modelling-based economic evaluation studies

Questions for critical appraisal	Blankenburg 2022[17]	Couto 2024[24]	Geuzinge 2021[20]	Gray 2017[27]	Hill 2023[28]	Kaiser 2021[26]	Lee 2015[18]	Ontario Health 2023[25]	Sprague 2015[19]	Tollens 2021[23]	Wang 2022[21]	Wang 2020[22]
Structure												
Statement of decision problem/objective												
Is there a clear statement of the decision problem?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Is the objective of the evaluation and model specified and consistent with the stated decision problem?	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y
Is the primary decision-maker specified?	Y	Y	Y	Y	Y	U	Y	Y	Y	N	N	N
Statement of scope/perspective												
Is the perspective of the model stated clearly?	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	N
Are the model inputs consistent with the stated perspective?	Y	Y	Y	Y	Y	U	Y	Y	Y	Y	N/A	N/A
Has the scope of the model been stated and justified?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	U
Are the outcomes of the model consistent with the perspective, scope and overall objective of the model?	Y	Y	Y	Y	Y	P	P	Y	P	P	N/A	N/A
Rationale for structure												
Has the evidence regarding the model structure been described?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Is the structure of the model consistent with a coherent theory of the health condition under evaluation?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	U
Have any competing theories regarding model structure been considered?	N	N	N	N	N	N	N	Y	N	N	N	N
Are the sources of data used to develop the structure of the model specified?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	U
Are the causal relationships described by the model structure justified appropriately?	Y	Y	Y	Y	P	Y	Y	Y	U	Y	U	U

Questions for critical appraisal	Blankenburg 2022[17]	Couto 2024[24]	Geuzinge 2021[20]	Gray 2017[27]	Hill 2023[28]	Kaiser 2021[26]	Lee 2015[18]	Ontario Health 2023[25]	Sprague 2015[19]	Tollens 2021[23]	Wang 2022[21]	Wang 2020[22]
Structural assumptions												
Are the structural assumptions transparent and justified?	Y	Y	Y	Y	Y	N	Y	Y	N	P	Y	Y
Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	P	Y	Y	Y	U	U	Y	P	U	U	P	P
Strategies/comparators												
Is there a clear definition of the options under evaluation?	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y
Have all feasible and practical options been evaluated?	Y	N	N	Y	N	N	Y	Y	Y	N	N	N
Is there justification for the exclusion of feasible options?	N	N	N	N	N/A	N	N	Y	N	N	N/A	N/A
Model type												
Is the chosen model type appropriate given the decision problem and specified causal relationships within the model?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N
Time horizon												
Is the time horizon of the model sufficient to reflect all important differences between options?	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y
Are the time horizon of the model, the duration of treatment and the duration of treatment effect described and justified?	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y
Has a lifetime horizon been used? If not, has a shorter time horizon been justified?	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y
Disease states/pathways												
Do the disease states (state transition model) or the pathways (decision tree model) reflect the underlying biological process of the disease in question and the impact of interventions?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	U

Questions for critical appraisal	Blankenburg 2022[17]	Couto 2024[24]	Geuzinge 2021[20]	Gray 2017[27]	Hill 2023[28]	Kaiser 2021[26]	Lee 2015[18]	Ontario Health 2023[25]	Sprague 2015[19]	Tollens 2021[23]	Wang 2022[21]	Wang 2020[22]
Cycle length												
Is the cycle length defined and justified in terms of the natural history of disease?	Y	Y	NA	N/A	NA	Y	N/A	N/A	N/A	P	N/A	N/A
Data												
Data identification												
Are the data identification methods transparent and appropriate given the objectives of the model?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Where choices have been made between data sources, are these justified appropriately?	N	Y	N	Y	Y	Y	Y	Y	Y	N	Y	Y
Has particular attention been paid to identifying data for the important parameters in the model?	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y
Has the process of selecting key parameters been justified and systematic methods used to identify the most appropriate data?	P	Y	P	Y	Y	P	P	Y	Y	Y	Y	Y
Has the quality of the data been assessed appropriately?	N/A	Y	N/A	N/A	Y	N/A	N/A	Y	N	N	N	N
Where expert opinion has been used, are the methods described and justified?	Y	N	Y	Y	U	N	N	Y	Y	N	N	Y
Premodel data analysis												
Is the data analysis (premodel) methodology based on justifiable statistical and epidemiological techniques?	U	U	Y	Y	Y	Y	Y	U	N	N	Y	N
Baseline data												
Is the choice of baseline data described and justified?	Y	N	N	Y	Y	N	Y	Y	N	N	Y	Y
Are transition probabilities calculated appropriately?	Y	Y	U	Y	Y	Y	N	U	N	N	U	U
Has a half-cycle correction been applied to both cost and outcome?	Y	N	N/A	N	N/A	N	N/A	N/A	N/A	N	N/A	N/A
If not, has this omission been justified?	N/A	N	N/A	N	N/A	N	N/A	N/A	N/A	N	N/A	N/A

Questions for critical appraisal	Blankenburg 2022[17]	Couto 2024[24]	Geuzinge 2021[20]	Gray 2017[27]	Hill 2023[28]	Kaiser 2021[26]	Lee 2015[18]	Ontario Health 2023[25]	Sprague 2015[19]	Tollens 2021[23]	Wang 2022[21]	Wang 2020[22]
Treatment effects												
If relative treatment effects have been derived from trial data, have they been synthesised using appropriate techniques?	U	Y	Y	Y	Y	N	Y	U	P	Y	Y	U
Have the methods and assumptions used to extrapolate shortterm results to final outcomes been documented and justified? Have alternative assumptions been explored through sensitivity analysis?	Y	Y	Y	Y	Y	N	Y	P	N	Y	N	N
Have assumptions regarding the continuing effect of treatment once treatment is complete been documented and justified? Have alternative assumptions been explored through sensitivity analysis?	N	Y	N	Y	U	N	Y	N	N	N	N	N
Quality of life weights (utilities)												
Are the utilities incorporated into the model appropriate? Is the source for the utility weights referenced?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N/A	N/A
Are the methods of derivation for the utility weights justified?	N	N	N	Y	Y	N	N	Y	Y	N	N/A	N/A
Data incorporation												
Have all data incorporated into the model been described and referenced in sufficient detail?	Y	Y	P	Y	Y	Y	Y	Y	Y	Y	Y	Y
Has the use of mutually inconsistent data been justified (i.e. are assumptions and choices appropriate)?	Y	Y	P	Y	U	Y	Y	P	U	U	Y	U
Is the process of data incorporation transparent?	Y	Y	Y	Y	Y	Y	Y	P	N	N	Y	P
If data have been incorporated as distributions, has the choice of distribution for each parameter been described and justified?	N	Y	N	Y	Y	Y	N	Y	N	N	Y	N

Questions for critical appraisal	Blankenburg 2022[17]	Couto 2024[24]	Geuzinge 2021[20]	Gray 2017[27]	Hill 2023[28]	Kaiser 2021[26]	Lee 2015[18]	Ontario Health 2023[25]	Sprague 2015[19]	Tollens 2021[23]	Wang 2022[21]	Wang 2020[22]
If data have been incorporated as distributions, is it clear that second order uncertainty is reflected?	N	Y	N	Y	Y	Y	N	Y	U	U	Y	N
Assessment of uncertainty												
Have the four principal types of uncertainty been addressed?	Y	Y	P	Y	Y	P	P	N	P	Y	Y	N
If not, has the omission of particular forms of uncertainty been justified?	N/A	N/A	N	N/A	N/A	N	N	N/A	N	N/A	N/A	N
Methodological												
Have methodological uncertainties been addressed by running alternative versions of the model with different methodological assumptions?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Structural												
Is there evidence that structural uncertainties have been addressed via sensitivity analysis?	Y	Y	Y	Y		Y	N	N	Y	Y	Y	N
Heterogeneity												
Has heterogeneity been dealt with by running the model separately for different subgroups?	N	Y	U	Y	U	N	N	Y	N	N	Y	Y
Parameter												
Are the methods of assessment of parameter uncertainty appropriate?	Y	Y	N	Y	Y	Y	Y	N	Y	Y	Y	N/A
Has probabilistic sensitivity analysis been done? If not, has this been justified?	Y	Y	N	Y	Y	Y	Y	N	Y	Y	Y	N
If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly and justified?	N	Y	N	Y	Y	Y	Y	Y	P	Y	Y	N/A

Questions for critical appraisal	Blankenburg 2022[17]	Couto 2024[24]	Geuzinge 2021[20]	Gray 2017[27]	Hill 2023[28]	Kaiser 2021[26]	Lee 2015[18]	Ontario Health 2023[25]	Sprague 2015[19]	Tollens 2021[23]	Wang 2022[21]	Wang 2020[22]
Model consistency												
Internal consistency												
Is there evidence that the mathematical logic of the model has been tested thoroughly before use?	N	N	N	N	Y	N	N	U	N	N	N	N
External consistency												
Are the conclusions valid given the data presented?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Are any counterintuitive results from the model explained and justified	Y	Y	U	Y	U	U	N	N	N	Y	Y	N
If the model has been calibrated against independent data, have any differences been explained and justified?	Y	N	Y	N	N	N	Y	N	Y	N	U	U
Have the results of the model been compared with those of previous models and any differences in results explained?	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y