Health Technology Assessment 2011; Vol. 15: No. 4

Positron emission tomography (PET) and magnetic resonance imaging (MRI) for the assessment of axillary lymph node metastases in early breast cancer: systematic review and economic evaluation

KL Cooper, Y Meng, S Harnan, SE Ward,
P Fitzgerald, D Papaioannou, L Wyld,
C Ingram, ID Wilkinson and E Lorenz

January 2011
10.3310/hta15040

Health Technology Assessment
NIHR HTA programme
www.hta.ac.uk

How to obtain copies of this and other HTA programme reports

An electronic version of this title, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (www.hta.ac.uk). A fully searchable DVD is also available (see below).

Printed copies of HTA journal series issues cost £20 each (post and packing free in the UK) to both public **and** private sector purchasers from our despatch agents.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per issue and for the rest of the world £3 per issue.

How to order:

– fax (with **credit card details**)
– post (with **credit card details** or **cheque**)
– phone during office hours (**credit card** only).

Additionally the HTA website allows you to either print out your order or download a blank order form.

Contact details are as follows:

Synergie UK (HTA Department)
Digital House, The Loddon Centre
Wade Road
Basingstoke
Hants RG24 8QW

Email: orders@hta.ac.uk

Tel: 0845 812 4000 – ask for 'HTA Payment Services' (out-of-hours answer-phone service)

Fax: 0845 812 4001 – put 'HTA Order' on the fax header

Payment methods

Paying by cheque
If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *University of Southampton* and drawn on a bank with a UK address.

Paying by credit card
You can order using your credit card by phone, fax or post.

Subscriptions

NHS libraries can subscribe free of charge. Public libraries can subscribe at a reduced cost of £100 for each volume (normally comprising 40–50 titles). The commercial subscription rate is £400 per volume (addresses within the UK) and £600 per volume (addresses outside the UK). Please see our website for details. Subscriptions can be purchased only for the current or forthcoming volume.

How do I get a copy of *HTA on DVD*?

Please use the form on the HTA website (www.hta.ac.uk/htacd/index.shtml). *HTA on DVD* is currently free of charge worldwide.

The website also provides information about the HTA programme and lists the membership of the various committees.

Positron emission tomography (PET) and magnetic resonance imaging (MRI) for the assessment of axillary lymph node metastases in early breast cancer: systematic review and economic evaluation

KL Cooper,[1]* Y Meng,[1] S Harnan,[1] SE Ward,[1] P Fitzgerald,[1] D Papaioannou,[1] L Wyld,[2] C Ingram,[2] ID Wilkinson[2] and E Lorenz[2]

[1]School of Health and Related Research (ScHARR), University of Sheffield, Sheffield, UK
[2]Sheffield Teaching Hospitals, University of Sheffield, Sheffield, UK

*Corresponding author

Published January 2011
DOI: 10.3310/hta15040

This report should be referenced as follows:

Cooper KL, Meng Y, Harnan S, Ward SE, Fitzgerald P, Papaioannou D, *et al*. Positron emission tomography (PET) and magnetic resonance imaging (MRI) for the assessment of axillary lymph node metastases in early breast cancer: systematic review and economic evaluation. *Health Technol Assess* 2011;**15**(4).

Health Technology Assessment is indexed and abstracted in *Index Medicus*/MEDLINE, *Excerpta Medica*/EMBASE, *Science Citation Index Expanded (SciSearch®)* and *Current Contents®*/Clinical Medicine.

NIHR Health Technology Assessment programme

The Health Technology Assessment (HTA) programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The research findings from the HTA programme directly influence decision-making bodies such as the National Institute for Health and Clinical Excellence (NICE) and the National Screening Committee (NSC). HTA findings also help to improve the quality of clinical practice in the NHS indirectly in that they form a key component of the 'National Knowledge Service'.

The HTA programme is needs led in that it fills gaps in the evidence needed by the NHS. There are three routes to the start of projects.

First is the commissioned route. Suggestions for research are actively sought from people working in the NHS, from the public and consumer groups and from professional bodies such as royal colleges and NHS trusts. These suggestions are carefully prioritised by panels of independent experts (including NHS service users). The HTA programme then commissions the research by competitive tender.

Second, the HTA programme provides grants for clinical trials for researchers who identify research questions. These are assessed for importance to patients and the NHS, and scientific rigour.

Third, through its Technology Assessment Report (TAR) call-off contract, the HTA programme commissions bespoke reports, principally for NICE, but also for other policy-makers. TARs bring together evidence on the value of specific technologies.

Some HTA research projects, including TARs, may take only months, others need several years. They can cost from as little as £40,000 to over £1 million, and may involve synthesising existing evidence, undertaking a trial, or other research collecting new data to answer a research problem.

The final reports from HTA projects are peer reviewed by a number of independent expert referees before publication in the widely read journal series *Health Technology Assessment*.

Criteria for inclusion in the HTA journal series

Reports are published in the HTA journal series if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

The research reported in this issue of the journal was commissioned and funded by the HTA programme on behalf of NICE as project number 08/35/01. The protocol was agreed in January 2009. The assessment report began editorial review in November 2009 and was accepted for publication in June 2010. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the referees for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

The views expressed in this publication are those of the authors and not necessarily those of the HTA programme or the Department of Health.

ISSN 1366-5278

Published by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk), on behalf of NETSCC, HTA.
Printed on acid-free paper in the UK by Henry Ling Ltd, The Dorset Press, Dorchester.

T

Abstract

Positron emission tomography (PET) and magnetic resonance imaging (MRI) for the assessment of axillary lymph node metastases in early breast cancer: systematic review and economic evaluation

KL Cooper,[1]* Y Meng,[1] S Harnan,[1] SE Ward,[1] P Fitzgerald,[1] D Papaioannou,[1] L Wyld,[2] C Ingram,[2] ID Wilkinson[2] and E Lorenz[2]

[1]School of Health and Related Research (ScHARR), University of Sheffield, Sheffield, UK
[2]Sheffield Teaching Hospitals, University of Sheffield, Sheffield, UK

*Corresponding author

Background: Breast cancer is the most common type of cancer in women. Evaluation of axillary lymph node metastases is important for breast cancer staging and treatment planning.
Objectives: To evaluate the diagnostic accuracy, cost-effectiveness and effect on patient outcomes of positron emission tomography (PET), with or without computed tomography (CT), and magnetic resonance imaging (MRI) in the evaluation of axillary lymph node metastases in patients with newly diagnosed early-stage breast cancer.
Data sources: A systematic review of literature and an economic evaluation were carried out. Key databases (including MEDLINE, EMBASE and nine others) plus research registers and conference proceedings were searched for relevant studies up to April 2009. A decision-analytical model was developed to determine cost-effectiveness in the UK.
Review methods: One reviewer assessed titles and abstracts of studies identified by the search strategy, obtained the full text of relevant papers and screened them against inclusion criteria. Data from included studies were extracted by one reviewer using a standardised data extraction form and checked by a second reviewer. Discrepancies were resolved by discussion. Quality of included studies was assessed using the quality assessment of diagnostic accuracy studies (QUADAS) checklist, applied by one reviewer and checked by a second.
Results: Forty-five citations relating to 35 studies were included in the clinical effectiveness review: 26 studies of PET and nine studies of MRI. Two studies were included in the cost-effectiveness review: one of PET and one of MRI. Of the seven studies evaluating PET/CT ($n = 862$), the mean sensitivity was 56% [95% confidence interval (CI) 44% to 67%] and mean specificity 96% (95% CI 90% to 99%). Of the 19 studies evaluating PET only ($n = 1729$), the mean sensitivity was 66% (95% CI 50% to 79%) and mean specificity 93% (95% CI 89% to 96%). PET performed less well for small metastases; the mean sensitivity was 11% (95% CI 5% to 22%) for micrometastases (≤ 2 mm; five studies; $n = 63$), and 57% (95% CI 47% to 66%) for macrometastases (> 2 mm; four studies; $n = 111$). The smallest metastatic nodes detected by PET measured 3 mm, while PET failed to detect some nodes measuring > 15 mm. Studies in which all patients were clinically node negative showed a trend towards lower sensitivity of PET compared with studies with a mixed population. Across five studies evaluating ultrasmall super-paramagnetic iron oxide (USPIO)-enhanced MRI ($n = 93$), the mean sensitivity was 98% (95% CI 61% to 100%) and mean specificity

96% (95% CI 72% to 100%). Across three studies of gadolinium-enhanced MRI ($n = 187$), the mean sensitivity was 88% (95% CI 78% to 94%) and mean specificity 73% (95% CI 63% to 81%). In the single study of in vivo proton magnetic resonance spectroscopy ($n = 27$), the sensitivity was 65% (95% CI 38% to 86%) and specificity 100% (95% CI 69% to 100%). USPIO-enhanced MRI showed a trend towards higher sensitivity and specificity than gadolinium-enhanced MRI. Results of the decision modelling suggest that the MRI replacement strategy is the most cost-effective strategy and dominates the baseline 4-node sampling (4-NS) and sentinel lymph node biopsy (SLNB) strategies in most sensitivity analyses undertaken. The PET replacement strategy is not as robust as the MRI replacement strategy, as its cost-effectiveness is significantly affected by the utility decrement for lymphoedema and the probability of relapse for false-negative (FN) patients.

Limitations: No included studies directly compared PET and MRI.

Conclusions: Studies demonstrated that PET and MRI have lower sensitivity and specificity than SLNB and 4-NS but are associated with fewer adverse events. Included studies indicated a significantly higher mean sensitivity for MRI than for PET, with USPIO-enhanced MRI providing the highest sensitivity. However, sensitivity and specificity of PET and MRI varied widely between studies, and MRI studies were relatively small and varied in their methods; therefore, results should be interpreted with caution. Decision modelling based on these results suggests that the most cost-effective strategy may be MRI rather than SLNB or 4-NS. This strategy reduces costs and increases quality-adjusted life-years (QALYs) because there are fewer adverse events for the majority of patients. However, this strategy leads to more FN cases at higher risk of cancer recurrence and more false-positive (FP) cases who would undergo unnecessary axillary lymph node dissection. Adding MRI prior to SLNB or 4-NS has little effect on QALYs, though this analysis is limited by lack of available data. Future research should include large, well-conducted studies of MRI, particularly using USPIO; data on the long-term impacts of lymphoedema on cost and patient utility; studies of the comparative effectiveness and cost-effectiveness of SLNB and 4-NS; and more robust UK cost data for 4-NS and SLNB as well as the cost of MRI and PET techniques.

Funding: This study was funded by the Health Technology Assessment programme of the National Institute of Health Research.

Contents

Glossary

4-node sampling (4-NS) An alternative staging technique to sentinel lymph node biopsy. A minimum of four lymph nodes (sometimes more) are surgically removed and examined histologically to determine the presence of axillary metastases.

Axillary lymph nodes Located in the armpit area, they receive lymph fluid from the arm, breast and ipsilateral upper torso.

Axillary lymph node dissection (ALND) Surgical removal of most or all axillary lymph nodes, to level I, II or III. Removed tissue is histologically examined to determine axillary spread of breast cancer.

Axillary sampling Within this assessment, axillary sampling refers to sentinel lymph node biopsy or 4-node sampling.

Computed tomography (CT) CT scans use a series of two-dimensional X-rays to generate a three-dimensional image of body structures. In modern positron emission tomography scanning, CT scans are often used alongside the positron emission tomography scan to allow concurrent visualisation of the anatomy of tissues and their metabolic activity.

False-negative (FN) A patient with a condition who is wrongly diagnosed as not having it.

False-positive (FP) A patient without a condition who is wrongly diagnosed as having it.

Lymph nodes Small glands that are part of the lymphatic system. White blood cells in the lymph nodes attack bacteria and viruses as they pass through the node.

Magnetic resonance imaging (MRI) MRI uses a magnetic field to provide a net alignment of hydrogen nuclei, predominantly in water molecules. These nuclei resonate when a non-ionising radiofrequency field is applied and give out a signal when it is turned off. An image is generated by encoding this signal using noisy switched magnetic field gradients. Nuclei in different tissues and pathology yield different image characteristics. In contrast-enhanced MRI, intravenously administered agents such as gadolinium chelates and USPIO (ultrasmall super-paramagnetic iron oxide) are used to provide localised image enhancement.

Positron emission tomography (PET) PET uses a positron-emitting radionucleotide tracer to create a three-dimensional image of the functional processes of the body.

Sensitivity The effectiveness of a diagnostic test in correctly identifying persons with a condition (true-positives divided by all persons with the condition).

Sentinel lymph node biopsy (SLNB) A surgical diagnostic technique used to determine metastatic spread in primary breast cancer. The sentinel node (the first draining node from the breast) is identified using a blue dye and/or radiotracer, and is dissected out and sent for histological examination. SLNB is a less invasive surgical procedure than axillary lymph node dissection as fewer nodes are removed.

Specificity The effectiveness of a diagnostic test in correctly diagnosing as negative persons who do not have a condition (true-negatives divided by all persons without the condition).

True-negative (TN) A patient without a condition who is correctly diagnosed as not having it.

True-positive (TP) A patient with a condition who is correctly diagnosed as having it.

List of abbreviations

4-NS	4-node sampling
ALND	axillary lymph node dissection
CI	confidence interval
CT	computed tomography
DCE-MRI	dynamic contrast-enhanced magnetic resonance imaging
DCIS	ductal carcinoma in situ
FDG	2-fluoro-2-deoxy-D-glucose
FN	false-negative
FNAC	fine-needle aspiration cytology
FP	false-positive
H&E	haematoxylin and eosin
LCIS	lobular carcinoma in situ
MeSH	medical subject heading
MR	magnetic resonance
MRI	magnetic resonance imaging
NICE	National Institute for Health and Clinical Excellence
PET	positron emission tomography
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSA	probabilistic sensitivity analysis
QALY	quality-adjusted life-year
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
ROC	receiver operating characteristic
SLNB	sentinel lymph node biopsy
SUV	standardised uptake value
TN	true-negative
TNM stage	tumour, node, metastases stage
TP	true-positive
USPIO	ultrasmall super-paramagnetic iron oxide

All abbreviations that have been used in this report are listed here unless the abbreviation is well known (e.g. NHS), or it has been used only once, or it is a non-standard abbreviation used only in figures/tables/appendices, in which case the abbreviation is defined in the figure legend or in the notes at the end of the table.

Executive summary

Background

Evaluation of axillary lymph node metastases is important for breast cancer staging and treatment planning. Axillary lymph node dissection (ALND) has been considered the gold standard for identifying axillary metastases and controlling spread, but has a high risk of morbidity including long-term lymphoedema. Sentinel lymph node biopsy (SLNB) and 4-node sampling (4-NS) have a lower risk of morbidity. Current guidelines from the National Institute for Health and Clinical Excellence recommend SLNB or 4-NS where ultrasound-guided biopsy is negative. ALND is indicated where ultrasound-guided biopsy, SLNB or 4-NS is positive.

Objectives

The objectives of this assessment were to evaluate the diagnostic accuracy, cost-effectiveness and effect on patient outcomes of positron emission tomography (PET), with or without computed tomography (CT), and magnetic resonance imaging (MRI) in the evaluation of axillary lymph node metastases in patients with newly diagnosed early-stage breast cancer. PET and MRI are assessed firstly as a replacement for SLNB or 4-NS, and secondly as an additional test prior to SLNB or 4-NS.

Methods

A systematic review was undertaken to identify studies reporting sensitivity and specificity of PET or MRI for the assessment of axillary lymph node metastases in early-stage breast cancer. The following databases were searched in April 2009: MEDLINE, MEDLINE In-Process & Other Non-Indexed Citations, EMBASE, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Cochrane Database of Systematic Reviews (CDSR), Cochrane Central Register of Controlled Trials (CENTRAL), Database of Abstracts of Reviews of Effects (DARE), NHS Economic Evaluation Database (NHS EED), HTA database, Science Citation Index and BIOSIS previews. Research registers and conference proceedings were also searched. PET studies with < 20 analysable patients were excluded, while MRI studies of all sizes were included (as there were fewer with a large sample size). Study quality was assessed using the quality assessment of diagnostic accuracy studies checklist. A bivariate random effects approach was used for the meta-analysis of pairs of sensitivity and specificity data.

A decision model was developed to investigate the benefits, harms and cost-effectiveness of PET and MRI, either as a replacement for SLNB or 4-NS or as an additional test prior to SLNB or 4-NS. Both SLNB and 4-NS are currently used in the UK. Comparison of SLNB with 4-NS was not part of the remit of this assessment; therefore, the two baseline strategies (SLNB and 4-NS) were considered separately. It was assumed that, prior to these investigations, all patients underwent clinical examination and axillary ultrasound (and ultrasound-guided biopsy where ultrasound was positive). The number of correct and incorrect diagnoses, costs and impact on quality-adjusted life-years (QALYs) due to cancer recurrence and adverse effects were determined for each strategy. Model results are presented in terms of net health benefit and cost per incremental QALY gained. Probabilistic sensitivity analyses were undertaken.

Results

Summary of clinical results

Diagnostic accuracy of positron emission tomography

Across 26 studies evaluating PET or PET/CT ($n = 2591$ patients), the mean sensitivity was 63% [95% confidence interval (CI) 52% to 74%; range 20%–100%] and mean specificity 94% (95% CI 91% to 96%; range 75%–100%). Of the seven studies evaluating PET/CT ($n = 862$), the mean sensitivity was 56% (95% CI 44% to 67%) and mean specificity 96% (95% CI 90% to 99%). Of the 19 studies evaluating PET only ($n = 1729$), the mean sensitivity was 66% (95% CI 50% to 79%) and mean specificity 93% (95% CI 89% to 96%). PET performed less well for small metastases; the mean sensitivity was 11% (95% CI 5% to 22%) for micrometastases (≤ 2 mm; five studies; $n = 63$), and 57% (95% CI 47% to 66%) for macrometastases (> 2 mm; four studies; $n = 111$). The smallest metastatic nodes detected by PET measured 3 mm, while PET failed to detect some nodes measuring > 15 mm. Studies in which all patients were clinically node negative showed a trend towards lower sensitivity of PET compared with studies with a mixed population.

Diagnostic accuracy of magnetic resonance imaging

There were nine studies evaluating MRI ($n = 307$ patients), many reporting more than one set of results according to different criteria for positivity. Based on the highest reported sensitivity and specificity per study, the mean sensitivity was 90% (95% CI 78% to 96%; range 65%–100%) and mean specificity 90% (95% CI 75% to 96%; range 54%–100%). Across five studies evaluating ultrasmall super-paramagnetic iron oxide (USPIO)-enhanced MRI ($n = 93$), the mean sensitivity was 98% (95% CI 61% to 100%) and mean specificity 96% (95% CI 72% to 100%). Across the three studies of gadolinium-enhanced MRI ($n = 187$), the mean sensitivity was 88% (95% CI 78% to 94%) and mean specificity 73% (95% CI 63% to 81%). In the single study of in vivo proton magnetic resonance spectroscopy ($n = 27$), the sensitivity was 65% (95% CI 38% to 86%) and specificity 100% (95% CI 69% to 100%). Therefore, USPIO-enhanced MRI showed a trend towards higher sensitivity and specificity than gadolinium-enhanced MRI. No data were presented according to size of metastases for MRI.

Adverse effects and contraindications

No adverse effects were reported for PET. Studies of MRI reported only mild-to-moderate adverse effects including mild rash following USPIO administration, claustrophobia and back pain. Cautions and contraindications exist for both PET (pregnancy) and MRI (allergy to contrast agents, renal or liver dysfunction, pacemakers and other metallic implants), and some patients are unable to undergo MRI due to claustrophobia.

Summary of cost-effectiveness and benefits versus risks

The results of the decision modelling suggest that the most cost-effective strategy is to replace axillary sampling (SLNB or 4-NS) with MRI. This strategy dominates the baseline SLNB and 4-NS strategies, generating higher QALYs and lower costs. SLNB and 4-NS are avoided for all patients, leading to fewer adverse effects, including lymphoedema, which has an assumed lifelong impact on quality of life. However, MRI has lower sensitivity than SLNB and 4-NS [leading to more false-negatives (FNs)] and a lower specificity [leading to more false-positives (FPs)]. Patients with a FN diagnosis will not receive ALND or adjuvant therapy, leading to a higher risk of cancer recurrence. Patients with a FP diagnosis will receive ALND unnecessarily, with the accompanying increased risk of adverse effects. At the population level, the model results suggest that the MRI replacement strategy costs less, and the health benefits gained by the majority of patients outweigh the negative impact on survival and quality of life of a small proportion of patients. This strategy may however be rejected on clinical grounds, owing to the increase in number of FP and FN cases compared with current practice.

If the replacement strategies are rejected on clinical grounds, the most cost-effective strategy is predicted to be the baseline 4-NS strategy (compared with 4-NS as the baseline) or the use of MRI prior to SLNB (compared with SLNB as the baseline). For strategies that place MRI or PET before axillary sampling, patients with true-positive PET or MRI results receive immediate ALND without the need to carry out a separate SLNB or 4-NS procedure (although the requirement for two procedures may also be averted through the use of intraoperative cytology following axillary sampling). The number of FN cases is also reduced owing to the use of two sequential tests. The disadvantage is that, due to the lower specificity of PET and MRI compared with SLNB and 4-NS, there are still more patients with FP results who receive ALND unnecessarily, with the accompanying increased risk of adverse effects. The total QALYs generated by the strategies of adding MRI prior to 4-NS or SLNB are very similar to those in the baseline strategies and these results are not considered to be robust, based on the quality of the data available.

The model results indicate that, as would be expected, patients with FN results have the lowest survival rates because they do not receive ALND or chemotherapy that may reduce the risk of recurrence. Compared with the two baseline strategies, the overall survival rates are lower for the MRI or PET replacement strategies (due to an increase in FNs) and higher for the strategies where MRI or PET are placed before 4-NS and SLNB (due to a decrease in FNs). However, the absolute differences in overall survival rate among tested diagnostic strategies are relatively small, since the absolute number of patients with FN results only accounts for a small proportion of all patients.

Sensitivity analysis suggests that the MRI replacement strategy remains the most cost-effective strategy in the majority of the one-way sensitivity analyses undertaken. The sensitivity analyses indicate that the cost-effectiveness of the PET replacement strategy is significantly affected by the assumption relating to long-term utility decrements for lymphoedema (for which few high-quality data exist) and the probability of relapse for FN patients. If the PET and MRI replacement strategies are excluded, the cost-effectiveness of the strategies of adding MRI prior to 4-NS or SLNB is affected by the assumed MRI sensitivity and specificity, the probability of relapse for FN patients, and the costs of SLNB. The results relating to addition of MRI prior to 4-NS or SLNB are not considered to be robust.

Conclusions

Implications for service provision

The included studies demonstrate a significantly higher sensitivity for MRI than for PET, with USPIO-enhanced MRI providing the highest sensitivity. However, as there were no studies directly comparing PET and MRI, caution should be taken when comparing these estimates. Sensitivity of PET was reduced for smaller metastases. Specificity was similar for PET and MRI. Sensitivity and specificity of PET and MRI varied widely between studies, and MRI studies were relatively small and varied in their methods; therefore, results should be interpreted with caution. All patients currently receive ultrasound prior to other investigations; the sensitivity of PET appears similar to that of ultrasound and so may provide little additional benefit, while the sensitivity of MRI appears slightly higher. Specificity of PET and USPIO-enhanced MRI appear slightly higher than for ultrasound.

PET and MRI have lower sensitivity and specificity than SLNB and 4-NS, but are associated with fewer adverse events. Decision modelling suggests that the most cost-effective strategy may be MRI rather than SLNB or 4-NS, reducing costs and increasing QALYs due to fewer adverse events for the majority of patients. However, this strategy may be clinically unacceptable due

to higher numbers of FN cases (leading to higher risk of recurrence) and FP cases (leading to unnecessary ALND). If this strategy is rejected on clinical grounds, the model suggests that the most cost-effective strategy may be the baseline 4-NS strategy (compared with 4-NS as the baseline) or the use of MRI prior to SLNB (compared with SLNB as the baseline). However, the results relating to addition of MRI prior to 4-NS or SLNB are not considered to be robust based on the quality of the input parameters available, and further work is required to provide more reliable estimates.

Suggested research priorities

If the use of MRI is deemed clinically acceptable (either to replace SLNB or 4-NS or as an additional test), then further large, well-conducted studies of MRI, particularly using USPIO, would be useful to obtain more robust data on sensitivity and specificity, adverse effects and the optimum criteria for defining a node as metastatic. In addition, further data on the long-term impacts of lymphoedema on cost and patient utility would be valuable, as well as studies of the comparative effectiveness and cost-effectiveness of SLNB and 4-NS. More robust UK cost data is needed for 4-NS and SLNB, as well as the cost of MRI and PET techniques.

Funding

This study was funded by the Health Technology Assessment programme of the National Institute of Health Research.

Chapter 1

Background

Description of health problem

Breast cancer and axillary metastases

Breast cancer is the most common type of cancer in women, with 38,048 new cases registered in women in England[1] and 2457 in Wales[2] in 2007. The usual site of spread outside the breast in newly diagnosed cases is to the lymph nodes in the axilla (underarm). The presence of axillary metastases and the extent of their spread are important prognostic factors for staging disease and planning treatment, whilst removal of any spread is essential to prevent recurrence and wider metastatic spread.

Aetiology, pathology and prognosis

Aetiology

A range of risk factors for breast cancer have been identified, including genetic, hormonal and lifestyle factors.[3] It has been estimated that 12% of women with breast cancer have one affected family member and 1% have two or more affected.[4] Genetic predisposition is mediated by high-penetrance genes such as breast cancer susceptibility gene 1 (*BRCA1*) and breast cancer susceptibility gene 2 (*BRCA2*) (responsible for around 80%–90% of hereditary cancers), and low-penetrance genes which confer increased and decreased risk.[3]

Environmental and lifestyle factors as well as genetic factors influence breast cancer risk. Asian migrants to the West have increased levels of risk compared with the indigenous population, while Asian-Americans born in the West have incidence rates approximating the US average.[5] Lifestyle and environmental factors thought to increase risk include hormonal factors such as taking the oral contraceptive pill[6] or hormone replacement therapy,[7] higher age at menopause, early age at menarche, late age at first birth and not giving birth. A recent systematic review suggests night work increases risk.[8] Factors which decrease risk include higher folate intake, higher number of pregnancies, breast feeding and younger age at first birth.[3] Obesity increases risk of breast cancer in post-menopausal women.[9,10] The picture is less clear for pre-menopausal women in whom the risk may be lower, but prognosis poorer.[11] Obesity may affect oestradiol or insulin levels, which is thought to account in part for the increased incidence. Obesity is also correlated to increased severity at presentation, which may be related to lower screening attendance for obese women.[12] Physical activity in adolescence and young adulthood confers a decreased risk of breast cancer,[13] which may be mediated hormonally. The protective effect is less evident in pre-menopausal cancers than post-menopausal.[14]

Other factors that confer risk include certain high- and low-penetrance genes, increasing age, height, dense breast tissue, alcohol consumption and exposure to ionising radiation in childhood. Among other protective factors are certain low-penetrance genes, fruit and vegetable consumption, chemopreventative agents and use of non-steroidal anti-inflammatory drugs.[3] For men, genetic predisposition, exposure to radiation (especially at a young age), obesity and high levels of oestrogen due to other conditions all affect risk.[15]

Pathology

Breast cancer starts with genetic changes in a single cell or small group of cells in the epithelia of the ducts or the lobules of the breast. The genetic change allows cells to reproduce uncontrollably, creating a tumour. Tumours that have not yet spread to surrounding tissue are known as 'carcinoma in situ' and may be ductal carcinoma in situ (DCIS) or lobular carcinoma in situ (LCIS). Once spread to surrounding tissue begins, a tumour is known as 'invasive'. More rapid growth and spread occurs once a blood supply is secured. Cancer spreads via the lymphatic system or the bloodstream. Lymphatic spread is usually first to the axillary lymph nodes. Spread via the bloodstream can lead to distant metastases in the bone or viscera, which are considered incurable.

The presence or absence of axillary metastases is a key indicator of stage of disease and prognosis, and adjuvant therapy is, in part, planned based on their presence and extent.[16] They are caused when a single cell or small numbers of cells detach from the main tumour, travel via the lymphatic system and establish themselves in the tissue of the lymph nodes. Axillary metastases occur in approximately 41%[17] of cases and prognosis is better where there is no axillary spread. Where metastases are present, surgical removal of axillary lymph nodes is indicated in order to prevent further spread and ensure local disease control.

Prognosis

Overall, breast cancer 5-year, age-standardised survival rates are around 80%.[18] Survival varies with age (*Table 1*) and stage of disease (*Table 2*).

Other factors can affect prognosis. Clinicians may use tools such as the Nottingham Prognostic Index,[20] which takes into account grade as well as size and spread, or Adjuvant Online,[21] which uses patient data such as age, tumour size, nodal involvement, hormonal receptor status and histological grade to predict disease course and treatment options. Good prognosis is associated with small tumour size, node-negative status and younger age, oestrogen receptor-positive and progesterone receptor-positive status. HER2 (human epidermal growth factor receptor 2) over-expression is associated with poor prognosis.

Epidemiology and incidence

Incidence varies most with gender. Women are far more likely to get breast cancer than men. For both men and women, incidence also varies with age (*Table 3*). Approximately 81% of cases occur in women aged 50 years and over.[22]

TABLE 1 Five-year survival rates according to age[a]

	Age (years)					
	15–39	40–49	50–59	60–69	70–79	80–99
5-year survival rate (%)	81	86	89	87	78	64

a Based on women in England diagnosed during 2001–6.[18]

TABLE 2 Five-year survival rates according to stage of disease[a]

	Stage of disease			
	I	II	III	IV
5-year survival (%)	88	69	43	12

a Based on women diagnosed in the West Midlands 1985–9 followed up to 1999.[19]

DOI: 10.3310/hta15040

TABLE 3 Breast cancer incidence for England[1] and Wales[2], 2007

		Age (years)															
Population	Country	15–19	20–24	25–29	30–34	35–39	40–44	45–49	50–54	55–59	60–64	65–69	70–74	75–79	80–84	85+	Total
Women	Wales			2	21	64	123	186	256	286	324	328	254	201	199	213	2457
	England	4	18	138	414	1178	2390	3239	4047	4347	5077	4637	3243	3406	2788	3122	38,048
Men	Wales							1	1	1	1	2	0	2	1	0	9
	England			1		2	4	5	14	25	29	23	36	45	34	25	243
Total		4	18	141	435	1244	2517	3431	4318	4659	5431	4990	3533	3654	3022	3360	40,757

Incidence varies with ethnicity. Asian, Chinese and Black ethnic groups and those with mixed heritage have a lower incidence than the white ethnic group in England. The rate ratios are 0.65, 0.75, 0.49 and 0.58, respectively, when compared with the white group.[23] Incidence is generally >80 in 100,000 in developed regions compared with <30 per 100,000 in developing regions of the world.[24] Similarly, within the developed world, incidence varies with socioeconomic status. In both England[25] and Wales[26] those who are classed as most deprived have a lower incidence of breast cancer. However, there is some evidence to suggest that the trend for mortality is reversed, with better survival for those from more affluent areas. It is unclear why this is, but may be due to lower levels of screening compliance, worse overall general health status and lower levels of treatment due to limited access to health care[27] and poorer compliance with treatment regimens.

Significance in terms of ill-health (burden of disease)

Breast cancer is a significant cause of death in England and Wales, ranking among the five leading causes when age-standardised rates are compared. It is currently the second biggest cause of cancer death in women after lung cancer, with an age-standardised mortality rate of 26 per 100,000 women. In 2008 this constituted 10,716 deaths for women in England and Wales.[28] Breast cancer deaths have been steadily falling since a peak in 1988.[29] The fall in mortality may be due to screening (earlier detection and more successful treatment), improvements in social awareness, diagnostic techniques and treatment options such as tamoxifen, chemotherapy and more recently trastuzumab.[30,31] As screening programmes detect more cancers at an earlier stage and women live longer after diagnosis, the long-term morbidity associated with treatment, such as lymphoedema, is becoming more significant.

Measurement of disease

Breast cancer has few obvious symptoms and can easily go undetected for many years. Among the more noticeable symptoms is a palpable lump in the breast, a change in breast shape and skin appearance or changes to the nipple such as inversion, a rash or discharge. Women are encouraged to be breast aware, and to seek medical advice if they notice anything unusual.[32] Screening was introduced in the UK in 1988.[33] Currently, women between the ages of 50 and 70 years are routinely invited to attend. Screening is thought to have reduced breast cancer deaths in the 55–69 years age category by an estimated 6.4% in addition to the effects of tamoxifen, chemotherapy and earlier presentation outside of screening.[30] Screening increases the proportion of tumours detected in the early, more curable stages.

A suspicious breast mass may be identified through screening, or via presentation to a general practitioner. The breast mass and axillary areas are investigated clinically through palpation and mammography or ultrasound for younger women, and the status of the tumour confirmed by histology of biopsied tissue. Staging of the disease depends on tumour size, the number of involved lymph nodes and the presence or absence of distant metastases. Tumour size and axillary metastases can be estimated by clinical examination and imaging techniques, but definitive status is achieved through surgery. Those with small tumours and no axillary metastases have the best prognosis, while those with distant metastases are considered incurable.

Current service provision

Current methods for staging of breast cancer

Three main factors are used to stage breast cancer. These are tumour size, metastases to the regional lymph nodes and distant metastases. The tumour/node/metastases (TNM) staging system was developed and is maintained by the American Joint Committee on Cancer and the Union for International Cancer Control.[34,35] T stage is classified according to size of the tumour and degree of local infiltration; N stage is classified according to the number and location of

metastases to the lymph nodes in the axilla, between the ribs (internal mammary nodes) and above or below the collarbone (supraclavicular and infraclavicular nodes); and M stage is classified by the presence of metastases beyond the breast and regional lymph nodes (*Table 4*). The overall TNM stage of the cancer is defined as in *Table 5*.

Early breast cancer is generally defined as cancer which has not spread beyond the breast or the ipsilateral axillary lymph nodes and is confined to stages I, II or IIIA.[36–38]

TABLE 4 Descriptions of T, N and M stages

Stage	Description
T: tumour stage	
Tx	Primary tumour cannot be assessed
T0	No evidence of primary tumour
Tis	Carcinoma in situ
T1	Tumour ≤ 2 cm across
T2	Tumour 2–5 cm across
T3	Tumour > 5 cm across
T4	Tumour of any size with direct extension to skin or chest wall, or inflammatory breast cancer
N: lymph node stage	
Nx	Nodal stage cannot be assessed
N0	No metastases to any ipsilateral lymph nodes
N1	Metastases to 1–3 axillary nodes or axillary nodes that are mobile
N2	Metastases to 4–9 axillary nodes, or axillary nodes that are fixed to one another or other structures, or clinically apparent metastases to internal mammary nodes
N3	Metastasis to nodes above or below collarbone (supraclavicular/infraclavicular), or to both axillary and internal mammary nodes, or to 10 + axillary nodes
M: metastasis stage	
Mx	Presence of metastases cannot be assessed
M0	No distant metastases
M1	Distant metastases

Source: Cancer Research UK[36] and American Cancer Society.[35,39]

TABLE 5 Summary of TNM stages

Stage	T	N	M
0 (DCIS/LCIS)	Tis	N0	M0
I	T1	N0	M0
IIA	T0–1	N1	M0
	T2	N0	M0
IIB	T2	N1	M0
	T3	N0	M0
IIIA	T0–2	N2	M0
	T3	N1–2	M0
IIIB	T4	N0–2	M0
IIIC	T(any)	N3	M0
IV	T(any)	N(any)	M1

Source: Cancer Research UK[40] and American Cancer Society.[39]

Current methods for assessment of axillary metastases

Axillary lymph nodes are the most common site of spread outside the breast (occurring in approximately 41% of cases)[17] and removal of lymph nodes affected by tumour is crucial to prevent recurrence. Assessment of whether cancer has spread to the lymph nodes is also important for staging, assessing prognosis and selection of adjuvant therapy. Clinical examination via palpation has a sensitivity of approximately 46% based on pooled data from a number of studies (in other words, of those patients with metastases, 46% can be detected via clinical examination).[41–46] Therefore, of all patients presenting with breast cancer, approximately 19% (41% × 46%) have axillary metastases that can be detected via clinical examination, while 22% (41% × 54%) have occult axillary metastases.

The following steps (see also *Figure 1*) for assessing the axilla are recommended in the 2009 National Institute for Health and Clinical Excellence (NICE) breast cancer guideline.[17] A clinical examination and an ultrasound scan are carried out. If these examinations suggest nodal metastases on the basis of size or abnormal morphology, ultrasound-guided needle biopsy [either fine-needle aspiration cytology (FNAC) or core biopsy] of abnormal nodes is undertaken, which detects 45% of axillary node metastases.[47] For patients shown by needle biopsy to have nodal metastases, standard management is surgery to remove all the lymph nodes in the axilla, known as axillary lymph node dissection (ALND) or axillary clearance.

As well as being the treatment for those with positive nodes, ALND has been considered the 'gold standard' procedure for staging the axilla. Typically, 10–15 lymph nodes are removed and at least one section from each is assessed via haematoxylin and eosin (H&E) staining. ALND is very accurate in establishing the presence of axillary disease and has the therapeutic advantage of being associated with a high long-term local disease control rate. However, ALND is associated with significant complications, including a 21% incidence of arm lymphoedema,[48–50] a 22% incidence of seromas[48,51] and a 14% infection rate.[48,52] In addition, insertion of a surgical drain during surgery is commonplace (79%) and usually necessitates prolongation of hospital stay.[52]

For patients with no evidence of lymph node involvement on ultrasound, or for whom the ultrasound-guided biopsy is negative, surgery to remove a sample of axillary lymph nodes is recommended, as opposed to full axillary clearance.[17] There are two axillary sampling techniques in current practice: sentinel lymph node biopsy (SLNB) and 4-node sampling (4-NS). SLNB is a procedure to identify and remove the sentinel lymph node, which is the first axillary lymph node to which lymphatic fluid drains from the breast, and therefore the most likely to be affected by metastases. The sentinel node may be identified via blue dye, radioactive isotope or a combination. Sentinel lymph nodes may be examined via H&E staining, and/or immunohistochemistry for epithelial or cytokeratin markers. 4-NS involves a more random surgical removal of a minimum of four lymph nodes (sometimes more) from the lower axilla,

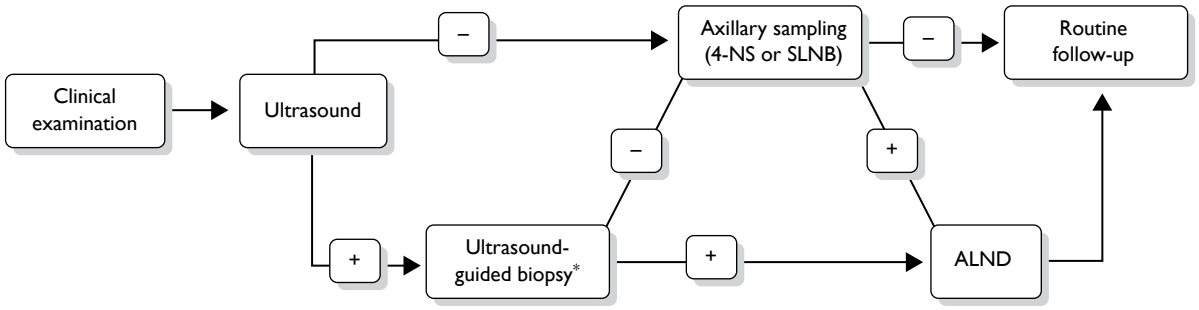


FIGURE 1 Diagnostic pathway for axillary metastases as recommended in the 2009 NICE breast cancer guidelines.[17]
*Either fine-needle aspiration cytology or core biopsy.

and may involve use of blue dye to aid in identification of nodal tissue in the axilla. The false-negative (FN) rate of SLNB and 4-NS have been estimated at between 0% and 10%.[17] As stated in the NICE guideline, data on 4-NS are limited and there are currently insufficient data to compare the diagnostic accuracy of SLNB and 4-NS.[17]

Sentinel lymph node biopsy and 4-NS involve shorter surgical procedures than ALND, and are associated with lower incidence of surgical complications and long-term adverse effects than ALND. Lymphoedema incidence falls from 21%[48–50] for ALND to 7%[53] for SLNB, seroma incidence falls from 22%[48,51] to 7%[48,51], surgical drain requirement from 79% to 2%[52,54] and infection incidence from 14%[48,52] to 2%.[48,52,54]

Patients in whom axillary lymph node metastases are identified via SLNB or 4-NS are advised to undergo ALND as a subsequent procedure to remove all axillary lymph nodes.[17] Some units are investigating the use of intraoperative cytology for assessment of axillary lymph nodes, whereby patients undergo axillary sampling (e.g. via SLNB) and the dissected lymph nodes are assessed while the surgical procedure is still ongoing. If any metastases are identified in the sampled nodes then the procedure is converted to a full ALND, bypassing the requirement for a second surgical procedure. However, it is currently unclear whether immediate or delayed ALND differ significantly in terms of adverse effects. Since intraoperative cytology is not currently used as standard, it is not included in this assessment.

Cost of current methods for assessment of axillary metastases

The costs of clinical examination, ultrasound, and ultrasound-guided biopsy are £86, £53 and £147, respectively[55,56] ALND when carried out as a stand-alone procedure costs £2448.[56] All costs have been adjusted to 2007 prices.

Sentinel lymph node biopsy and 4-NS are normally performed at the same time as the main breast cancer surgery while ALND may be performed at the same time as the breast surgery or as a stand-alone surgical procedure. No UK costs are identified from previous studies regarding the combined costs of SLNB, 4-NS and ALND when performed at the same time as the breast surgery. However, it is widely accepted that SLNB has a significantly higher cost than 4-NS.

Relevant national guidelines

The 2009 NICE guideline *Early and locally advanced breast cancer: diagnosis and treatment* includes recommendations on methods for assessment of the axilla.[17]

Variation in services and uncertainty about best practice

Until recently, ALND was the gold standard technique for assessing axillary lymph node status. A 2006 audit of 271 UK breast surgeons, asked how they would manage a woman with small, clinically node-negative breast cancer, reported that 27% performed ALND, 21% used 4-NS and 52% used SLNB.[57] The 2009 NICE guideline recommends that axillary sampling techniques (SLNB or 4-NS) should be used instead of ALND for patients with clinically node-negative disease, stating that SLNB is the preferred technique.[17] However, there are insufficient data to allow comparison of SLNB and 4-NS, either in terms of diagnostic accuracy or complication rate.[17] Therefore, there is likely to be variation in practice within the UK in the use of ALND, SLNB or 4-NS to assess axillary status.

Description of technology under assessment

Summary of diagnostic tests under assessment (index tests)

This review assesses two imaging techniques: positron emission tomography (PET) and magnetic resonance imaging (MRI).

Positron emission tomography

Positron emission tomography is a nuclear medicine imaging technique that produces a three-dimensional image or map of functional processes in the body. The system detects pairs of gamma rays emitted indirectly by a positron-emitting radioactive tracer, which is introduced into the body attached to a biologically active molecule. Images of tracer concentration in three-dimensional space within the body are then reconstructed by computer analysis. A tracer commonly used for PET scanning is ^{18}F-FDG, which is a glucose analogue (2-fluoro-2-deoxy-D-glucose; FDG) attached to the radioactive isotope fluorine-18. When FDG is injected into a patient, a PET scanner can form images of the distribution of FDG within the body. This provides a picture of the glucose uptake, i.e. metabolic activity, in different areas of the body. As cancer cells frequently have a higher glucose requirement and uptake than many normal body cells, areas of cancerous activity can be detected.[58,59]

The definition of increased uptake on a PET scan may be based on the reader's qualitative visual impression, or more formally by using indices such as the standardised uptake value (SUV) (tissue radioactivity concentration divided by the total injected dose, normalised to body size).[58]

Positron emission tomography/computed tomography

In recent practice, PET scans are commonly undertaken as PET/CT scans, in which a computed tomography (CT) scan is taken alongside the PET scan. Modern PET scanners often perform a reduced-dose CT scan and a PET scan on the patient during the same session, using the same machine. CT scans use a large series of two-dimensional X-rays to generate a three-dimensional image of body structures. This modality combination allows concurrent visualisation of both the anatomy of tissues and their metabolic activity.[58–60] This review includes both studies of PET only and studies of PET/CT.

Cautions and contraindications for positron emission tomography and computed tomography scanning

In terms of safety, PET and CT scanning are non-invasive, but do involve exposure to ionising radiation. The effective radiation dose from a CT scan is approximately 10 millisieverts (mSv), which is about the same as the average person receives from background radiation in 3–4 years.[61] Patients with small children may be advised to limit proximity to them for several hours following the completion of a PET scan. CT scanning is not generally recommended for pregnant women unless it is essential, owing to the potential risk to the baby. Nursing mothers are recommended to wait for 24 hours after contrast material injection before resuming breast-feeding. Serious allergic reaction to contrast materials containing iodine is rare.[61]

Magnetic resonance imaging

Magnetic resonance imaging provides detailed images of the body in any plane. MRI provides much greater contrast between the different soft tissues of the body than does CT. MRI scanning uses a powerful magnetic field to align the nuclear magnetisation of (usually) hydrogen atoms in water in the body. Radiofrequency fields are used to systematically alter the alignment of this magnetisation, causing the hydrogen nuclei to produce a rotating magnetic field detectable by the scanner. This signal can be manipulated by additional, time-varying magnetic field gradients to introduce spatial information that is used to construct two- or three-dimensional images of

the area of interest. Different imaging 'sequences' are used in which the relative timings of these and other sequence components are altered to produce various contrasts between different tissue types, states and pathologies. This is often referred to as defining the degree of proton density, T1-, T2- and T2*-weighting that is present in the resultant set of images.

An MRI scan may provide information on whether a lesion is suspicious for metastasis, based on criteria such as size, morphology and enhancement characteristics following administration of a contrast agent.[62,63]

Variations in method of magnetic resonance imaging

Magnetic resonance imaging techniques may vary in terms of factors such as the field strength, image sequence parameters and type of coil used. Also, a contrast agent is often administered via intravenous injection during the MR procedure. Images obtained post contrast are usually compared with those obtained pre contrast in order to determine where the contrast agent has accrued. Gadolinium-based agents form one class of contrast agent. Another class of contrast agent is ultrasmall super-paramagnetic iron oxide (USPIO), also known as ferumoxtran-10. Dynamic-contrast-enhanced MRI (DCE-MRI) uses repeated imaging to track the entrance of contrast agents into tissue over time. The pattern of uptake over time can be used to assess whether or not cancerous tissue is present.

Magnetic resonance (MR) spectroscopy is used to measure the levels of different metabolites (other than water) in body tissues. Studies of in vitro and in vivo proton MR spectroscopy of the breast have shown that high levels of choline-containing compounds may indicate metastatic lesions.[64]

Cautions and contraindications for magnetic resonance imaging scanning

Magnetic resonance imaging scanning is non-invasive, but use of contrast agents may need to be reviewed for patients with reduced renal function. Different chelating agents may be used to minimise the associated risk. These compounds can have differential effects on resultant image contrast. Allergic reactions to contrast agents have been observed and these agents may be contraindicated in patients with a strong allergic disposition. MRI is contraindicated in people with pacemakers, some artificial heart valves, electronically or magnetically-activated implants and some types of metallic implants and foreign bodies. Full safety guidelines for MRI are available from the Medicines and Healthcare products Regulatory Agency.[65]

Identification of important subgroups

Subgroup analyses were undertaken according to the different methods of PET and MRI, the reference standard test used, and the clinical characteristics of included patients (see *Chapter 2, Decision Problem* for a full list). Sensitivity analyses were also undertaken according to study quality criteria.

Current usage in the NHS

In recent years, imaging techniques such as PET and MRI have been increasingly used for diagnosis and staging in various types of cancer. However, at present they are not routinely used for staging the axilla in breast cancer.

Anticipated costs associated with index tests

The cost of MRI is £232.[56] The cost of PET is £978 according to a study based on UK hospital.[66] All costs have been adjusted to 2007 prices.

Chapter 2

Definition of the decision problem

Decision problem

Population and relevant subgroups

Population

The relevant population consists of patients newly diagnosed with early-stage invasive primary breast cancer.

In this review, early-stage breast cancer is defined as TNM stage I, II or IIIA. The included studies recruited patients before full staging investigations had been undertaken, as would be the case in practice, and therefore tended to include patients with a range of cancer stages. Where sufficient data were reported in the primary study, patients with advanced, metastatic or recurrent cancer were excluded from the analysis. Patients with carcinoma in situ (DCIS or LCIS) were also excluded where possible, as these patients do not generally undergo diagnostic axillary surgery. Where separate data were not presented, studies were included if at least 80% of the study population had early-stage, newly diagnosed breast cancer (a sensitivity analysis was undertaken including early-stage patients only). Only patients with histologically confirmed breast cancer were included in this analysis.

Since there are several studies of PET in this setting with large sample sizes, PET studies with < 20 analysable patients were excluded. MRI studies of all sizes were included as there are few with a large sample size.

Subgroups

Subgroup analyses were undertaken according to the following variables:

- PET alone or PET/CT
- MRI using different contrast agents and methods
- criteria for defining a node as metastatic
- reference standard test used
- whether the included patients were all early stage and newly diagnosed
- size of axillary metastases and nodal stage
- prevalence of patients with axillary metastases within the study
- clinical axillary nodal status (positive or negative)
- study quality.

Diagnostic tests under assessment (index tests)

The index tests assessed in this review are:

- PET, including:
  - PET alone
  - PET/CT

- MRI, including:
  - gadolinium-enhanced MRI
  - dynamic gadolinium-enhanced MRI
  - USPIO-enhanced MRI
  - MR spectroscopy.

Reference standard tests (comparator tests)

The most relevant reference standard tests were considered to be ALND, SLNB and 4-NS. ALND is the 'gold standard' method of staging the axilla. SLNB and 4-NS were also thought to be acceptable comparators, since these techniques are now recommended by NICE and are not associated with a survival detriment in the long term.[17] Studies using other reference standard tests, or where the reference standard is not clear, were included, but a sensitivity analysis excluding these studies was also undertaken.

Outcomes

Relevant outcomes include:

- sensitivity and specificity [or data required to calculate these, i.e. numbers of true-positive (TP), false-negative (FN), true-negative (TN) and false-positive (FP) results]
- adverse effects and withdrawals
- health-related quality of life
- cost-effectiveness and cost–utility.

Studies were only included if they report the numbers of TP, FN, TN and FP results for PET or MRI scanning in comparison with a reference standard test. These values can be used to calculate measures of diagnostic accuracy such as sensitivity and specificity.

Patients are classified by the index test (the diagnostic test being investigated) as either positive (axillary metastases) or negative (no axillary metastases). The reference standard (an established diagnostic test) is also undertaken to identify patients' true health status. The reference standard is assumed to have 100% sensitivity and specificity; however, subgroup analyses are undertaken to test the effect of using different reference standards. Patients fall into one of four groups. Where the index test is positive, patients may be TP where both tests agree that they have metastases, or FP where the index test indicates that they have metastases but the reference standard does not. Where the index test is negative, patients may be TN where both tests agree they are metastasis free, or FN where the index test incorrectly classifies them as metastasis free. This can be represented in a 2×2 table (*Table 6*). In the clinical setting, FPs can result in patients receiving unnecessary treatment, while FNs can result in people not receiving treatment they require. Sensitivity indicates the effectiveness of the index test in correctly identifying metastases (TPs divided by all persons with metastases). Specificity indicates the effectiveness of the index test in correctly classifying people as metastasis free (TNs divided by all persons without metastases). Sensitivity and specificity can be calculated as simple percentages, as shown in *Table 6*. In practice, diagnostic tests often have a high sensitivity at the expense of a low specificity and vice versa. Ideally, a test would have both high sensitivity and high specificity.

Study design

Studies of a cohort design (prospective or retrospective) were included. Studies that were both prospective and consecutive were examined in a separate sensitivity analysis as part of the assessment of study quality. Case–control studies (where the test is evaluated in a group of patients already known to have the outcome and a separate group of patients without the outcome) were excluded; however, no studies of this type were identified within this review.

TABLE 6 Calculation of sensitivity and specificity

Index test result	Reference standard-positive	Reference standard-negative
Index test positive	TP	FP
Index test negative	FN	TN
	Sensitivity = [TP/(TP + FN)] × 100	Specificity = [TN/(TN + FP)] × 100

Overall aims and objectives of assessment

The aim of this assessment was to assess the diagnostic accuracy, cost-effectiveness, and effect on patient outcomes of PET and MRI in the evaluation of axillary lymph node metastases in patients with newly diagnosed, early-stage invasive primary breast cancer. PET and MRI are assessed, firstly, as a replacement for SLNB or 4-NS and, secondly, as an additional test prior to SLNB or 4-NS.

The objectives of the assessment are:

- To conduct a systematic review of the published evidence on the diagnostic accuracy and cost-effectiveness of PET and MRI for the assessment of axillary lymph node metastases in newly diagnosed, early-stage breast cancer.
- To develop a decision model to investigate the benefits, harms, and cost-effectiveness of PET and MRI, either as a replacement for surgical assessment of the axilla or as an additional test in the diagnostic pathway for assessing the axilla. Outcomes from the model will be expressed in terms of net health benefit and cost per quality-adjusted life-year (QALY).

Chapter 3

Assessment of clinical effectiveness

Methods for reviewing effectiveness

A systematic review was undertaken according to the general principles recommended in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.[67,68]

Identification of studies

Search strategy

The search strategy comprised the following elements:

- searching of electronic databases
- scrutiny of bibliographies of retrieved papers and previous reviews
- contact with experts in the field.

The MEDLINE search strategy is included in *Appendix 1*, and comprised medical subject heading (MeSH) terms and free-text terms as follows: terms for breast cancer, terms for PET and MRI, terms for the axilla or lymph nodes and terms to identify diagnostic studies. No restrictions were used according to language or date of publication. Searches were undertaken in April 2009.

Databases

The following databases were searched:

- MEDLINE
- MEDLINE In-Process & Other Non-Indexed Citations
- EMBASE
- Cumulative Index to Nursing and Allied Health Literature (CINAHL)
- Cochrane Library including Cochrane Database of Systematic Reviews (CDSR), Cochrane Central Register of Controlled Trials (CENTRAL), Database of Abstracts of Reviews of Effects (DARE), NHS Economic Evaluation Database (NHS EED) and HTA databases
- Science Citation Index (via Web of Science)
- BIOSIS Previews
- National Research Register archive (www.nrr.nhs.uk; searched until 2007)
- UK NIHR Clinical Research Network (www.ukcrn.org.uk; searched post-2007)
- ClinicalTrials.gov (www.clinicaltrials.gov)
- current controlled trials (www.controlled-trials.com)
- American Society of Clinical Oncology abstracts (conference proceedings)
- European Society for Medical Oncology abstracts (conference proceedings).

Titles and abstracts were screened for inclusion by two reviewers. Full text relevant papers were screened against the inclusion criteria by two reviewers and any disagreements resolved by consensus.

Inclusion and exclusion criteria

Inclusion criteria

Studies were included if they assessed the diagnostic accuracy of PET or MRI for use in the assessment of axillary metastases in patients newly diagnosed with early-stage invasive primary breast cancer. Studies were only included if they reported the numbers of TP, FN, TN and FP results for PET or MRI scanning in comparison with a reference standard test.

Exclusion criteria

As there are several studies of PET in this setting with large sample sizes, PET studies with < 20 analysable patients were excluded, since they were thought to add little to the overall estimates of accuracy. MRI studies of all sizes were included as there are few with a large sample size. Studies were also excluded if > 20% of the study population had breast cancer that was non-early stage, non-newly diagnosed or DCIS. Animal models, pre-clinical and biological studies, narrative reviews, editorials, opinion papers and non-English-language papers were excluded. Case–control studies were excluded (although no studies of this type were identified within this review).

Data extraction strategy

Data were extracted by one reviewer using a standardised data extraction form and checked by a second reviewer. Discrepancies were resolved by discussion.

Critical appraisal strategy

Study quality was considered with the aid of the quality assessment of diagnostic accuracy studies (QUADAS) checklist.[69] Quality assessment was performed by one reviewer and checked by a second. The quality assessment criteria were scored as detailed in *Appendix 2*. Three items from the published checklist were not used within our assessment, as follows. The 'description of selection criteria' item was omitted as this was thought to be covered by the 'representative patient spectrum' item, where studies with insufficient description of the selection criteria scored 'unclear'. The 'partial verification bias' item was omitted because, in almost all studies included in this review, all patients received a reference standard. The 'incorporation bias' item was also omitted because, in this review, the reference standard was always independent of the index test (it was noted by Whiting *et al.*[69] that the above two items are not relevant in every review).

Methods of data synthesis

Sensitivity and specificity are presented for each study. Meta-analysis was undertaken to calculate a mean sensitivity and specificity across studies. Sensitivity and specificity are linked, so that changing the threshold at which a test is considered positive will tend to increase the sensitivity but decrease the specificity, or vice versa. Therefore, sensitivity and specificity were meta-analysed using a bivariate random effects method within STATA (StataCorp©, College Station, TX, USA). This approach assumes a bivariate normal distribution for the logits of sensitivity and specificity, which allows the correlation between them to be accounted for in the meta-regression model; covariates may be used to adjust the (marginal) logits of both sensitivity and specificity.[70,71] Where significant heterogeneity was observed, the random effects method was used in order to account for variation both within and between studies. Forest plots and receiver operating characteristic (ROC) plots were generated within REVIEW MANAGER (version 5.0, Cochrane Collaboration©, Copenhagen, Denmark).[72] To explore possible sources of bias, all study quality variables were added as covariates in univariate regression models for sensitivity and specificity for PET and MRI, in order to test whether any variables had a significant effect ($p < 0.10$) on sensitivity or specificity.

Results

Quantity and quality of research available

The search identified 658 citations (646 from the literature search and 12 from other sources such as relevant reviews; *Figure 2*). Of these, 520 were excluded at the title/abstract stage and 138 were obtained for examination of the full text. Ninety-three citations were excluded at the full text stage (*Appendix 3*). In total, 45 citations relating to 35 studies were included in the review: 26 studies of PET and nine studies of MRI.

Study characteristics

Positron emission tomography studies

The characteristics of the 26 included studies of PET for assessment of the axilla are described in *Table 7*. There were seven studies assessing PET/CT[73–79] and 19 studies assessing PET alone.[44,80–97]

Eight PET studies used ALND as the reference standard for all patients,[44,73,77,84,87,89,92,95,98–101] 12 studies used a mixture of ALND and SLNB,[74–76,78–83,86,88,91,102–105] and six studies did not specify the reference standard or used a method other than ALND/SLNB for some of the patients. Nine studies recruited patients both prospectively and consecutively.[85,90,93,94,96,97,106,107] The number of analysed patients (relevant to this review) ranged from 24 to 308. Fourteen studies presented data such that the patients analysed in this review were entirely early stage (stage I, II or IIIA), newly diagnosed and non-DCIS,[44,75,76,80,83–89,92,94,97] while the remaining 12 studies comprised up to 20% of patients who did not meet these criteria.[73,74,77–79,81,82,90,91,93,95,96] Eight studies consisted entirely of patients who were clinically node negative,[73,76,79,81,83,84,87,88] 13 studies included a mixture of node-negative and node-positive patients,[44,74,77,78,80,85,86,91–94,96,97] and in five studies nodal status was not reported.[75,82,89,90,95] The mean age of the included patients ranged from 49 to 67 years (mean across studies was 56 years), and the majority of patients were female. The prevalence of axillary

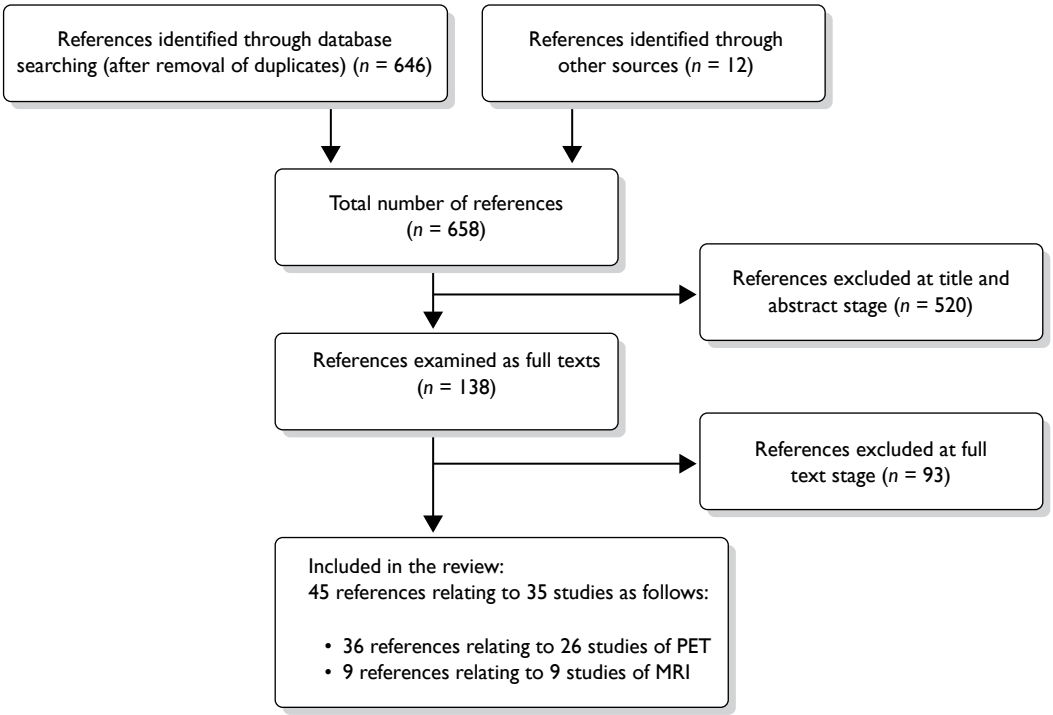


FIGURE 2 Preferred reporting items for systematic reviews and meta-analyses flow chart of included and excluded studies.

TABLE 7 Characteristics of included PET studies

Study	Country	Index test	Reference standard	Prospective/ retrospective?	Consecutive?	*n* met criteria?	*n* analysed	Mean age (range) (years)	Gender	Cancer stage	Clinical nodal status	Prevalence of axillary metastases (%)	Confirmation of breast cancer	Other inclusion and exclusion criteria
Studies of PET/CT														
Chae 2009[73]	South Korea	PET/ CT	100% ALND (plus SLNB)	Retrospective	NR	108	108	49 (27–75)	NR	T1 = 71% T2 = 27% T3 = 2% N0 = 69% N1 = 17% N2 = 9% N3 = 5%	100% negative; micrometastases/ ITCs excluded	31	CNB or FNAC	NR
Heusner 2009[74]	Germany	PET/ CT	ALND and/ or SLNB	Retrospective	Consecutive	54	54	56 (28–78)	Female	T1 = 44% T2 = 56% N0 = 59% N1 = 26% N2 = 7% N3 = 7%	Positive and negative (% NR)	41	Histopathology	NR
Kim 2009[75]	South Korea	PET/ CT	ALND and/ or SLNB (plus non-SLNB)	Prospective	Consecutive	137	137	51 (27–85)	99% female	T1/T2 = 100%, no DCIS	NR	26	Biopsy	*Exclusion*: diabetes, neoadjuvant chemotherapy, excisional biopsy
Taira 2009[76]	Japan	PET/ CT	ALND (if PET or SLNB positive) and/or SLNB	Retrospective	NR	92 axilla	92 axilla	55 (21–82)[a]	Female	T1 = 68% T2 = 29% T3 = 2%, no DCIS	100% negative	29	Six EB; remainder NR	*Exclusion*: prior chemotherapy, hormone therapy, radiotherapy

Study	Country	Index test	Reference standard	Prospective/ retrospective?	Consecutive?	*n* met criteria?	*n* analysed	Mean age (range) (years)	Gender	Cancer stage	Clinical nodal status	Prevalence of axillary metastases (%)	Confirmation of breast cancer	Other inclusion and exclusion criteria
Fuster 2008[77]	Spain	PET/ CT	100% ALND	Prospective	Consecutive	60	52	57 ± 13	NR	T1 = 0% T2 = 21% T3 = 68% T4 = 11% (all primary tumour > 3 cm)	97% negative, 3% positive	38	CNB	*Exclusion*: IBC, breast surgery, chemotherapy or radiotherapy, pregnant, diabetes, age < 18 years
Ueda 2008[78]	Japan	PET/ CT	ALND and/ or SLNB	Prospective	Unclear (states 'series')	183	183	57 (32–81)	NR	Tis = 5% T1 = 50% T2 = 37% T3 = 8%	Positive and negative (% NR)	32	CNB	*Exclusion*: distant metastases, systemic therapy, excisional biopsy, diabetes, pregnancy
Veronesi 2007[79]	Italy	PET/ CT	ALND (if PET or SLNB positive) and/or SLNB	Prospective	Consecutive	236	236	49 (24–79)[a]	99.6% female	T1 = 58% T2 = 37% T3 = 6%, no DCIS N0 = 56% N1 = 32% N2 = 8% N3 = 4%	100% negative	44	FNAC, CNB or EB	*Exclusion*: neoadjuvant chemotherapy
Studies of PET only														
Cermik 2008,[80] Kumar 2006[102]	USA	PET only	ALND and/ or SLNB	Prospective	Consecutive	188 axilla	188 axilla	51 (24–80)	Female (confirmed by author)	N0 = 61% N1 = 31% N2 = 8%	Positive and negative (% NR)	39	CNB 67.5%, EB 32.5%	NR

continued

TABLE 7 Characteristics of included PET studies *(continued)*

Study	Country	Index test	Reference standard	Prospective/ retrospective?	Consecutive?	*n* met criteria?	*n* analysed	Mean age (range) (years)	Gender	Cancer stage	Clinical nodal status	Prevalence of axillary metastases (%)	Confirmation of breast cancer	Other inclusion and exclusion criteria
Gil-Rendo 2006,[81] Zornoza 2004[103]	Spain	PET only	ALND (first *n*=150 and next *n*=125 if PET/SLNB-positive) and/or SLNB	Prospective	Unclear (states 'series')	275	275	51 (24–87)	Female	Stage I–II=100% T1=49% extensive DCIS=4% neoadjuvant chemotherapy =2%	100% negative	52	CNB	*Exclusion*: stage III/ IV, biopsy or surgery to breast or axilla, uncontrolled diabetes
Weir 2005[82]	Canada	PET only	ALND and/ or SLNB	Retrospective	NR	40	40	52 (30–88)[a]	Female	Stage not reported. Newly diagnosed	NR	45	Histology (no further detail)	*Exclusion*: non-newly diagnosed, axillary surgery, chemotherapy
Agresti 2004,[83] Agresti 2001[104]	Italy	PET only	ALND (if PET or SLNB-positive) and/or SLNB	NR	NR	71	71	55 (24–78)	NR	T1=86% T2=14%	100% negative	44	NR	*Exclusion*: primary tumour >2.5 cm
Fehr 2004[84]	Switzer-land	PET only	100% ALND (plus SLNB)	NR	NR	24	24	56±10.8	NR	T1=14 (58%) T2=9 (38%) T3=1 (4%)	100% negative	42	FNAC	*Exclusion*: tumour >3 cm, IBC, multifocality, pregnancy, lactation, diabetes, radiotherapy, breast/axilla surgery
Inoue 2004,[85] Yutani 2000,[106] Yutani 1999[107]	Japan	PET only	All confirmed by histology, 'almost all' ALND (via author)	Retrospective (via author)	NR	81	81	53 (32–78)[a]	Female	T1=41% T2=46% T3=14% N0=23% N1–N2=77%	23% negative, 77% positive	43	Histology after surgery (via author)	*Exclusion*: distant metastases

Study	Country	Index test	Reference standard	Prospective/ retrospective?	Consecutive?	*n* met criteria?	*n* analysed	Mean age (range) (years)	Gender	Cancer stage	Clinical nodal status	Prevalence of axillary metastases (%)	Confirmation of breast cancer	Other inclusion and exclusion criteria
Lovrics 2004,[86] Lovrics 2002[105]	Canada	PET only	ALND and/ or SLNB	Prospective	Consecutive	115	90	56 (SD 11)	Female	T1 = 81% T2 = 18% T3 = 1%	Positive and negative (% NR)	28	Histology: 14 EB, NR for remainder	*Exclusion*: stage III/ IV, multiple, multicentric, IBC, male, prior ALND, pregnancy, uncontrolled diabetes
Wahl 2004[44]	USA	PET only	100% ALND (some SLNB also)	Prospective	Consecutive (via author)	330 axilla	308 axilla	52 (27–82)	Female	T1 = 65% T2 = 28% T3 = 2% Tx = 1% missing = 4%	92% negative, 8% positive	35	NR	*Exclusion*: non-invasive, distant metastases, diabetes, infections, serious organ dysfunction, neoadjuvant chemotherapy, SLNB only
Barranger 2003[87]	France	PET only	100% ALND (plus SLNB)	Prospective	Consecutive	32	32	58 (29–77)	NR	T0 = 28% T1 = 56% T2 = 16%	100% negative	47	FNAC or CNB	*Exclusion*: neoadjuvant chemotherapy, pregnancy, diabetes
Guller 2002[88]	Switzer-land	PET only	ALND and/ or SLNB	Prospective	Consecutive	31	31	65 (47–88)	Female	T1 = 61% T2 = 39%	100% negative	45	Histopathology (no further detail)	*Exclusion*: pregnancy, diabetes, inability to lie still in PET scanner
Nakamoto 2002[89]	USA	PET only	100% ALND	Prospective	NR	30	30	51 (28–78)	94% female	T1 = 67% T2 = 37% T3 = 7%	NR	43	Histology (no further detail)	NR

continued

TABLE 7 Characteristics of included PET studies *(continued)*

Study	Country	Index test	Reference standard	Prospective/ retrospective?	Consecutive?	*n* met criteria?	*n* analysed	Mean age (range) (years)	Gender	Cancer stage	Clinical nodal status	Prevalence of axillary metastases (%)	Confirmation of breast cancer	Other inclusion and exclusion criteria
Rieber 2002[90]	Germany	PET only	Histology (no further details)	NR	NR	43	40	53 (27–84)	NR	DCIS = 3% Tmic = 3% T1 = 21% T2 = 49% T3 = 10% T4 = 15%	NR	50	Histology (no further detail)	NR
van der Hoeven 2002[91]	Nether-lands	PET only	ALND and/ or SLNB (depending on tumour size)	Prospective	Unclear	70	70	58 (SD 13)	NR	T0 = 6% T1 = 53% T2 = 26% T3 = 6% T4 = 4% Unknown = 6%	71% negative, 29% positive	46	EB 17%, FNAC 83%	*Exclusion*: diabetes
Greco 2001,[92] Crippa 1998[98] Bombardieri 1998,[99] Crippa 1997[100]	Italy	PET only	100% ALND	Prospective	Consecutive	167	167	54 (28–84)	NR	T1 = 59% T2 = 41%	77% negative, 23% positive	43	Histology (no further detail)	*Exclusion*: primary tumour > 5 cm, abnormal blood glucose
Noh 1998[93]	South Korea	PET only	Histology (no further details)	NR	NR	31 axilla	27 axilla	NR	NR	Stage not reported DCIS = 12%; not clear if included in analysis	70% negative, 30% positive	56	Histology (no further detail), 1 FNAC	NR
Smith 1998[94]	UK	PET only	ALND 90%; FNAC 10% (large/locally advanced)	NR	NR	38	38	67 (26–89)	Female	T1 = 21% T2 = 55% T3 = 24%	70% negative, 30% positive	42	FNAC	*Exclusion*: age < 18 years, pregnant, diabetes, unable to lie still on PET scanner

Study	Country	Index test	Reference standard	Prospective/ retrospective?	Consecutive?	*n* met criteria?	*n* analysed	Mean age (range) (years)	Gender	Cancer stage	Clinical nodal status	Prevalence of axillary metastases (%)	Confirmation of breast cancer	Other inclusion and exclusion criteria
Adler 1997,[95] Adler 1996[101]	USA	PET only	100% ALND	Part of prospective study	NR	54 axilla	52 axilla	36–79	NR	Benign = 2% T1 = 60% T2 = 33% T3 = 6%	NR	38	NR	*Exclusion*: primary tumour < 0.5 cm, < level II ALND, < 10 nodes dissected, age < 30 years, prior ALND, neoadjuvant chemotherapy
Avril 1996[96]	Germany	PET only	ALND 90%; clinical examination 10% (locally advanced)	NR	NR	41	41	50 (18–74)	Female	T1 = 44% T2 + = 56% Locally advanced = 10% Distant or non-axillary metastases ≥ 12%	Positive and negative (% NR)	59	Histology (no further detail) after surgery	*Exclusion*: age < 18 years, pregnancy, diabetes
Utech 1996[97]	USA/ Germany	PET only	ALND 44%; modified radical mastectomy 56%	Prospective	NR	124	124	59 (32–94)	NR	T1 = 67% T2 = 29% T3 = 4% N0 = 64% N1 = 35% N2 = 2%	92% negative, 8% positive	35	CNB 42%, EB 54%, partial mastectomy 4%	NR

CNB, core needle biopsy; EB, excisional biopsy; IBC, inflammatory breast cancer; ITC, isolated tumour cell; NR, not reported; SD, standard deviation.
a Median age (range).

metastases (as measured via the reference standard) ranged from 26% to 59%, the average across studies being 41%. Further details of the methods of PET scanning and the reference standard are described in *Appendix 4*.

Magnetic resonance imaging studies

The characteristics of the nine[41,42,64,108–113] included studies of MRI for assessment of the axilla are described in *Table 8*. There were five studies of USPIO-enhanced MRI,[108–112] two studies of dynamic gadolinium-enhanced MRI,[41,113] one study of (non-dynamic) gadolinium-enhanced MRI,[42] and one study of in vivo proton MR spectroscopy.[64] Several of the studies reported more than one set of results on diagnostic accuracy, according to different criteria for defining whether axillary metastases were present (e.g. based on USPIO or gadolinium uptake, size, morphology, or combinations of these criteria).

Eight MRI studies used ALND as the reference standard for all patients,[41,42,64,109–113] while one study used a mixture of ALND and SLNB.[108] Five studies recruited patients both prospectively and consecutively.[64,108–110,112] The number of analysed patients (relevant to this review) ranged from 10 to 75. Three studies presented data such that the patients analysed in this review were entirely early stage (stage I, II or IIIA), newly diagnosed and non-DCIS,[108,110,113] while the remaining six studies comprised up to 20% of patients who did not meet these criteria.[41,42,64,109,111,112] One study consisted entirely of patients who were clinically node negative,[108] two studies included a mixture of node-negative and node-positive patients,[41,64] and in six studies nodal status was not reported.[42,109–113] The mean age of the included patients ranged from 53 to 66 years (mean across studies was 59 years), and the majority of patients were female. The prevalence of axillary metastases (as measured via the reference standard) ranged from 20% to 70%, the average across studies being 45%. Further details of the methods of MRI scanning and the reference standard are described in *Appendix 4*.

Study quality

Figures 3 and *4* provide an overview of the methodological quality of the 35 included studies.[41,42,44,64,73–97,108–113] Of the PET studies, quality items that scored poorly overall were representative patient spectrum, availability of relevant clinical information, handling/reporting of uninterpretable results and interpretation of the reference standard with blinding to index test results. Patient spectrum scored negatively either because the study included up to 20% participants who were not early stage or newly diagnosed (12 studies)[73,74,77–79,81,82,90,91,93,95,96] or because the patients were not recruited both prospectively and consecutively (two studies).[76,85] It was unclear if participants were recruited prospectively and consecutively in 12 cases.[73,78,81,83,84,89,91,93–97] The index test was interpreted blind to reference standard results in most studies. The index test was often interpreted blind to other clinical data in addition to the reference standard, possibly to ensure a robust evaluation of PET; however, as this is likely to differ from expected clinical practice, studies where this occurred were scored negatively for the 'relevant clinical information' item. Uninterpretable results were dealt with well in only eight studies,[75,77,79,80,86–88,92] and were not discussed in 16 cases.[73,74,76,78,81–85,87,89–91,93,94,96] Similarly, blinding of the reference standard results was under-reported, with 20 studies scoring 'unclear'.[73–75,77,78,80–85,87–89,92–96] It is difficult to know what impact these have on study quality and transferability to real life practice. The reference standard was adequate in nearly all cases, with only three studies failing to give sufficient detail,[85,90,93] and two not performing ALND on those patients with large or locally advanced disease.[94,96] The delay between reference standard and index test was acceptable in 14 studies[44,76,78,79,83–88,91,92,94,95] and not reported in 12.[73–75,77,80–82,87,89,90,93,96] Where patients were given a different reference standard depending on the index test results (six studies),[75,76,78,79,81,83] this was either ALND or SLNB. SLNB is thought to have slightly lower accuracy so differential verification bias may occur in these studies. The index test was usually very well described, while the reference standard was often only partially described, probably due

TABLE 8 Characteristics of included MRI studies

Study	Country	Index test	Reference standard	Prospective/ retrospective?	Consecutive?	*N* met criteria?	*N* analysed	Mean age (range) (years)	Gender	Cancer stage	Clinical nodal status	Prevalence of axillary metastases (%)	Confirmation of breast cancer	Other inclusion and exclusion criteria
Kimura 2010[108]	Japan	USPIO-enhanced	ALND and/or SLNB	Prospective	Consecutive	10	10	66 (35 to 79)	Female	100% clinically T2 N0 M0 (stage IIA)	100% negative	20	Pathology (no further detail)	*Exclusion*: strong allergic disposition, liver dysfunction
Harada 2007[109]	Japan	USPIO-enhanced	100% ALND	Prospective	Consecutive	33	33	58 (36–77)	97% female	Stage II = 73% Stage IIIA = 24% Stage IIIB = 3%	NR	70	Pathology (no further detail)	*Exclusion*: stage I, strong allergic disposition, liver dysfunction
Memarsadeghi 2006[110]	Austria	USPIO-enhanced	100% ALND	Prospective	Consecutive	24	22	60 (40–79)	Female	T1 = 59% T2 = 41%	NR	27	CNB	*Exclusion*: contraindication to MRI, allergy to dextran or iron salts, chemotherapy or radiotherapy, no ALND, pregnancy, lactation, unable to cooperate, other trial, under care of guardian
Stadnik 2006[111]	Belgium	USPIO-enhanced	100% ALND	Prospective	NR	10	10	56 (41–74)	Female	Stage not reported. Included pts scheduled for mastectomy	NR	50	NR	*Exclusion*: not scheduled for mastectomy, contraindication for MRI, strong allergic disposition to gadolinium, dextrans or iron salts, unable to obtain PET (for technical or accessibility reasons)
Michel 2002[112]	Switzerland	USPIO-enhanced	100% ALND	Prospective	Consecutive	18	18	53 (22–76)	Female	T1 = 56% T2 = 39% T4 = 6%	NR	61	Cytology 95%, histology 5%	*Exclusion*: strong allergic disposition, contraindication to MRI
Murray 2002[113]	UK	Dynamic gadolinium-enhanced	100% ALND	NR	NR	47	47	63 (50–87)	Female	T1/T2 = 100%	NR	21	Histology (no further detail)	*Exclusion*: primary tumour < 0.5 cm or > 3.1 cm.

continued

TABLE 8 Characteristics of included MRI studies *(continued)*

Study	Country	Index test	Reference standard	Prospective/ retrospective?	Consecutive?	*N* met criteria?	*N* analysed	Mean age (range) (years)	Gender	Cancer stage	Clinical nodal status	Prevalence of axillary metastases (%)	Confirmation of breast cancer	Other inclusion and exclusion criteria
Kvistad 2000[41]	Norway	Dynamic gadolinium-enhanced	100% ALND	NR	NR	67	65	59 (38–79)	NR	T1 = 58% T2 = 31% T3/T4 = 11% (neoadjuvant chemotherapy)	Positive and negative (% NR)	37	Histology or FNAC	NR
Mumtaz 1997[42]	UK	Gadolinium-enhanced	100% ALND	NR	NR	92 axilla	75 axilla	49 (29–80)[a]	NR	T1 = 11% T2 = 72% T3 = 3% T4 = 3% Tx = 11% DCIS = 4%	NR	53	FNAC 90%, CNB 10% (if equivocal)	NR
Yeung 2002[64]	Hong Kong	MR spectroscopy	100% ALND	Prospective	Consecutive	32	27	53 (26–82)	NR	Stage not reported	52% negative 48% positive	63	CNB	*Exclusion*: receiving chemotherapy

CNB, core needle biopsy; NR, not reported.
a Median age (range).

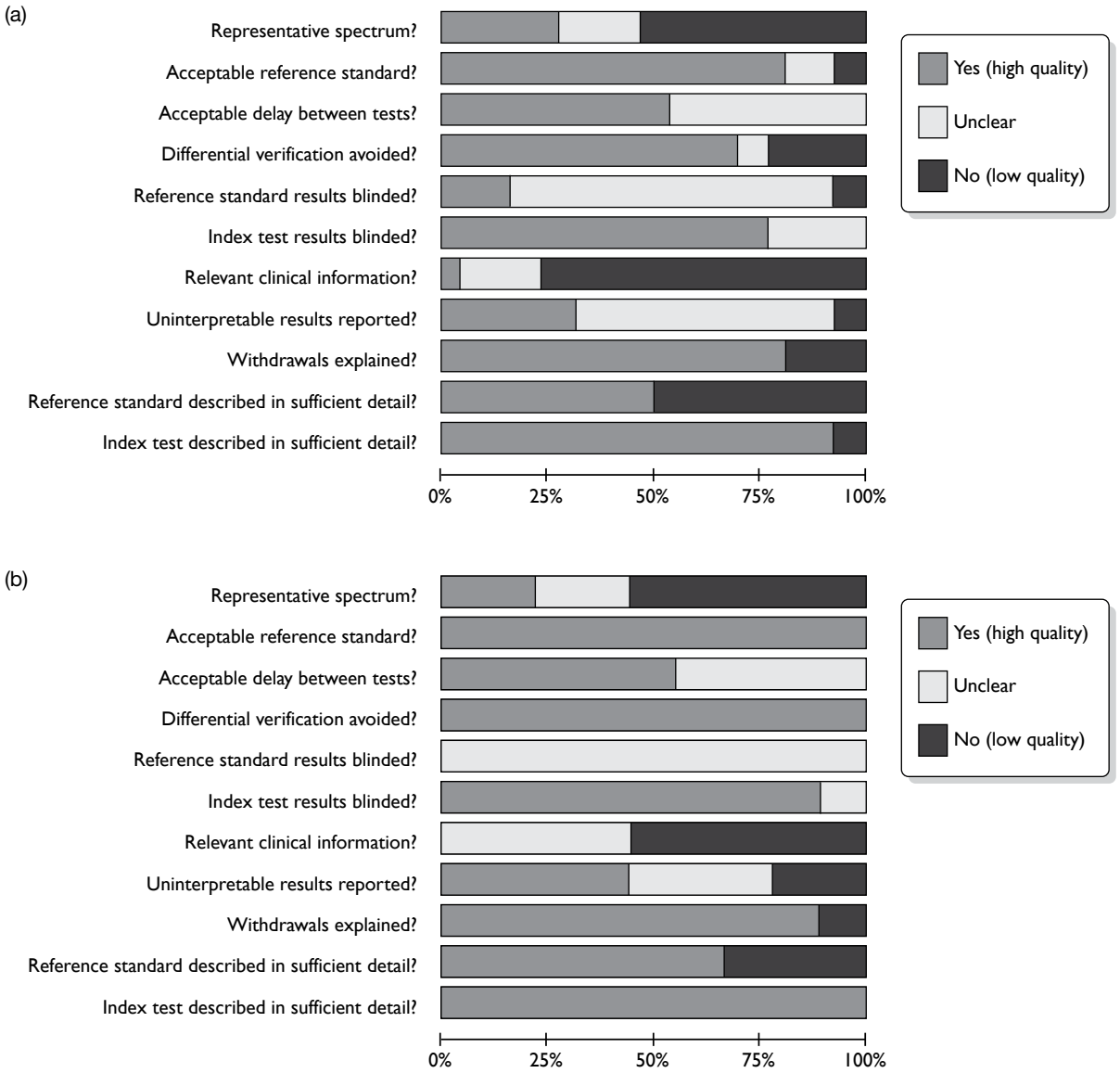


FIGURE 3 Methodological quality: summary across all studies. (a) PET studies; (b) MRI studies.

to the widespread routine use of these techniques. Most studies did not have any withdrawals to explain, so this item scored well overall.

All nine MRI studies used an acceptable reference standard and avoided differential verification bias.[41,42,64,108–113] All described the index test in detail, probably due to the novel nature of the techniques, and usually described the reference standard well. Withdrawals were explained or did not occur in eight cases.[41,42,64,108,109,111–113] However, as with PET studies, the spectrum of patients was mixed, with over half not recruiting only early stage and newly diagnosed, or not recruiting prospectively or consecutively. Relevant clinical information was not available to the image interpreter in five cases[41,42,110,111,112] and not reported in the other four,[64,108,109,113] and uninterpretable results were dealt with well by only four studies.[108–110,112] None of the studies reported whether the reference standard results were interpreted blindly.

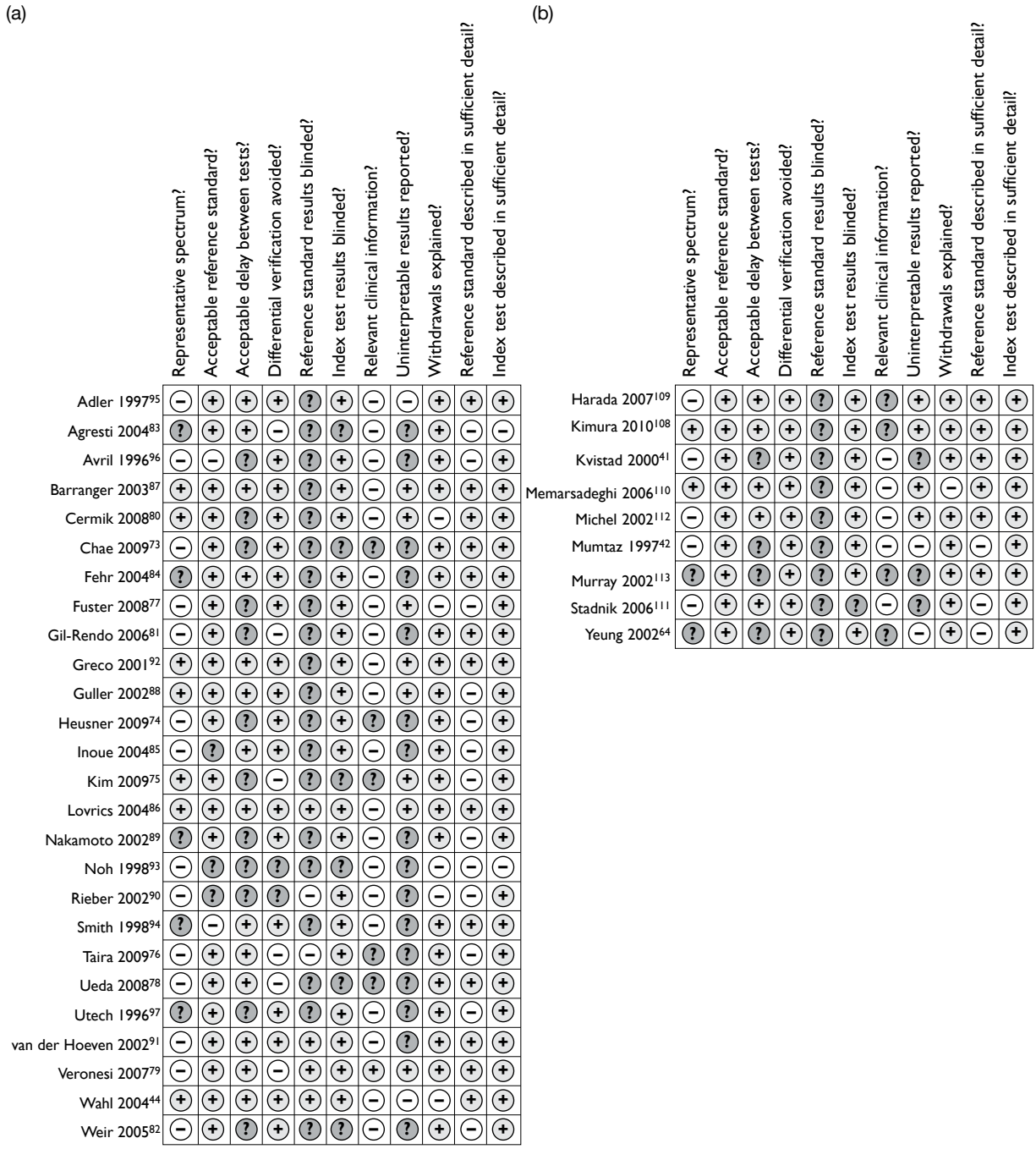


FIGURE 4 Methodological quality: summary for each individual study. (a) PET studies; (b) MRI studies.

Assessment of diagnostic accuracy

Positron emission tomography studies: diagnostic accuracy results

Summary findings

Across 26 studies of PET or PET/CT ($n = 2591$ patients),[44,73–97] the mean sensitivity was 63% [95% confidence interval (CI) 52% to 74%; range 20%–100%] and the mean specificity was 94% (95% CI 91% to 96%; range 75%–100%) (*Table 9* and *Figures 5* and *6*). For the seven studies of PET/CT ($n = 862$),[73–79] the mean sensitivity was 56% (95% CI 44% to 67%) and the mean specificity was 96% (95% CI 90% to 99%). For the 19 studies of PET only ($n = 1729$),[44,80–97] the mean sensitivity was 66% (95% CI 50% to 79%) and the mean specificity was 93% (95% CI 89%

TABLE 9 Summary of pooled sensitivities and specificities for PET studies

Diagnostic test	No. of studies	No. of patients	Sensitivity, % (95% CI)	Specificity, % (95% CI)
All PET studies				
All PET studies[44,73–97]	26	2591	63 (52 to 74)	94 (91 to 96)
PET studies with or without CT				
PET/CT[73–79]	7	862	56 (44 to 67)	96 (90 to 99)
PET (no CT)[44,80–97]	19	1729	66 (50 to 79)	93 (89 to 96)

to 96%). Therefore PET/CT gave a slightly lower mean sensitivity than PET only. The reason for this is not clear, as a concurrent CT scan is generally thought to enhance the accuracy of PET. As sensitivity varied widely between studies, this finding may have been due to chance. It may also have been due to differences in populations or study methods; for example, PET/CT studies used ALND/SLNB for all patients, whereas some PET-only studies used another reference standard for some patients, which may have led to overestimates of sensitivity (see *Effect of reference standard*). Also, three of seven PET/CT studies[73,76,79] (compared with 5[,83,84,87,88] of 19 PET-only studies) were restricted to clinically node-negative patients, which may have decreased the overall sensitivity estimate for the PET/CT studies (see *Effect of clinical nodal status*).

Effect of date of publication

Although the majority of studies do not show a clear trend in results according to date of publication, the earliest six studies (published between 1996 and 2001)[92–97] do appear to report higher sensitivities than later studies (see *Figure 5*). This may reflect differences in methodology; for example, four[93,94,96,97] of six of these early studies did not report using ALND or SLNB for all patients, which may have led to overestimates of sensitivity. Also, none of these early studies restricted inclusion by clinical nodal status, whereas a subset of the later studies included clinically node-negative patients only, which may have decreased the overall sensitivity estimate for the later studies.

Effect of size and number of axillary metastases

A few studies analysed the sensitivity according to the size and number of axillary metastases. Thresholds for size were not consistent across studies and the number of analysed patients was small. However, there was a trend for lower sensitivities where metastatic lymph nodes were smaller or fewer in number (*Table 10*; *Figure 7*). Axillary micrometastases (≤ 2 mm) were associated with a mean sensitivity of 11% (95% CI 5% to 22%) based on data from five studies ($n = 63$ patients),[79,80,86–88] while macrometastases (> 2 mm) were associated with a mean sensitivity of 57% (95% CI 47% to 66%) based on data from four studies ($n = 111$ patients).[73,80,87,88] In addition, some studies reported the mean size and number of axillary metastases in TP patients (i.e. those detected by PET) and FN patients (i.e. those which PET failed to detect). As shown in *Table 11*, the cases which PET failed to detect tended to have smaller and fewer axillary metastases, although there was variation between studies. The smallest metastatic nodes detected by PET measured 3 mm,[87,88] while PET failed to detect some nodes measuring > 15 mm[86,94] (including one case of 25 mm in one study[94]).

Effect of clinical nodal status

Studies in which all patients were clinically node negative, which generally referred to non-palpable axillary nodes, had a trend towards lower sensitivity and similar specificity when compared with studies including both clinically node-negative and node-positive patients (*Table 12*). This may reflect the fact that clinically negative axillary metastases are likely to be smaller and more difficult to detect via PET. This analysis was limited by the fact that, even in

PET/CT

Study	TP	FP	FN	TN	Sensitivity	Specificity
Chae 2009[73]	16	12	17	63	0.48 (0.31 to 0.66)	0.84 (0.74 to 0.91)
Heusner 2009[74]	13	3	9	29	0.59 (0.36 to 0.79)	0.91 (0.75 to 0.98)
Kim 2009[75]	27	0	8	102	0.77 (0.60 to 0.90)	1.00 (0.96 to 1.00)
Taira 2009[76]	13	5	14	60	0.48 (0.29 to 0.68)	0.92 (0.83 to 0.97)
Fuster 2008[77]	14	0	6	32	0.70 (0.46 to 0.88)	1.00 (0.89 to 1.00)
Ueda 2008[78]	34	6	25	118	0.58 (0.44 to 0.70)	0.95 (0.90 to 0.98)
Veronesi 2007[79]	38	5	65	128	0.37 (0.28 to 0.47)	0.96 (0.91 to 0.99)

PET only

Study	TP	FP	FN	TN	Sensitivity	Specificity
Cermik 2008[80]	34	11	39	104	0.47 (0.35 to 0.59)	0.90 (0.84 to 0.95)
Gil-Rendo 2006[81]	120	2	22	131	0.85 (0.77 to 0.90)	0.98 (0.95 to 1.00)
Weir 2005[82]	5	3	13	19	0.28 (0.10 to 0.53)	0.86 (0.65 to 0.97)
Agresti 2004[83]	20	3	11	37	0.65 (0.45 to 0.81)	0.93 (0.80 to 0.98)
Fehr 2004[84]	2	1	8	13	0.20 (0.03 to 0.56)	0.93 (0.66 to 1.00)
Inoue 2004[85]	21	2	14	44	0.60 (0.42 to 0.76)	0.96 (0.85 to 0.99)
Lovrics 2004[86]	9	2	16	63	0.36 (0.18 to 0.57)	0.97 (0.89 to 1.00)
Wahl 2004[44]	66	40	43	159	0.61 (0.51 to 0.70)	0.80 (0.74 to 0.85)
Barranger 2003[87]	3	0	12	17	0.20 (0.04 to 0.48)	1.00 (0.80 to 1.00)
Guller 2002[88]	6	1	8	16	0.43 (0.18 to 0.71)	0.94 (0.71 to 1.00)
Nakamoto 2002[89]	6	1	7	16	0.46 (0.19 to 0.75)	0.94 (0.71 to 1.00)
Rieber 2002[90]	16	1	4	19	0.80 (0.56 to 0.94)	0.95 (0.75 to 1.00)
van der Hoeven 2002[91]	8	1	24	37	0.25 (0.11 to 0.43)	0.97 (0.86 to 1.00)
Greco 2001[92]	68	13	4	82	0.94 (0.86 to 0.98)	0.86 (0.78 to 0.93)
Noh 1998[93]	14	0	1	12	0.93 (0.68 to 1.00)	1.00 (0.74 to 1.00)
Smith 1998[94]	13	1	2	22	0.87 (0.60 to 0.98)	0.96 (0.78 to 1.00)
Adler 1997[95]	16	6	4	26	0.80 (0.56 to 0.94)	0.81 (0.64 to 0.93)
Avril 1996[96]	19	0	5	17	0.79 (0.58 to 0.93)	1.00 (0.80 to 1.00)
Utech 1996[97]	44	20	0	60	1.00 (0.92 to 1.00)	0.75 (0.64 to 0.84)

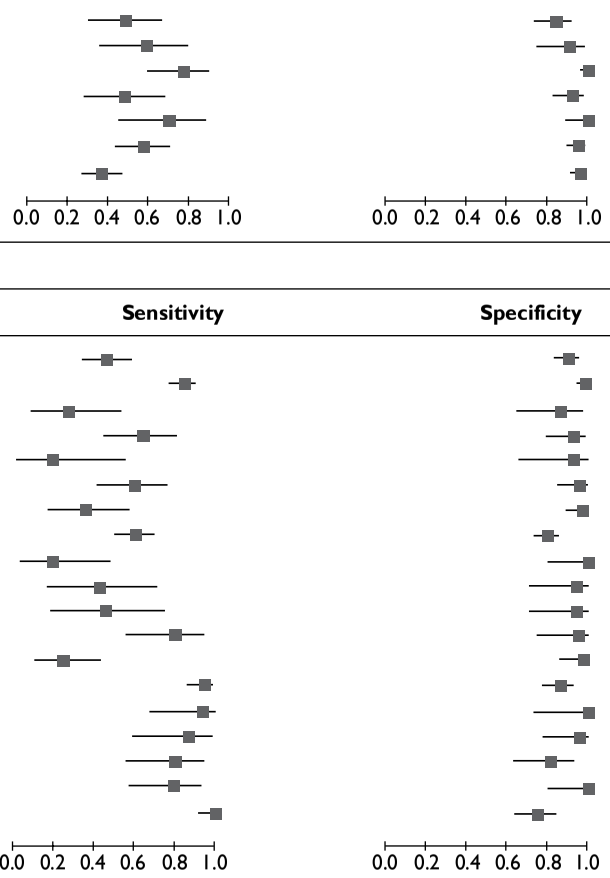


FIGURE 5 Forest plot of all PET studies. Brackets show 95% CIs. The figure shows the sensitivity and specificity for each study (squares) and 95% CIs (horizontal lines).

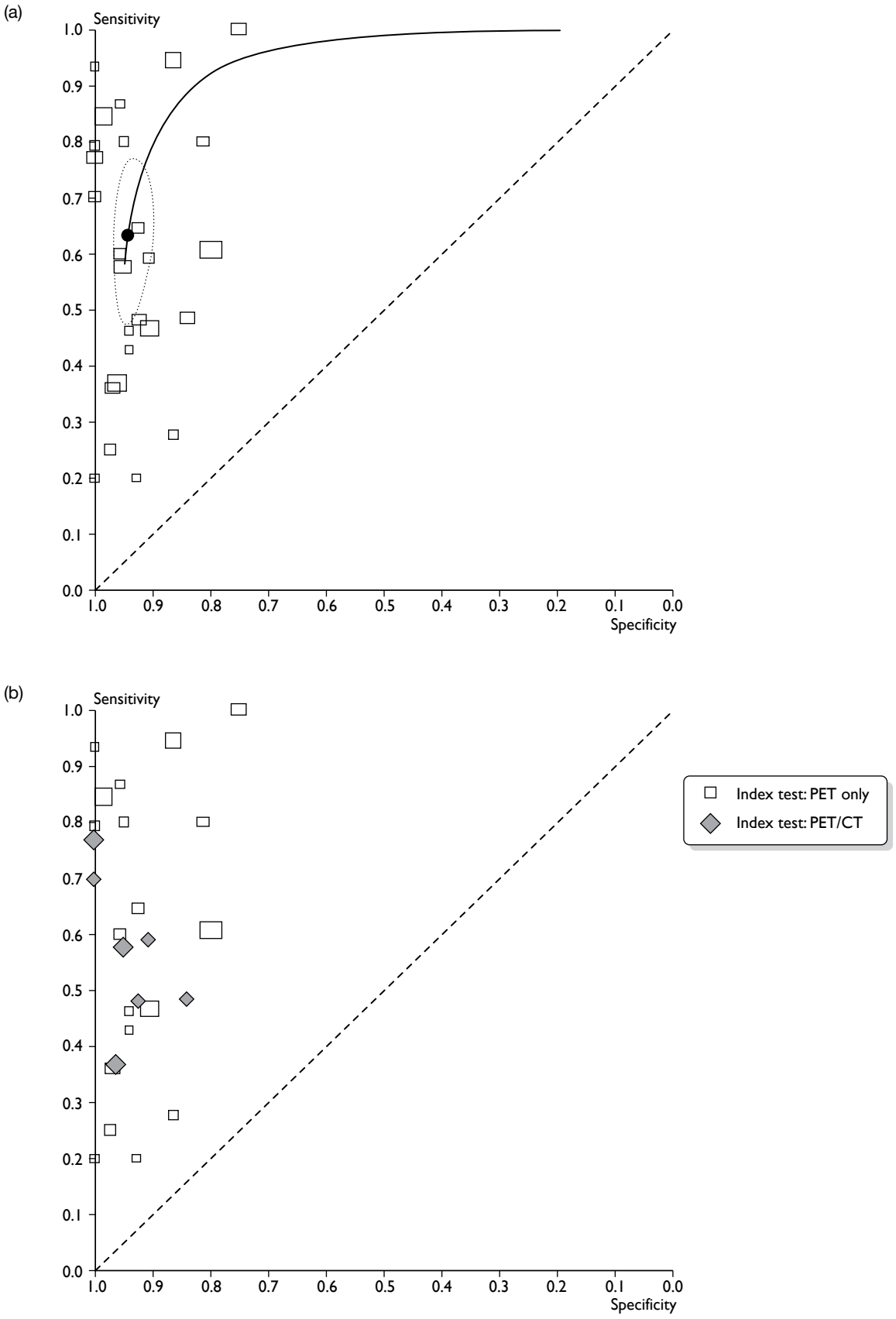


FIGURE 6 Receiver operating characteristic plots for PET. (a) All PET studies, showing ROC curve (solid line), mean sensitivity/specificity (black spot) and 95% confidence region (dashed ellipse); (b) studies of PET only (rectangles) and PET/CT (diamonds). Rectangles/diamonds indicate sensitivity and specificity for individual studies (size proportional to sample sizes).

TABLE 10 Positron emission tomography sensitivity by size and number of axillary metastases and nodal stage

Subgroup	No. of studies	No. of patients	Sensitivity, % (95% CI)	Specificity[a], % (95% CI)
Size of axillary metastases				
≤ 2 mm[79,80,86–88]	5	63	11 (5 to 22)	Not calculable
< 5 mm[78]	1	15	33 (15 to 59)	Not calculable
> 2 mm[73,80,87,88]	4	111	57 (47 to 66)	Not calculable
> 5 mm[79]	1	51	57 (43 to 70)	Not calculable
> 10 mm[78]	1	28	79 (60 to 90)	Not calculable
Number of axillary metastases				
1 metastatic node[78,91,96]	3	52	27 (7 to 63)	Not calculable
Multiple metastatic nodes[91]	1	12	50 (21 to 79)	Not calculable
2–5 metastatic nodes[78,96]	2	23	61 (40 to 78)	Not calculable
> 5 metastatic nodes[78,96]	2	24	77 (53 to 91)	Not calculable
Nodal stage				
pN1[80,81]	2	147	63 (24 to 91)	Not calculable
pN2[80,81]	2	53	83 (51 to 96)	Not calculable
pN3[80,81]	2	21	88 (66 to 97)	Not calculable

a Specificity not calculable as these patients all had axillary metastases.

TABLE 11 Size and number of metastatic nodes in TP and FN cases

	TP cases (PET detected)			FN cases (PET failed to detect)		
Study	No. of patients in analysis	Mean	Range	No. of patients in analysis	Mean	Range
Size of largest metastatic node per patient (mm)						
Heusner 2009[74]	13	9.0	4–22	9	3.0	0.8–6
Kim 2009[75]				8	2.6	1–7
Weir 2005[82]				13		All ≤ 10
Fehr 2004[84]	2	11.0	10–12	8	7.5	Micro–15
Lovrics 2004[86]	9	22.0		16	13.0	NR (> 15 in 7 cases)
Wahl 2004[44]	66	15.6		43	11.5	
Barranger 2003[87]	3	12.0	3–20	12	4.2	0.1–10
Guller 2002[88]	6	13.5	3–30	8	2.9	ITCs–13
Smith 1998[94]	19		8–NR	2	18.5[a]	12–25
Number of metastatic nodes						
Heusner 2009[74]	13	4.3	1–11	9	1.8	1–4
Kim 2009[75]				8	1.4	1–3
Fuster 2008[77]	14	5.2	1–19	6	1.8	1–3
Lovrics 2004[86]	9	5.8		16	2.1	NR (≥ 3 in 5 cases)
Wahl 2004[44]	66	5.0		43	2.7	
Greco 2001[92]				4	1.3	1–2
Avril 1996[96]				5		1–4

ITCs, isolated tumour cells; micro, micrometastasis; NR, not reported.
a Smith *et al.*[94] suggest that the two FN cases may have been due to patient positioning with arms at sides rather than above head.

TABLE 12 Positron emission tomography results according to clinical variables

Subgroup	No. of studies	No. of patients	Sensitivity, % (95% CI)	Specificity, % (95% CI)
Clinical nodal status				
All clinically negative[73,76,79,81,83,84,87,88]	8	869	48 (32 to 64)	94 (90 to 97)
Mix of positive and negative[44,74,77,78,80,85,86,91–94,96,97]	13	1423	72 (54 to 85)	93 (88 to 96)
Nodal status not reported[75,82,89,90,95]	5	299	65 (44 to 81)	95 (82 to 99)
Patient sample				
All patients early stage and newly diagnosed[44,75,76,80,83–89,92,94,97]	14	1413	63 (44 to 79)	94 (89 to 96)
Not all patients early stage and newly diagnosed[73,74,77–79,81,82,90,91,93,95,96]	12	1178	63 (49 to 76)	95 (91 to 98)
T stage (size) of breast tumour				
T1 (≤ 2 cm)[74,79,81,84,88,89,92,94,96,97]	10	451	53 (34 to 72)	87 (82 to 91)
T2 (> 2 cm, ≤ 5 cm)[74,77,79,81,84,88,89,92,94]	9	343	67 (40 to 86)	86 (78 to 92)
T2 or above (> 2 cm)[96,97]	2	82	96 (84 to 99)	82 (13 to 99)
T3 (> 5 cm)[74,77,79,84,89,94]	6	41	65 (44 to 82)	88 (58 to 98)
T4 (tumour spread/fixed to skin/chest wall, or IBC)[94]	1	10	100 (54 to 100)	100 (40 to 100)
Reference standard				
100% ALND[44,73,77,84,87,89,92,95]	8	773	59 (36 to 78)	90 (80 to 95)
ALND and/or SLNB[74–76,78–83,86,88,91]	12	1467	52 (40 to 63)	95 (93 to 97)
Not all ALND/SLNB, or not reported[85,90,93,94,96,97]	6	351	88 (68 to 96)	94 (85 to 98)
Prevalence of axillary metastases				
< 40%[44,73,75–78,80,86,95,97]	10	1334	66 (49 to 80)	92 (85 to 96)
40%–49%[74,79,82–85,87–89,91,92,94]	12	874	51 (34 to 68)	94 (91 to 96)
≥ 50%[81,90,93,96]	4	383	84 (78 to 88)	98 (94 to 99)

IBC, inflammatory breast cancer.

studies with a mixed population, the majority of patients were clinically node negative (see *Table 7*). One study reported separate results according to clinical nodal status: among clinically node-negative patients the sensitivity was 39/42 (93%) and the specificity 76/87 (87%), while among clinically node-positive patients the sensitivity was 29/30 (97%) and the specificity 6/8 (75%).[92] The mix of clinically node-positive and node-negative patients within the included studies was thought to be representative of clinical practice.

Effect of patient sample

The mean sensitivity and specificity were not affected by whether all analysed patients were early stage (stage I, II or IIIA), newly diagnosed and non-DCIS; this may be explained by the fact that all included studies had a majority of patients who were early stage, newly diagnosed and non-DCIS (see *Table 12*).

Effect of T stage (size) of the breast tumour

Some studies reported sensitivity and specificity according to the T stage (size) of the primary breast tumour (see *Table 12*; *Figure 8*). The pattern was difficult to interpret due to the wide variation in sensitivity and specificity between studies and small patient numbers per subgroup. Data from some of the individual studies suggest a trend for lower sensitivity in patients with smaller breast tumours (e.g. between T1 and T2[74,81,92] and between T2 and T3[79,94]), although the pattern for the meta-analysed data is less clear.

(a)

Axillary metastases < 2 mm

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Cermik 2008[80]	2	0	16	0	0.11 (0.01 to 0.35)	Not estimable		
Veronesi 2007[79]	3	0	25	0	0.11 (0.02 to 0.28)	Not estimable		
Lovrics 2004[86]	1	0	6	0	0.14 (0.00 to 0.58)	Not estimable		
Barranger 2003[87]	0	0	6	0	0.00 (0.00 to 0.46)	Not estimable		
Guller 2002[88]	0	0	4	0	0.00 (0.00 to 0.60)	Not estimable		

Axillary metastases < 5 mm

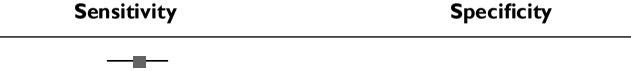
Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Ueda 2008[78]	5	0	10	0	0.33 (0.12 to 0.62)	Not estimable		

Axillary metastases > 2 mm

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Chae 2009[73]	16	0	17	0	0.48 (0.31 to 0.66)	Not estimable		
Cermik 2008[80]	38	0	23	0	0.62 (0.49 to 0.74)	Not estimable		
Barranger 2003[87]	3	0	5	0	0.38 (0.09 to 0.76)	Not estimable		
Guller 2002[88]	6	0	3	0	0.67 (0.30 to 0.93)	Not estimable		

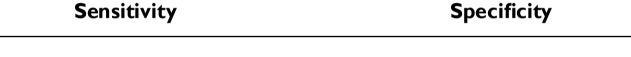
Axillary metastases > 5 mm

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Veronesi 2007[79]	29	0	22	0	0.57 (0.42 to 0.71)	Not estimable		



Axillary metastases > 10 mm

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Ueda 2008[78]	22	0	6	0	0.79 (0.59 to 0.92)	Not estimable		



DOI: 10.3310/hta15040

(b)

1 metastatic node

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Ueda 2008[78]	14	0	14	0	0.50 (0.31 to 0.69)	Not estimable		
van der Hoeven 2002[91]	2	0	18	0	0.10 (0.01 to 0.32)	Not estimable		
Avril 1996[96]	1	0	3	0	0.25 (0.01 to 0.81)	Not estimable		

Multiple metastatic nodes

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
van der Hoeven 2002[91]	6	0	6	0	0.50 (0.21 to 0.79)	Not estimable		

2–5 metastatic nodes

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Ueda 2008[78]	10	0	7	0	0.59 (0.33 to 0.82)	Not estimable		
Avril 1996[96]	4	0	2	0	0.67 (0.22 to 0.96)	Not estimable		

>5 metastatic nodes

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Ueda 2008[78]	10	0	4	0	0.71 (0.42 to 0.92)	Not estimable		
Avril 1996[96]	10	0	0	0	1.00 (0.69 to 1.00)	Not estimable		

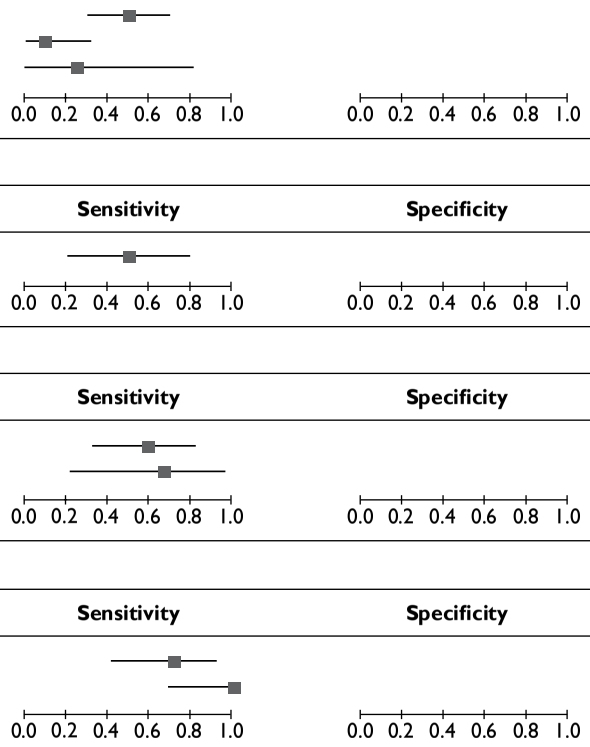


FIGURE 7 Forest plot of sensitivity of PET according to (a) size of axillary metastases; (b) number of axillary metastases; and (c) nodal stage. Brackets show 95% CIs. The figure shows the sensitivity and specificity for each study (squares) and 95% CIs (horizontal lines).

continued

continued

(c)

Nodal stage pN1

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Cermik 2008[80]	24	0	34	0	0.41 (0.29 to 0.55)	Not estimable		
Gil-Rendo 2006[81]	72	0	17	0	0.81 (0.71 to 0.88)	Not estimable		

Nodal stage pN2

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Cermik 2008[80]	10	0	5	0	0.67 (0.38 to 0.88)	Not estimable		
Gil-Rendo 2006[81]	35	0	3	0	0.92 (0.79 to 0.98)	Not estimable		

Nodal stage pN3

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Cermik 2008[80]	6	0	0	0	1.00 (0.54 to 1.00)	Not estimable		
Gil-Rendo 2006[81]	13	0	2	0	0.87 (0.60 to 0.98)	Not estimable		

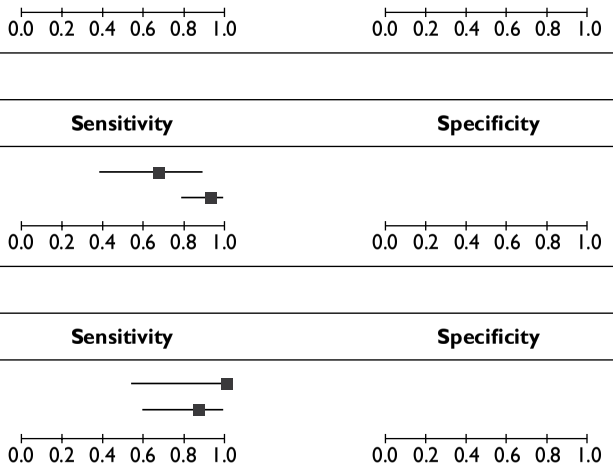


FIGURE 7 Forest plot of sensitivity of PET according to (a) size of axillary metastases; (b) number of axillary metastases; and (c) nodal stage. Brackets show 95% CIs. The figure shows the sensitivity and specificity for each study (squares) and 95% CIs (horizontal lines).

DOI: 10.3310/hta15040

T1 (breast tumour up to 2 cm)

Study	TP	FP	FN	TN	Sensitivity	Specificity
Heusner 2009[74]	5	1	5	13	0.50 (0.19 to 0.81)	0.93 (0.66 to 1.00)
Veronesi 2007[79]	16	3	30	87	0.35 (0.21 to 0.50)	0.97 (0.91 to 0.99)
Gil-Rendo 2006[81]	36	0	13	0	0.73 (0.59 to 0.85)	Not estimable
Fehr 2004[84]	1	0	3	10	0.25 (0.01 to 0.81)	1.00 (0.69 to 1.00)
Guller 2002[88]	2	1	5	11	0.29 (0.04 to 0.71)	0.92 (0.62 to 1.00)
Nakamoto 2002[89]	2	0	5	13	0.29 (0.04 to 0.71)	1.00 (0.75 to 1.00)
Greco 2001[92]	20	10	3	65	0.87 (0.66 to 0.97)	0.87 (0.77 to 0.93)
Smith 1998[94]	1	0	0	7	1.00 (0.03 to 1.00)	1.00 (0.59 to 1.00)
Avril 1996[96]	2	0	4	12	0.33 (0.04 to 0.78)	1.00 (0.74 to 1.00)
Utech 1996[97]	12	13	0	40	1.00 (0.74 to 1.00)	0.75 (0.62 to 0.86)

T2 (breast tumour >2 cm, up to 5 cm)

Study	TP	FP	FN	TN	Sensitivity	Specificity
Heusner 2009[74]	8	2	4	16	0.67 (0.35 to 0.90)	0.89 (0.65 to 0.99)
Fuster 2008[77]	7	0	5	0	0.58 (0.28 to 0.85)	Not estimable
Veronesi 2007[79]	13	6	30	37	0.30 (0.17 to 0.46)	0.86 (0.72 to 0.95)
Gil-Rendo 2006[81]	84	0	9	0	0.90 (0.82 to 0.95)	Not estimable
Fehr 2004[84]	1	1	4	3	0.20 (0.01 to 0.72)	0.75 (0.19 to 0.99)
Guller 2002[88]	4	0	3	5	0.57 (0.18 to 0.90)	1.00 (0.48 to 1.00)
Nakamoto 2002[89]	3	1	3	4	0.50 (0.12 to 0.88)	0.80 (0.28 to 0.99)
Greco 2001[92]	48	3	1	17	0.98 (0.89 to 1.00)	0.85 (0.62 to 0.97)
Smith 1998[94]	7	1	2	11	0.78 (0.40 to 0.97)	0.92 (0.62 to 1.00)

T2 or above (breast tumour >2 cm)

Study	TP	FP	FN	TN	Sensitivity	Specificity
Avril 1996[96]	17	0	1	5	0.94 (0.73 to 1.00)	1.00 (0.48 to 1.00)
Utech 1996[97]	32	7	0	20	1.00 (0.89 to 1.00)	0.74 (0.54 to 0.89)

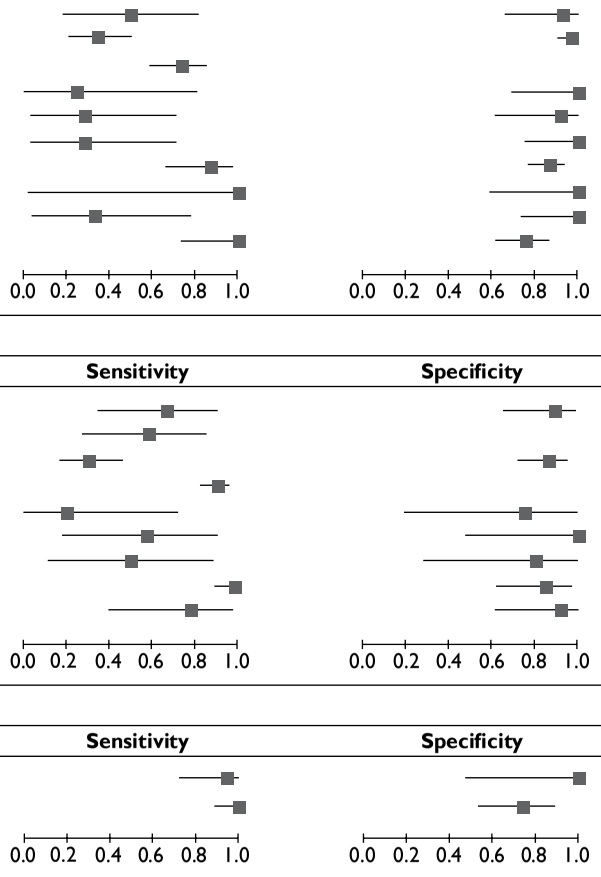


FIGURE 8 Forest plot of sensitivity of PET according to T stage (size) of breast tumour. Brackets show 95% CIs. The figure shows the sensitivity and specificity for each study (squares) and 95% CIs (horizontal lines).

continued

continued

T3 (breast tumour >5 cm)

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Heusner 2009[74]	1	0	1	5	0.50 (0.01 to 0.99)	1.00 (0.48 to 1.00)		
Fuster 2008[77]	7	0	1	0	0.88 (0.47 to 1.00)	Not estimable		
Veronesi 2007[79]	5	0	5	4	0.50 (0.19 to 0.81)	1.00 (0.40 to 1.00)		
Fehr 2004[86]	0	0	1	0	0.00 (0.00 to 0.97)	Not estimable		
Nakamoto 2002[89]	2	0	0	0	1.00 (0.16 to 1.00)	Not estimable		
Smith 1998[94]	5	0	0	4	1.00 (0.48 to 1.00)	1.00 (0.40 to 1.00)		
							0.0 0.2 0.4 0.6 0.8 1.0	0.0 0.2 0.4 0.6 0.8 1.0

T4 (breast tumour spread/fixed to skin/chest wall, or IBC)

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Smith 1998[94]	6	0	0	4	1.00 (0.54 to 1.00)	1.00 (0.40 to 1.00)		
							0.0 0.2 0.4 0.6 0.8 1.0	0.0 0.2 0.4 0.6 0.8 1.0

FIGURE 8 Forest plot of sensitivity of PET according to T stage (size) of breast tumour. Brackets show 95% CIs. The figure shows the sensitivity and specificity for each study (squares) and 95% CIs (horizontal lines). IBC, inflammatory breast cancer.

Effect of reference standard

In terms of reference standard, studies in which all patients received ALND had similar mean sensitivity and specificity values to studies in which some patients received ALND and some SLNB. Studies in which the reference standard was not stated (or some patients did not receive ALND or SLNB) had a higher sensitivity, possibly due to the poorer-quality reference standard in these studies (see *Table 12*).

Effect of prevalence of axillary metastases

There was no clear correlation between prevalence of axillary metastases and sensitivity or specificity (see *Table 12*).

Sensitivity analysis for study quality

Eleven quality items from the QUADAS checklist[69] were used to assess study quality (see *Appendix 2*). For PET, study quality variables which had a significant effect ($p < 0.10$) on either sensitivity or specificity were as follows. Studies with a representative patient spectrum (all early stage, newly diagnosed and non-DCIS, with prospective and consecutive recruitment) had a higher sensitivity than those in which the patient spectrum was unclear ($p = 0.001$). Studies that used the same reference standard regardless of index test results (differential verification avoided) had a lower sensitivity than studies in which differential verification occurred or those in which this was unclear ($p < 0.001$). Studies in which the index test was interpreted blind to reference standard results had a higher sensitivity than studies in which this was unclear ($p < 0.001$). Studies in which the reference standard was interpreted blind to index test results had a lower sensitivity than studies in which this was unclear, but a higher sensitivity than studies where this was unblinded (although there were only two studies[76,90] known to be unblinded) ($p < 0.001$). Studies in which uninterpretable test results did not occur, or occurred and were included in the analysis, had a lower sensitivity than studies in which this was unclear ($p < 0.001$). No study quality variables had a significant effect on the specificity of PET. These results show a mixed pattern. Studies which score 'unclear' on reporting of quality variables may be expected to be of poorer quality, but this could theoretically lead to either underestimates or overestimates of diagnostic accuracy.

Magnetic resonance imaging studies: diagnostic accuracy results

Summary findings

Of the nine studies evaluating MRI ($n = 307$ patients),[41,42,64,108–113] several reported more than one set of results on diagnostic accuracy, according to different criteria for defining whether axillary metastases were present. Results are first presented across all studies using the highest reported sensitivity and specificity for each study (*Table 13* and *Figures 9* and *10*). Results are then presented according to each criterion for positivity (*Table 14* and *Figures 11* and *12*).

Across all nine MRI studies, using the highest sensitivity and specificity for each study, the mean sensitivity was 90% (95% CI 78% to 96%; range 65%–100%) and the mean specificity was 90% (95% CI 75% to 96%; range 54%–100%) (see *Table 13* and *Figures 9* and *10*). According to the type of MRI, the mean estimates of sensitivity and specificity were as follows (see *Table 13*). Across five studies of USPIO-enhanced MRI ($n = 93$),[108–112] the mean sensitivity was 98% (95% CI 61% to 100%)and specificity 96% (95% CI 72% to 100%). Across three studies of gadolinium-enhanced MRI ($n = 187$),[41,42,113] the mean sensitivity was 88% (95% CI 78% to 94%) and specificity 73% (95% CI 63% to 81%). In the single study of in vivo proton MR spectroscopy ($n = 27$),[64] the sensitivity was 65% (95% CI 38% to 86%) and the specificity 100% (95% CI 69% to 100%). Therefore, USPIO-enhanced MRI showed a trend towards higher sensitivity and specificity than gadolinium-enhanced MRI.

In addition, the diagnostic accuracy data were analysed according to the criteria for defining whether axillary metastases were present (e.g. based on USPIO or gadolinium uptake, size, morphology, or combinations of these criteria) (see *Table 14* and *Figures 11* and *12*). Within this analysis, some studies appear more than once. The exact combinations of criteria were often not consistent across studies. The use of contrast uptake pattern as the main criterion for defining a node as metastatic appeared to give better combined sensitivity and specificity than size and morphology, although many studies used criteria based on both uptake and size/morphology, and the methods of interpreting uptake patterns varied within and between studies.

TABLE 13 Summary of pooled sensitivities and specificities for MRI studies[a]

Diagnostic test	No. of studies	No. of patients	Sensitivity, % (95% CI)	Specificity, % (95% CI)
All MRI studies				
All MRI studies[41,42,64,108–113]	9	307	90 (78 to 96)	90 (75 to 96)
MRI studies by type of MRI				
USPIO-enhanced MRI[108–112]	5	93	98 (61 to 100)	96 (72 to 100)
Gadolinium-enhanced MRI[41,42,113]	3	187	88 (78 to 94)	73 (63 to 81)
MR spectroscopy[64]	1	27	65 (38 to 86)	100 (69 to 100)

a Where studies report results using more than one set of criteria for positivity, these analyses use data corresponding to the criteria with the highest reported estimates of diagnostic accuracy per study.

TABLE 14 Magnetic resonance imaging results according to criteria for positivity

Criteria for positivity	No. of studies	No. of patients	Sensitivity, % (95% CI)	Specificity, % (95% CI)
USPIO-based criteria				
USPIO uptake[108–111]	4	75	98 (63 to 100)	94 (69 to 99)
USPIO uptake, size > 10 mm, round shape[112] (not clear if 'and' or 'or')	1	18	82 (48 to 98)	100 (59 to 100)
Gadolinium-based criteria				
Gd uptake, size > 5 mm[42] (not clear if 'and' or 'or')	1	75	90 (76 to 97)	83 (66 to 93)
Dynamic Gd signal intensity increase[41,113]	2	112	86 (68 to 94)	59 (45 to 72)
Dynamic Gd + positive washout[41]	1	65	71 (49 to 87)	90 (77 to 97)
Dynamic Gd + size > 4 sq-mm[113]	1	47	100 (69 to 100)	54 (37 to 71)
Dynamic Gd + size > 5 mm + abnormal morphology[41]	1	65	63 (41 to 81)	93 (80 to 98)
MR spectroscopy				
MR spectroscopy[64]	1	27	65 (38 to 86)	100 (69 to 100)
Size and/or morphological criteria				
Size > 4 sq-mm[113]	1	47	100 (69 to 100)	19 (8 to 35)
Size > 5 mm[109]	1	33	100 (85 to 100)	10 (0 to 45)
Size > 10 mm[109]	1	33	43 (23 to 66)	80 (44 to 97)
Abnormal morphology[109]	1	33	96 (78 to 100)	20 (3 to 56)
Size > 5 mm + abnormal morphology[109]	1	65	63 (41 to 81)	80 (65 to 91)
Size > 10 mm and/or round shape[110]	1	22	83 (36 to 100)	31 (11 to 59)

Gd, gadolinium.

DOI: 10.3310/hta15040

USPIO-enhanced MRI

Study	TP	FP	FN	TN	MRI criteria	Sensitivity	Specificity	Sensitivity	Specificity
Kimura 2010[108]	2	0	0	8	USPIO uptake	1.00 (0.16 to 1.00)	1.00 (0.63 to 1.00)		
Harada 2007[109]	23	2	0	8	USPIO uptake	1.00 (0.85 to 1.00)	0.80 (0.44 to 0.97)		
Memarsadeghi 2006[110]	6	0	0	16	USPIO uptake	1.00 (0.54 to 1.00)	1.00 (0.79 to 1.00)		
Stadnik 2006[111]	5	1	0	4	USPIO uptake	1.00 (0.48 to 1.00)	0.80 (0.28 to 0.99)		
Michel 2002[112]	9	0	2	7	USPIO + >10 mm + round	0.82 (0.48 to 0.98)	1.00 (0.59 to 1.00)		
								0.0 0.2 0.4 0.6 0.8 1.0	0.0 0.2 0.4 0.6 0.8 1.0

Gadolinium-enhanced MRI

Study	TP	FP	FN	TN	MRI criteria	Sensitivity	Specificity	Sensitivity	Specificity
Mumtaz 1997[42]	36	6	4	29	Gd uptake + >5 mm	0.90 (0.76 to 0.97)	0.83 (0.66 to 0.93)		
								0.0 0.2 0.4 0.6 0.8 1.0	0.0 0.2 0.4 0.6 0.8 1.0

Dynamic gadolinium-enhanced MRI

Study	TP	FP	FN	TN	MRI criteria	Sensitivity	Specificity	Sensitivity	Specificity
Murray 2002[113]	10	17	0	20	Dynamic Gd + >4 mm^2	1.00 (0.69 to 1.00)	0.54 (0.37 to 0.71)		
Kvistad 2000[41]	20	4	4	37	Dynamic Gd	0.83 (0.63 to 0.95)	0.90 (0.77 to 0.97)		
								0.0 0.2 0.4 0.6 0.8 1.0	0.0 0.2 0.4 0.6 0.8 1.0

MR spectroscopy (in vivo)

Study	TP	FP	FN	TN	MRI criteria	Sensitivity	Specificity	Sensitivity	Specificity
Yeung 2002[64]	11	0	6	10	MR spectroscopy (in vivo)	0.65 (0.38 to 0.86)	1.00 (0.69 to 1.00)		
								0.0 0.2 0.4 0.6 0.8 1.0	0.0 0.2 0.4 0.6 0.8 1.0

FIGURE 9 Forest plot of all MRI studies. Brackets show 95% CIs. The figure shows the sensitivity and specificity for each study (squares) and 95% CIs (horizontal lines). Where studies report results using more than one set of criteria for positivity, these analyses use data corresponding to the criteria with the highest reported estimates of diagnostic accuracy per study. The criteria used for each study are shown on the plot.

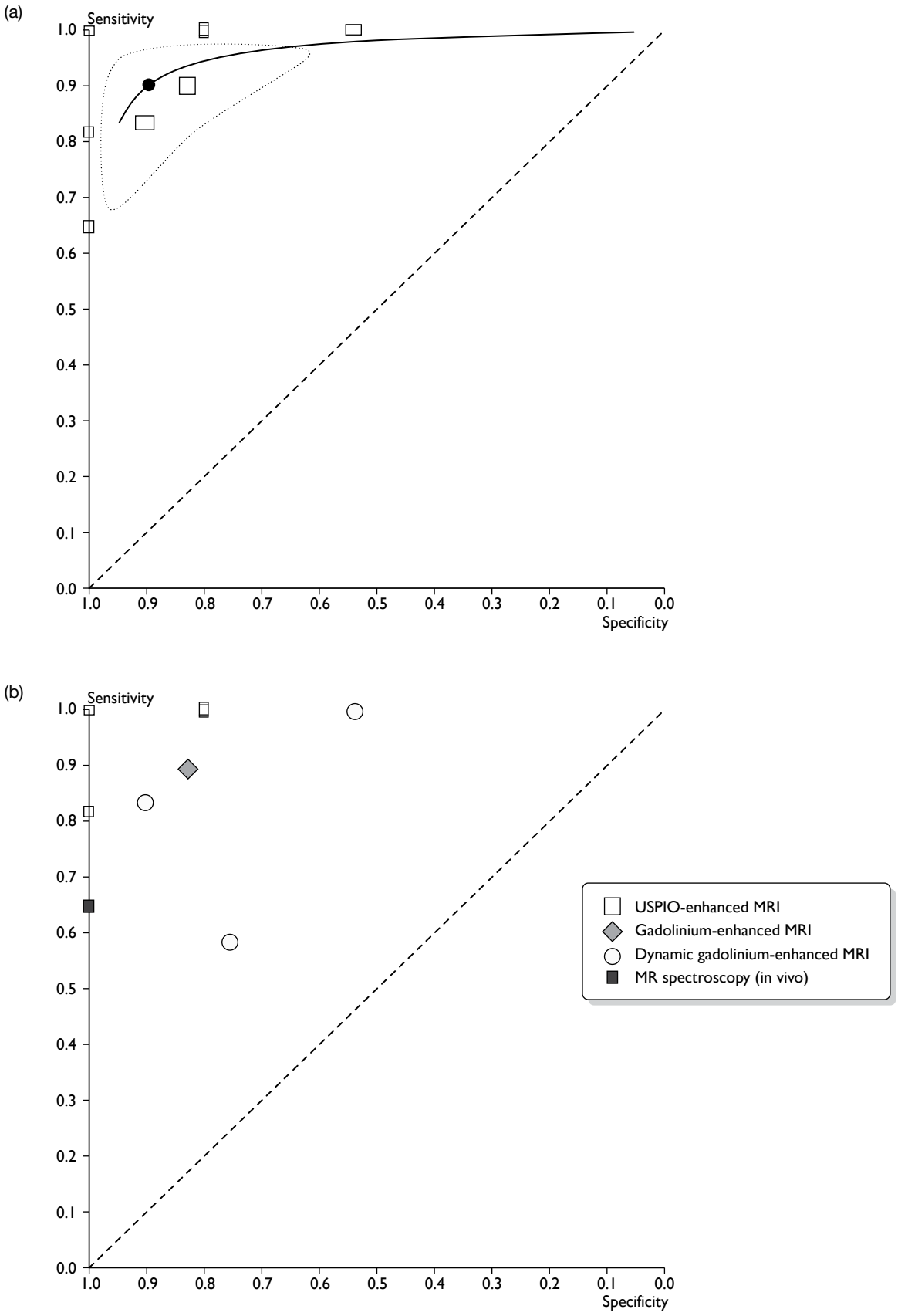


FIGURE 10 Receiver operating characteristic plots for MRI. (a) All MRI studies, showing ROC curve (solid line), mean sensitivity/specificity (black spot) and 95% confidence region (dashed ellipse); (b) all MRI studies, showing type of MRI. Shapes indicate sensitivity and specificity for individual studies (size proportional to sample sizes, legend shows type of MRI). Studies appear only once, using the highest reported sensitivity and specificity per study.

DOI: 10.3310/hta15040

(a)

USPIO uptake

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Kimura 2010[108]	2	0	0	8	1.00 (0.16 to 1.00)	1.00 (0.63 to 1.00)		
Harada 2007[109]	23	2	0	8	1.00 (0.85 to 1.00)	0.80 (0.44 to 0.97)		
Memarsadeghi 2006[110]	6	0	0	16	1.00 (0.54 to 1.00)	1.00 (0.79 to 1.00)		
Stadnik 2006[111]	5	1	0	4	1.00 (0.48 to 1.00)	0.80 (0.28 to 0.99)		
							0.0 0.2 0.4 0.6 0.8 1.0	0.0 0.2 0.4 0.6 0.8 1.0

USPIO + >10 mm + round

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Michel 2002[112]	9	0	2	7	0.82 (0.48 to 0.98)	1.00 (0.59 to 1.00)		
							0.0 0.2 0.4 0.6 0.8 1.0	0.0 0.2 0.4 0.6 0.8 1.0

FIGURE 11 Forest plots of MRI studies (using various criteria for defining a node as metastatic; many studies appear more than once). (a) Based on ultrasmall superparamagnetic iron oxide uptake (and size/morphology); (b) Based on gadolinium uptake (and size/morphology); (c) Based on size and/or morphology; (d) based on MR spectroscopy. Brackets show 95% CIs. The figure shows the sensitivity and specificity for each study (squares) and 95% CIs (horizontal lines).

continued

continued

(b)

Gadolinium uptake + >5 mm

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Mumtaz 1997[42]	36	6	4	29	0.90 (0.76 to 0.97)	0.83 (0.66 to 0.93)	0.0 0.2 0.4 0.6 0.8 1.0	0.0 0.2 0.4 0.6 0.8 1.0

Dynamic Gadolinium

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Murray 2002[113]	10	22	0	15	1.00 (0.69 to 1.00)	0.41 (0.25 to 0.58)	0.0 0.2 0.4 0.6 0.8 1.0	0.0 0.2 0.4 0.6 0.8 1.0
Kvistad 2000[41]	20	4	4	37	0.83 (0.63 to 0.95)	0.90 (0.77 to 0.97)		

Dynamic Gadolinium + washout

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Kvistad 2000[41]	17	4	7	37	0.71 (0.49 to 0.87)	0.90 (0.77 to 0.97)	0.0 0.2 0.4 0.6 0.8 1.0	0.0 0.2 0.4 0.6 0.8 1.0

Dynamic Gadolinium + >5mm + abnormal shape

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Kvistad 2000[41]	15	3	9	38	0.63 (0.41 to 0.81)	0.93 (0.80 to 0.98)	0.0 0.2 0.4 0.6 0.8 1.0	0.0 0.2 0.4 0.6 0.8 1.0

Dynamic Gadolinium + >4mm^2

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Murray 2002[113]	10	17	0	20	1.00 (0.69 to 1.00)	0.54 (0.37 to 0.71)	0.0 0.2 0.4 0.6 0.8 1.0	0.0 0.2 0.4 0.6 0.8 1.0

DOI: 10.3310/hta15040

(c)

>4 mm²

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Murray 2002[113]	10	30	0	7	1.00 (0.69 to 1.00)	0.19 (0.08 to 0.35)	0.0 0.2 0.4 0.6 0.8 1.0	0.0 0.2 0.4 0.6 0.8 1.0

>5 mm

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Harada 2007[109]	23	9	0	1	1.00 (0.85 to 1.00)	0.10 (0.00 to 0.45)	0.0 0.2 0.4 0.6 0.8 1.0	0.0 0.2 0.4 0.6 0.8 1.0

>10 mm

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Harada 2007[109]	10	2	13	8	0.43 (0.23 to 0.66)	0.80 (0.44 to 0.97)	0.0 0.2 0.4 0.6 0.8 1.0	0.0 0.2 0.4 0.6 0.8 1.0

FIGURE 11 Forest plots of MRI studies (using various criteria for defining a node as metastatic; many studies appear more than once). (a) Based on ultrasmall superparamagnetic iron oxide uptake (and size/morphology); (b) based on gadolinium uptake (and size/morphology); (c) based on size and/or morphology; (d) based on MR spectroscopy. Brackets show 95% CIs. The figure shows the sensitivity and specificity for each study (squares) and 95% CIs (horizontal lines).

continued

(c) *continued*

Abnormal morphology

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Harada 2007[109]	22	8	1	2	0.96 (0.78 to 1.00)	0.20 (0.03 to 0.56)	0.0 0.2 0.4 0.6 0.8 1.0	0.0 0.2 0.4 0.6 0.8 1.0

>5 mm + abnormal morphology

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Kvistad 2000[41]	15	8	9	33	0.63 (0.41 to 0.81)	0.80 (0.65 to 0.91)	0.0 0.2 0.4 0.6 0.8 1.0	0.0 0.2 0.4 0.6 0.8 1.0

>10 mm ± round

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Memarsadeghi 2006[110]	5	11	1	5	0.83 (0.36 to 1.00)	0.31 (0.11 to 0.59)	0.0 0.2 0.4 0.6 0.8 1.0	0.0 0.2 0.4 0.6 0.8 1.0

(d)

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Yeung 2002[64]	11	0	6	10	0.65 (0.38 to 0.86)	1.00 (0.69 to 1.00)	0.0 0.2 0.4 0.6 0.8 1.0	0.0 0.2 0.4 0.6 0.8 1.0

FIGURE 11 Forest plots of MRI studies (using various criteria for defining a node as metastatic; many studies appear more than once). (a) Based on ultrasmall superparamagnetic iron oxide uptake (and size/morphology); (b) based on gadolinium uptake (and size/morphology); (c) Based on size and/or morphology; (d) based on MR spectroscopy. Brackets show 95% CIs. The figure shows the sensitivity and specificity for each study (squares) and 95% CIs (horizontal lines).

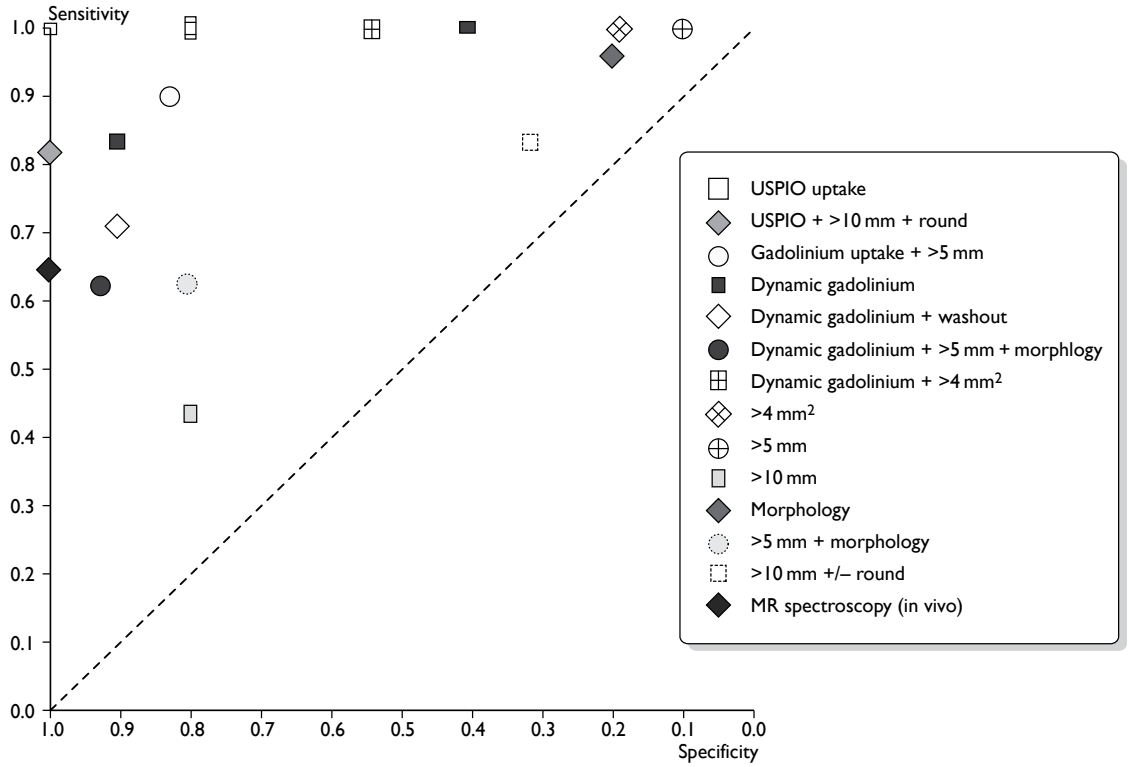


FIGURE 12 Receiver operating characteristic plot for MRI (using various criteria for defining a node as metastatic). Shapes indicate sensitivity and specificity for individual studies (size proportional to sample sizes, legend shows criteria used for defining a node as metastatic). Many studies appear more than once.

Effect of size and number of axillary metastases

No MRI studies presented data allowing calculation of sensitivity according to size and number of axillary metastases.

Effect of clinical nodal status

It was difficult to assess the effect of clinical nodal status as six[42,109–113] of nine studies did not report these data (*Table 15*).

Effect of patient sample

Studies in which all analysed patients were early stage (stage I, II or IIIA), newly diagnosed and non-DCIS had a trend towards a higher sensitivity, and a significantly lower specificity, than studies in which not all patients were early stage, newly diagnosed and non-DCIS; however, there was wide variation in results between studies (see *Table 15*).

Effect of T stage (size) of the breast tumour

There were insufficient consistently reported data to assess the relationship between individual size or stage categories for the primary tumour, and the sensitivity and specificity of MRI in detecting axillary metastases.

Effect of reference standard

It was not possible to assess the effect of reference standard as, in eight[41,42,64,109–113] of nine studies, all patients received ALND (see *Table 15*).

TABLE 15 Magnetic resonance imaging results according to clinical variables

Subgroup	No. of studies	No. of patients	Sensitivity, % (95% CI)	Specificity, % (95% CI)
Reference standard				
100% ALND[41,42,64,109–113]	8	297	91 (78 to 97)	88 (73 to 95)
ALND and/or SLNB[108]	1	10	100 (16 to 100)	100 (63 to 100)
Not all ALND/SLNB or not reported	0	0	N/A	N/A
Cancer stage				
All patients early stage and newly diagnosed[108,110,113]	3	79	90 (63 to 98)	61 (46 to 74)
Not all patients early stage and newly diagnosed[41,42,64,109,111,112]	6	228	83 (74 to 89)	86 (78 to 92)
Clinical nodal status				
All clinically negative[108]	1	10	100 (16 to 100)	100 (63 to 100)
Mix of positive and negative[41,64]	2	92	75 (69 to 79)	91 (86 to 94)
Nodal status not reported[42,109–113]	6	205	94 (83 to 98)	85 (63 to 95)
Prevalence of axillary metastases				
< 40%[41,108,110,113]	4	144	86 (71 to 94)	71 (59 to 81)
40%–49%	0	0	N/A	N/A
≥ 50%[42,64,109,111,112]	5	163	94 (89 to 97)	84 (73 to 91)

N/A, not applicable.

Effect of prevalence of axillary metastases

There was no clear correlation between prevalence of axillary metastases and sensitivity or specificity (see *Table 15*).

Sensitivity analysis for study quality

Eleven quality items from the QUADAS checklist[69] were used to assess study quality (see *Appendix 2*). For MRI, no study quality variables had a significant effect on sensitivity or specificity, although this analysis was limited by the small number of studies and the fact that there was little variation in scoring between studies for some quality variables.

Withdrawal rates, adverse events and contraindications

Positron emission tomography studies

In 20[73–76,78–85,87–89,91,92,94,96,97] of the 26 PET studies, no withdrawals were reported (however, studies that were not prospective and consecutive may have included only patients with complete data for both tests). The remaining six studies reported that between 4% and 22% of patients withdrew.[44,77,86,90,93,95] Reasons for withdrawal included no ALND, anxiety and inconvenience, unavailability of PET scanner, and unusable PET scans (*Table 16*). No adverse events were reported for any of the 26 PET studies. In addition, many PET studies excluded patients who were pregnant or had diabetes mellitus.

Magnetic resonance imaging studies

In five[108,109,111–113] of nine MRI studies, no withdrawals were reported (however, studies that were not prospective and consecutive may have included only patients with complete data for both tests). The remaining four studies reported that between 3% and 18% of patients withdrew.[41,42,64,110] Reasons for withdrawal included no ALND, inadequate MRI data, and claustrophobia or poor health (*Table 17*). No serious adverse effects were reported in any of the MRI studies. Mild-to-moderate adverse effects included mild rash following USPIO

TABLE 16 Withdrawals and adverse events: PET studies

Study	Index test	No. met criteria?	Withdrawals (%)	Reasons (%)	Adverse effects
Chae 2009[73]	PET/CT	108	None reported		None reported
Heusner 2009[74]	PET/CT	54	None reported		None reported
Kim 2009[75]	PET/CT	137	None reported		None reported
Taira 2009[76]	PET/CT	92 axilla	None reported		None reported
Fuster 2008[77]	PET/CT	60	8/60 (13)	Eight (13) no ALND (reason not reported)	None reported
Ueda 2008[78]	PET/CT	183	None reported		None reported
Veronesi 2007[79]	PET/CT	236	None reported		None reported
Cermik 2008[80]	PET only	188 axilla	None reported		None reported
Gil-Rendo 2006[81]	PET only	275	None reported		None reported
Weir 2005[82]	PET only	40	None reported		None reported
Agresti 2004[83]	PET only	71	None reported		None reported
Fehr 2004[84]	PET only	24	None reported		None reported
Inoue 2004[85]	PET only	81	None reported		None reported
Lovrics 2004[86]	PET only	115	25/115 (22)	Seventeen (15) withdrew (anxiety and inconvenience), six (5) no PET (machine unavailable), two (2) no PET (anxiety)	None reported
Wahl 2004[44]	PET only	330 axilla	22/330 (7)	Five (1.5) no surgery (reason not reported), 15 (5) SLNB, only two (0.6) unusable PET scans	None reported
Barranger 2003[87]	PET only	32	None reported		None reported
Guller 2002[88]	PET only	31	None reported		None reported
Nakamoto 2002[89]	PET only	30	None reported		None reported
Rieber 2002[90]	PET only	43	3/43 (7)	Three (7) reason not reported	None reported
van der Hoeven 2002[91]	PET only	70	None reported		None reported
Greco 2001[92]	PET only	167	None reported		None reported
Noh 1998[93]	PET only	31 axilla	4/31 (13)	Four (13) reason not reported	None reported
Smith 1998[94]	PET only	38	None reported		None reported
Adler 1997[95]	PET only	54 axilla	2/54 (4)	Two (4) PET scans uninterpretable due to high FDG accumulation in myocardium	None reported
Avril 1996[96]	PET only	41	None reported		None reported
Utech 1996[97]	PET only	124	None reported		None reported

administration (recovered without treatment or following antihistamine treatment) and inability to complete the MRI scan due to claustrophobia or back pain as a result of holding the same position for some time (see *Table 17*). In addition, many of the studies excluded patients with contraindications to MRI, such as strong allergic disposition, allergy to contrast agents or liver dysfunction (see *Table 8*).

Discussion for clinical effectiveness

Diagnostic accuracy of positron emission tomography

Across all 26 studies ($n = 2591$ patients)[44,73–97] evaluating PET or PET/CT for assessment of axillary metastases, the mean sensitivity was 63% (95% CI 52% to 74%; range 20%–100%) and the mean specificity was 94% (95% CI 91% to 96%; range 75%–100%). For the seven studies ($n = 862$)[73–79] evaluating PET/CT, the mean sensitivity was 56% (95% CI 44% to 67%) and the

TABLE 17 Withdrawals and adverse events: MRI studies

Study	Index test	No. met criteria?	Withdrawals (%)	Reasons (%)	Adverse effects
Kimura 2010[108]	USPIO-enhanced	10	None reported		No serious adverse effects. Mild rash on four limbs of one patient after USPIO; recovered without further treatment
Harada 2007[109]	USPIO-enhanced	33	None reported		No serious adverse effects. 2/33 (6%) patients had minor adverse event of mild rash; one (3%) required oral antihistamine and one (3%) recovered without further treatment
Memarsadeghi 2006[110]	USPIO-enhanced	24	2/24 (8)	2/24 (8) no ALND (reason not reported)	No discomfort or adverse reactions were observed (0/22, 0%)
Stadnik 2006[111]	USPIO-enhanced	10	None reported		None reported
Michel 2002[112]	USPIO-enhanced	18	None reported		No serious adverse effects. Of 22 examinations in 20 patients: mild adverse effects in 3/22 (14%); moderate adverse effect in 1/22 (5%); and antihistamines administered. Symptoms: rash ($n=2$), pruritis ($n=2$), abdominal and or lumbar pain ($n=1$), chest pain ($n=3$), orthostatic reaction ($n=1$). MRI of axilla performed in all 20 patients, but three wished to terminate scan before MRI of breast due to claustrophobia or back pain as a result of holding same position
Murray 2002[113]	Dynamic gadolinium-enhanced	47	None reported		None reported
Kvistad 2000[41]	Dynamic gadolinium-enhanced	67	2/67 (3)	One (1.5) no ALND due to old age, one (1.5) died before surgery	None reported
Mumtaz 1997[42]	Gadolinium-enhanced	92 axilla	17/92 (18)	Six (7) image obscured by cardiac flow or not in field of view, eight (9) inadequate MR data and three (3) no ALND data	None reported
Yeung 2002[64]	MR spectroscopy	32	5/32 (16)	One (3) no MR spectroscopy due to machine breakage, three (9) no MR spectroscopy due to claustrophobia or poor general health, and one (3) refused surgery	3/39 (8%) could not complete procedure due to claustrophobia or poor general health

mean specificity was 96% (95% CI 90% to 99%). For the 19 studies ($n=1729$)[44,80–97] evaluating PET only, the mean sensitivity was 66% (95% CI 50% to 79%) and the mean specificity was 93% (95% CI 89% to 96%). PET performed less well in terms of identifying small metastases; micrometastases (≤ 2 mm) were associated with a mean sensitivity of 11% (95% CI 5% to 22%) based on data from five studies ($n=63$),[79,80,86–88] while macrometastases (> 2 mm) were associated with a mean sensitivity of 57% (95% CI 47% to 66%) based on data from four studies ($n=111$).[73,80,87,88] The smallest metastatic nodes reported as being detected by PET measured 3 mm,[87,88] while PET failed to detect some nodes measuring > 15 mm.[86,94] Current PET cameras are thought to achieve spatial resolutions of 4–7 mm (around 4–5 mm in the centre of the field of view).[115] PET studies in which all patients were clinically node negative showed a trend towards lower sensitivity compared with studies that included both clinically node-negative and node-positive patients, which may reflect the fact that clinically-negative axillary metastases are likely to be smaller.

Diagnostic accuracy of magnetic resonance imaging

The review identified nine studies ($n = 307$ patients) evaluating MRI.[41,42,64,108,–113] Several MRI studies reported more than one set of diagnostic accuracy results, according to different criteria for defining whether axillary metastases were present. Based on the highest reported sensitivity and specificity per study, the mean sensitivity across all nine MRI studies was 90% (95% CI 78% to 96%; range 65%–100%) and the mean specificity was 90% (95% CI 75% to 96%; range 54%–100%). Across the five studies ($n = 93$ patients) evaluating USPIO-enhanced MRI,[108–112] the mean sensitivity was 98% (95% CI 61% to 100%) and specificity 96% (95% CI 72% to 100%). Across three studies of gadolinium-enhanced MRI ($n = 187$),[41,42,113] the mean sensitivity was 88% (95% CI 78% to 94%) and specificity 73% (95% CI 63% to 81%). In the single study of in vivo proton MR spectroscopy ($n = 27$),[64] the sensitivity was 65% (95% CI 38% to 86%) and specificity 100% (95% CI 69% to 100%). USPIO-enhanced MRI showed a trend towards higher sensitivity and specificity than gadolinium-enhanced MRI. However, studies of USPIO-enhanced MRI are currently small and in the experimental stages. The use of contrast uptake pattern as the main criterion for defining a node as metastatic appeared to give better combined sensitivity and specificity than size and morphology, although some studies used criteria based on both uptake and size/morphology. No data were presented according to size of metastases for the MRI studies. However, current MRI methods are thought to achieve a resolution of approximately 1 mm using modern scanners and based on the methods used in the included papers.

Adverse effects and contraindications

Both PET and MRI appeared to be relatively safe in this setting, with no adverse effects reported for any of the 26 PET studies,[44,73–97] and no serious adverse effects reported in any of the MRI studies. Mild-to-moderate adverse effects of MRI included mild rash following USPIO administration and inability to complete the MRI scan due to claustrophobia or back pain. Previous studies of USPIO have suggested an adverse event rate > 2%; most events are minor and include lumbar pain, rash, transient decrease in blood pressure, and arrhythmia.[116–118] The mechanism of lumbar pain is unknown, but this is also observed with other particulate agents, and usually disappears when the infusion is stopped or its speed is reduced.[116] Cautions and contraindications exist for both PET (pregnancy) and MRI (allergy to contrast agents, renal or liver dysfunction, pacemakers and other metallic implants). Several PET studies excluded patients who had diabetes mellitus (or high serum glucose levels). In addition, some patients are unable to complete an MRI scan due to claustrophobia. These factors may limit the applicability of PET and MRI for some patients.

Internal and external validity

The sensitivity and specificity of both PET and MRI varied widely between studies, so the pooled accuracy data should be interpreted with caution. The included MRI studies were relatively small and there was variation between and within studies in terms of the MRI method used (MRI or MR spectroscopy; contrast agent; field strength; image sequence parameters; type of coil) and the criteria for defining a node as positive.

Study quality was assessed using the QUADAS checklist[69] and was mixed. In the majority of studies, the reference standard was adequate (ALND or SLNB). Studies using ALND reported similar sensitivity and specificity to studies using a combination of ALND and SLNB, while studies in which not all patients had ALND or SLNB (or in which the reference standard was not stated) had a higher mean sensitivity, which may represent an overestimate. Differential verification bias was avoided in most studies, but in six PET studies[75,76,78,79,81,83] patients received a different reference standard (ALND or SLNB) depending on the PET results. Nineteen studies were not considered to have a representative patient spectrum,[73,74,76–79,81,82–85,89–91,93–97] either because recruitment was not prospective and consecutive or because some of the included patients were non-early stage, non-newly diagnosed or had DCIS. Most studies reported few or

no withdrawals. The index test was interpreted blind to reference standard results in most studies. Interpretation of the reference standard with blinding to index test results and handling of uninterpretable results were often not reported. Index test details were very well reported, while reference standard details were sometimes incomplete, probably due to the widespread routine use of these techniques. For PET, two quality variables (representative patient spectrum and index test interpreted blind to reference standard results) were associated with higher sensitivity in studies scoring 'yes' than in studies scoring 'unclear', while three quality variables (differential verification avoided, reference standard interpreted blind to index test and uninterpretable test results did not occur or occurred and were included in the analysis) were associated with lower sensitivity in studies scoring 'yes' than in studies scoring 'unclear'. For MRI, no study quality variables had a significant effect on sensitivity or specificity, although this analysis was limited by the small number of studies. These results show a mixed pattern. Studies that score 'unclear' on reporting of quality variables may be expected to be of poorer quality, but this could theoretically lead to either underestimates or overestimates of diagnostic accuracy.

The external validity of the included studies relates to the extent to which the populations and methods are generalisable to clinical practice in the UK. The average age of the patients in the included studies was 56 years, while the average age at diagnosis using published data was 64 years.[1,2] This may reflect slightly younger patients being recruited into clinical studies or being candidates for ALND. However, this is unlikely to substantially affect the accuracy estimates of PET and MRI. The majority of patients in the included studies had early-stage, newly diagnosed breast cancer; patients with DCIS were excluded where possible as these patients would not usually undergo surgical assessment of the axilla in the UK (with the exception of patients with extensive DCIS). Within PET studies, sensitivity and specificity were not affected by whether or not all analysed patients were early stage, newly diagnosed and non-DCIS; for MRI, sensitivity was slightly higher and specificity slightly lower where all patients were early stage, newly diagnosed and non-DCIS. The mean prevalence of axillary metastases across all studies was 42% (range 20%–70%); there was no clear relationship with sensitivity and specificity.

Comparison of positron emission tomography and magnetic resonance imaging with current techniques for assessing the axilla

Positron emission tomography and MRI have lower sensitivity and specificity than SLNB and 4-NS, but are associated with fewer adverse effects. The potential use of PET or MRI, either as a replacement for SLNB or 4-NS or as an additional test prior to SLNB or 4-NS, was evaluated using a decision model. The model was used to assess both the cost-effectiveness and clinical acceptability of each potential diagnostic strategy, via an estimation of the number of correct and incorrect diagnoses, costs, and impact on QALYs due to cancer recurrence and adverse effects associated with each strategy. The results of this analysis are described further within the cost-effectiveness section of this report (see *Chapter 4*).

Ultrasound is currently recommended prior to surgical assessment of the axilla for all patients with early-stage breast cancer.[17] In order to be useful, any additional imaging technique would need a higher sensitivity and/or specificity than ultrasound. A systematic review estimated the average sensitivity of ultrasound at 69%–71% in all patients and 44%–61% in patients with non-palpable axillary nodes, and the specificity at 75%–86% in all patients and 77%–92% in patients with non-palpable axillary nodes.[47] In terms of sensitivity, PET appears similar to the above estimates for ultrasound, while MRI appears slightly higher. As ultrasound is not currently considered sensitive enough to completely replace SLNB or 4-NS, it is unlikely that PET could fulfil this role either. In terms of specificity, PET and USPIO-enhanced MRI appear slightly higher than ultrasound.

Conclusions for clinical effectiveness

The studies in this review demonstrate a significantly higher sensitivity for MRI than for PET, with USPIO-enhanced MRI providing the highest sensitivity. However, since none of the included studies directly compared PET and MRI, caution should be taken when comparing these estimates. Sensitivity of PET is reduced for smaller metastases. Specificity was similar for PET and MRI. However, this analysis is limited by the small number and size of MRI studies, and the wide variations between and within studies in terms of the MRI method used and the criteria for defining a node as positive. The sensitivity and specificity of both PET and MRI vary widely between studies, which is likely to reflect differences in imaging methods and interpretation. Therefore, caution should be taken when interpreting these results, particularly for MRI.

Positron emission tomography and MRI have lower sensitivity and specificity than the current surgical diagnostic techniques of SLNB and 4-NS, but are associated with fewer adverse events. The potential use of PET or MRI, either as a replacement for SLNB or 4-NS or as an additional test prior to SLNB or 4-NS, was evaluated using a decision model (described further within *Assessment of cost-effectiveness*). As the sensitivity of PET is only moderate and is similar to that of ultrasound, it appears unlikely that PET could entirely replace SLNB or 4-NS in the assessment of axillary metastases.

As data on MRI are currently in the experimental stage, further large, well-conducted studies using up-to-date MRI methods are required to obtain more accurate data on the sensitivity and specificity of MRI in the assessment of axillary metastases. Further studies of USPIO-enhanced MRI would be valuable in order to gain more robust data on sensitivity and specificity, adverse effects and which are the best criteria for defining a node as metastatic.

Chapter 4

Assessment of cost-effectiveness

This section of the assessment focuses on the health economics of enhanced imaging techniques in the assessment of axillary lymph node metastases in comparison with standard diagnostic techniques. It includes a brief review of existing economic evaluations of the relevant imaging techniques in the assessment of axillary lymph node metastases and a detailed explanation of the methodologies and results of the economic model.

Review of existing cost-effectiveness evidence presents the results of the review of economic literature. The modelling approach adopted for this study is discussed in *Independent economic assessment – methods*, with the results of the analysis being presented in *Independent economic assessment – results*.

Review of existing cost-effectiveness evidence

The primary objective of this review was to identify and evaluate studies exploring the cost-effectiveness of enhanced imaging techniques in the assessment of axillary lymph node metastases. The secondary objective was to evaluate methodologies used to inform our own economic evaluation.

The literature was searched using the strategy described in *Chapter 3, Methods for reviewing effectiveness* and *Appendix 1*. Published economic evaluations of PET or MRI in the assessment of axillary lymph node metastases in breast cancer were included in the review. The search identified 245 citations. Of these, 242 were excluded at the title/abstract stage and one was excluded at the full-text stage. In total, two studies were included in the review: one of PET[119] and one of MRI.[120]

One published economic evaluation of PET in the assessment of axillary lymph node metastases in breast cancer was identified.[119] The model studied breast cancer in general, rather than early-stage breast cancer in particular. A decision tree model was built which did not include the lifetime of the patient or breast cancer recurrence. Furthermore, patient utilities and QALYs were not used in the model. The second paper compared the cost-effectiveness of MR lymphangiography-based strategies with that of SLNB in the axillary staging of early breast cancer.[120] However, the model did not consider the short- and long-term adverse events (e.g. lymphoedema) that are associated with SLNB. The diagnostic pathway modelled did not represent the typical pathway in the UK, where ultrasound and ultrasound-guided biopsy are also used. The disease pathway in the model did not include the locoregional relapse and subsequent remission states which are important health states for breast cancer patients. All costs of the study were based in the USA and are unlikely to represent the costs in the UK, given the significant difference in the organisation and funding of health services between the two countries. The literature review confirmed the need for new published economic evaluations in this area.

Independent economic assessment – methods

Objective

The aim of the model was to evaluate the effects on patient outcomes and cost-effectiveness of enhanced imaging techniques (MRI and PET) compared with standard techniques in the assessment of axillary lymph node metastases in women with early-stage breast cancer. Two axillary sampling techniques, 4-NS and SLNB, are used currently in the UK. It is beyond the remit of this assessment to compare 4-NS and SLNB. The enhanced imaging techniques (PET and MRI) are therefore compared with the two baseline sampling techniques separately.

Diagnostic methods

The diagnostic methods that were evaluated in the model (MRI, PET, 4-NS and SLNB) were discussed in detail in *Chapter 1, Current methods for assessment of axillary metastases* and *Chapter 1, Description of technology under assessment*.

Structure of the model

A probabilistic discrete-event simulation model has been developed in SIMUL8 (SIMUL8®, Boston, MA, USA) to explore the costs and health outcomes associated with the assessment of axillary lymph node metastases and the treatment of women with early breast cancer.

Discrete-event simulation concerns the modelling of a system as it evolves over time by a representation in which the state variables change instantaneously at separate and countable points in time.[121] In the context of health-care modelling, patients are individually represented in a discrete-event simulation model, and normally have associated attributes indicating their distinctive demographical information and diagnostic and/or disease history.

The model differs from classical Markov state transition models, in which the state transition probabilities of patients are evaluated during fixed time intervals and patients may remain in one state after each evaluation. In discrete-event simulation, the time spent in each health state is sampled from a distribution when an individual patient enters this state. If one state can transit to multiple states, then the timing to each subsequent state is sampled and compared, and the patient will transit to the state that has the shortest transition delay.

The discrete-event simulation model consists of two main parts: the diagnostic pathway, which represents current and alternative diagnostic strategies (including PET and MRI), and the treatment pathway, which represents the disease progression among various health states and the management of patients with early breast cancer after the diagnosis of axillary lymph node metastases. A hypothetical cohort of 5000 early breast cancer patients was modelled. Each individual patient follows a specific diagnostic path and will obtain one of the four diagnostic results: TN, FP, TP or FN. Short- and long-term adverse events associated with sampling diagnostic techniques (4-NS and SLNB) and the associated cost and utility implications are also determined for each patient. The diagnostic results will influence the time spent in each of the subsequent health states, which is sampled from exponential distributions based on yearly transition probabilities. The starting age of patients was 56 years, which is based on the clinical effectiveness review within this assessment. The model was run for the remaining lifetime of patients.

Resource use and utilities are mainly taken from published literature. Input parameters are assigned probability distributions to reflect their imprecision and Monte Carlo techniques are performed to reflect this uncertainty in the results. Results are presented in terms of net health benefit and cost per incremental QALY gained.

Diagnostic pathway

The status of axillary lymph node metastases was assessed for all early breast cancer patients. Current methods of assessment consist of clinical examination followed by ultrasound.[17] If the result of ultrasound is positive, an ultrasound-guided fine-needle or core-needle biopsy will be conducted. If axillary metastases are identified via the biopsy (TP or FP), then the patients will be managed as node-positive patients (i.e. patients who have axillary lymph node metastases) and ALND will be performed, normally at the same time as the main breast cancer surgery is carried out.

For those women whose axillary metastases have not been identified by ultrasound or biopsy (i.e. either ultrasound or biopsy result is negative), current management involves the surgical removal of only some of the axillary lymph nodes (axillary sampling via either 4-NS or SLNB) for histological examination. Axillary sampling is normally performed at the same time as the main breast cancer surgery. Depending on the results of axillary sampling, patients will be managed either as node-negative patients (i.e. patients who do not have axillary lymph node metastases), who will receive no further investigation at that time, or as node-positive patients, who will undergo ALND to remove all axillary lymph nodes. *Figure 13* illustrates the current standard diagnostic pathway.

Apart from the two baseline 4-NS and SLNB strategies, six alternative strategies that involve either MRI or PET were evaluated. Two alternative strategies are to replace the axillary sampling methods with either MRI or PET (scenarios 1 and 2). Another four alternative strategies are to add MRI or PET before 4-NS (scenarios 3 and 4), and to add MRI or PET before SLNB (scenarios 5 and 6) in the diagnostic pathway. In theory, a biopsy (fine-needle or core-needle biopsy) could be undertaken in the event of a positive MRI or PET result in the alternative diagnostic pathways. However, MRI-guided or PET-guided biopsy is not currently available in most centres in the UK and no clinical studies were identified to provide data on this. Therefore, these techniques were not included in our assessment. The eight diagnostic strategies were as follows:

- baseline 1: 4-NS
- baseline 2: SLNB
- scenario 1: replace sampling with MRI
- scenario 2: replace sampling with PET
- scenario 3: add MRI before 4-NS
- scenario 4: add PET before 4-NS
- scenario 5: add MRI before SLNB
- scenario 6: add PET before SLNB.

The cost-effectiveness of the alternative MRI or PET replacement strategies was evaluated using the diagnostic pathway illustrated in *Figure 13*, in which MRI or PET replaces the current axillary sampling procedures. In order to evaluate the other four alternative strategies, the standard pathway was modified to create an alternative diagnostic pathway (*Figure 14*). In the alternative pathway, enhanced imaging techniques will be carried out for patients who have negative ultrasound or biopsy results. If the results of the imaging techniques are positive, no axillary sampling is performed and the patients are regarded as node positive and go on to receive ALND. If the results of imaging techniques are negative, further axillary sampling (4-NS or SLNB) is still performed as in the standard pathway.

Disease pathway

The diagnostic results will affect the choice of adjuvant therapies and the probability of locoregional relapse and developing metastatic diseases (e.g. patients with FN diagnoses, who have metastatic nodes which are not detected and removed, are more likely to suffer from

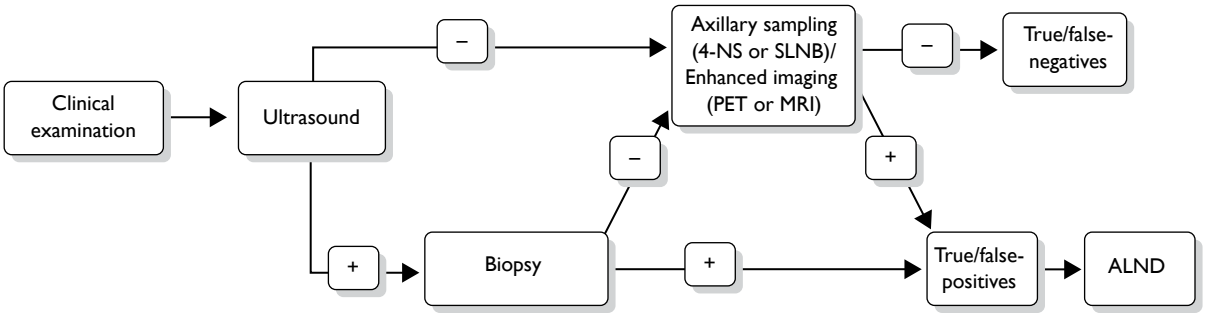


FIGURE 13 Standard and alternative imaging replacement diagnostic pathway in the School of Health and Related Research (ScHARR) model.

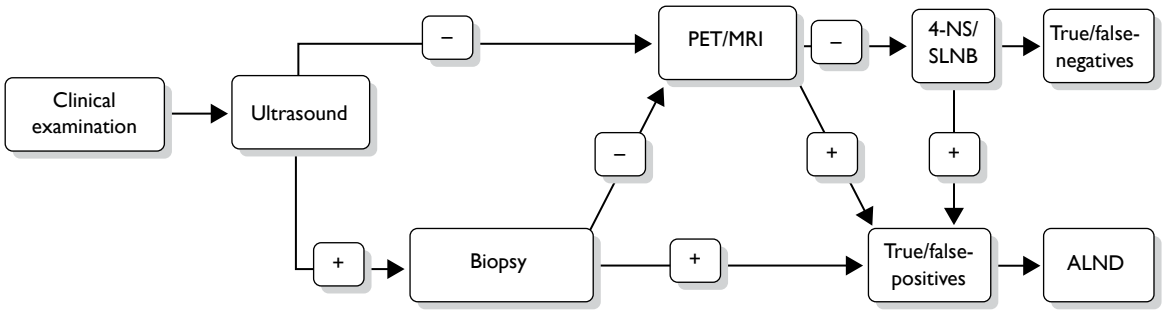


FIGURE 14 Alternative imaging addition diagnostic pathway in the School of Health and Related Research (ScHARR) model.

relapse). Certain diagnostic techniques are associated with long-term adverse events, such as lymphoedema, which may have lifetime cost and utility implications. Therefore, the disease pathway of patients with early breast cancer is also modelled.

All patients with early breast cancer receive adjuvant therapy after the main breast cancer surgery. Node-positive patients normally receive chemotherapy followed by hormonal therapy (where appropriate) while node-negative patients normally receive only hormonal therapy (where appropriate).[17] The aim of the adjuvant therapies is to reduce the risk of cancer recurrences. Following the adjuvant therapy, patients may enter into a disease-free survival state of post-adjuvant therapy and may potentially stay in this state for the rest of their lives (i.e. cured). Some patients may, however, experience locoregional relapse during or after the therapy. Patients who are in the post-adjuvant therapy state may experience locoregional or metastatic relapse. Patients experiencing locoregional relapse receive further treatment (e.g. surgical removal of lymph nodes, chemotherapy, radiotherapy, hormonal therapy). The patients may then enter a further remission period without evidence of cancer until death or further relapse to metastatic disease. Metastatic/distant relapse is not considered curable. Patients experiencing a metastatic relapse receive active palliative treatment to control symptoms and improve quality of life, a period of supportive care and ultimately a period of intensive end-of-life care for the last few days/weeks of life. Patients may also die owing to other causes in any health state.

In the UK, women with breast cancer receive follow-up examinations for a number of years following their treatment. The aim of this follow-up is to detect any recurrences earlier and therefore the frequency and effectiveness of follow-up may affect the overall effectiveness of the diagnostic strategies assessed in this study. In theory, axillary metastases in patients misdiagnosed as FN may be identified by either follow-up or self-presentation. However, in practice it is

difficult for a clinician to determine whether axillary metastases identified by follow-up or self-presentation are actually due to previous misdiagnosis or due to recurrence. Because the issue of follow-up is outside the scope of the assessment and not enough published evidence is available on the effectiveness of follow-up (especially for patients with FN results), the model did not explicitly represent follow-up.

Health states

There are 10 health states within the disease pathway part of the model:

- adjuvant therapy – node (true) negative patients
- adjuvant therapy – node (false) positive patients
- adjuvant therapy – node (true) positive patients
- adjuvant therapy – node (false) negative patients
- post-adjuvant therapy
- locoregional relapse
- remission (post-locoregional relapse)
- metastatic relapse
- death from breast cancer
- death from other causes.

The disease pathway is shown in *Figure 15*. Each individual patient starts from one of the four adjuvant therapy states, depending on the previous diagnostic result.

Model state transitions

The following health-state transitions are possible in the model:

1. adjuvant therapy – patients can move to:
   i. post-adjuvant therapy
   ii. locoregional relapse
   iii. death from other causes

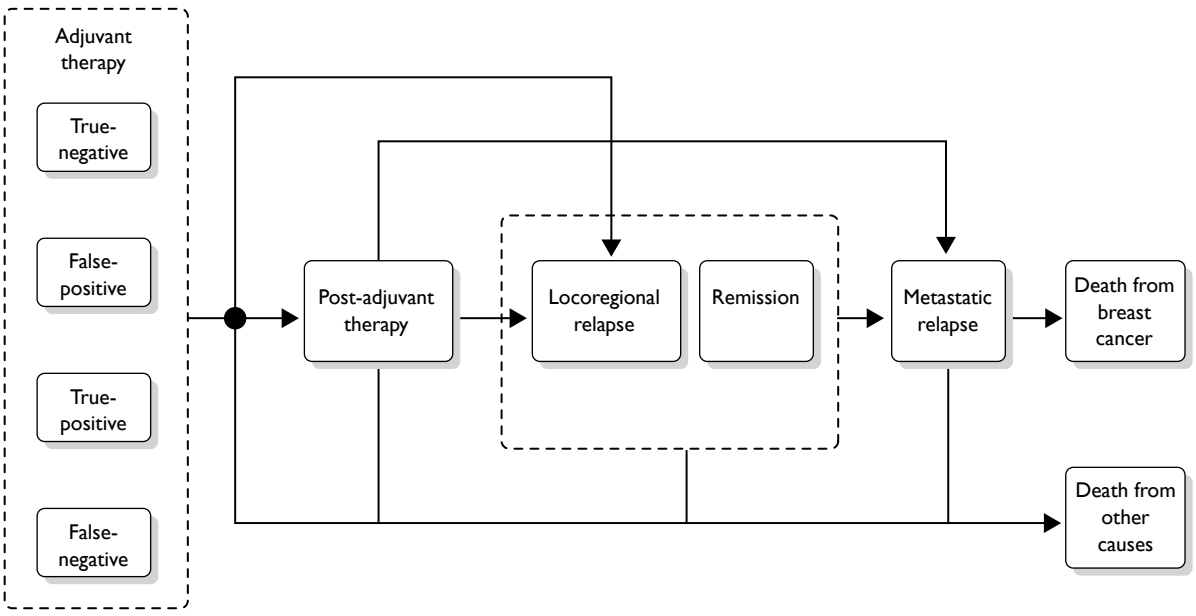


FIGURE 15 Treatment pathways in the School of Health and Related Research (ScHARR) model.

2. post-adjuvant therapy – patients can move to:
    i. locoregional relapse
    ii. metastatic relapse
    iii. death from other causes
3. locoregional relapse – patients can move to:
    i. remission
    ii. metastatic relapse
    iii. death from other causes
4. remission – patients can move to:
    i. metastatic relapse
    ii. death from other causes
5. metastatic relapse – patients can move to:
    i. death from breast cancer
    ii. death from other causes
6. death from breast cancer – absorbing state
7. death from other causes – absorbing state.

Model assumptions

The model employs a number of simplifying assumptions, which are detailed below.

- The sensitivity and specificity of biopsy, axillary sampling (4-NS and SLNB), and imaging techniques (MRI and PET) are independent of preceding diagnostic results.
- The sensitivity and specificity of MRI is based on all identified studies, with no distinction made between different types of MRI (see *Chapter 3, Quantity and quality of research available*). This assumption is tested in the sensitivity analysis.
- Seroma, surgical drain and infection are the short-term adverse events associated with diagnostic techniques (4-NS, SLNB and ALND) considered by the model.
- Lymphoedema is the only long-term adverse event considered. Lymphoedema is classified as either mild/moderate or severe.
- Studies have reported adverse events associated with SLNB, whereas no studies were identified which quantify the short-term adverse events associated with 4-NS or compare the probability of adverse events between SLNB and 4-NS. Therefore, the probability of adverse events is assumed to be equal for 4-NS and SLNB.
- Short-term adverse events increase the costs, but do not affect quality of life.
- Long-term adverse events (i.e. lymphoedema) affect both costs and quality of life for the rest of the patient's life.
- During the adjuvant therapy period, node-positive patients receive chemotherapy plus hormonal therapy (where appropriate) and node-negative patients receive hormonal therapy (where appropriate).
- Patients receive adjuvant therapy for a fixed 5-year period. Node-positive patients receive chemotherapy for half a year, followed by hormonal therapy for 4.5 years. Node-negative patients receive hormonal therapy for 5 years. This is the maximum time patients may stay in this state. Patients may, owing to model dynamics, spend < 5 years in the adjuvant therapy state if the sampled time to locoregional relapse or death from other causes is < 5 years.
- Following locoregional relapse patients cannot experience further locoregional relapse; they can only experience metastatic relapse.
- Death rates for non-breast cancer causes are based on UK mortality statistics and applied across all health states. These are not adjusted to exclude breast cancer mortality, and so may overestimate the risk of dying due to non-breast cancer causes.

Model inputs: accuracy and costs of diagnostic techniques

The model inputs of accuracy and costs of diagnostic methods were summarised in *Table 18*. The sensitivity and specificity of clinical examination and biopsy were based on previous published studies.[41,42,47,94] The sensitivity and specificity of ultrasound for clinically negative patients were calculated based on either the size criterion (60.9% and 77.3%, respectively) or morphology criterion (43.9% and 92.4%, respectively) from a systematic review by Alvarez *et al.*[47] We used the averages to represent overall sensitivity and specificity assuming both size and morphology are used to assess axillary lymph node metastases. The sensitivity and specificity of ultrasound for clinically-positive patients were estimated based on expert opinion, as no data were identified from published studies.

The sensitivity of 4-NS was based on two identified studies with data from 335 patients.[122,123] The sensitivity of SLNB was based on a systematic review and meta-analysis of 69 studies undertaken by Kim *et al.*[125] with data from over 8000 patients. All studies were included in the NICE guideline.[17] The mean sensitivity of 4-NS is slightly higher than that of SLNB (94.5% vs 93%) according to literature reviewed within the NICE guideline. However, it is important to note that the sensitivity of SLNB is based on a significantly larger sample size and should be more robust. The specificities of 4-NS and SLNB are set at 100% as, by definition, there should be no FP cases when histological methods are used.

TABLE 18 Accuracy and costs of diagnostic methods

Diagnostic technique	Parameter	Value	Distribution	Source
Clinical examination	Sensitivity (%)	45.9	Beta	Mumtaz *et al.* 1997;[42] Smith *et al.* 1998;[94] Kvistad *et al.* 2000[41]
	Specificity (%)	78.1		
	Costs (£)	86	Normal	*NHS Reference Costs 2008*[56]
Ultrasound	Sensitivity for clinically positive (%)	90.0		Expert opinion
	Specificity for clinically positive (%)	70.0		
	Sensitivity for clinically negative (%)	52.4	Beta	Alvarez *et al.* 2006[47]
	Specificity for clinically negative (%)	84.9		
	Costs (£)	53	Normal	*NHS Reference Costs 2008*[56]
Biopsy	Sensitivity (%)	45.4	Beta	Alvarez *et al.* 2006[47]
	Specificity (%)	99.6		
	Costs (£)	147	Normal	*NHS Reference Costs 2003*[65] (adjusted to 2007 prices)
4-NS	Sensitivity (%)	94.5	Beta	Sato *et al.* 2001;[124] Tanaka *et al.* 2006[123]
	Specificity (%)	100.0		
	Costs (breast surgery and 4-NS) (£)	2099		Assumption
SLNB	Sensitivity (%)	93.0	Beta	Kim *et al.* 2006[125]
	Specificity (%)	100.0		
	Costs (breast surgery and SLNB) (£)	2728		Pandharipande *et al.* 2008[120]
MRI	Sensitivity (%)	90.3	Beta	Systematic review by this assessment
	Specificity (%)	89.7		
	Costs (£)	232	Normal	*NHS Reference Costs 2008*[56]
PET	Sensitivity (%)	63.4	Beta	Systematic review within this assessment
	Specificity (%)	94.2		
	Costs (£)	978		Jacklin *et al.* 2002[66] (adjusted to 2007 prices)
ALND	Costs of ALND alone (£)	2448	Normal	*NHS Reference Costs 2008*[56]
	Costs of ALND and breast surgery (£)	3186		Pandharipande *et al.* 2008[120]
Costs of breast surgery (£)		1908		*NHS Reference Costs 2008*[56]

The costs of clinical examination, ultrasound, biopsy, MRI and ALND were obtained from NHS reference costs.[55,56] Apart from scanning the axilla, MRI is currently also used in the pre-surgical evaluation of breast tumours for some patients. However, the procedure is performed in a small proportion of women with breast cancer (e.g. lobular cancers and patients having neoadjuvant chemotherapy) and it is likely that axillary and breast MRI scans would be undertaken separately, as pre-contrast scans would be required for both procedures. Therefore marginal cost savings of axillary MRI scans due to breast MRI scans are not considered and we assumed the full NHS reference costs for MRI in the model.

The costs of PET were calculated based on a published UK study.[66] The costs of breast surgery were calculated based on NHS costs of total mastectomy surgery [Healthcare Resource Group (HRG) codes are JA07A, JA07B and JA07C] and intermediate breast surgery (HRG codes are JA09A and JA09B),[56] assuming that two-thirds of patients receive intermediate breast surgery and one-third of patients receive mastectomy.

Sentinel lymph node biopsy and 4-NS are assumed to be carried out at the same time as the breast surgery. ALND is assumed to be carried out at the same time as the breast surgery, if no sampling is performed beforehand (e.g. patients are diagnosed as node positive by biopsy). No UK costs of SLNB, 4-NS and ALND (at the same time as breast surgery) were identified from published studies. One study from the USA was identified which reported the costs of breast surgery alone, SLNB together with breast surgery and ALND together with breast surgery.[20] The relative ratios between breast surgery and SLNB together with breast surgery, and between breast surgery and ALND together with breast surgery, can be calculated. These ratios were used to adjust the UK costs of breast surgery to obtain the costs of SLNB and ALND together with breast surgery. The procedure of 4-NS together with breast surgery was assumed to increase the costs of breast surgery alone by 10%.

Model inputs: probability and costs of short-term adverse events

The model inputs of probability and costs of short-term adverse events were summarised in *Table 19*. The probabilities of developing short-term adverse events due to SLNB and ALND were estimated based on published studies. No studies were identified that quantify the short-term adverse events associated with 4-NS. Based on expert opinion, it is assumed that 4-NS is associated with the same probabilities of developing short-term adverse events as SLNB. Sensitivity analysis was carried out to test the assumption that patients are more likely to develop adverse events for 4-NS than SLNB. The costs of short-term adverse events were based on a study in the US, as no UK costs were identified.

Model inputs: costs and utilities of health states

The model inputs of costs and utilities of health states are summarised in *Table 20*. Patients with node-negative results (TN and FN) are assumed to receive hormonal therapy for 5 years (where appropriate). For patients with FP results, the TN nodal status will be picked up by the following ALND. Therefore, patients with FP diagnoses also receive hormonal therapy. It is assumed that 81% of patients are oestrogen receptor positive, which means that they will respond to hormonal therapy.[126] Among these patients, it is assumed that 90% receive aromatase inhibitors [anastrozole (Armidex®, AstraZeneca), exemestane (Aromasin®, Pfizer) or letrozole (Femara®, Novartis)] and 10% receive tamoxifen (Nolvadex, Istubal, Valodex). It is also assumed that each patient has one clinic visit and one mammogram per year in the adjuvant therapy state. The annual cost of aromatase inhibitors, tamoxifen and mammography is estimated to be £1002 based on a published study.[129]

Patients with TP results are assumed to receive chemotherapy for 6 months (£8788) followed by hormonal therapy for 4.5 (£1002 per year). It is assumed docetaxel (Taxotere®,

TABLE 19 Probabilities and costs of short-term adverse events

Short-term adverse events	Parameter	Value	Distribution	Source
Seroma	Probability (4-NS)	0.072	Beta	Expert opinion – assumed same as SLNB
	Probability (SLNB)	0.072 (95% CI 0.065 to 0.08)	Beta	Blanchard *et al.* 2003;[48] Purushotham *et al.* 2005;[51] Wilke *et al.* 2006[54]
	Probability (ALND)	0.221 (95% CI 0.174 to 0.278)	Beta	Blanchard *et al.* 2003;[48] Purushotham *et al.* 2005[51]
	Costs (£)	292		Jeruss *et al.* 2006[127] (adjusted to 2007 prices)
Surgical drains	Probability (4-NS)	0.02	Beta	Expert opinion – assumed same as SLNB
	Probability (SLNB)	0.02 (95% CI 0.016 to 0.025)	Beta	Mansel *et al.* 2006;[128] Wilke *et al.* 2006[54]
	Probability (ALND)	0.792 (95% CI 0.753 to 0.827)	Beta	Mansel *et al.* 2006[128]
	Costs (£)	292		Jeruss *et al.* 2006[127] (adjusted to 2007 prices)
Infection	Probability (4-NS)	0.021	Beta	Expert opinion – assumed same as SLNB
	Probability (SLNB)	0.021 (95% CI 0.018 to 0.026)	Beta	Blanchard *et al.* 2003;[48] Mansel *et al.* 2006;[128] Wilke *et al.* 2006[54]
	Probability (ALND)	0.142 (95% CI 0.115 to 0.1730)	Beta	Blanchard *et al.* 2003;[48] Mansel *et al.* 2006[128]
	Costs (£)	292		Jeruss *et al.* 2006[127] (adjusted to 2007 prices)

TABLE 20 Costs and utilities of each health state

Health state	Parameter	Value	Distribution	Source
Adjuvant therapy (TN and FN, FP)	Annual costs (£)	1002		Ward *et al.* 2007[128] (adjusted to 2007 prices)
	Utility	0.82	Beta	Tengs and Wallace 2000[129]
Adjuvant therapy (TP)	Costs of chemotherapy for 6 months (£)	8788		Ward *et al.* 2007[128] (adjusted to 2007 prices)
	Annual costs after chemotherapy (£)	1002		
	Utility	0.74	Beta	Tengs and Wallace 2000[129]
Post-adjuvant therapy	Annual costs (£)	0		Assumption
	Utility	0.94	Beta	Tengs and Wallace 2000[129]
Locoregional relapse	Annual costs (£)	4737		Ward *et al.* 2007[128] (adjusted to 2007 prices)
	Utility	0.70	Beta	Tengs and Wallace 2000[129]
Remission	Annual costs for the first 5 years (£)	1002		Ward *et al.* 2007[128] (adjusted to 2007 prices)
	Annual costs after 5 years (£)	0		
	Utility	0.85	Beta	Tengs and Wallace 2000[129]
Metastatic relapse	Annual costs (£)	10,196		Ward *et al.* 2007[128] (adjusted to 2007 prices)
	Utility	0.40	Beta	Tengs and Wallace 2000[129]
Death	Costs (£)	3321		Ward *et al.* 2007[128] (adjusted to 2007 prices)
	Utility	0		By definition

Sanofi-aventis)-based chemotherapy is used.[17] The costs of chemotherapy were calculated based on a published study.[129]

Annual costs for the post-adjuvant therapy state are assumed to be £0. Annual costs for the locoregional and metastatic relapse states and the death state were calculated based on a published study and are £4737, £10,196 and £3321, respectively.[129] For patients in the remission state after locoregional relapse, the annual costs for the first 5 years were assumed to be the same

as for patients in the adjuvant therapy state receiving hormonal therapy (£1002). The costs are assumed to be £0 after 5 years in the remission state.

The source of utility data used in the model is from Tengs and Wallace,[130] which is a systematic review of health-related quality of life estimates from publicly available source documents. Given that the health-related quality of life in the general population decreases with age, it is important to take this into account in the model. General population utility estimates from Ara and Brazier[131] were applied using a regression analysis of utility versus age. The age-related utility is calculated by the following formula (*Equation 1*):

$$\text{Utility} = A \times (\text{Age}) + B \times (\text{Age} \times \text{Age}) + C \qquad \text{[Equation 1]}$$

where $A = -0.0001728$, $B = -0.000034$ and $C = 0.9584588$.

The utilities for all health states are multiplied by this age-related utility value for each year of the model.

Model inputs: transition probabilities between health states

Transition probabilities between health states are summarised in *Table 21*. When a patient starts adjuvant therapy, the expected life expectancy of the patient is determined by the life table.[132] The patient will die from other causes once the expected life expectancy is met. The transition probabilities from the adjuvant therapy state to the locoregional recurrence state depend on the diagnostic results for lymph node metastases (i.e. TN, FP, TP and FN). Node-negative patients, including TN and FP, will have lower transition probabilities and node-positive patients, including TP and FN, will have higher transition probabilities. In particular, patients with FN results will have the highest transition probability because they have been denied ALND and chemotherapy, needed to reduce the risk of recurrence. The annual transition probabilities for locoregional recurrence were based on a study by Orr *et al.*,[133] which provided the estimates of annual transition probabilities of recurrence in patients with negative results (0.03), TP results (0.09) and FN results (0.14). The study assumed that patients with TP results receive chemotherapy and patients with FN results do not receive chemotherapy (and only receive hormonal therapy). The same assumption is used in this assessment.

When a patient enters the adjuvant therapy state, the model uses exponential distributions to sample the time to locoregional relapse according to the above annual transition probabilities. The sampled time to locoregional relapse will then be compared with the time to death (according to the life table) and the 5-year maximum period for adjuvant therapy. Depending on which event happens first (locoregional relapse, death due to other causes or finishing the adjuvant therapy), the patient will transit to the corresponding state after the time delay. The methodology used to determine which state the patient transits to is the same for other health states.

When patients enter the post-adjuvant therapy state, they may experience locoregional or metastatic relapse, or they may die from other causes. The annual transition probabilities of locoregional relapse and the time to death from other causes are assumed to be the same as under the adjuvant therapy state. The transition probabilities to metastatic relapse are 0.0023 for node-negative patients (TN and FP), 0.0052 for TP patients, and 0.0094 for FN patients.[120]

If patients enter the locoregional relapse state, they may enter a subsequent remission state, may have metastatic relapse, or may die from other causes. It is assumed that patients can stay in the locoregional relapse state for a maximum of 1 year. The annual probability of developing

TABLE 21 Transition probabilities between health states

Parameter	Value	Source
Annual probability of locoregional relapse (TN and FP)	0.03	Orr *et al.* 1999[133]
Annual probability of locoregional relapse (TP)	0.09	
Annual probability of locoregional relapse (FN)	0.14	
Annual probability of metastatic relapse (TN and FP)	0.0023	Pandharipande *et al.* 2008[120]
Annual probability of metastatic relapse (TP)	0.0052	
Annual probability of metastatic relapse (FN)	0.0094	
Annual probability of metastatic relapse from locoregional relapse	0.18	Kamby and Sengelov 1997[134]
Annual probability of metastatic relapse from remission	0.13	
Annual probability of death from locoregional relapse	0.30	Orr *et al.* 1999[133]
Annual probability of death from metastatic relapse	0.37	Ward *et al.* 2007[129]

metastatic cancer in the first year of locoregional relapse is 0.18,[134] which is much higher compared with disease-free survival states (i.e. adjuvant therapy and post-therapy states). Orr *et al.*[133] suggested that the annual probability of death for patients with locoregional relapse is 0.30. This includes death due to both breast cancer and other causes. The model does not distinguish between death from breast cancer and death from other causes once a patient enters the locoregional relapse state. The maximum lifetime of a patient is still bounded by the life expectancy of the patient (i.e. death due to other causes).

If patients enter the remission state, they may still experience metastatic relapse before death. The average annual probability of developing metastatic cancer from the remission state is 0.13[134] and the probability of death (for all reasons) is assumed to be the same as in the locoregional relapse state, which is 0.30.

Metastatic cancer is assumed not to be curable and the annual probability of death from metastatic relapse is 0.37.[129]

Model inputs: probability, costs and utilities of long-term adverse events

The model inputs of probability and costs of long-term adverse events (i.e. lymphoedema) are summarised in *Table 22*. The probabilities of developing lymphoedema due to SLNB and ALND were estimated based on published studies.[48–50,53] The probability of having lymphoedema due to 4-NS was assumed to be the same as SLNB, as no data were identified for 4-NS.

Lymphoedema was classified as either mild/moderate or severe. A literature search was undertaken, but no studies reporting utility for patients with lymphoedema were identified. The proportion of patients within each category and the utility decrements of each category were therefore estimated from a published study reporting quality of life using the FACT-B + 4 (Functional Assessment of Cancer Therapy for Breast Cancer, adding a four-item arm subscale) quality-of-life instrument.[135] The study reported the data regarding the quality of life of breast cancer patients who suffer from different degrees of lymphoedema, using the FACT-B + 4 quality-of-life instrument. The utility decrements were estimated based on these quality-of-life data, therefore the decrements do not represent the true utility decrements due to lymphoedema. Sensitivity analyses were carried out to explore the impact on the cost-effectiveness results caused by changing the estimated utility decrements. The annual additional costs due to lymphoedema were based on expert opinions from the Sheffield Lymphoedema Service.

The utility decrement represents the reduced quality of life for patients with lymphoedema.

TABLE 22 Probabilities and costs of long-term adverse events

Long-term adverse events	Parameter	Value	Distribution	Source
Lymphoedema	Probability (4-NS)	0.068	Beta	Expert opinion – assumed same as SLNB
	Probability (SLNB)	0.068 (95% CI 0.062 to 0.074)	Beta	Liu *et al.* 2009[53]
	Probability (ALND)	0.214 (95% CI 0.18 to 0.252)	Beta	Blanchard *et al.* 2003;[48] Crane-Okada *et al.* 2008;[49] McLaughlin *et al.* 2008[50]
	Proportion of mild/moderate lymphoedema (%)	66.3		Mak *et al.* 2009[135]
	Proportion of severe lymphoedema (%)	33.7		
	Additional costs of mild/moderate lymphoedema (£)	66.50		Expert opinions from the Sheffield Lymphoedema Service
	Additional costs of severe lymphoedema (£)	1180.00		
	Assumed utility decrement due to mild/moderate lymphoedema (%)	9.9		Mak *et al.* 2009[135]
	Assumed utility decrement due to severe lymphoedema (%)	12.3		

Discounting

The economic analysis assumes that both costs and QALYs are discounted at 3.5% per annum, in line with current recommendations from Her Majesty's Treasury.[136]

Univariate sensitivity analysis

In order to explore the impact on the cost-effectiveness results of changes to individual parameters and assumptions, a number of sensitivity analyses were performed.

Sensitivity and specificity of magnetic resonance imaging and positron emission tomography

The analysis is limited by the small number and size of MRI studies, and the wide variations between and within studies in terms of the MRI method used. A sensitivity analysis was carried out to decrease the mean sensitivity of MRI to the lower CI (from 90% to 78%) and maintain the mean specificity of MRI. Another sensitivity analysis was carried out to decrease the mean specificity of MRI to the lower CI (from 90% to 75%) and maintain the mean sensitivity of MRI.

In order to test the sensitivity of model results to increased MRI accuracy, a sensitivity analysis was carried out to increase both the sensitivity and specificity of MRI to the levels for USPIO-enhanced MRI. USPIO-enhanced MRI is a subtype of MRI that appears to have higher sensitivity and specificity (98% and 96%, respectively). In this sensitivity analysis the cost of MRI was assumed to increase by £100 to take account of the additional cost of the contrast agent used in USPIO-enhanced MRI.

Regarding PET, a sensitivity analysis was carried out to increase the sensitivity of PET to the higher CI (from 63% to 74%) and maintain the mean specificity.

Utility decrements and additional costs for lymphoedema

The main advantage of imaging techniques is to reduce short- and long-term adverse events including lymphoedema. Therefore, the utility decrements and additional costs for lymphoedema will impact on the cost-effectiveness of imaging techniques compared with sampling methods. Data on the long-term costs and utility impact of lymphoedema are, however, limited. Two

sensitivity analyses were carried out to increase/decrease the utility decrements for lymphoedema by 50%. Another two sensitivity analyses were performed to increase/decrease the additional costs due to lymphoedema by 20%.

Probabilities of relapse for false-negative patients

Two sensitivity analyses were carried out to increase/decrease the probabilities of locoregional relapse for patients with FN diagnoses by 20%. Due to lower sensitivity, imaging techniques, especially PET, produce more FN cases than 4-NS and SLNB. The probabilities of relapse for patients with FN diagnoses were changed so that the impact on model results can be assessed.

Costs of sampling methods

High-quality UK cost data for 4-NS and SLNB procedures have not been identified. The costs used in the model were derived from non-UK studies. Sensitivity analyses were carried out to increase/decrease the costs of 4-NS and SLNB by 20%.

Probabilistic sensitivity analysis

Probabilistic sensitivity analysis (PSA) was undertaken to demonstrate the impact of uncertainty in the key model parameters and to generate information on the likelihood that each of the diagnostic strategies is optimal.

The sensitivity and specificity of a diagnostic test are generally correlated. To maintain this correlation, the sensitivities of each diagnostic method are sampled from a beta distribution, while the specificities of the test are derived based on the sampled sensitivity. The sensitivity and specificity are linked by the prevalence and the overall accuracy of the test, which are assumed to be constants. Therefore, when the sensitivity is sampled, the specificity can be calculated deterministically.

Equation 2 is the formula for calculating the test accuracy. After rearranging *Equation 2*, the model applies *Equation 3* to derive test specificity from sensitivity. The overall accuracy of each diagnostic method is presented in *Table 23*, where the prevalence of early breast cancer among breast cancer patients is assumed to be fixed at 41.2%. The beta distributions representing the sensitivity of each diagnostic method are based on trial data from previous literature and the systematic review within this assessment (MRI and PET).

$$\text{Accuracy} = \text{Sensitivity} \times \text{Prevalence} + \text{Specificity} \times (1 - \text{Prevalence}) \quad \text{[Equation 2]}$$

$$\text{Specificity} = (\text{Accuracy} - \text{Sensitivity} \times \text{Prevalence})/(1 - \text{Prevalence}) \quad \text{[Equation 3]}$$

Health utilities and the probabilities of developing short- and long-term adverse events were modelled using beta distributions. Costs were sampled from normal distributions where standard errors can be calculated. Transition probabilities among health states and other costs were sampled from uniform distributions bounded by a 10% increase or decrease in the mean value.

The PSA was carried out by allowing the key model input parameters to vary according to the uncertainty specified in their probability distributions, with 500 sets of random numbers used to generate 500 parameter configurations, which produce 500 sets of model outputs. All model results were based on the PSA model outputs.

To demonstrate that 500 replications is enough to obtain accurate model outputs, the cumulative mean total costs and total QALYs of the baseline 4-NS diagnostic strategy based on 2000 replications were calculated. A significance level of 5% was used to construct the CI around the

TABLE 23 Overall accuracies of each diagnostic method

Diagnostic method	Overall accuracy (%)
Clinical examination	64.8
Ultrasound for clinically negative	78.2
Ultrasound for clinically positive	71.5
Biopsy	88.7
4-NS	97.7
SLNB	97.1
MRI	89.9
PET	81.5
ALND	100.0

cumulative means. The analysis suggests that the CI was sufficiently narrow and the cumulative mean stabilises after 500 replications are performed.

Independent economic assessment – results

This section details the results of the health economic model. The cost-effectiveness results of imaging techniques are presented as marginal estimates when compared against the standard sampling methods of 4-NS and SLNB. All results are presented in terms of net benefits and incremental cost per QALY gained.

Base-case estimates of cost-effectiveness

The base-case estimates given below are the mean estimates from the 500 runs of the PSA. For each strategy, results are presented in terms of the diagnostic results, the total number of diagnostic and surgical procedures performed, the total costs and QALYs, the net benefit and incremental cost-effectiveness.

Diagnostic results

The proportions of patients whose lymph node diagnostic results are TN, FP, TP and FN are presented in *Table 24* and *Figure 16*.

Key findings from the diagnostic results

Replacing sampling

1. The number of FP cases increases significantly when sampling is replaced with imaging techniques (from 0.2% to 6.3% and 3.6% for MRI and PET, respectively). The main reason is that imaging techniques have lower specificity (89.7% and 94.2% for MRI and PET, respectively).
   Patients with FP diagnoses will receive ALND, as a stand-alone procedure (if detected by 4-NS or SLNB where the breast surgery is performed at the same time) or at the same time as the breast surgery (if detected by imaging techniques), which they actually do not need. The surgery is associated with short- and long-term adverse events which will both increase costs and affect quality of life of the patients. Since ALND is carried out for patients with FP diagnoses, the negative nodal status will be confirmed and the patients will be managed as node-negative patients afterwards.
2. The number of FN cases increases slightly when the sampling methods are replaced with MRI. FN cases increase significantly when the sampling methods are replaced with PET (from around 1% to 7.2%) due to the low sensitivity of PET (63.4%).

TABLE 24 Diagnostic results of each strategy

Diagnostic strategy	TN (%)	FP (%)	TP (%)	FN (%)	Total (%)
Baseline 1: 4-NS	58.6	0.2	40.1	1.1	100.0
Baseline 2: SLNB	58.6	0.2	39.9	1.3	100.0
Scenario 1: replace sampling with MRI	52.5	6.3	39.3	1.9	100.0
Scenario 2: replace sampling with PET	55.2	3.6	34.0	7.2	100.0
Scenario 3: add MRI before 4-NS	52.5	6.3	41.1	0.1	100.0
Scenario 4: add PET before 4-NS	55.2	3.6	40.8	0.4	100.0
Scenario 5: add MRI before SLNB	52.5	6.3	41.1	0.1	100.0
Scenario 6: add PET before SLNB	55.2	3.6	40.7	0.5	100.0

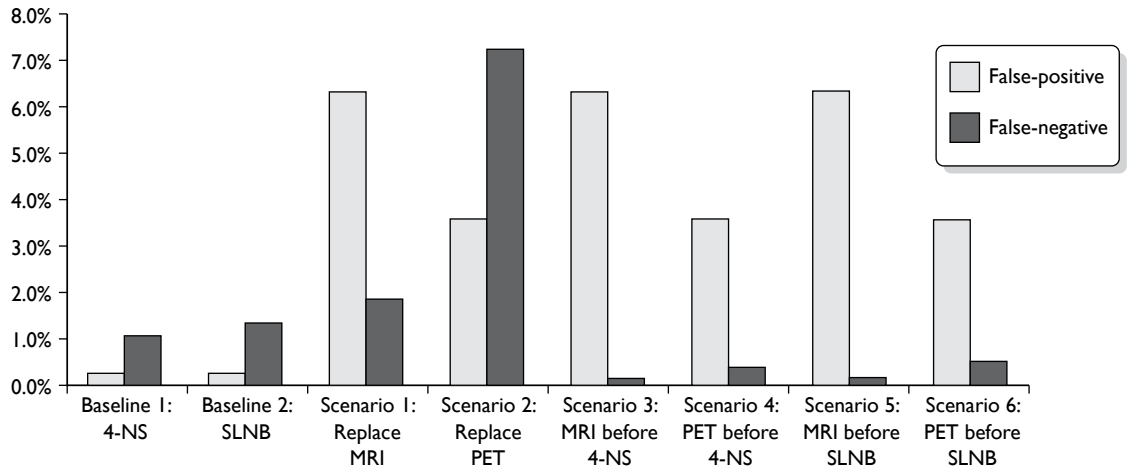


FIGURE 16 Proportion of people with FP and FN diagnostic results under each strategy.

For patients with FN diagnoses, as no ALND will be performed, the true nodal status will remain unknown. The FN patients will miss the ALND which they actually need, resulting in a higher risk of locoregional and distant relapse which have significant costs and quality of life implications.

3. As MRI has a lower specificity and a higher sensitivity than PET, the strategy that involves MRI will have more FP cases and fewer FN cases than strategies involving PET. However, the evidence on the accuracy of MRI is less robust, given that there are fewer studies on MRI than PET. The systematic review in the assessment also demonstrated that the sensitivity and specificity of both PET and MRI vary significantly between studies.
4. The two strategies of replacing axillary sampling with imaging techniques produce higher levels of FP cases and (particularly for PET) FN cases. These strategies may be considered unacceptable on clinical grounds, even if they are found to be more cost-effective.

Adding imaging techniques before sampling

1. When the imaging techniques are placed before the sampling methods (scenarios 3–6), the number of FP cases is increased from baseline to the same extent as in the corresponding replacement strategies (as expected according to the diagnostic pathway).
2. The benefit of putting imaging techniques before sampling methods is that they will identify a proportion of TP patients (depending on the sensitivity of the tests) who will undergo the ALND straight away (at the same time as the main breast surgery), rather than undergoing

both sampling (at the same time as the breast surgery) and a separate ALND procedure. This will reduce costs and possibly morbidity associated with having two sequential operations.

3. As expected, the number of FN cases is reduced when imaging techniques are placed before sampling due to the use of two sequential diagnostic tests. When MRI is placed before the sampling methods, the proportion of FN cases drops to about 0.1%.
4. Strategies that place imaging techniques before sampling methods produce fewer FN cases. However, they also produce significantly more FP cases, especially in the case of PET. This may be considered unacceptable on clinical grounds.

Resource use for diagnosis and surgical procedures

Table 25 presents the number of main diagnostic tests and surgical procedures carried out under each diagnostic strategy based on 5000 patients. Breast surgery is performed as a stand-alone procedure only if the negative nodal status is obtained by an imaging technique and no sampling methods are used (in scenario 1 or 2).

The key findings from the resource use results

Replacing sampling

When sampling methods are replaced with imaging techniques, all ALND procedures will be performed at the same time as the breast surgery as node-positive patients are diagnosed by either biopsy or non-invasive imaging techniques. It is also possible that breast surgery is performed as a stand-alone procedure if negative nodal status is obtained by the imaging techniques.

The same numbers of imaging techniques are carried out in scenarios 1 and 2, in which sampling methods are replaced with MRI and PET. There are more ALND procedures in scenario 1 than in scenario 2 (2283 vs 1877) as MRI is associated with more TP and FP cases than PET (see *Table 24*) due to its higher sensitivity and lower specificity. For the same reason, there are fewer stand-alone breast surgeries in scenario 1 than in scenario 2 (2717 vs 3123).

Adding imaging techniques before sampling

When imaging techniques are introduced before the sampling methods, ALND will be performed as a stand-alone procedure if positive nodal status is obtained from sampling procedures. ALND will be carried out at the same time as the breast surgery if positive nodal status is obtained from either imaging techniques or biopsy. As in the two baseline scenarios, it is

TABLE 25 Number of diagnostic tests and surgical procedures carried out for each diagnostic strategy (based on 5000 patients)

Diagnostic strategy	MRI	PET	4-NS (with breast surgery)	SLNB (with breast surgery)	ALND (stand-alone)	ALND (with breast surgery)	ALND – subtotal	Breast surgery (stand-alone)
Baseline 1: 4-NS	N/A	N/A	3913	N/A	930	1087	2017	0
Baseline 2: SLNB	N/A	N/A	N/A	3913	916	1087	2003	0
Scenario 1: replace sampling with MRI	3913	N/A	N/A	N/A	0	2283	2283	2717
Scenario 2: replace sampling with PET	N/A	3913	N/A	N/A	0	1877	1877	3123
Scenario 3: add MRI before 4-NS	3913	N/A	2716	N/A	88	2284	2372	0
Scenario 4: add PET before 4-NS	N/A	3914	3123	N/A	341	1877	2219	0
Scenario 5: add MRI before SLNB	3914	N/A	N/A	2716	87	2284	2371	0
Scenario 6: add PET before SLNB	N/A	3914	N/A	3123	336	1877	2213	0

N/A, not applicable.

not possible to have a stand-alone breast surgery when the imaging techniques are placed before the sampling methods.

As the imaging techniques are performed before the sampling methods, the numbers of imaging techniques performed are the same as the replacement scenarios. Compared with the two baseline scenarios, the numbers of sampling procedures carried out are reduced (from 3913 to 2716 if MRI is placed before sampling and to 3123 if PET is placed before sampling). The reduction is more evident when MRI is placed before sampling than PET, as MRI is associated with more positive cases (both true and false), which do not require further sampling methods.

Adverse events

The numbers of short- and long-term adverse events associated with each diagnostic strategy are presented in *Table 26* and *Figure 17*. The numbers of short- and long-term adverse events are proportional to the number of sampling and ALND procedures undertaken (see *Table 25*). The model assumes that the probabilities of short-term adverse events are the same for both 4-NS and SLNB, while ALND is associated with much higher probabilities of developing adverse events.

The key findings from the adverse event results

Replacing sampling

The two replacement strategies (scenarios 1 and 2) have both fewer short- and long-term adverse events than the two baseline strategies, despite more ALND procedures being undertaken. This is due to the number of sampling procedures avoided (see *Table 25*). The PET replacement strategy has fewer adverse events than the MRI replacement strategy as fewer ALND procedures are performed.

Adding imaging techniques before sampling

In general, there are more short- and long-term adverse events when imaging techniques are placed before sampling methods (scenarios 3–6). Under these scenarios, some sampling procedures are avoided (if positive nodal status is obtained from imaging techniques). However, more patients will undergo ALND (either stand-alone or with breast surgery) in these scenarios as the number of patients with both TP and FP diagnoses will increase (see *Table 24*). The model suggests that the adverse events associated with increased ALND procedures outnumber the adverse events associated with avoided sampling procedures (ALND is associated with significantly higher probability of developing adverse events than sampling methods).

Survival results

The 5-year survival rates for patients with different diagnostic results are presented in *Figure 18*. Patients with axillary lymph node metastases (both TP and FN) have lower survival rates than

TABLE 26 Adverse event cases associated with each diagnostic strategy

Diagnostic strategy	Short-term adverse events	Long-term adverse events (lymphoedema)
Baseline 1: 4-NS	2781	633
Baseline 2: SLNB	2764	631
Scenario 1: replace sampling with MRI	2642	487
Scenario 2: replace sampling with PET	2172	401
Scenario 3: add MRI before 4-NS	3055	684
Scenario 4: add PET before 4-NS	2925	662
Scenario 5: add MRI before SLNB	3055	684
Scenario 6: add PET before SLNB	2919	661

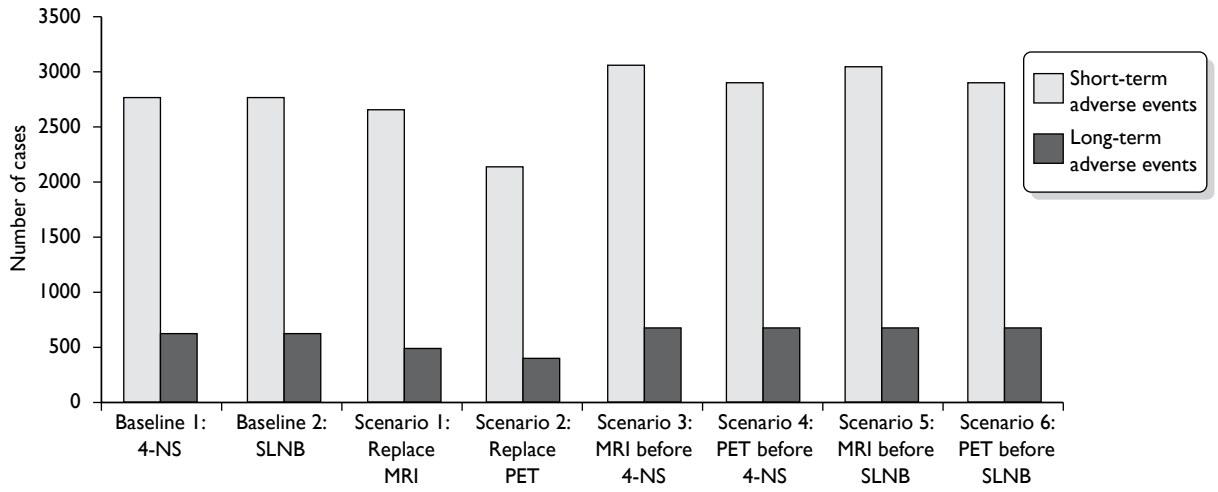


FIGURE 17 Adverse event cases associated with each diagnostic strategy.

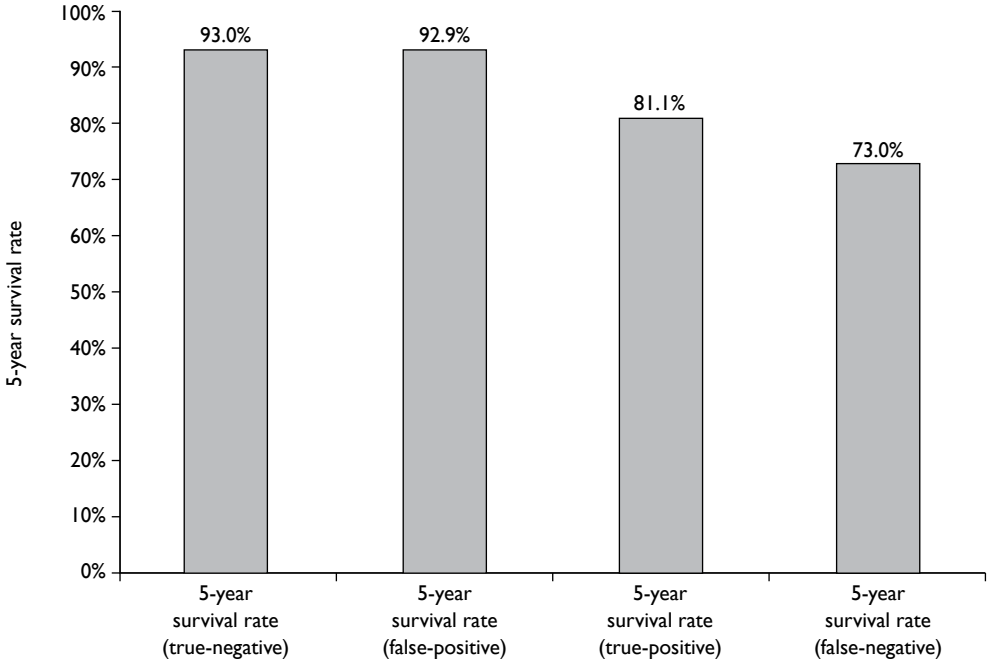


FIGURE 18 The 5-year survival rates for patients with different diagnostic results.

patients with negative nodal status. In particular, patients with FN results are associated with the lowest survival rate, because they do not receive ALND or chemotherapy that may reduce the chance of recurrence.

The 5-year survival rates of all patients with early breast cancer and the absolute number of deaths per 10,000 patients presenting with early breast cancer are presented in *Table 27* for each tested diagnostic strategy. The overall survival rate of each strategy is dependent on the proportion of patients with each diagnostic result (see *Table 24*) and the individual survival rate

TABLE 27 The 5-year survival rates and the absolute number of deaths in England for patients with early breast cancer (comparing different diagnostic strategies)

Diagnostic strategy	5-year survival rate (all patients) (%)	Number of deaths during first 5 years (per 10,000 patients)	Absolute difference in number of deaths during first 5 years (per 10,000 patients)
Baseline 1: 4-NS	88.03	1197	0
Baseline 2: SLNB	88.00	1200	3
Scenario 1: replace MRI	87.96	1204	7
Scenario 2: replace PET	87.52	1248	51
Scenario 3: MRI before 4-NS	88.10	1190	–7
Scenario 4: PET before 4-NS	88.08	1192	–5
Scenario 5: MRI before SLNB	88.10	1190	–7
Scenario 6: PET before SLNB	88.07	1193	–4

for patients with different diagnostic results (see *Figure 18*). As patients with FN results have the lowest survival rate, diagnostic strategies that are associated with more FN results have lower overall survival rate (e.g. PET replacement strategy). The absolute differences in overall survival rates among tested diagnostic strategies are relatively small (range from 87.52% to 88.1%). This is because the absolute numbers of patients with FN results only account for a small proportion of patients (range from 0.1% to 7.2% as shown in *Table 24*). The absolute difference in the number of deaths during the first 5 years per 10,000 early-stage breast cancer patients is also presented in *Table 27*. This shows the change in absolute mortality for each tested diagnostic strategy, using the 4-NS baseline as the reference strategy.

Cost-effectiveness results (net benefit analysis)

Positron emission tomography and MRI are assumed to be associated with no short- and long-term adverse events. However, due to the lower accuracy of the imaging techniques, more FP and FN cases will be produced, which will lead to increased costs, worse quality of life due to adverse events, and in some cases higher probability of recurrence and subsequent death from breast cancer. Economic modelling provides a systematic way to understand and quantify the complex trade-offs between the advantages and disadvantages of the imaging techniques, so that the overall cost-effectiveness of alternative strategies can be determined.

Net benefit is the increase in effectiveness (ΔE), multiplied by the amount the decision-maker is willing to pay per unit of increased effectiveness (R_T), less the increase in cost (ΔC). The formula for calculating net benefit is:

$$\text{Net benefit} = R_T \Delta E - \Delta C > 0 \qquad \text{[Equation 4]}$$

A strategy is most cost-effective if it has the highest positive net benefit. Thresholds of willingness to pay per QALY of £10,000, £20,000 and £30,000 were used to calculate the net benefit of each strategy.

Two baseline strategies (4-NS and SLNB) are considered. Both 4-NS and SLNB are currently used in the UK. It is not within the scope of this assessment to compare these sampling methods. *Tables 28* and *29* summarise the net benefits of each strategy using either 4-NS or SLNB as the baseline strategy. Note that scenarios 1 and 2 appeared in both tables because they are comparable to both 4-NS and SLNB strategies. The total costs and QALYs of each diagnostic strategy are plotted in *Figures 19* and *20*.

TABLE 28 Total costs and QALYs and the net benefit of each diagnostic strategy using 4-NS as baseline

	Costs associated with diagnostic tests or short-term adverse events (£)	Costs associated with health states or long-term adverse events (£)	Total costs (£)	Total QALYs	Net benefit (£10,000 per QALY)	Net benefit (£20,000 per QALY)	Net benefit (£30,000 per QALY)
Baseline 1: 4-NS	3157	16,571	19,728	8.122		Baseline	
Scenario 1: replace sampling with MRI	3030	16,295	19,325	8.174	**919**	**1435**	**1952**
Scenario 2: replace sampling with PET	3484	15,835	19,319	8.126	450	490	531
Scenario 3: add MRI before 4-NS	3204	16,655	19,859	8.125	–104	–78	–51
Scenario 4: add PET before 4-NS	3818	16,623	20,440	8.126	–681	–649	–618

TABLE 29 Total costs and QALYs and the net benefit of each diagnostic strategy using SLNB as baseline

	Costs associated with diagnostic tests or short-term adverse events (£)	Costs associated with health states or long-term adverse events (£)	Total costs (£)	Total QALYs	Net benefit (£10,000 per QALY)	Net benefit (£20,000 per QALY)	Net benefit (£30,000 per QALY)
Baseline 2: SLNB	3642	16,547	20,189	8.119		Baseline	
Scenario 1: replace sampling with MRI	3030	16,295	19,325	8.174	**1412**	**1959**	**2507**
Scenario 2: replace sampling with PET	3484	15,835	19,319	8.126	942	1014	1085
Scenario 5: add MRI before SLNB	3546	16,655	20,201	8.124	35	82	129
Scenario 6: add PET before SLNB	4208	16,614	20,822	8.125	–577	–520	–464

The horizontal axis of *Figures 19* and *20* represents the total QALYs accrued by each diagnostic strategy and the vertical axis represents the total costs associated with each strategy. The lines in the figures denote the cost-effective frontiers if PET and MRI replacement strategies are excluded based on clinical grounds.

The total and breakdown costs (costs associated with diagnostic tests or short-term adverse events and costs associated with health states or long-term adverse events) of each strategy are illustrated in *Figure 21*.

Key findings from the net benefit results

Replacing sampling

1. When 4-NS is used as the baseline (see *Table 28* and *Figure 19*), the most cost-effective strategy is to replace sampling with MRI (scenario 1), which has the highest net benefits under all willingness-to-pay thresholds tested. The next most cost-effective strategy is to replace sampling with PET (scenario 2). The baseline 4-NS strategy is dominated by both scenario 1 and scenario 2, as they have lower total costs and higher total QALYs.
2. When SLNB is used as the baseline (see *Table 29* and *Figure 20*), the most cost-effective strategy is still to replace sampling with MRI (scenario 1), which has the highest net benefits

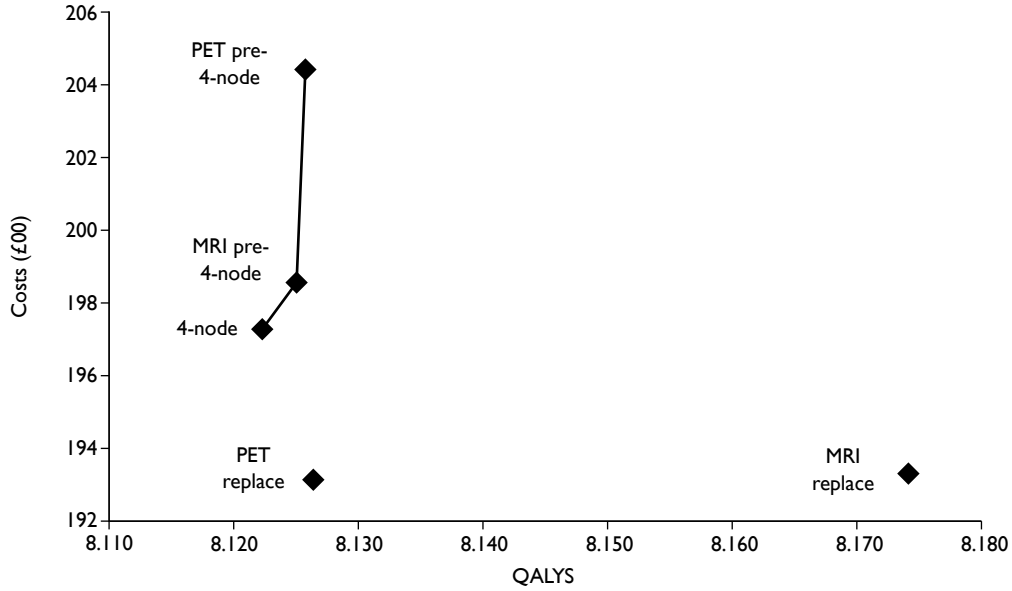


FIGURE 19 Total costs and QALYs of diagnostic strategies using 4-node sampling as baseline.

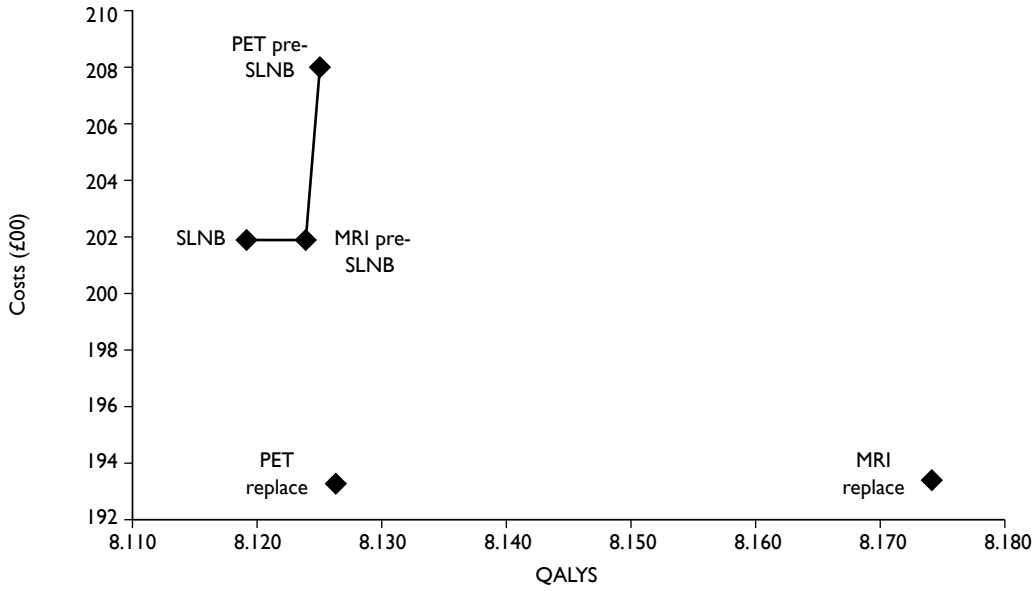


FIGURE 20 Total costs and QALYs of diagnostic strategies using sentinel lymph node biopsy as baseline.

under all willingness-to-pay thresholds tested. The next most cost-effective strategy is also to replace sampling with PET (scenario 2). The baseline SLNB strategy is dominated by both scenarios 1 and 2, as they have lower total costs and higher total QALYs.

3. MRI has reasonably good sensitivity and specificity and lower cost than PET, 4-NS and SLNB. Compared with the baseline 4-NS and SLNB strategies, the disadvantages of the MRI replacement strategy are that it is associated with more FP cases (increased from 0.2% to 6.3%) resulting in many node-negative patients undergoing unnecessary ALND, and more FN cases (increased from around 1.0% to 1.9%) who are left at higher risk of cancer recurrence. The advantage of the MRI replacement strategy, compared with the two baseline strategies, is that many node-positive patients will be correctly diagnosed by MRI and

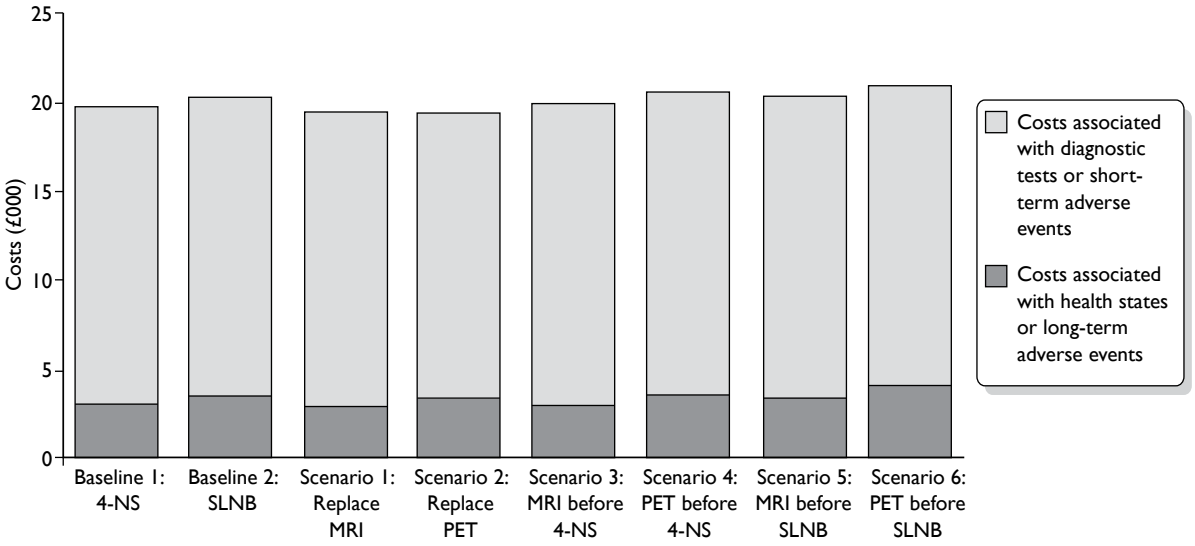


FIGURE 21 Costs associated with each diagnostic strategy.

undergo ALND (at the same time as the breast surgery) rather than undergoing two surgical procedures (4-NS or SLNB followed by ALND). For TN patients, the advantage of the MRI replacement strategy is that these patients will be correctly diagnosed without the need for a sampling procedure. The sampling procedures are associated with increased costs and risk of short- and long-term adverse events. The model suggests that the advantages of the MRI replacement strategy outweigh the disadvantages in relation to benefits as expressed by QALYs.

4. The advantages and disadvantages of MRI compared with sampling methods also applies to PET. Compared with the baseline strategies, the model also suggests that the advantages of PET outweigh the disadvantages for both costs and QALYs. Compared with the MRI replacement strategy, the PET replacement strategy has similar total costs but significantly lower total QALYs. This is because PET is associated with more FN cases due to lower sensitivity and patients with FN diagnoses are more likely to experience locoregional and metastatic relapse.

Adding magnetic resonance imaging or positron emission tomography before sampling

The MRI and PET replacement strategies may be deemed unacceptable on clinical grounds due to higher numbers of patients with FN and FP diagnoses. When the replacement strategies are excluded, the baseline 4-NS and SLNB strategies were only compared with the strategies of adding MRI or PET before sampling methods.

1. The most cost-effective strategy is to retain the baseline 4-NS strategy (when 4-NS is used as the baseline) or to place MRI before SLNB (when SLNB is used as the baseline).
2. The advantages of the strategies of adding MRI or PET before sampling methods are that there are fewer FN cases (reduced from around 1.0% to 0.1% for MRI) due to the use of two sequential tests, and fewer sampling procedures performed (because sampling methods are avoided if MRI or PET results are positive). The disadvantages of these strategies are that there are more FP cases because the specificities of MRI and PET are lower than those of SLNB and 4-NS (FPs increase from 0.2% to 6.3% for MRI prior to SLNB, which is the same as for the MRI replacement strategy). Overall, the cost-effectiveness results suggest that there

are both higher costs and higher QALYs associated with strategies of adding MRI or PET before sampling methods compared with the baseline 4-NS and SLNB strategies.

3. In terms of cost-effectiveness, the model results suggest that adding MRI prior to SLNB is cost-effective, whereas adding MRI prior to 4-NS is not. This is because the addition of MRI means that fewer sampling procedures are required, and the cost saving associated with this is greater for SLNB than for 4-NS because SLNB is more costly than 4-NS.
4. The absolute differences in QALYs among the baseline 4-NS and SLNB strategies and the strategies of adding MRI or PET before sampling are very small. When the MRI and PET replacement strategies are excluded, the total QALYs range from 8.122 to 8.126 when 4-NS is used as the baseline and from 8.119 to 8.125 when SLNB is used as the baseline. This implies that there is no significant absolute improvement in QALYs when MRI or PET are added before 4-NS or SLNB. For example, although the model results suggest the strategy of adding MRI before SLNB is cost-effective, the change from the baseline SLNB strategy to the alternative strategy of MRI before SLNB only increases the QALYs by 0.005. The main reason that the alternative strategy is cost-effective is that there is an even smaller increase in total costs (relative to the QALY increase).

Cost-effectiveness results (incremental analysis)

The incremental cost-effectiveness analyses were performed assuming that MRI and PET replacement strategies are deemed unacceptable on clinical grounds. *Tables 30* and *31* show the incremental cost-effectiveness analyses where 4-NS and SLNB are used as the baseline strategy.

The cost-effectiveness plane for the MRI before 4-NS strategy versus the baseline 4-NS strategy is presented in *Figure 22*. The cost-effectiveness plane for the MRI before SLNB strategy versus the baseline SLNB strategy is presented in *Figure 23*.

Both *Figures 22* and *23* demonstrate that MRI before sampling strategy may lead to either an increase or decrease in QALYs compared with the baseline strategy. The advantage of this strategy in terms of QALYs is therefore uncertain. *Figure 22* shows that the MRI before 4-NS typically generates higher total costs than the 4-NS strategy. *Figure 23* shows that MRI before SLNB strategy may lead to either an increase or a decrease in total costs compared with the SLNB strategy. The benefits offered by these strategies are not clear-cut.

TABLE 30 Incremental cost-effectiveness analysis of diagnostic strategies using 4-NS as baseline (excluding MRI and PET replacement strategies)

	Total cost (£)	Total QALYs	Incremental cost (£)	Incremental QALY	ICER (£)
Baseline 1: 4-NS	19,728	8.122		Baseline	
Scenario 3: add MRI before 4-NS	19,859	8.125	131	0.003	48,986
Scenario 4: add PET before 4-NS	20,440	8.126	581	0.000	1,200,212

ICER, incremental cost-effectiveness ratio.

TABLE 31 Incremental cost-effectiveness analysis of diagnostic strategies using SLNB as baseline (excluding MRI and PET replacement strategies)

	Total cost (£)	Total QALYs	Incremental cost (£)	Incremental QALY	ICER (£)
Baseline 2: SLNB	20,189	8.119		Baseline	
Scenario 5: add MRI before SLNB	20,201	8.124	11	0.005	2451
Scenario 6: add PET before SLNB	20,822	8.125	622	0.001	647,415

ICER, incremental cost-effectiveness ratio.

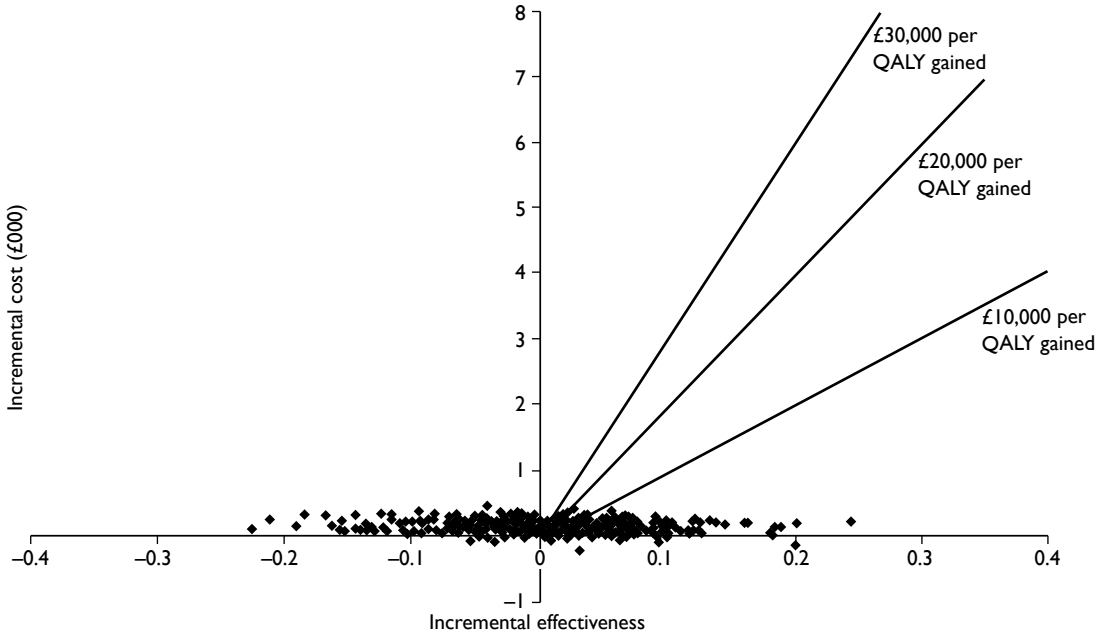


FIGURE 22 Cost-effectiveness plane for MRI before 4-node sampling (4-NS) strategy versus baseline 4-NS strategy.

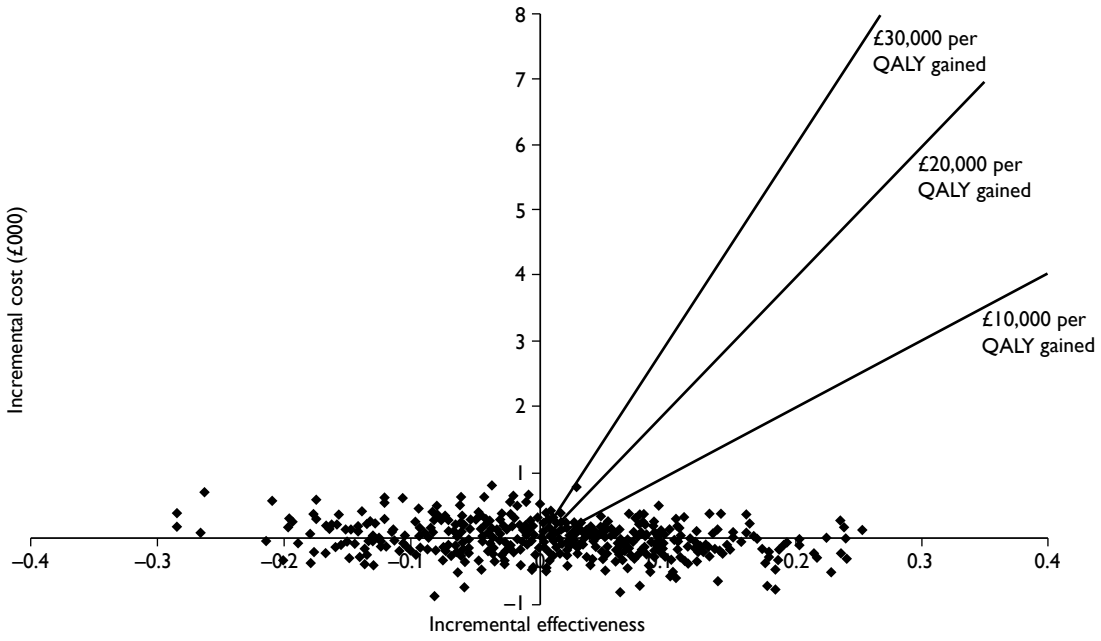


FIGURE 23 Cost-effectiveness plane for MRI before SLNB strategy versus baseline SLNB strategy.

Univariate sensitivity analysis

Sensitivity and specificity of magnetic resonance imaging

When the sensitivity of MRI is reduced from 90% to 78% (the lower CI) while maintaining the mean specificity of 90%, the strategy of MRI replacement is still the most cost-effective strategy when either 4-NS or SLNB is used as baseline. If the MRI and PET replacement strategies are

rejected on clinical grounds, then the conclusions differ from the baseline results. When SLNB is used as the baseline, the most cost-effective strategy is to retain the baseline SLNB strategy when a willingness-to-pay threshold of £20,000 per QALY is used. This differs from the baseline results, where the most cost-effective strategy is the addition of MRI before SLNB.

When the specificity of MRI is reduced from 90% to 75% (the lower CI), while maintaining the mean sensitivity of 90%, the PET replacement strategy becomes the most cost-effective strategy when a willingness-to-pay threshold of £20,000 per QALY is used. This conclusion differs from the baseline results where the MRI replacement strategy is most cost-effective. If the MRI and PET replacement strategies are rejected on clinical grounds, then the strategy of adding MRI before SLNB is dominated by the baseline SLNB strategy. This means it is cost-effective to retain the baseline SLNB strategy.

When the sensitivity and specificity of MRI are increased to the levels of USPIO-enhanced MRI (from 90% to 98% and from 90% to 96%, respectively), the MRI replacement strategy remains the most cost-effective strategy. If the MRI and PET replacement strategies are rejected based on clinical grounds, then the most cost-effective strategy is to add MRI before 4-NS (when 4-NS is used as baseline) and to add MRI before SLNB (when SLNB is used as baseline).

The sensitivity analyses demonstrate that the sensitivity and specificity of MRI have a significant impact on the cost-effectiveness results.

Sensitivity of positron emission tomography

When the sensitivity of PET is increased from 63% to 74% (the upper CI), while maintaining the mean specificity of 94%, the baseline cost-effectiveness results do not change. The total QALYs of the three strategies that involve PET, including the PET replacement strategy and strategies to add PET before 4-NS and SLNB, were all increased. However, the increase is not significant enough to alter the conclusions relating to cost-effectiveness. The sensitivity analysis demonstrates that the sensitivity of PET does not appear to have a significant impact on cost-effectiveness results.

Utility decrement for lymphoedema

When the assumed utility decrement for lymphoedema is reduced by 50% (i.e. from 9.9% to 5.0% for mild/moderate lymphoedema and from 12.3% to 6.2% for severe lymphoedema), the MRI replacement strategy is still the most cost-effective strategy. However, the total QALYs for the PET replacement strategy, which is the second most effective strategy, becomes the smallest among all diagnostic strategies modelled. The PET replacement strategy is associated with the fewest cases of lymphoedema (see *Table 26*) and is therefore most affected. The cost-effectiveness results remain unchanged if the MRI and PET replacement strategies are rejected on clinical grounds.

When the assumed utility decrement for lymphoedema is increased by 50% (i.e. from 9.9% to 14.9% for mild/moderate lymphoedema and from 12.3% to 18.5% for severe lymphoedema), the baseline cost-effectiveness results do not change. Although the total QALYs for all strategies are reduced due to a higher utility decrement for lymphoedema, the sensitivity analysis shows that the decrease is smallest for the PET replacement strategy, as this strategy has the smallest number of lymphoedema cases. However, the MRI replacement strategy is still the most cost-effective.

The sensitivity analyses demonstrate that the PET replacement strategy is most affected by the change of utility decrements for lymphoedema. However, the overall cost-effectiveness results do not appear to be significantly affected by this parameter.

Additional costs of lymphoedema

When the annual additional cost of lymphoedema is decreased by 20% (i.e. from £66.50 to £53.20 for mild/moderate lymphoedema and from £1180 to £944 for severe lymphoedema) or increased by 20% (i.e. from £66.50 to £79.80 for mild/moderate lymphoedema and from £1180 to £1416 for severe lymphoedema), the baseline cost-effectiveness results do not change. The sensitivity analysis demonstrates that the additional cost of lymphoedema does not appear to have a significant impact on cost-effectiveness results.

Probability of relapse for false-negative patients

When the probability of relapse for patients with FN diagnoses is reduced by 20% (i.e. from 0.140 to 0.112 for locoregional relapse and from 0.0094 to 0.0075 for metastatic relapse), the MRI and PET replacement strategies remain the most and second most effective strategies. However, if the MRI or PET replacement strategies are rejected on clinical grounds, then the most cost-effective strategy is either the baseline 4-NS strategy or the SLNB strategy (depending on which baseline is used) which dominate the strategies of adding MRI before 4-NS and SLNB, respectively.

When the probability of relapse for patients with FN diagnoses is increased by 20% (i.e. from 0.14 to 0.17 for locoregional relapse and from 0.0094 to 0.0113 for metastatic relapse), baseline cost-effectiveness results do not change. The total QALYs for all strategies are reduced due to the increased probability of relapse. The sensitivity analysis shows that the decrease in QALYs is most significant for the PET replacement strategy. The total QALYs for this strategy, which dominates the baseline 4-NS and SLNB strategies in the baseline results, becomes the lowest among all strategies. The PET replacement strategy is affected the most because it is associated with the largest number of patients with FN diagnoses.

The sensitivity analysis shows that cost-effectiveness results are affected by the change in probability of relapse for FN patients. Among all strategies, the PET replacement strategy is most affected.

Cost of 4-node sampling

When the cost of 4-NS is decreased by 20% (i.e. from £2099 to £1679) or increased by 20% (i.e. from £2099 to £2518), the baseline cost-effectiveness results do not change. The sensitivity analyses demonstrate that the cost of 4-NS does not appear to have a significant impact on cost-effectiveness results.

Cost of sentinel lymph node biopsy

When the cost of SLNB is decreased by 20% (i.e. from £2728 to £2182), the MRI replacement strategy remains the most cost-effective strategy. However, if the MRI and PET replacement strategies are rejected on clinical grounds and SLNB is used as the baseline, then the most cost-effective strategy is to retain the baseline SLNB strategy when a willingness-to-pay threshold of £20,000 is used. This differs from the baseline results where the most cost-effective strategy is the addition of MRI before SLNB. When the cost of SLNB is increased by 20% (i.e. from £2728 to £3274), the baseline cost-effectiveness results do not change.

The sensitivity analyses demonstrate that the cost of SLNB influences the cost-effectiveness results, when the cost of SLNB is decreased.

Discussion of cost-effectiveness and modelling results

Diagnostic results

The baseline (4-NS and SLNB) strategies produce the smallest number of FP cases among all strategies. The number of FN cases produced by the two baseline strategies is also very small (1.1% for 4-NS and 1.3% for SLNB). In the MRI replacement strategy, the number of FP cases is increased significantly from 0.2% to 6.3% and the number of FN cases is increased to a lesser extent, from around 1.0% to 1.9%. In the PET replacement strategy the numbers of both FP and FN cases are increased significantly, from 0.2% to 3.6% for FP cases and from around 1.0% to 7.2% for FN cases. Overall, the PET replacement strategy produces the largest number of FP/FN cases among all the diagnostic strategies tested.

If MRI or PET is placed before sampling methods, the number of FN cases is reduced from around 1.0% to 0.1% if MRI is placed before sampling and to around 0.5% if PET is placed before sampling. The number of FP cases remains the same as for the MRI and PET replacement strategies.

The baseline strategies produce the lowest number of FP cases. The strategies of adding MRI before 4-NS and SLNB produce the lowest number of FN cases. Overall, the baseline strategies produce the smallest number of combined FP and FN cases. The MRI and PET replacement strategies may be considered unacceptable on clinical grounds, as they both generate more FP and FN cases than current standard practice. The strategies of adding MRI or PET before sampling also produces more FP cases, although the number of FN cases is reduced.

Number of procedures

Under the baseline sampling strategies a proportion of patients will need two separate surgical procedures: a sampling procedure (4-NS or SLNB) and subsequent ALND. In the replacement strategies, no sampling procedures are needed and the ALND procedures are carried out for TP and FP cases. Compared with the baseline strategies, the total number of ALND procedures carried out is increased for the MRI replacement strategy (due to more FP cases) and decreased for the PET replacement strategy (due to less TP cases).

If MRI or PET is placed before the sampling methods, the numbers of sampling procedures carried out are reduced compared with the baseline 4-NS and SLNB strategies because both TP and FP cases detected by MRI or PET will receive ALND surgery without sampling procedures. The reduction is more evident when MRI rather than PET is placed before sampling, as MRI is associated with more positive cases (both true and false). However, due to the increase in FP cases, the strategies of adding MRI or PET before sampling are associated with more ALND procedures than the baseline 4-NS and SLNB strategies.

Adverse events

The number of short- and long-term adverse events is proportional to the number of 4-NS, SLNB and ALND surgical procedures carried out. Adverse events are more frequent for ALND than 4-NS and SLNB. Among all diagnostic strategies modelled, the PET replacement strategy is associated with the lowest number of adverse events, followed by the MRI replacement strategy. The PET replacement strategy is associated with the smallest number of ALND procedures. The strategies of adding MRI or PET before sampling produce more adverse events than the baseline 4-NS and SLNB strategies, because more ALND procedures are carried out.

Survival rates

Patients with axillary lymph node metastases have lower survival rates (both TP and FN) than patients with negative nodal status (both TN and FP). Patients with FN results have the lowest survival rates because they do not receive ALND or chemotherapy that may reduce the risk of recurrence.

Compared with the two baseline strategies, the overall survival rates of early-stage breast cancer patients are lower for the MRI or PET replacement strategies and higher for the strategies where MRI or PET is placed before 4-NS and SLNB. The absolute differences in overall survival rate among tested diagnostic strategies are relatively small, as the absolute number of patients with FN results only account for a small proportion of all patients.

Cost-effectiveness analyses – baseline results

The PET and MRI strategies are compared with two baseline techniques – SLNB and 4-NS – which are both used currently in the UK. It is beyond the remit of this assessment to compare 4-NS and SLNB.

The MRI replacement strategy is the most cost-effective strategy and dominates the two baseline strategies. The higher QALYs of the MRI replacement strategy are driven by fewer cases of lymphoedema, which has a lifelong impact on quality of life. The PET replacement strategy is the next most cost-effective strategy and also dominates the two baseline strategies. Compared with the MRI replacement strategy, the PET replacement strategy has significantly lower QALYs, which is driven by more FN cases who are more likely to experience locoregional and metastatic relapse.

The cost-effectiveness results demonstrate that, on the population level, it is beneficial to replace invasive sampling methods with the non-invasive imaging techniques of MRI or PET. If the MRI or PET replacement strategies are used, a small proportion of patients will be wrongly diagnosed as FP or FN, which will impact on life-years gained and quality of life for these patients. However, the majority of patients will be correctly diagnosed by MRI or PET, without the need for sampling procedures. This will improve their quality of life owing to avoidance of short- and long-term adverse events such as lymphoedema. The model results suggest that the health benefits gained by the majority of patients outweigh the negative impact on reduced survival and lower quality of life of a small proportion of patients. The imaging replacement strategies, especially for MRI, also cost less than the baseline 4-NS and SLNB strategies. Overall, the analysis predicts that it is cost-effective to replace 4-NS or SLNB with MRI or PET.

Despite the cost-effectiveness results, the MRI and PET replacement strategies may be considered unacceptable on clinical grounds, due to higher numbers of FP and FN cases. If this is the case, then the most cost-effective strategy is the 4-NS strategy (if 4-NS is used as the baseline) or the addition of MRI before SLNB (if SLNB is used as the baseline). However, these results are less robust than the results for the replacement strategies. The differences in costs and QALYs between strategies are small and therefore the results are sensitive to changes in the input parameters. More robust evidence is needed on the costs of 4-NS and SLNB and the costs of MRI and PET. In addition more robust evidence on the sensitivity and specificity of 4-NS is also needed.

The strategies of placing MRI or PET before sampling may also be rejected on the clinical grounds that they are associated with more FP cases. In order to have a similar level of FP cases to the sampling methods, the specificity of MRI and PET needs to be improved to be close to 100% which, by definition, is the specificity of 4-NS and SLNB. For the MRI or PET replacement strategies, in order to have similar levels of FP and FN cases to the sampling methods, both

the sensitivity and specificity have to be improved to the levels of 4-NS and SLNB. The most promising technique is the USPIO-enhanced MRI, which has a mean sensitivity of 98% and specificity of 96%. These figures are, however, based on a limited number of small studies. Further studies are needed to confirm the robustness of these figures.

In general, based on the estimates of sensitivity and specificity of MRI and PET in this assessment, the analysis predicts that it is cost-effective to replace 4-NS or SLNB with MRI or PET which will accrue more QALYs and cost less at the population level. Within the two replacement strategies, it is more cost-effective to replace sampling with MRI than PET.

Cost-effectiveness analyses – sensitivity analysis results

The sensitivity and specificity of MRI have a significant impact on the cost-effectiveness results, and change the results in terms of the most cost-effective strategy. Further evidence on the sensitivity and specificity of MRI is therefore needed.

The utility decrement for lymphoedema has a significant impact on the cost-effectiveness of the PET replacement strategy, as the strategy is associated with the smallest number of lymphoedema cases. If the assumed utility decrement for lymphoedema is reduced, the PET replacement strategy no longer dominates the baseline 4-NS and SLNB strategies. Further studies on the costs and quality of life of lymphoedema are needed.

The probabilities of relapse for FN patients also have a significant impact on the cost-effectiveness results, particularly for the PET replacement strategy. The PET replacement strategy no longer dominates the baseline 4-NS and SLNB strategies when the probabilities of relapse for FN patients are increased.

The change in cost of SLNB has an impact on the cost-effectiveness of the strategy of adding MRI before SLNB. When the cost of SLNB is decreased, the strategy of adding MRI before SLNB becomes less cost-effective than the baseline SLNB strategy.

Sensitivity analyses also suggest that cost-effectiveness results appear to be robust when the model inputs of sensitivity of PET, the additional costs of lymphoedema and the costs of 4-NS are changed.

In general, the MRI replacement strategy remains the most cost-effective strategy and dominates the baseline 4-NS and SLNB strategies in most of the sensitivity analyses undertaken. The PET replacement strategy is not as robust as the MRI replacement strategy, as its cost-effectiveness is significantly affected by the utility decrement for lymphoedema and the probability of relapse for FN patients. When the imaging replacement strategies are excluded, the cost-effectiveness of adding MRI before 4-NS or SLNB is affected by the change of model inputs of MRI sensitivity and specificity, the probabilities of relapse for FN patients and the cost of SLNB.

Limitations of the analysis

The main limitations of the cost-effectiveness analyses include:

- Evidence on the sensitivity and specificity of 4-NS is limited and is less robust than evidence on SLNB. The adverse event rates of 4-NS are assumed to be the same as rates for SLNB, but this may underestimate the adverse event rates of 4-NS.
- The sensitivity and specificity of MRI and PET vary significantly across different studies. The evidence for MRI is less robust than the evidence for PET, given that it based on a limited number of small studies.

- The sensitivity and specificity of MRI and PET are assumed to be independent of previous tests.
- The quality of evidence on the cost of lymphoedema and the impact of lymphoedema on quality of life is poor.
- The quality of evidence on the costs of short-term adverse events is poor.
- More robust costing information for the baseline sampling procedures and for MRI and PET are also needed for the UK setting.

Chapter 5

Assessment of factors relevant to the NHS and other parties

Positron emission tomography and MRI are used in the management of many cancers, for staging and for assessing response to treatment. In the context of breast cancer, PET scanning at the time of diagnosis may have other advantages in addition to detection of axillary metastases. For example, whole-body PET scans may detect undiagnosed distant metastases or synchronous tumours.[137,138]

Availability of PET and MRI scanning facilities may be limited. If PET or MRI were recommended as part of the routine screening pathway for all patients with early breast cancer, this would mean a large increase in the number of PET or MRI procedures required. This issue would require careful consideration if PET or MRI were to be used in this setting. Also, a requirement for additional diagnostic techniques may add to the delay between diagnosis and treatment.

As there are contraindications associated with both PET and MRI, and some patients are unable to complete an MRI scan owing to claustrophobia, a subset of patients may be unable to undergo these techniques and may require alternative investigations.

Chapter 6

Discussion

Statement of principal findings

Summary of clinical results

Diagnostic accuracy of positron emission tomography

Across all 26 studies ($n = 2591$ patients) evaluating PET or PET/CT for assessment of axillary metastases, the mean sensitivity was 63% (95% CI 52% to 74%; range 20%–100%) and the mean specificity was 94% (95% CI 91% to 96%; range 75%–100%). For the seven studies ($n = 862$) evaluating PET/CT,[73–79] the mean sensitivity was 56% (95% CI 44% to 67%) and the mean specificity was 96% (95% CI 90% to 99%). For the 19 studies ($n = 1729$) evaluating PET only,[44,80–97] the mean sensitivity was 66% (95% CI 50% to 79%) and the mean specificity was 93% (95% CI 89% to 96%). Therefore PET/CT gave a slightly lower mean sensitivity than PET only; this may have been due to chance variation or to differences in the populations or study methods.

Positron emission tomography performed less well in terms of identifying small metastases; micrometastases (≤ 2 mm) were associated with a mean sensitivity of 11% (95% CI 5% to 22%) based on data from five studies ($n = 63$),[79–80,86–88] while macrometastases (> 2 mm) were associated with a mean sensitivity of 57% (95% CI 47% to 66%) based on data from four studies ($n = 111$).[73,80,87,88] The smallest metastatic nodes detected by PET measured 3 mm,[87,88] while PET failed to detect some nodes measuring > 15 mm.[86,94] Current PET cameras are thought to achieve spatial resolutions of 4–7 mm (around 4–5 mm in the centre of the field of view).[115] PET studies in which all patients were clinically node negative showed a trend towards lower sensitivity compared with studies which included both clinically node-negative and node-positive patients, which may reflect the fact that clinically negative axillary metastases are likely to be smaller. This mix of node-positive and node-negative patients is likely to reflect clinical practice. Studies using ALND reported similar sensitivity and specificity to studies using a combination of ALND and SLNB, while studies in which not all patients had ALND or SLNB (or in which the reference standard was not stated) had a higher mean sensitivity which may represent an overestimate.

Diagnostic accuracy of magnetic resonance imaging

The review identified nine studies ($n = 307$ patients) evaluating MRI.[41,42,64,108–113] Several MRI studies reported more than one set of diagnostic accuracy results, according to different criteria for defining whether axillary metastases were present. Based on the highest reported sensitivity and specificity per study, the mean sensitivity across all nine MRI studies was 90% (95% CI 78% to 96%; range 65%–100%) and the mean specificity was 90% (95% CI 75% to 96%; range 54%–100%). Across the five studies ($n = 93$ patients) evaluating USPIO-enhanced MRI,[108–112] the mean sensitivity was 98% (95% CI 61% to 100%) and specificity 96% (95% CI 72% to 100%). Across three studies of gadolinium-enhanced MRI ($n = 187$),[41,42,113] the mean sensitivity was 88% (95% CI 78% to 94%) and specificity 73% (95% CI 63% to 81%). In the single study of in vivo proton MR spectroscopy ($n = 27$),[64] the sensitivity was 65% (95% CI 38% to 86%) and specificity 100% (95% CI 69% to 100%). Therefore USPIO-enhanced MRI showed a trend towards higher sensitivity and specificity than gadolinium-enhanced MRI. No data were presented according to the size of metastases for the MRI studies. However, current MRI methods are thought to achieve a resolution of approximately 1 mm using modern scanners and based on the methods used in the included papers.

Adverse effects and contraindications

Both PET and MRI appeared to be relatively safe in this setting, with no adverse effects reported for PET, and only mild rash, claustrophobia and back pain reported as adverse effects of MRI. There are some contraindications to both PET (pregnancy) and MRI (allergy to contrast agents, renal or liver dysfunction, pacemakers and other metallic implants). In addition, claustrophobia may prevent scanning in some patients. These factors may limit the applicability of PET and MRI for some patients.

Comparison between positron emission tomography and magnetic resonance imaging

The mean sensitivity of MRI for the studies included here was significantly higher than for PET. However, as none of the included studies directly compared PET and MRI, caution should be taken when comparing these estimates. In particular, USPIO-enhanced MRI showed high sensitivity and specificity in these preliminary studies. However, these results should be interpreted with caution, as the included MRI studies were relatively small and there was variation between and within studies in terms of the MRI method used, the criteria for defining a node as positive, and the sensitivity and specificity results for individual studies. The included studies for both PET and MRI were heterogeneous in their results. This may reflect differences in conduct and interpretation of the index test and reference standard between studies, not all of which will have been captured in the study reports. Therefore, the mean sensitivity and specificity data should be interpreted with caution.

Summary of cost-effectiveness and benefits versus risks for each strategy

The decision model evaluated the benefits, harms and cost-effectiveness of PET and MRI, either as a replacement for SLNB or 4-NS or as an additional test prior to SLNB or 4-NS. Comparison of SLNB with 4-NS was not part of the remit of this assessment, but both techniques are in current use. Therefore, two baseline strategies (SLNB and 4-NS) were modelled and strategies involving SLNB were assessed separately to strategies involving 4-NS.

The results of the decision modelling suggest that the most cost-effective strategy is to replace sampling (SLNB or 4-NS) with MRI. This strategy dominates the baseline SLNB and 4-NS strategies, generating higher QALYs and lower costs. Axillary sampling (SLNB and 4-NS) is avoided for all patients, leading to fewer adverse effects, especially lymphoedema, which has an assumed lifelong impact on quality of life. However, MRI has lower sensitivity than SLNB and 4-NS (leading to more FNs) and a lower specificity (leading to more FPs). Patients with a FN diagnosis will not receive ALND or adjuvant therapy, leading to a higher risk of cancer recurrence. Patients with a FP diagnosis will receive ALND unnecessarily, with the accompanying increased risk of adverse effects. On the population level, the model results suggest that the MRI replacement strategy costs less, and also the health benefits gained by the majority of patients outweigh the negative impact on survival and quality of life of a small proportion of patients. This strategy may, however, be rejected on clinical grounds, owing to the increase in FP and FN cases compared with current practice.

If the replacement strategies are rejected on clinical grounds, the most cost-effective strategy is predicted to be the baseline 4-NS strategy (compared with 4-NS as the baseline) or the use of MRI prior to SLNB (compared with SLNB as the baseline). For strategies that place MRI or PET before sampling, patients with TP PET or MRI results receive immediate ALND without the need to carry out a separate SLNB or 4-NS procedure. However, the need for a second surgical procedure may also be averted through the use of intraoperative cytology, whereby the axillary sampling procedure may be converted immediately to full ALND for patients with positive nodes. Intraoperative cytology is currently under investigation at a number of units.[139] The number of FN cases is also reduced due to the use of two sequential tests. The disadvantage is

that, due to the lower specificity of PET and MRI compared with SLNB and 4-NS, there are still more patients with FP results who receive ALND unnecessarily, with the accompanying increased risk of adverse effects. The total QALYs generated by the strategies of adding MRI prior to 4-NS or SLNB are very similar to those in the baseline strategies and these results are not considered to be robust, based on the quality of the data available.

The model results indicate that, as would be expected, patients with axillary lymph node metastases (either TP or FN) have lower survival rates than patients with negative nodal status. Patients with FN results have the lowest survival rates because they do not receive ALND or chemotherapy that may reduce the risk of recurrence. Compared with the two baseline strategies, the overall survival rates are lower for the MRI or PET replacement strategies (due to an increase in FNs) and higher for the strategies where MRI or PET are placed before 4-NS and SLNB (due to a decrease in FNs). However, the absolute differences in overall survival rate among tested diagnostic strategies are relatively small, as the absolute number of patients with FN results only account for a small proportion of all patients.

Sensitivity analysis suggests that the MRI replacement strategy remains the most cost-effective strategy in the majority of the one-way sensitivity analyses undertaken. The sensitivity analyses indicate that the cost-effectiveness of the PET replacement strategy is significantly affected by the assumption relating to utility decrements for lymphoedema and the probability of relapse for FN patients. The quality of data on long-term utility decrements generated by lymphoedema is particularly poor. If the PET and MRI replacement strategies are excluded, the cost-effectiveness of the strategies of adding MRI prior to 4-NS or SLNB is affected by the assumed MRI sensitivity and specificity, the probability of relapse for FN patients, and the costs of SLNB. The results relating to addition of MRI prior to 4-NS or SLNB are not considered to be robust.

Strengths and limitations of the assessment

This assessment provides a comprehensive evaluation of PET and MRI for the assessment of axillary metastases. Mean values for sensitivity and specificity have been calculated using the bivariate random effects approach, which accounts for variation within and between studies as well as taking into account the correlation between sensitivity and specificity.[70,71]

Uncertainties

The included studies for both PET and MRI were heterogeneous in their results, implying that their sensitivities and specificities may vary in practice according to the methods used. For MRI there are fewer studies with smaller numbers of patients, and there are wide variations between studies in terms of the MRI methods used and the criteria for defining a node as positive. Therefore, caution should be taken when interpreting these results, particularly for MRI. The sensitivity and specificity data for 4-NS are not as robust as for SLNB due to the fewer studies identified and the smaller patient sample sizes within studies.

There are a number of key uncertainties and limitations relating to the cost-effectiveness modelling analyses. The evidence for MRI is based on a limited number of small studies. The sensitivity and specificity of PET and MRI are assumed to be independent of previous tests. There are insufficient high-quality data to estimate the costs of SLNB, 4-NS and ALND procedures, the costs of short-term adverse events and the impact of lymphoedema on patient utility. Results for the replacement strategies are considered to be more robust than for strategies in which MRI or PET are added in before sampling.

Other relevant factors

As there are contraindications associated with both PET and MRI, and some patients are unable to complete an MRI scan due to claustrophobia, a subset of patients may be unable to undergo these techniques and may require alternative investigations. Additional diagnostic techniques may add to the delay between diagnosis and treatment. Conversely, PET and MRI may have additional advantages; for example, whole-body PET scanning may detect undiagnosed distant metastases or synchronous tumours. Availability of PET and MRI scanning facilities would need to be considered if PET or MRI were recommended as part of the routine screening pathway for all patients with early breast cancer.

Chapter 7

Conclusions

Implications for service provision

The studies in this review demonstrate a significantly higher sensitivity for MRI than for PET, with USPIO-enhanced MRI providing the highest sensitivity. However, as none of the included studies directly compared PET and MRI, caution should be taken when comparing these estimates. Sensitivity of PET is reduced for smaller metastases. Specificity was similar for PET and MRI. However, this analysis is limited by the small number and size of MRI studies, and the wide variations between and within studies in terms of the MRI method used and the criteria for defining a node as positive. The sensitivity and specificity of both PET and MRI vary widely between studies, which is likely to reflect differences in imaging methods and interpretation. Therefore, caution should be taken when interpreting these results, particularly for MRI.

All patients currently receive ultrasound prior to other investigations; the sensitivity of PET appears similar to that of ultrasound while the sensitivity of MRI appears slightly higher. As ultrasound is not currently considered sensitive enough to completely replace SLNB or 4-NS, it is unlikely that PET could fulfil this role either. Specificity of PET and USPIO-enhanced MRI appear slightly higher than for ultrasound.

Positron emission tomography and MRI have lower sensitivity and specificity than the current surgical diagnostic techniques of SLNB and 4-NS, but are associated with fewer adverse events. Decision modelling suggests that MRI has the potential to offer an alternative to current sampling techniques. The current analysis suggests that the most cost-effective strategy may be MRI rather than SLNB or 4-NS, reducing costs and increasing QALYs due to fewer adverse events for the majority of patients. More robust data on the sensitivity and specificity of MRI techniques, particularly USPIO-enhanced MRI, are required, along with more accurate UK costs for the diagnostic and sampling procedures and high-quality utility data on the impact of lymphoedema. If the replacement strategy is considered clinically unacceptable due to higher numbers of FN cases (leading to higher risk of recurrence) and FP cases (leading to unnecessary ALND), then the most cost-effective strategy appears to be to retain the baseline 4-NS strategy (if 4-NS is used as the baseline) or to use MRI prior to SLNB (if SLNB is used as the baseline). However, the results relating to addition of MRI prior to 4-NS or SLNB are not considered to be robust, based on the quality of the input parameters available and further work is required to provide more reliable estimates.

Suggested research priorities

If the use of MRI is deemed clinically acceptable (either to replace SLNB or 4-NS or as an additional test prior to SLNB or 4-NS), then further large, well-conducted studies using up-to-date MRI methods are required to obtain more accurate data on the sensitivity and specificity of MRI in this setting. Further studies of USPIO-enhanced MRI would be valuable in order to gain more robust data on sensitivity and specificity, adverse effects and which are the best criteria for defining a node as metastatic. In addition, further data on the long-term impacts of differing severities of lymphoedema on cost and patient utility would be valuable. More robust UK cost

data are needed for 4-NS and SLNB, as well as the cost of MRI and PET techniques. It would also be useful to identify the dependencies of different diagnostic methods in the pathway; in particular, whether and to what extent the accuracy of MRI or PET depends on the diagnostic results of ultrasound and ultrasound-guided biopsy. One potential use of MRI or PET might be to triage patients with a positive result to immediate ALND, avoiding a prior SLNB or 4-NS procedure. However, the need for two sequential surgical procedures may also be averted through the use of intraoperative cytology. Therefore, further studies on this technique would be of relevance when assessing the potential benefits of PET or MRI in this setting. A major limitation of imaging techniques such as PET and MRI is their limited ability to detect small metastases. It is therefore important that future research into diagnostic techniques should assess accuracy in detecting metastases of different sizes, including micrometastases. It may also be useful for future studies to report diagnostic accuracy according to subgroups of patients with different sizes and stages of primary breast tumour, in order to inform management decisions for these different patient groups.

Acknowledgements

Contributions of authors

Katy Cooper coordinated the review; Diana Papaioannou undertook the literature searches; Katy Cooper and Sue Harnan extracted data from papers, undertook data analyses and wrote the clinical review; Patrick Fitzgerald provided statistical support and undertook the meta-analyses; Yang Meng and Sue Ward undertook the health economic review and economic modelling and wrote the cost-effectiveness assessment; and, Lynda Wyld, Christine Ingram, Iain Wilkinson and Eleanor Lorenz provided clinical advice.

About ScHARR

The School of Health and Related Research (ScHARR) is one of the nine departments that comprise the Faculty of Medicine, Dentistry and Health at the University of Sheffield. ScHARR specialises in health services and public health research, and the application of health economics and decision science to the development of health services and the improvement of the public health.

The ScHARR Technology Assessment Group (ScHARR-TAG) synthesises research on the clinical and cost-effectiveness of healthcare interventions for the National Institute for Health Research (NIHR) Health Technology Assessment Programme on behalf of a range of policy-makers, including the National Institute for Health and Clinical Excellence. ScHARR-TAG is part of a wider collaboration of six units from other regions. The other units are: Southampton Health Technology Assessment Centre (SHTAC), University of Southampton; Aberdeen Health Technology Assessment Group (Aberdeen HTA Group), University of Aberdeen; Liverpool Reviews & Implementation Group (LRiG), University of Liverpool; Peninsular Technology Assessment Group (PenTAG), University of Exeter; NHS Centre for Reviews and Dissemination, University of York; and West Midlands Health Technology Assessment Collaboration (WMHTAC), University of Birmingham.

References

1. Office for National Statistics. *Cancer registration statistics England 2007: MB1 No 38.* URL: www.statistics.gov.uk/StatBase/Product.asp?vlnk=7720&Pos=1&ColRank=1&Rank=272 (accessed 29 November 2010).
2. Welsh Cancer Intelligence & Surveillance Unit (WCISU). *Cancer incidence in Wales 2003–2007*. URL: www.wales.nhs.uk/sites3/Documents/242/Cancer%20Incidence%20in%20Wales%202003–2007.pdf (accessed 23 January 2009).
3. Dumitrescu RG, Cotarla I. Understanding breast cancer risk – where do we stand in 2005? *J Cell Mol Med* 2005;**9**(1):208–21.
4. Collaborative Group on Hormonal Factors in Breast Cancer. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58 209 women with breast cancer and 101 986 women without the disease. *Lancet* 2001;**358**(9291):1389–99.
5. Ziegler RG, Hoover RN, Pike MC, Hildesheim A, Nomura AM, West DW, *et al.* Migration patterns and breast cancer risk in Asian-American women. *J Natl Cancer Inst* 1993;**85**(22):1819–27.
6. Kahlenborn C, Modugno F, Potter DM, Severs WB. Oral contraceptive use as a risk factor for premenopausal breast cancer: a meta-analysis. *Mayo Clin Proc* 2006;**81**(10):1290–302.
7. Beral V. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52, 705 women with breast cancer and 108, 411 women without breast cancer. *Lancet* 1997;**350**(9084):1047–59.
8. Megdal SP, Kroenke CH, Laden F, Pukkala E, Schernhammer ES. Night work and breast cancer risk: a systematic review and meta-analysis. *Eur J Cancer* 2005;**41**(13):2023–32.
9. Guh DP, Zhang W, Bansback N, Amarsi Z, Birmingham CL, Anis AH. The incidence of co-morbidities related to obesity and overweight: a systematic review and meta-analysis. *BMC Public Health* 2009;**9**:88.
10. Renehan AG, Tyson M, Egger M, Heller RF, Zwahlen M. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. *Lancet* 2008;**371**(9612):569–78.
11. Loi S, Milne RL, Friedlander ML, McCredie MRE, Giles GG, Hopper JL, *et al.* Obesity and outcomes in premenopausal and postmenopausal breast cancer. *Cancer Epidemiol Biomarkers Prev* 2005;**14**(7):1686–91.
12. Kerlikowske K, Walker R, Miglioretti DL, Desai A, Ballard-Barbash R, Buist DSM, *et al.* Obesity, mammography use and accuracy, and advanced breast cancer risk. *J Natl Cancer Inst* 2008;**100**(23):1724–33.
13. Lagerros YT, Hsieh SF, Hsieh CC. Physical activity in adolescence and young adulthood and breast cancer risk: a quantitative review. *Eur J Cancer Prev* 2004;**13**(1):5–12.
14. Monninkhof EM, Elias SG, Vlems FA, van der Tweel I, Schuit AJ, Voskuil DW, *et al.* Physical activity and breast cancer: a systematic review. *Epidemiology* 2007;**18**(1):137–57.
15. Cancer Research UK. *Breast cancer in men*. URL: www.cancerhelp.org.uk/help/default.asp?page=5075 (accessed 1 May 2009).

16. Fisher B, Bauer M, Wickerham DL, Redmond CK, Fisher ER, Cruz AB, *et al.* Relation of number of positive axillary nodes to the prognosis of patients with primary breast cancer. An NSABP update. *Cancer* 1983;**52**(9):1551–7.

17. National Institute for Health and Clinical Excellence (NICE), National Collaborating Centre for Cancer. *Early and locally advanced breast cancer: diagnosis and treatment*. Cardiff: National Collaborating Centre for Cancer; 2009.

18. Office for National Statistics. *Survival Rates in England, patients diagnosed 2001–2006 followed up to 2007*. URL: www.statistics.gov.uk/StatBase/Product.asp?vlnk=14007. 2007.

19. West Midlands Cancer Intelligence Unit. *0–10 year relative survival for cases of breast cancer by stage diagnosed in the West Midlands 1985–1989 followed up to the end of 1999, as at January 2002*. Birmingham: West Midlands Cancer Intelligence Unit; 2009.

20. Galea MH, Blamey RW, Elston CE, Ellis IO. The Nottingham Prognostic Index in primary breast cancer. *Breast Cancer Res Treat* 1992;**22**(3):207–19.

21. Adjuvant! Online. *Adjuvant! Online. Decision making tools for health care professionals*. URL: www.adjuvantonline.com/index.jsp (accessed 29 November 2010).

22. Cancer Research UK. *Cancer statistics: breast cancer UK*. URL: http://publications.cancerresearchuk.org/WebRoot/crukstoredb/CRUK_PDFs/CSBREA09.pdf (accessed 1 May 2009).

23. Cancer Research UK, National Cancer Intelligence Network. *Cancer incidence and survival by major ethnic group, England, 2002–2006*. URL: http://publications.cancerresearchuk.org/WebRoot/crukstoredb/CRUK_PDFs/CR-UKNCIN_ETHNIC.pdf (acessed 1 June 2009).

24. Parkin DM, Fernandez LM. Use of statistics to assess the global burden of breast cancer. *Breast J* 2006;**12**(Suppl. 1):S70–S80.

25. National Cancer Intelligence Network. *Cancer incidence by deprivation England, 1995–2004*. URL: www.ncin.org.uk/publications/reports/default.apsx.

26. Welsh Cancer Intelligence & Surveillance Unit. *Cancer incidence, mortality and survival by deprivation in Wales*. URL: www.wales.nhs.uk/sites3/Documents/242/Deprivation%20in%20Wales%201993–2007.pdf (accessed 1 July 2009).

27. Cancer Research UK. *Cancer survival and deprivation statistics*. URL: http://info.cancerresearchuk.org/cancerstats/survival/survivaldeprivation/ (accessed 4 April 2005).

28. Office for National Statistics. *Death registrations in England and Wales, 2008, causes*. URL: www.statistics.gov.uk/pdfdir/dthreg0809.pdf (accessed 26 August 2009).

29. South East England Public Health Observatory, Oxford University. *Breast cancer in England 1996 to 2004. Mortality trends*. URL: www.uhce.ox.ac.uk/epidembase/temp_epi/trends_mortality.php.

30. Blanks RG, Moss SM, McGahan CE, Quinn MJ, Babb PJ. Effect of NHS breast screening programme on mortality from breast cancer in England and Wales, 1990–8: comparison of observed with predicted mortality. *BMJ* 2000;**321**(7262):665–9.

31. Kobayashi S. What caused the decline in breast cancer mortality in the United Kingdom? *Breast Cancer* 2004;**11**(2):156–9.

32. NHS Cancer screening programme. *Breast awareness*. URL: www.cancerscreening.nhs.uk/breastscreen/breastawareness.html (accessed 29 November 2010).

33. NHS Cancer screening programme. *Breast cancer screening*. URL: www.cancerscreening.nhs.uk/breastscreen/index.html (accessed 29 November 2010).

34. International Union Against Cancer. *TNM classification of malignant tumours*. 5th edn. New York, NY: Wiley-Liss; 1997.

35. Singletary SE, Connolly JL. Breast cancer staging: working with the sixth edition of the AJCC Cancer Staging Manual. *CA Cancer J Clin* 2006;**56**(1):37–47.

36. Cancer Research UK. *TNM breast cancer staging*. URL: www.cancerhelp.org.uk/help/default.asp?page=3316 (accessed 29 November 2010).

37. Agency for Healthcare Research and Quality (AHRQ). *Surgery choices for women with early-stage breast cancer*. URL: www.ahrq.gov/consumer/brcanchoice.htm. (accessed 29 November 2010).

38. US National Cancer Institute (NCI). *Surgery choices for women with early-stage breast cancer*. URL: www.cancer.gov/cancertopics/breast-cancer-surgery-choices. 2009.

39. American Cancer Society. *How is breast cancer staged*? URL: www.cancer.org/docroot/CRI/content/CRI_2_4_3X_How_is_breast_cancer_staged_5.asp. (accessed 29 November 2010).

40. Cancer Research UK. *Number stages of breast cancer*. URL: www.cancerhelp.org.uk/help/default.asp?page=3315. (accessed 29 November 2010).

41. Kvistad KA, Rydland J, Smethurst HB, Lundgren S, Fjosne HE, Haraldseth O. Axillary lymph node metastases in breast cancer: preoperative detection with dynamic contrast-enhanced MRI. *Eur Radiol* 2000;**10**(9):1464–71.

42. Mumtaz H, Hall-Craggs MA, Davidson T, Walmsley K, Thurell W, Kissin MW, *et al.* Staging of symptomatic primary breast cancer with MR imaging. *AJR* 1997;**169**(2):417–24.

43. Smith IC, Welch AE, Chilcott F, Heys SD, Sharp P, Eremin O. Gamma emission imaging in the management of breast disorders. *Eur J Surg Oncol* 1998;**24**(4):320–9.

44. Wahl RL, Siegel BA, Coleman RE, Gatsonis CG, PET Study Group. Prospective multicenter study of axillary nodal staging by positron emission tomography in breast cancer: a report of the staging breast cancer with PET Study Group. *J Clin Oncol* 2004;**22**(2):277–85.

45. Wallace IW, Champton HR. Axillary nodes in breast cancer. *Lancet* 1972;**1**(7744):217–18.

46. Fisher B, Wolmark N, Bauer M, Redmond C, Gebhardt M. The accuracy of clinical nodal staging and of limited axillary dissection as a determinant of histologic nodal status in carcinoma of the breast. *Surg Gynecol Obstet* 1981;**152**(6):765–72.

47. Alvarez S, Anorbe E, Alcorta P, Lopez F, Alonso I, Cortes J. Role of sonography in the diagnosis of axillary lymph node metastases in breast cancer: a systematic review. *AJR Am J Roentgenol* 2006;**186**(5):1342–8.

48. Blanchard DK, Donohue JH, Reynolds C, Grant CS. Relapse and morbidity in patients undergoing sentinel lymph node biopsy alone or with axillary dissection for breast cancer. *Arch Surg* 2003;**138**(5):482–7.

49. Crane-Okada R, Wascher RA, Elashoff D, Giuliano AE. Long-term morbidity of sentinel node biopsy versus complete axillary dissection for unilateral breast cancer. *Ann Surg Oncol* 2008;**15**(7):1996–2005.

50. McLaughlin SA, Wright MJ, Morris KT, Giron GL, Sampson MR, Brockway JP, *et al.* Prevalence of lymphedema in women with breast cancer 5 years after sentinel lymph node biopsy or axillary dissection: objective measurements. *J Clin Oncol* 2008;**26**(32):5213–19.

51. Purushotham AD, Upponi S, Klevesath MB, Bobrow L, Millar K, Myles JP, *et al.* Morbidity after sentinel lymph node biopsy in primary breast cancer: results from a randomized controlled trial. *J Clin Oncol* 2005;**23**(19):4312–21.

52. Mansel RE, Fallowfield L, Kissin M, Goyal A, Newcombe RG, Dixon JM, *et al.* Randomized multicenter trial of sentinel node biopsy versus standard axillary treatment in operable breast cancer: the ALMANAC Trial. *J Natl Cancer Inst* 2006;**98**(9):599–609.

53. Liu CQ, Guo Y, Shi JY, Sheng Y. Late morbidity associated with a tumour-negative sentinel lymph node biopsy in primary breast cancer patients: a systematic review. *Eur J Cancer* 2009;**45**(9):1560–8.

54. Wilke LG, McCall LM, Posther KE, Whitworth PW, Reintgen DS, Leitch AM, *et al.* Surgical complications associated with sentinel lymph node biopsy: results from a prospective international cooperative group trial. *Ann Surg Oncol* 2006;**13**(4):491–500.

55. Department of Health. *NHS Reference Costs 2003*. London: Department of Health; 2004.

56. Department of Health. *NHS Reference Costs 2008*. London: Department of Health; 2009.

57. Mansfield L, Sosa I, Dionello R, Subramanian A, Devalia H, Mokbel K. Current management of the axilla in patients with clinically node-negative breast cancer: a nationwide survey of United Kingdom breast surgeons. *Int Semin Surg Oncol* 2007;**4**:4.

58. Facey K, Bradbury I, Laking G, Payne E. Overview of the clinical effectiveness of positron emission tomography imaging in selected cancers. *Health Technol Assess* 2007;**11**(44).

59. Cherry SR. *The* 2006 Henry N. Wagner lecture: of mice and men (and positrons)--advances in PET imaging technology. *J Nucl Med* 2006;**47**(11):1735–45.

60. Messa C, Bettinardi V, Picchio M, Pelosi E, Landoni C, Gianolli L, *et al.* PET/CT in diagnostic oncology. *Q J Nucl Med Mol Imaging* 2004;**48**(2):66–75.

61. Radiology Info. URL: www.radiologyinfo.org/en/info.cfm?pg=abdominct. (accessed 29 November 2010).

62. Erguvan-Dogan B, Whitman GJ, Kushwaha AC, Phelps MJ, Dempsey PJ. BI-RADS-MRI: a primer. *AJR Am J Roentgenol* 2006;**187**(2):W152–W160.

63. Lehman CD, Schnall MD. Imaging in breast cancer: magnetic resonance imaging. *Breast Cancer Res* 2005;**7**(5):215–19.

64. Yeung DK, Yang WT, Tse GM. Breast cancer: in vivo proton MR spectroscopy in the characterization of histopathologic subtypes and preliminary observations in axillary node metastases. *Radiology* 2002;**225**(1):190–7.

65. Medicines and Healthcare products Regulatory Agency (MHRA). *Safety guidelines for magnetic resonance imaging equipment in clinical use: device bulletin*. URL: www.mhra.gov.uk/Publications/Safetyguidance/DeviceBulletins/CON2033018. (accessed 29 November 2010).

66. Jacklin PB, Barrington SF, Roxburgh JC, Jackson G, Sariklis D, West PA, *et al.* Cost-effectiveness of preoperative positron emission tomography in ischemic heart disease. *Ann Thorac Surg* 2002;**73**(5):1403–9.

67. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009;**151**(4):264–9, W64.

68. PRISMA statement website. URL: www.prisma-statement.org. (accessed 29 November 2010).

69. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;**3**:25.

70. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005;**58**(10):982–90.

71. Leeflang MM, bets-Ossenkopp YJ, Visser CE, Scholten RJ, Hooft L, Bijlmer HA, *et al.* Galactomannan detection for invasive aspergillosis in immunocompromized patients. *Cochrane Database Syst Rev* 2008;**4**:CD007394.

72. Review Manager (revman). Version 5.0. 2008. Copenhagen, The Nordic Cochrane Centre, The Cochrane Collaboration.

73. Chae BJ, Bae JS, Kang BJ, Kim SH, Jung SS, Song BJ. Positron emission tomography-computed tomography in the detection of axillary lymph node metastasis in patients with early stage breast cancer. *Jpn J Clin Oncol* 2009;**39**(5):284–9.

74. Heusner TA, Kuemmel S, Hahn S, Koeninger A, Otterbach F, Hamami ME, *et al.* Diagnostic value of full-dose FDG PET/CT for axillary lymph node staging in breast cancer patients. *Eur J Nucl Med Mol Imaging* 2009;**36**(10):1543–50.

75. Kim J, Lee J, Chang E, Kim S, Suh K, Sul J, *et al.* Selective sentinel node plus additional non-sentinel node biopsy based on an FDG-PET/CT scan in early breast cancer patients: single institutional experience. *World J Surg* 2009;**33**(5):943–9.

76. Taira N, Ohsumi S, Takabatake D, Hara F, Takashima S, Aogi K, *et al.* Determination of indication for sentinel lymph node biopsy in clinical node-negative breast cancer using preoperative ^{18}F-fluorodeoxyglucose positron emission tomography/computed tomography fusion imaging. *Jpn J Clin Oncol* 2009;**39**(1):16–21.

77. Fuster D, Duch J, Paredes P, Velasco M, Munoz M, Santamaria G, *et al.* Preoperative staging of large primary breast cancer with [^{18}F]fluorodeoxyglucose positron emission tomography/computed tomography compared with conventional imaging procedures. *J Clin Oncol* 2008;**26**(29):4746–51.

78. Ueda S, Tsuda H, Asakawa H, Omata J, Fukatsu K, Kondo N, *et al.* Utility of ^{18}F-fluoro-deoxyglucose emission tomography/computed tomography fusion imaging (^{18}F-FDG PET/CT) in combination with ultrasonography for axillary staging in primary breast cancer. *BMC Cancer* 2008;**8**:165.

79. Veronesi U, De CC, Galimberti VE, Fernandez JR, Rotmensz N, Viale G, *et al.* A comparative study on the value of FDG-PET and sentinel node biopsy to identify occult axillary metastases. *Ann Oncol* 2007;**18**(3):473–8.

80. Cermik TF, Mavi A, Basu S, Alavi A. Impact of FDG PET on the preoperative staging of newly diagnosed breast cancer. *Eur J Nucl Med Mol Imaging* 2008;**35**(3):475–83.

81. Gil-Rendo A, Zornoza G, Garcia-Velloso MJ, Regueira FM, Beorlegui C, Cervera M. Fluorodeoxyglucose positron emission tomography with sentinel lymph node biopsy for evaluation of axillary involvement in breast cancer. *Br J Surg* 2006;**93**(6):707–12.

82. Weir L, Worsley D, Bernstein V. The value of FDG positron emission tomography in the management of patients with breast cancer. *Breast J* 2005;**11**(3):204–9.

83. Agresti R, Crippa F, Gerali A, Maccauro M, Giovanazzi R, Bombardieri E, *et al.* Lymph node metastases detection by FDG-PET and sentinel node biopsy in breast cancer patients: comparison of these different approaches. *EJC Supplements* 2004;**2**(3):81, abstract 86.

84. Fehr MK, Hornung R, Varga Z, Burger D, Hess T, Haller U, *et al.* Axillary staging using positron emission tomography in breast cancer patients qualifying for sentinel lymph node biopsy. *Breast J* 2004;**10**(2):89–93.

85. Inoue T, Yutani K, Taguchi T, Tamaki Y, Shiba E, Noguchi S. Preoperative evaluation of prognosis in breast cancer patients by [(^{18})F]2-Deoxy-2-fluoro-D-glucose-positron emission tomography. *J Cancer Res Clin Oncol* 2004;**130**(5):273–8.

86. Lovrics PJ, Chen V, Coates G, Cornacchi SD, Goldsmith CH, Law C, *et al.* A prospective evaluation of positron emission tomography scanning, sentinel lymph node biopsy, and standard axillary dissection for axillary staging in patients with early stage breast cancer. *Ann Surg Oncol* 2004;**11**(9):846–53.

87. Barranger E, Grahek D, Antoine M, Montravers F, Talbot JN, Uzan S. Evaluation of fluorodeoxyglucose positron emission tomography in the detection of axillary lymph node metastases in patients with early-stage breast cancer. *Ann Surg Oncol* 2003;**10**(6):622–7.

88. Guller U, Nitzsche EU, Schirp U, Viehl CT, Torhorst J, Moch H, *et al.* Selective axillary surgery in breast cancer patients based on positron emission tomography with ^{18}F-fluoro-2-deoxy-D-glucose: not yet! *Breast Cancer Res Treat* 2002;**71**(2):171–3.

89. Nakamoto Y, Chang AE, Zasadny KR, Wahl RL, Nakamoto Y, Chang AE, *et al.* Comparison of attenuation-corrected and non-corrected FDG-PET images for axillary nodal staging in newly diagnosed breast cancer. *Mol Imaging Biol* 2002;**4**(2):161–9.

90. Rieber A, Schirrmeister H, Gabelmann A, Nuessle K, Reske S, Kreienberg R, *et al.* Pre-operative staging of invasive breast cancer with MR mammography and/or PET: boon or bunk? *Br J Radiol* 2002;**75**(898):789–98.

91. van der Hoeven JJ, Hoekstra OS, Comans EF, Pijpers R, Boom RP, van Geldere D, *et al.* Determinants of diagnostic performance of [F-18]fluorodeoxyglucose positron emission tomography for axillary staging in breast cancer. *Ann Surg* 2002;**236**(5):619–24.

92. Greco M, Crippa F, Agresti R, Seregni E, Gerali A, Giovanazzi R, *et al.* Axillary lymph node staging in breast cancer by 2-fluoro-2-deoxy-D-glucose-positron emission tomography: clinical evaluation and alternative management. *J Natl Cancer Inst* 2001;**93**(8):630–5.

93. Noh DY, Yun IJ, Kim JS, Kang HS, Lee DS, Chung JK, *et al.* Diagnostic value of positron emission tomography for detecting breast cancer. *World J Surg* 1998;**22**(3):223–8.

94. Smith IC, Ogston KN, Whitford P, Smith FW, Sharp P, Norton M, *et al.* Staging of the axilla in breast cancer – accurate in vivo assessment using positron emission tomography with 2-(fluorine-18)-fluoro-2-deoxy-D-glucose. *Ann Surg* 1998;**228**(2):220–7.

95. Adler LP, Faulhaber PF, Schnur KC, Al-Kasi NL, Shenk RR. Axillary lymph node metastases: screening with [F-18]2-deoxy-2-fluoro-D-glucose (FDG) PET. *Radiology* 1997;**203**(2):323–7.

96. Avril N, Dose J, Janicke F, Ziegler S, Romer W, Weber W, *et al.* Assessment of axillary lymph node involvement in breast cancer patients with positron emission tomography using radiolabeled 2-(fluorine-18)-fluoro-2-deoxy-D-glucose. *J Natl Cancer Inst* 1996;**88**(17):1204–9.

97. Utech CI, Young CS, Winter PF. Prospective evaluation of fluorine-18 fluorodeoxyclucose positron emission tomography in breast cancer for staging of the axilla related to surgery and immunocytochemistry. *Eur J Nucl Med* 1996;**23**(12):1588–93.

98. Crippa F, Agresti R, Seregni E, Greco M, Pascali C, Bogni A, *et al.* Prospective evaluation of fluorine-18-FDG PET in presurgical staging of the axilla in breast cancer. *J Nucl Med* 1998;**39**(1):4–8.

99. Bombardieri E, Crippa F, Agresti R, Leutner M, Delledonne V, Greco M, *et al.* FDG-PET in the assessment of axillary metastases in breast cancer. Clinical application and economic implications. *J Nucl Med* 1998;**39**(5):1119.

100. Crippa F, Agresti R, Donne VD, Pascali C, Bogni A, Chiesa C, *et al.* The contribution of positron emission tomography (PET) with ^{18}F-fluorodeoxyglucose (FDG) in the preoperative detection of axillary metastases of breast cancer: the experience of the National Cancer Institute of Milan. *Tumori* 1997;**83**(2):542–3.

101. Adler LP, Schnur KC, Faulhaber PF, Shenk RR. High sensitivity interpretation PET of the axilla: screening test for axillary lymph node metastases in patients with breast cancer. *Radiology* 1996;**201**:551.

102. Kumar R, Zhuang H, Schnall M, Conant E, Damia S, Weinstein S, *et al.* FDG PET positive lymph nodes are highly predictive of metastasis in breast cancer. *Nucl Med Commun* 2006;**27**(3):231–6.

103. Zornoza G, Garcia-Velloso MJ, Sola J, Regueira FM, Pina L, Beorlegui C. F-18-FDG PET complemented with sentinel lymph node biopsy in the detection of axillary involvement in breast cancer. *Eur J Surg Oncol* 2004;**30**(1):15–19.

104. Agresti R, Crippa F, Gerali A, Maccauro M, Giovanazzi R, Guida V, *et al.* Lymph node metastases detection by FDG-PET and sentinel node biopsy in breast cancer patients: comparison of these different approaches. *Eur J Cancer* 2001;**37**(Suppl. 6):S184, abstract 670.

105. Lovrics PJ, Chen V, Coates G, Cornacchi SD, Goldsmith C, Law C, *et al.* A prospective comparison of positron emission tomography scanning, sentinel lymph node biopsy and axillary dissection in staging patients with early stage breast cancer. *Breast Cancer Res Treat* 2002;**76**(Suppl. 1):S129, abstract 514.

106. Yutani K, Shiba E, Kusuoka H, Tatsumi M, Uehara T, Taguchi T, *et al.* Comparison of FDG-PET with MIBI-SPECT in the detection of breast cancer and axillary lymph node metastasis. *J Comput Assist Tomogr* 2000;**24**(2):274–80.

107. Yutani K, Tatsumi M, Shiba E, Kusuoka H, Nishimura T. Comparison of dual-head coincidence gamma camera FDG imaging with FDG PET in detection of breast cancer and axillary lymph node metastasis. *J Nucl Med* 1999;**40**(6):1003–8.

108. Kimura K, Tanigawa N, Matsuki M, Nohara T, Iwamoto M, Sumiyoshi K, *et al.* High-resolution MR lymphography using ultrasmall superparamagnetic iron oxide (USPIO) in the evaluation of axillary lymph nodes in patients with early stage breast cancer: preliminary results. *Breast Cancer* 2010;**17**(4)241–6.

109. Harada T, Tanigawa N, Matsuki M, Nohara T, Narabayashi I. Evaluation of lymph node metastases of breast cancer using ultrasmall superparamagnetic iron oxide-enhanced magnetic resonance imaging. *Eur J Radiol* 2007;**63**(3):401–7.

110. Memarsadeghi M, Riedl CC, Kaneider A, Galid A, Rudas M, Matzek W, *et al.* Axillary lymph node metastases in patients with breast carcinomas: assessment with nonenhanced versus uspio-enhanced MR imaging. *Radiology* 2006;**241**(2):367–77.

111. Stadnik TW, Everaert H, Makkat S, Sacre R, Lamote J, Bourgain C. Breast imaging. Preoperative breast cancer staging: comparison of USPIO-enhanced MR imaging and ^{18}F-fluorodeoxyglucose (FDC) positron emission tomography (PET) imaging for axillary lymph node staging – initial findings. *Eur Radiol* 2006;**16**(10):2153–60.

112. Michel SC, Keller TM, Frohlich JM, Fink D, Caduff R, Seifert B, *et al.* Preoperative breast cancer staging: MR imaging of the axilla with ultrasmall superparamagnetic iron oxide enhancement. *Radiology* 2002;**225**(2):527–36.

113. Murray AD, Staff RT, Redpath TW, Gilbert FJ, Ah-See AK, Brookes JA, *et al.* Dynamic contrast enhanced MRI of the axilla in women with breast cancer: comparison with pathology of excised nodes. *Br J Radiol* 2002;**75**(891):220–8.

114. Lind PG. Radioimmunoscintigraphy with Tc-99m labeled monoclonal antibody 170H.82 in suspected primary, recurrent, or metastatic breast cancer. *Clin Nucl Med* 1997;**22**(1):30–4.

115. Pichler BJ, Wehrl HF, Judenhofer MS. Latest advances in molecular imaging instrumentation. *J Nucl Med* 2008;**49**(Suppl. 2):5S–23S.

116. Bellin MF. Non-gadolinium-based contrast agents. In Thomsen HS, Webb JAW, editors. *Contrast media: safety issues and ESUR guidelines*. 2nd edn. Berlin: Springer-Verlag; 2009. pp. 205–12.

117. Bellin MF, Roy C, Kinkel K, Thoumas D, Zaim S, Vanel D, *et al.* Lymph node metastases: safety and effectiveness of MR imaging with ultrasmall superparamagnetic iron oxide particles--initial clinical experience. *Radiology* 1998;**207**(3):799–808.

118. Harisinghani MG, Barentsz J, Hahn PF, Deserno WM, Tabatabaei S, van de Kaa CH, *et al.* Noninvasive detection of clinically occult lymph-node metastases in prostate cancer. *N Engl J Med* 2003;**348**(25):2491–9.

119. Sloka JS, Hollett PD, Mathews M. Cost-effectiveness of positron emission tomography in breast cancer. *Mol Imaging Biol* 2005;**7**(5):351–60.

120. Pandharipande PV, Harisinghani MG, Ozanne EM, Specht MC, Hur C, Lee JM, *et al.* Staging MR lymphangiography of the axilla for early breast cancer: cost-effectiveness analysis. *AJR Am J Roentgenol* 2008;**191**(5):1308–19.

121. Law AM. *Simulation modeling and analysis*. 4th edn. New York, NY: McGraw Hill; 2007.

122. Sato T, Yamaguchi K, Morita Y, Noikura T, Sugihara K, Matsune S. Lymphoscintigraphy for interpretation of changes of cervical lymph node function in patients with oral malignant tumors: comparison of Tc-99m-Re and Tc-99m-HSA-D. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2000;**90**(4):525–37.

123. Tanaka K, Yamamoto D, Kanematsu S, Okugawa H, Kamiyama Y. A four node axillary sampling trial on breast cancer patients. *Breast* 2006;**15**(2):203–9.

124. Sato K, Tamaki K, Takeuchi H, Tsuda H, Kosuda S, Kusano S, *et al.* Management of the axilla in breast cancer: a comparative study between sentinel lymph node biopsy and four-node sampling procedure. *Jpn J Clin Oncol* 2001;**31**(7):318–21.

125. Kim T, Giuliano AE, Lyman GH. Lymphatic mapping and sentinel lymph node biopsy in early-stage breast carcinoma: a metaanalysis. *Cancer* 2006;**106**(1):4–16.

126. Hind D, Ward S, De Nigris E, Simpson E, Carroll C, Wyld L. Hormonal therapies for early breast cancer: systematic review and economic evaluation. *Health Technol Assess* 2007;**11**(26).

127. Jeruss JS, Hunt KK, Xing Y, Krishnamurthy S, Meric-Bernstam F, Cantor SB, *et al.* Is intraoperative touch imprint cytology of sentinel lymph nodes in patients with breast cancer cost effective? *Cancer* 2006;**107**(10):2328–36.

128. Mansel RE, Fallowfield L, Kissin M, Goyal A, Newcombe RG, Dixon JM, *et al.* Randomized multicenter trial of sentinel node biopsy versus standard axillary treatment in operable breast cancer: the ALMANAC Trial. *J Natl Cancer Inst* 2006;**98**(9):599–609.

129. Ward S, Simpson E, Davis S, Hind D, Rees A, Wilkinson A. Taxanes for the adjuvant treatment of early breast cancer: systematic review and economic evaluation. *Health Technol Assess* 2007;**11**(40).

130. Tengs TO, Wallace A. One thousand health-related quality-of-life estimates. *Med Care* 2000;**38**(6):583–637.

131. Ara R, Brazier J. Populating an economic mode with health state utility values: moving towards better practice. *Value Health* 2010;**13**(5):S09–18.

132. Office for National Statistics. *Interim Life Tables, United Kingdom, 1980–82 to 2005–07*. London: Office for National Statistics; 2008.

133. Orr RK, Col NF, Kuntz KM. A cost-effectiveness analysis of axillary node dissection in postmenopausal women with estrogen receptor-positive breast cancer and clinically negative axillary nodes. *Surgery* 1999;**126**(3):568–76.

134. Kamby C, Sengelov L. Pattern of dissemination and survival following isolated locoregional recurrence of breast cancer. A prospective study with more than 10 years of follow up. *Breast Cancer Res Treat* 1997;**45**(2):181–92.

135. Mak SS, Mo KF, Suen JJ, Chan SL, Ma WL, Yeo W. Lymphedema and quality of life in Chinese women after treatment for breast cancer. *Eur J Oncol Nurs* 2009;**13**(2):110–15.

136. Her Majesty's Treasury. *The green book: appraisal and evaluation in central government*. London: Her Majesty's Treasury; 2003.

137. Escalona S, Blasco JA, Reza MM, Andradas E, Gomez N. A systematic review of FDG-PET in breast cancer. *Med Oncol* 2010;**27**:114–29.

138. Samson D, Redding FC, Aronson N. FDG positron emission tomography for evaluating breast cancer. Chicago, IL: Blue Cross and Blue Shield Association, 2001.

139. Olson JA, Jr, McCall LM, Beitsch P, Whitworth PW, Reintgen DS, Blumencranz PW, *et al.* Impact of immediate versus delayed axillary node dissection on surgical outcomes in breast cancer patients with positive sentinel nodes: results from American College of Surgeons Oncology Group Trials Z0010 and Z0011. *J Clin Oncol* 2008;**26**(21):3530–5.

140. Bland KI. Re: axillary staging using positron emission tomography in breast cancer patients qualifying for sentinel lymph node biopsy. *Breast J* 2004;**10**(2):83–4.

141. Guller U, Nitzsche E, Moch H, Zuber M. Is positron emission tomography an accurate non-invasive alternative to sentinel lymph node biopsy in breast cancer patients? *J Natl Cancer Inst* 2003;**95**(14):1040–3.

142. Wahl RS. Prospective multicenter study of axillary nodal staging by positron emission tomography in breast cancer: a report of the staging of breast cancer with PET study group. *Womens Oncol Rev* 2004;**4**(2):161–2.

143. Bahri S, Chen JH, Yu HJ, Kuzucan A, Nalcioglu O, Su MY. Can dynamic contrast-enhanced MRI (DCE-MRI) predict tumor recurrence and lymph node status in patients with breast cancer? *Ann Oncol* 2008;**19**(4):822–4.

144. Hsiang DJ, Yamamoto M, Mehta RS, Su MY, Baick CH, Lane KT, *et al.* Predicting nodal status using dynamic contrast-enhanced magnetic resonance imaging in patients with locally advanced breast cancer undergoing neoadjuvant chemotherapy with and without sequential trastuzumab. *Arch Surg* 2007;**142**(9):855–61.

145. Zytoon AA. Clinical relevance of the standardized uptake value (Suv) in staging breast cancer with FDG-PET/CT. *Dokkyo J Med Sci* 2007;**34**(3):193–204.

146. Adler LP, Crowe JP, al-Kaisi NK, Sunshine JL. Evaluation of breast masses and axillary lymph nodes with [F-18] 2-deoxy-2-fluoro-D-glucose PET. *Radiology* 1993;**187**(3):743–50.

147. Bassa P, Kim EE, Inoue T, Wong FC, Korkmaz M, Yang DJ, *et al.* Evaluation of preoperative chemotherapy using PET with fluorine-18-fluorodeoxyglucose in breast cancer. *J Nucl Med* 1996;**37**(6):931–8.

148. Beer AJ, Niemeyer M, Carlsen J, Sarbia M, Nahrig J, Watzlowik P, *et al.* Patterns of alphavbeta3 expression in primary and metastatic human breast cancer as shown by ^{18}F-Galacto-RGD PET. *J Nucl Med* 2008;**49**(2):255–9.

149. Bombardieri E, Crippa F, Maffioli L, Chiti A, Castellani MR, Greco M, *et al.* Axillary lymph node metastases detection with nuclear medicine approaches in patients with newly diagnosed breast cancer: can positron emission tomography (PET) with ^{18}F-FDG be considered as the best method? *Int J Oncol* 1996;**8**(4):693–9.

150. Burcombe RJ, Makris A, Pittam M, Lowe J, Emmott J, Wong WL. Evaluation of good clinical response to neoadjuvant chemotherapy in primary breast cancer using [^{18}F]-fluorodeoxyglucose positron emission tomography. *Eur J Cancer* 2002;**38**(3):375–9.

151. Crowe JP, Jr, Adler LP, Shenk RR, Sunshine J. Positron emission tomography and breast masses: comparison with clinical, mammographic, and pathological findings. *Ann Surg Oncol* 1994;**1**(2):132–40.

152. Danforth DN, Aloj L, Carrasquillo JA, Bacharach SL, Chow C, Zujewski J, *et al.* The role of F-18-FDG-PET in the local/regional evaluation of women with breast cancer. *Breast Cancer Res Treat* 2002;**75**(2):135–46.

153. Harisinghani MG, Saksena MA, Hahn PF, King B, Kim J, Torabi MT, *et al.* Ferumoxtran-10-enhanced MR lymphangiography: does contrast-enhanced imaging alone suffice for accurate lymph node characterization? *AJR Am J Roentgenol* 2006;**186**(1):144–8.

154. Harisinghani MG, Weissleder R. Sensitive, noninvasive detection of lymph node metastases. *PLoS Med* 2004;**1**(3):202–9.

155. Hoh CK, Hawkins RA, Glaspy JA, Dahlbom M, Tse NY, Hoffman EJ, *et al.* Cancer-detection with whole-body Pet using 2-[F-18]Fluoro-2-Deoxy-D-Glucose. *J Comput Assist Tomogr* 1993;**17**(4):582–9.

156. Kelemen PR, Lowe V, Phillips N. Positron emission tomography and sentinel lymph node dissection in breast cancer. *Clin Breast Cancer* 2002;**3**(1):73–77.

157. Luciani A, Pigneur F, Ghozali F, Dao TH, Cunin P, Meyblum E, *et al.* Ex vivo MRI of axillary lymph nodes in breast cancer. *Eur J Radiol* 2009;**69**(1):59–66.

158. Noh DY, Yun IJ, Kang HS, Kim YC, Kim JS, Chung JK, *et al.* Detection of cancer in augmented breasts by positron emission tomography. *Eur J Surg* 1999;**165**(9):847–51.

159. Palmedo H, Bender H, Grunwald F, Mallmann P, Zamora P, Krebs D, *et al.* Comparison of fluorine-18 fluorodeoxyglucose positron emission tomography and technetium-99m methoxyisobutylisonitrile scintimammography in the detection of breast tumours. *Eur J Nucl Med* 1997;**24**(9):1138–45.

160. Scheidhauer K, Scharl A, Pietrzyk U, Wagner R, Gohring UJ, Schomacker K, *et al.* Qualitative [^{18}F]FDG positron emission tomography in primary breast cancer: clinical relevance and practicability. *Eur J Nucl Med* 1996;**23**(6):618–23.

161. Smyczek-Gargya B, Fersis N, Dittmann H, Vogel U, Reischl G, Machulla HJ, *et al.* PET with [^{18}F]fluorothymidine for imaging of primary breast cancer: a pilot study. *Eur J Nucl Med Mol Imaging* 2004;**31**(5):720–4.

162. Sun X-CL. Comparative analysis of multi-slice spiral CT and positron emission tomography-CT in evaluation of axillary lymph nodes in breast cancer patients. *Zhonghua Fang She Xian Yi Xue Za Zhi* 2008;**42**(1):68–72.

163. Tse NY, Hoh CK, Hawkins RA, Zinner MJ, Dahlbom M, Choi Y, *et al.* The application of positron emission tomographic imaging with fluorodeoxyglucose to the evaluation of breast disease. *Ann Surg* 1992;**216**(1):27–34.

164. Vansant J, Balfore J, Monticciolo D, Votaw D, Carlson G, Garcia E, *et al.* (F-18)-FDG PET qualitative and quantitative assessment of breast cancer. *J Nucl Med* 1996;**37**(Suppl. 5):254P, abstract 1137.

165. Yang J-HN. Comparison of intraoperative frozen section analysis of sentinel node with preoperative positron emission tomography in the diagnosis of axillary lymph node status in breast cancer patients. *Jpn J Clin Oncol* 2001;**31**(1):1–6.

166. Black D, Specht M, Lee JM, Dominguez F, Gadd M, Hughes K, *et al.* Detecting occult malignancy in prophylactic mastectomy: preoperative MRI versus sentinel lymph node biopsy. *Ann Surg Oncol* 2007;**14**(9):2477–84.

167. Bollet MA, Thibault F, Bouillon K, Meunier M, Sigal-Zafrani B, Savignoni A, *et al.* Role of dynamic magnetic resonance imaging in the evaluation of tumor response to preoperative concurrent radiochemotherapy for large breast cancers: a prospective phase II study. *Int J Radiat Oncol Biol Phys* 2007;**69**(1):13–18.

168. Dose J, Bleckmann C, Bachmann S, Bohuslavizki KH, Berger J, Jenicke L, *et al.* Comparison of fluorodeoxyglucose positron emission tomography and "conventional diagnostic procedures" for the detection of distant metastases in breast cancer patients. *Nucl Med Commun* 2002;**23**(9):857–64.

169. Gil-Rendo A, Martinez-Regueira F, Zornoza G, Garcia-Velloso MJ, Beorlegui C, Rodriguez-Spiteri N, *et al.* Association between [^{18}F]fluorodeoxyglucose uptake and prognostic parameters in breast cancer. *Br J Surg* 2009;**96**(2):166–70.

170. Kim S-JK. Predictive value of [^{18}F]FDG PET for pathological response of breast cancer to neo-adjuvant chemotherapy. *Ann Oncol* 2004;**15**(9):1352–7.

171. Kubota K, Ogawa Y, Nishioka A, Murata Y, Itoh S, Hamada N, *et al.* Radiological imaging features of invasive micropapillary carcinoma of the breast and axillary lymph nodes. *Oncol Rep* 2008;**20**(5):1143–7.

172. Kumar R, Chauhan A, Zhuang HM, Chandra P, Schnall M, Alavi A. Clinicopathologic factors associated with false negative FDG-PET in primary breast cancer. *Breast Cancer Res Treat* 2006;**98**(3):267–74.

173. Perman WH, Heiberg EV, Herrmann VM. Half-Fourier, three-dimensional technique for dynamic contrast-enhanced MR imaging of both breasts and axillae: initial characterization of breast lesions. *Radiology* 1996;**200**(1):263–9.

174. Alberini JL, Lerebours F, Wartski M, Fourme E, Le SE, Gontier E, *et al.* ^{18}F-fluorodeoxyglucose positron emission tomography/computed tomography (FDG-PET/CT) imaging in the staging and prognosis of inflammatory breast cancer. *Cancer* 2009;**115**(21):5038–47.

175. Bathen TF, Jensen LR, Sitter B, Fjosne HE, Halgunset J, Axelson DE, *et al.* MR-determined metabolic phenotype of breast cancer in prediction of lymphatic spread, grade, and hormone status. *Breast Cancer Res Treat* 2007;**104**(2):181–9.

176. Bleckmann C, Dose J, Bohuslavizki KH, Buchert R, Klutmann S, Mester J, *et al.* Effect of attenuation correction on lesion detectability in FDG PET of breast cancer. *J Nucl Med* 1999;**40**(12):2021–4.

177. Buchmann I, Riedmuller K, Hoffner S, Mack U, Aulmann S, Haberkorn U. Comparison of 99mtechnetium-pertechnetate and 123iodide SPECT with FDG-PET in patients suspicious for breast cancer. *Cancer Biother Radiopharm* 2007;**22**(6):779–89.

178. Carkaci S, Macapinlac HA, Cristofanilli M, Mawlawi O, Rohren E, Gonzalez Angulo AM, *et al.* Retrospective study of ^{18}F-FDG PET/CT in the diagnosis of inflammatory breast cancer: preliminary data. *J Nucl Med* 2009;**50**(2):231–8.

179. Chen JH, Baek HM, Nalcioglu O, Su MY. Estrogen receptor and breast MR imaging features: a correlation study. *J Magn Reson Imaging* 2008;**27**(4):825–33.

180. Chung A, Liou D, Karlan S, Waxman A, Fujimoto K, Hagiike M, *et al.* Preoperative FDG-PET for axillary metastases in patients with breast cancer. *Arch Surg* 2006;**141**(8):783–8.

181. Dizendorf EV, Baumert BG, von Schulthess GK, Lutolf UM, Steinert HC. Impact of whole-body ^{18}F-FDG PET on staging and managing patients for radiation therapy. *J Nucl Med* 2003;**44**(1):24–9.

182. Groheux D, Moretti JL, Baillet G, Espie M, Giacchetti S, Hindie E, *et al.* Effect of (^{18})F-FDG PET/CT imaging in patients with clinical Stage II and III breast cancer. *Int J Radiat Oncol Biol Phys* 2008;**71**(3):695–704.

183. Heusner TA, Kuemmel S, Umutlu L, Koeninger A, Freudenberg LS, Hauth EA, *et al.* Breast cancer staging in a single session: whole-body PET/CT mammography. *J Nucl Med* 2008;**49**(8):1215–22.

184. Matsushima S, Sasaki F, Yamaura H, Iwata H, Ohsaki H, Era S, *et al.* Equivalent cross-relaxation rate imaging for sentinel lymph node biopsy in breast carcinoma. *Magn Reson Med* 2005;**54**(5):1300–4.

185. Matsushima S, Nishiofuku H, Iwata H, Era S, Inaba Y, Kinosada Y. Equivalent cross-relaxation rate imaging of axillary lymph nodes in breast cancer. *J Magn Reson Imaging* 2008;**27**(6):1278–83.

186. Mavi A, Cermik TF, Urhan M, Zhuang H, Alavi A. Dual time point imaging method may increase the sensitivity and the specificity of FDG PET for detecting axillary metastasis from primary breast cancer. *Eur J Nucl Med Mol Imaging* 2006;**33**:S167.

187. Mortellaro VE, Marshall J, Singer L, Hochwald SN, Chang M, Copeland EM, *et al.* Magnetic resonance imaging for axillary staging in patients with breast cancer. *J Magn Reson Imaging* 2009;**30**(2):309–12.

188. Mussurakis S, Buckley DL, Horsman A. Dynamic MR imaging of invasive breast cancer: correlation with tumour grade and other histological factors. *Br J Radiol* 1997;**70**(833):446–51.

189. Nieweg OE, Kim EE, Wong WH, Broussard WF, Singletary SE, Hortobagyi GN, *et al.* Positron emission tomography with fluorine-18-deoxyglucose in the detection and staging of breast cancer. *Cancer* 1993;**71**(12):3920–5.

190. Rotaru N, Luciani A, Rotaru N, Luciani A. Magnetic resonance imaging of the breast: potential for lesion characterization. *J BUON* 2004;**9**(1):77–82.

191. Schirrmeister H, Kuhn T, Guhlmann A, Santjohanser C, Horster T, Nussle K, *et al.* Fluorine-18 2-deoxy-2-fluoro-D-glucose PET in the preoperative staging of breast cancer: comparison with the standard staging procedures. *Eur J Nucl Med* 2001;**28**(3):351–8.

192. Schirrmeister HH, Kuehn T, Buck AC, Santjohanser C, Reske SN. FDG-PET in preoperative assessment of newly diagnosed breast cancer. *Breast Cancer Res Treat* 2000;**64**(1):118, abstract 510.

193. Song B, Kim M, Lee J, Seo Y, Park W, Oh S, *et al.* Evaluation of PET-CT for axillary lymph node staging in patients with early stage breast cancer. *EJC Supplements* 2006;**4**(2):65, abstract 70.

194. Stets C, Brandt S, Wallis F, Buchmann J, Gilbert FJ, Heywang-Kobrunner SH. Axillary lymph node metastases: a statistical analysis of various parameters in MRI with USPIO. *J Magn Reson Imaging* 2002;**16**(1):60–8.

195. Torrenga HL. Re: Axillary lymph node staging in breast cancer by 2-fluoro-2-deoxy-D-glucose-positron emission tomography: clinical evaluation and alternative management. *J Natl Cancer Inst* 2001;**93**(21):1659–61.

196. Ueda S, Tsuda H, Asakawa H, Shigekawa T, Fukatsu K, Kondo N, *et al.* Clinicopathological and prognostic relevance of uptake level using ^{18}F-fluorodeoxyglucose positron emission tomography/computed tomography fusion imaging (^{18}F-FDG PET/CT) in primary breast cancer. *Jpn J Clin Oncol* 2008;**38**(4):250–8.

197. Yang SK, Cho N, Moon WK. The role of PET/CT for evaluating breast cancer. *Korean J Radiol* 2007;**8**(5):429–37.

198. Bock E, Bock C, Belli P, Campioni P, Manfredi R, Pastore G. [Role of diagnostic imaging of the breast in patients treated with postsurgical radiotherapy or presurgical radiotherapy or chemotherapy] [Italian]. *Radiologia Medica* 1998;**95**(1–2):38–43.

199. Chen SH, Ma QH, Li ZJ, Hu CW, Chen Sh, Ma Qh, *et al.* [Diagnosis of breast diseases by mammography in combination with MRI] [Chinese]. *Zhongguo Yiliao Qixie Zazhi* 2007;**31**(3):216–18.

200. Dose-Schwarz J, Mahner S, Schirrmacher S, Jenicke L, Muller V, Habermann CR, *et al.* [Detection of metastases in breast cancer patients: comparison of FDG PET with chest X-ray, bone scintigraphy and ultrasound of the abdomen] [German]. *Nuclear-Medizin* 2008;**47**(3):97–103.

201. Enya M, Goto H, Nandate Y, Kiryu T, Kanematsu M, Hoshi H. [Usefulness of dynamic MR mammography for diagnosis of axillary lymph node status in breast cancer patient]. [Japanese]. *Nippon Igaku Hoshasen Gakkai Zasshi* 2000;**60**(13):764–70.

202. Koga S, Nakano S, Honma Y, Ogasawara N. [FDG-PET (positron emission tomography) in the detection of primary breast cancer and lymph node involvement]. [Japanese]. *Nippon Rinsho* 2007;**65**(Suppl. 6):379–84.

203. Nakayama S, Kakizaki D, Kaise H, Kusama M, Ishikawa A, Amino M, *et al.* [Three-dimentional volumetric interpolated breath-hold magnetic resonance imaging for the diagnosis of breast tumors] [Japanese]. *Nippon Rinsho* 2004;**62**(4):790–8.

204. Romer W, Avril N, Schwaiger M. [Possible uses of positron emission tomography in breast carcinoma]. *Acta Medica Austriaca* 1997;**24**(2):60–2.

205. Tiutin L, Fadeev N, Ryzhkova D, Kostenikiv N, Arzumanov A, Savello V, *et al.* [Clinical experience with positron emission tomography with 2-fluorine, ^{18}F-2-deoxy-D-glucose used for the diagnosis of malignant breast neoplasms]. [Russian]. *Vestn Rentgenol Radio* 2001;**6**:14–18.

206. Zhao TT, Li JG, Li YM. [Performance of ^{18}F-FDG PET/CT in the detection of primary breast cancer and staging of the regional lymph nodes] [Chinese]. *Chung-Hua Chung Liu Tsa Chih* 2007;**29**(3):206–9.

207. al-Yasi AR, Carroll MJ, Ellison D, Granowska M, Mather SJ, Wells CA, *et al.* Axillary node status in breast cancer patients prior to surgery by imaging with Tc-99m humanised anti-PEM monoclonal antibody, hHMFG1. *Br J Cancer* 2002;**86**(6):870–8.

208. Altinyollar H, Dingil G, Berberoglu U. Detection of infraclavicular lymph node metastases using ultrasonography in breast cancer. *J Surg Oncol* 2005;**92**(4):299–303.

209. Altomare V, Guerriero G, Carino R, Battista C, Primavera A, Altomare A, *et al.* Axillary lymph node echo-guided fine-needle aspiration cytology enables breast cancer patients to avoid a sentinel lymph node biopsy. Preliminary experience and a review of the literature. *Surg Today* 2007;**37**(9):735–9.

210. Holle LH, Trampert L, Lung-Kurt S, Villena-Heinsen CE, Puschel W, Schmidt S, *et al.* Investigations of breast tumors with fluorine-18-fluorodeoxyglucose and SPECT. *J Nucl Med* 1996;**37**(4):615–22.

211. Lumachi FT. Accuracy of ultrasonography and ^{99m}Tc-sestamibi scintimammography for assessing axillary lymph node status in breast cancer patients. A prospective study. *Eur J Surg Oncol* 2006;**32**(9):933–6.

212. Mankoff DA, Dunnwald LK, Gralow JR, Ellis GK, Drucker MJ, Livingston RB. Monitoring the response of patients with locally advanced breast carcinoma to neoadjuvant chemotherapy using [technetium 99m]-sestamibi scintimammography. *Cancer* 1999;**85**(11):2410–23.

213. Mansi L, Rambaldi PF, Procaccini E, DiGregorio F, Laprovitera A, Pecori B, *et al.* Scintimammography with technetium-99m tetrofosmin in the diagnosis of breast cancer and lymph node metastases. *Eur J Nucl Med* 1996;**23**(8):932–9.

214. Myslivecek M, Koranda P, Kaminek M, Husak V, Hartlova M, Duskova M, *et al.* Technetium-99m-MIBI scintimammography by planar and SPECT imaging in the diagnosis of breast carcinoma and axillary lymph node involvement. *Nucl Med Rev* 2004;**7**(2):151–5.

215. Paredes P, Vidal-Sicart S, Zanon G, Pahisa J, Fernandez PL, Velasco M, *et al.* Clinical relevance of sentinel lymph nodes in the internal mammary chain in breast cancer patients. *Eur J Nucl Med Mol Imaging* 2005;**32**(11):1283–7.

216. Schillaci O, Scopinaro F, Danieli R, Tavolaro R, Cannas P, Picardi V, *et al.* Technetium-99m sestamibi imaging in the detection of axillary lymph node involvement in patients with breast cancer. *Anticancer Res* 1997;**17**(3B):1607–10.

217. Shien T, kashi-Tanaka S, Yoshida M, Hojo T, Iwamoto E, Miyakawa K, *et al.* Evaluation of axillary status in patients with breast cancer using thin-section CT. *Int J Clin Oncol* 2008;**13**(4):314–19.

218. Spanu A, Chessa F, Sanna D, Cottu P, Manca A, Nuvoli S, *et al.* Breast cancer axillary lymph node metastasis detection by a high-resolution dedicated breast camera: a comparative study with SPECT and pinhole SPECT. *Cancer Biother Radiopharm* 2007;**22**(6):799–811.

219. Spanu A, Dettori G, Chessa F, Porcu A, Cottu P, Solinas P, *et al.* Tc-99m-tetrofosmin pinhole-SPECT (P-SPECT) and radioguided sentinel node (SN) biopsy and in breast cancer axillary lymph node staging. *Cancer Biother Radiopharm* 2001;**16**(6):501–13.

220. Spanu A, Dettori G, Chiaramida P, Cottu P, Falchi A, Porcu A, *et al.* The role of Tc-99m-Tetrofosmin Pinhole-SPECT in breast cancer axillary lymph node staging. *Cancer Biother Radiopharm* 2000;**15**(1):81–91.

221. Spanu A, Tanda F, Dettori G, Manca A, Chessa F, Porcu A, *et al.* The role of Tc-99m-tetrofosmin pinhole-SPECT in breast cancer non palpable axillary lymph node metastases detection. *Q J Nucl Med* 2003;**47**(2):116–28.

222. Sperber FW. Preoperative clinical, mammographic and sonographic assessment of neoadjuvant chemotherapy response in breast cancer. *Isr Med Assoc J* 2006;**8**(5):342–6.

223. Taillefer R, Robidoux A, Turpin S, Lambert R, Cantin J, Leveille J. Metastatic axillary lymph node technetium-99m-MIBI imaging in primary breast cancer. *J Nucl Med* 1998;**39**(3):459–64.

224. Taillefer R, Robidoux A, Lambert R, Turpin S, Laperriere J. Technetium-99M-sestamibi prone scintimammography to detect primary breast-cancer and axillary lymph-node involvement. *J Nucl Med* 1995;**36**(10):1758–65.

225. Dose J, Avril N, Graeff H, Jaenicke F. Positron emission tomography for diagnosis of breast tumors. *Onkologie* 1997;**20**(3):190–5.

226. Prati R, Minami CA, Gornbein JA, Debruhl N, Chung D, Chang HR, *et al.* Accuracy of clinical evaluation of locally advanced breast cancer in patients receiving neoadjuvant chemotherapy. *Cancer* 2009;**115**(6):1194–202.

227. Smith IC, Welch AE, Hutcheon AW, Miller ID, Payne S, Chilcott F, *et al.* Positron emission tomography using [F-18]-fluorodeoxy-D-glucose to predict the pathologic response of breast cancer to primary chemotherapy. *J Clin Oncol* 2000;**18**(8):1676–88.

228. Bradley AJ, Carrington BM, Hammond CL, Swindell R, Magee B. Accuracy of axillary MR imaging in treated breast cancer for distinguishing between recurrent tumour and treatment effects: does intravenous Gd-DTPA enhancement help in cases of diagnostic dilemma? *Clin Radiol* 2000;**55**(12):921–8.

229. Lin W-YT. Fluorine-18 FDG-PET in detecting local recurrence and distant metastases in breast cancer – Taiwanese experiences. *Cancer Invest* 2002;**20**(5–6):725–9.

230. Mustafa S, Marshall C, Griffiths PA, Chuthapisith S, Wheatley DC, Eremin JM, *et al.* ^{18}F-FDG dual-headed gamma camera PET in detection of axillary nodal disease in patients with large or locally advanced breast cancer: possible alternative staging of axilla. *Oncol Rep* 2007;**18**(6):1545–9.

231. Ohta M, Tokuda Y, Saitoh Y, Suzuki Y, Okumura A, Kubota M, *et al.* Comparative efficacy of positron emission tomography and ultrasonography in preoperative evaluation of axillary lymph node metastases in breast cancer. *Breast Cancer* 2000;**7**(1):99–103.

232. Rostom AY, Powe J, Kandil A, Ezzat A, Bakheet S, el-Khwsky F, *et al.* Positron emission tomography in breast cancer: a clinicopathological correlation of results. *Br J Radiol* 1999;**72**(863):1064–8.

Appendix 1

Literature search strategies

The numbers in the brackets are the number of citations identified by each search term.

MEDLINE search strategy for clinical effectiveness studies

1. exp Breast Neoplasms/ (160,544)
2. exp Neoplasms/ (1,998,053)
3. exp Carcinoma/ (378,056)
4. exp Adenocarcinoma/ (215,523)
5. 4 or 3 or 2 (1,998,053)
6. exp Breast/ (25,097)
7. 6 and 5 (13,438)
8. ((breast$or mamma$) adj5 (cancer$or carcin$or tumour$or tumor$or neoplasm$or malignan$)).tw. (162,727)
9. 1 or 7 or 8 (203,102)
10. Positron-Emission Tomography/ (11,502)
11. positron emission tomography.tw. (20,247)
12. PET.tw. (29,636)
13. exp Magnetic Resonance Imaging/ (198,356)
14. magnetic resonance.tw. (129,805)
15. mri.tw. (79,788)
16. or/10–15 (290,335)
17. 9 and 16 (4802)
18. exp "Sensitivity and Specificity"/ (276,863)
19. sensitivity.tw. (353,675)
20. specificity.tw. (227,637)
21. ((pre-test or pretest) adj probability).tw. (755)
22. post-test probability.tw. (217)
23. predictive value$.tw. (43,259)
24. likelihood ratio$.tw. (4749)
25. or/18–24 (702,339)
26. 25 and 17 (1471)
27. Axilla/ (7809)
28. axilla$.tw. (19,578)
29. exp lymphatic system/or exp lymph nodes/ (197,948)
30. Lymphatic Metastasis/ (52,995)
31. lymphatic system/or exp lymphatic vessels/or exp lymphoid tissue/ (197,847)
32. lymph$.tw. (548,454)
33. or/27–32 (691,123)
34. 33 and 26 (299)

MEDLINE search strategy for cost-effectiveness studies

1. exp Breast Neoplasms/ (160,809)
2. exp Neoplasms/ (2,000,086)
3. exp Carcinoma/ (378,481)
4. exp Adenocarcinoma/ (215,800)
5. 4 or 3 or 2 (2,000,086)
6. exp Breast/ (25,139)
7. 6 and 5 (13,472)
8. ((breast$or mamma$) adj5 (cancer$or carcin$or tumour$or tumor$or neoplasm$or malignan$)).tw. (163,002)
9. 1 or 7 or 8 (203,430)
10. Positron-Emission Tomography/ (11,569)
11. positron emission tomography.tw. (20,284)
12. PET.tw. (29,688)
13. exp Magnetic Resonance Imaging/ (198,724)
14. magnetic resonance.tw. (129,992)
15. mri.tw. (79,951)
16. or/10–15 (290,851)
17. 9 and 16 (4816)
18. Axilla/ (7823)
19. axilla$.tw. (19,613)
20. exp lymphatic system/or exp lymph nodes/ (198,061)
21. Lymphatic Metastasis/ (53,075)
22. lymphatic system/or exp lymphatic vessels/or exp lymphoid tissue/ (197,960)
23. lymph$.tw. (548,945)
24. or/18–23 (691,694)
25. 24 and 17 (855)
26. Economics/ (25,278)
27. "costs and cost analysis"/ (36,971)
28. Cost allocation/ (1848)
29. Cost–benefit analysis/ (44,801)
30. Cost control/ (17,850)
31. cost savings/ (6180)
32. Cost of illness/ (11,331)
33. Cost sharing/ (1482)
34. "deductibles and coinsurance"/ (1202)
35. Health care costs/ (17,597)
36. Direct service costs/ (860)
37. Drug costs/ (8947)
38. Employer health costs/ (998)
39. Hospital costs/ (5768)
40. Health expenditures/ (10,378)
41. Capital expenditures/ (1841)
42. Value of life/ (5003)
43. exp economics, hospital/ (15,808)
44. exp economics, medical/ (11,638)
45. Economics, nursing/ (3779)
46. Economics, pharmaceutical/ (2001)
47. exp "fees and charges"/ (23,905)
48. exp budgets/ (10,044)

49. (low adj cost).mp. (11,582)
50. (high adj cost).mp. (5361)
51. (health?care adj cost$).mp. (1899)
52. (fiscal or funding or financial or finance).tw. (49,856)
53. (cost adj estimate$).mp. (918)
54. (cost adj variable).mp. (23)
55. (unit adj cost$).mp. (941)
56. (economic$or pharmacoeconomic$or price$or pricing).tw. (108,763)
57. or/26–56 (327,682)
58. 25 and 57 (21)

Appendix 2

Quality assessment

Quality assessment of diagnostic accuracy studies (QUADAS)[69] and details of criteria for scoring studies.

1. Was the spectrum of patients representative of the patients who will receive the test in practice?	
Yes	All patients were early stage, newly diagnosed and non-DCIS, and patients were recruited both prospectively and consecutively
No	Not all patients were early stage/newly diagnosed, or some had DCIS, or patients were studied retrospectively or non-consecutively
Unclear	Insufficient details given about stage or recruitment methods to make a judgement about whether the patient spectrum would be scored 'yes'
2. Is the reference standard likely to correctly classify the target condition?	
Yes	All patients received either ALND or SLNB
No	Some or all patients received any other reference standard, including FNAC, or no reference standard
Unclear	Reference standard is not stated
3. Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	
Yes	Reference standard was performed within 3 months of the index test
No	Reference standard was performed more than 3 months after the index test
Unclear	The time between reference standard and index test is not stated
4. Did patients receive the same reference standard regardless of the index test result?	
Yes	Selection of reference standard was not determined by the index text result
No	Selection of reference standard was determined by the index test result
Unclear	It is not clear whether selection of reference standard was determined by the index test result
5. Was the execution of the index test described in sufficient detail to permit replication of the test?	
6. Was the execution of the reference standard described in sufficient detail to permit replication of the test?	
Yes	Sufficient details of the test execution are reported
No	Sufficient details are not reported
Unclear	Not applicable
7. Were the index test results interpreted without knowledge of the results of the reference standard?	
8. Were the reference standard results interpreted without knowledge of the results of the index test?	
Yes	The index test was interpreted without knowledge of the results of the reference standard or vice versa. If the test was clearly interpreted before the results of the other test were available then this was scored as 'yes'
No	The person interpreting the index test was aware of the results of the reference standard or vice versa
Unclear	No information is provided regarding whether tests were interpreted blindly
9. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	
Yes	PET studies included a CT scan; reference standard interpreted using usual methods; and clinical data relating to the primary tumour (examination, ultrasound, biopsy results) were available to the interpreting radiologist
No	PET studies did NOT include a CT scan; and/or reference standard was NOT interpreted using usual methods; and/or clinical data relating to the primary tumour (examination, ultrasound, biopsy results) were NOT available to the interpreting radiologist
Unclear	Insufficient data reported regarding the text methods and availability of clinical data
10. Were uninterpretable/intermediate test results reported?	
Yes	
No	
Unclear	

11. Were withdrawals from the study explained?	
Yes	All patients recruited to the study were accounted for
No	There appear to be patients who were recruited into the study who are not accounted for
Unclear	It is not clear whether any withdrawals occurred

Appendix 3

Summary of excluded studies

Study	Number	Reason for exclusion
Bland 2004,[140] Guller 2003,[141] Wahl 2004[142]	3	Comment, not original research
Bahri 2008,[143] Hsiang 2007,[144] Zytoon 2007[145]	3	Imaging of breast not axilla
Adler 1993,[146] Bassa 1996,[147] Beer 2008,[148] Bombardieri 1996,[149] Burcombe 2002,[150] Crowe 1994,[151] Danforth 2002,[152] Harisinghani 2006,[153] Harisinghani 2004,[154] Hoh 1993,[155] Kelemen 2002,[156] Luciani 2009,[157] Noh 1999,[158] Palmedo 1997,[159] Scheidhauer 1996,[160] Smyczek-Gargya 2004,[161] Sun 2008,[162] Tse 1992,[163] Vansant 1996,[164] Yang 2001[165]	20	$n < 20$, PET
Black 2007,[166] Bollet 2007,[167] Dose 2002,[168] Gil-Rendo 2009,[169] Kim 2004,[170] Kubota 2008,[171] Kumar 2006,[172] Perman 1996[173]	8	No data on axilla
Alberini 2009,[174] Bathen 2007,[175] Bleckmann 1999,[176] Buchmann 2007,[177] Carkaci 2009,[178] Chen 2008,[179] Chung 2006,[180] Dizendorf 2003,[181] Groheux 2008,[182] Heusner 2008,[183] Matsushima 2005,[184] Matsushima 2008,[185] Mavi 2006,[186] Mortellaro 2009,[187] Mussurakis 1997,[188] Nieweg 1993,[189] Rotaru 2004,[190] Schirrmeister 2001,[191] Schirrmeister 2000,[192] Song 2006,[193] Stets 2002,[194] Torrenga 2001,[195] Ueda 2008,[196] Yang 2007[197]	24	No relevant data
Bock 1998,[198] Chen 2007,[199] Dose-Schwarz 2008,[200] Enya 2000,[201] Koga 2007,[202] Nakayama 2004,[203] Romer 1997,[204] Tiutin 2001,[205] Zhao 2007[206]	9	Non-English language
Al-Yasi 2002,[207] Altinyollar 2005,[208] Altomare 2007,[209] Holle 1996,[210] Lumachi 2006,[211] Mankoff 1999,[212] Mansi 1996,[213] Myslivecek 2004,[214] Paredes 2005,[215] Schillaci 1997,[216] Shien 2008,[217] Spanu 2007,[218] Spanu 2001,[219] Spanu 2000,[220] Spanu 2003,[221] Sperber 2006,[222] Taillefer 1998,[223] Taillefer 1995[224]	18	Not PET or MRI
Dose 1997[225]	1	Review
Prati 2009,[226] Smith 2000,[227] Bradley 2000,[228] Lin 2002,[229] Mustafa 2007,[230] Ohta 2000,[231] Rostom 1999[232]	7	> 20% not correct population (not early stage, had neoadjuvant chemotherapy or not newly diagnosed)

Appendix 4

Details of index test and reference standard

Positron emission tomography studies

Study	PET or PET/CT	PET details	Reference standard (summary)	Reference standard (detail)
Chae 2009[73]	PET/CT	Fasted for at least 6 hours. Serum glucose levels <120 mg/dl. 5 mCi FDG injected, 1 hour later whole body scan and abdominal tomoscintigraphy performed in supine position, thoracic tomoscintigraphy in prone position with arms extended. SUV_{max} measurement was corrected for body weight. Scans interpreted quantitatively using an SUV cut-off ≥0.64. Not reported who interpreted scans	100% ALND (all SLNB in addition; ALND results used in analysis)	*SLNB*: each half-sentinel node sectioned at 3 mm intervals; each section analysed by four levels of 150 µm and four parallel sections. One level used for H&E staining. Immunohistochemical staining was NOT performed routinely. *ALND*: following SLNB, level I and II dissection performed
Heusner 2009[74]	PET/CT	Fasted 6 hours before intravenous injection of 271 ± 35 MBq FDG (210–360 MBq). Blood glucose <150 mg/dl. 1500 ml water-based oral contrast agent applied for bowel marking. Whole body PET imaging 1 hour after FDG injection. CT acquired before PET with 100 mA/s at 130 kV; 140 ml of an iodinated contrast material administered by injection. All images reconstructed with 5 mm slice thickness and 2.4 mm increment. PET emission time adapted to the patient's body weight: <65 kg = 4 minutes per bed position, 65–85 kg = 5 minutes per bed position, >85 kg = 6 minutes per bed position Image reconstruction with two iterations and eight subsets. Data filtered and corrected for scatter. Experienced radiologist and nuclear medicine physician analysed PET/CT data and diagnosis was made by consensus. Scans evaluated in three orthogonal planes (axial, coronal and sagittal). Visual analysis considered positive if abnormally high uptake corresponding with a lymph node	ALND and/or SLNB	SLNB for those with cT1 or cT2 and no clinical suspicion of lymph node metastases. ALND for everyone else and those with positive SLNB. *SLNB*: intraoperative frozen section, remove perinodal fatty tissue, section into 2 mm slices. Only most suspicious slice snap frozen and one superficial cryostat section examined histologically. Standard fixation overnight and paraffin embedded; H&E staining at three serial 500 µm levels, reviewed by two gynaecological pathologists. *ALND*: no details reported
Kim 2009[75]	PET/CT	Fasted for 6 hours, serum glucose ≤150 mg/dl, whole-body scan. CT scan (non-contrast) 1 hour after injection of 7.4 MBq/kg (0.2 mCi/kg) of 18-FDG for attenuation correction. Transaxial, coronal, sagittal planes. Interpreted by two experienced nuclear medicine physicians. Visual assessment of nodes; considered positive if more uptake than background. SUV_{max} measurement corrected for body weight	ALND and/or SLNB (SLNB included additional non-SLNB)	*SLNB + ADD*: injection of blue dye followed by gentle breast massage for 3 minutes. 5 minutes after injection, incision made over lower axilla to remove sentinel nodes and additional non-SLNB enlarged at lower axilla. 2 mm slices examined in frozen section intraoperatively, H&E staining also performed on paraffin-embedded permanent sections. ALND for those with positive PET/CT scan or positive intraoperative sample on SLNB or ADD. No details reported
Taira 2009[76]	PET/CT	Four-hour fast, 3 MBq/kg body weight 18-F-FDG injected, scan performed 60 minutes later. Whole-body scan in supine position with elevation of arms. Analysed by two expert radiologists. Visual analysis; positive if more uptake than background. Quantitative SUV_{max} measurement if abnormal uptake	ALND and/or SLNB (ALND if PET or SLNB positive or due to patient preference)	*SLNB*: blue dye method. 3 ml injected, massaged 5 minutes, then excised. Rapid pathological diagnosis with H&E. ALND performed if SLNB positive. *SLNB post-operative histology*: fixed with formalin and embedded in paraffin. 200 µm whole cross section, with H&E staining and immunostaining with anti-cytokeratin antibody. If positive, secondary ALND performed. Non-NO sentinel nodes 1–2 sections prepared from each lymph node, H&E staining
Fuster 2008[77]	PET/CT	Fast 4 hours prior. Blood glucose required to be <7 mmol/l. 740 MBq FDG injected, scan 1 hour later. No oral or i.v. contrast used. Whole-body PET. Reconstructed with and without CT data for attenuation correction. Interpreted separately by two nuclear medicine physicians. Visual analysis and SUV measurement	100% ALND	None reported

Study	PET or PET/CT	PET details	Reference standard (summary)	Reference standard (detail)
Ueda 2008[78]	PET/CT	Fast 4 hours prior. Blood glucose required to be ≤120 mg/dl. 3.7 Mbq/kg FDG injected, scan 1 hour later. Transmission scan with CT for attenuation correction and anatomical imaging for 90 seconds. i.v. contrast not administered for CT. Gaussian filter, back projection image obtained, spacial resolution 6–7 mm cranio-caudal, 6.3–7.1 mm right–left, 6.3 and 7.1 mm in anterior-posterior directions. Two experienced radiologists interpreted scans. PET results determined via visual assessment (abnormal axillary uptake) and also via various SUV cut-offs. Results using visual assessment presented here	ALND and/or SLNB	ALND for those positive at ultrasound or due to patient preference. SLNB for those negative at ultrasound, with full ALND if intraoperative histology positive *SLNB histology*: 2 mm thick slices, haematoxylin and eosin stain, two pathologists examined
Veronesi 2007[79]	PET/CT	Fast 6 hours prior. Blood glucose required to be <140 mg/dl. 5.3 MBq/kg FDG injected, three glasses of water, empty bladder. Scan 45 minutes later. Supine position, arms over head. CT scan before PET, attenuation-corrected images reconstructed in transaxial, coronal and sagittal planes. Interpreted independently by three experienced nuclear medicine physicians. Initial qualitative assessment of images, quantitative analysis of those equivocal, probably abnormal or abnormal. Any focal uptake greater than normal considered, definitely abnormal if SUV (normalised to body weight) >1.2	ALND and/or SLNB (ALND if PET or SLNB positive)	*SLNB only if PET/CT negative*: SN dissected along major axis if >5 mm, uncut if <5 mm, then sectioned at 50/100 µm intervals. H&E stain. *Suspicious sections*: untreated portion of section stained with cytokeratins. Micrometastases ≤2 mm diameter. ALND if SLNB or PET/CT positive. Level I, II and III cleared – SNs removed during ALND were treated differently, sectioned 100–500 µm apart, stained with H&E only
Cermik 2008[80]	PET only	Fast 4 hours prior. Serum glucose levels required to be <140 mg/dl. Intravenous FDG (5.2 MBq/kg), PET scan 1 hour later. Whole-body PET scans. Transmission images interleaved between multiple emission scans. Images interpreted with and without attenuation correction. Interpreted by single experienced nuclear medicine physician. Visual analysis and SUV measurement	ALND and/or SLNB	*Kumar 2006*:[102] radiotracer and blue dye injected (subareolar) to identify sentinel lymph nodes. SLNs dissected. Dissected lymph nodes formalin fixed and paraffin embedded. SLNs evaluated by H&E and by immunohistochemistry (cytokeratin antibody). Non-sentinel nodes evaluated by H&E *Cermik 2008*:[80] no details reported
Gil-Rendo 2006[81]	PET only	Fast 6 hours prior. Blood glucose required to be <120 mg/dl. 370 MBq FDG injected, scan 40–60 minutes later. Whole-body emission-transmission scan. Reconstructed with and without attenuation correction. From Zornoza paper:[103] supine position, arms separated from side by roller. Four-to-five overlapping emission-transmission images obtained in 2D acquisition mode. Two experienced nuclear medicine physicians interpreted scans. Visual analysis; considered positive if axillary uptake higher than surrounding tissue. SUVmax also measured	ALND and/or SLNB (first $n=150$: ALND; second $n=125$: ALND if PET or SLNB positive; SLNB otherwise)	Lymphatic mapping using blue dye. SLNs studied intraoperatively (sections stained with H&E) and all patients with SLN metastases received completion ALND. If SLN negative on initial analysis, further sections analysed via immunocytochemistry with anti-cytokeratin antibodies. *ALND*: level I and II. Nodes sectioned and stained with H&E. *Note*: nodes undergoing SLNB and immunocytochemistry are sectioned multiple times compared to ALND nodes therefore may have detected more micrometastases
Weir 2005[82]	PET only	Six-hour fast. 5.5 MBq/Kg of FDG given intravenously on contralateral arm, 40–60 minutes before scan. Three-to-six bed position interleaved emission and transmission scans. Images reconstructed using iterative methods. Images interpreted by 'experienced nuclear medicine physicians'	ALND and/or SLNB	ALND level I and II
Agresti 2004[83]	PET only	No details reported	ALND and/or SLNB (ALND if PET or SLNB positive)	*SLNB*: radioguided using lymphoscintigraphy. ALND/SLNB evaluated on sections

DOI: 10.3310/hta15040

Study	PET or PET/CT	PET details	Reference standard (summary)	Reference standard (detail)
Fehr 2004[84]	PET only	Four-hour fast. 600–700 MBq of FDG given intravenously, 40–50 minutes before scan. Whole body scan in supine position plus scan of breasts and axilla in prone position. Images reconstructed using standard back-projections and with segmented attenuation correction. Images interpreted by two nuclear medicine physicians. Considered positive if axilla showed more uptake than surrounding tissue	100% ALND (all SLNB in addition)	*SLNB*: radioguided using lymphoscintigraphy (technetium 99 m colloidal albumin) and blue dye. SLNs formalin fixed, paraffin embedded, half the sections stained with H&E and half via immunohistochemistry using pan-cytokeratin staining. No frozen sections. *ALND*: level I and II. Formalin fixed, paraffin embedded, H&E stained (and immunohistochemistry using pan-cytokeratin staining for cases of invasive lobular carcinoma)
Inoue 2004[85]	PET only	Fast 4 hours prior. Blood glucose required to be < 150 mg/dl. 300–510 MBq FDG injected, scan 1 hour later. Reconstructed with and without attenuation correction using filtered back-projection and Butterworth filter. From Yutani 2000:[107] emission then transmission scans. Images interpreted by two experienced radiologists. Considered positive if focal FDG uptake present in axilla (SUV also measured for primary tumour)	All confirmed by histology, 'almost all' ALND (information from author)	No details reported
Lovrics 2004[86]	PET only	Patients fasted for at least 4 hours and had two glasses of water 30–60 minutes before scan. If plasma glucose levels were > 10 mmol/l, given insulin at nuclear physician's discretion. Intravenous FDG (5 mCi) in contralateral hand, PET scan 45 minutes later. Whole-body PET scans, arms over head. Forty-seven slices per bed position (mean of five positions), overlap of 5 cm, resolution 6 mm, axial field of view 15 cm. No attenuation correction. Visual and semi-quantitative analysis. Interpreted by single nuclear physician	ALND and/or SLNB	*ALND*: morning of surgery or afternoon of day before surgery, 1.0 mCi technetium 99 sulfur colloid injected peritumorally/at excision site (in cases of excision biopsy?) Lymphoscintigraphy not routinely performed. Intraoperatively vital blue dye injected peritumorally, Sentinel node identified with hand-held gamma probe. *Histology*: level I and II and routine histology: sentinal nodes > 0.5 cm sectioned transversely, < 0.5 cm bisected. The H&E SLNB as above. Where routine histology negative, micrometastasis histology performed: Cam 5.2 (anti-cytokeratin) stain and serial sections at 50 µm
Wahl 2004[44]	PET only	Four-hour fast, checked fasting blood sugar. Injected with 17 ± 3 mCi of FDG in opposite arm or foot if bilateral disease, 40 minutes before scanning. Nearly all cases, arms above head for entire scanning period. Emission and transmission scans at two levels. Images reconstructed by filtered back-projection using a Hann filter. Quality control committee assessed image quality. Three experienced nuclear medicine physicians independently interpreted results. Visual analysis; probably or definitely abnormal PET was considered positive. Also SUV measurement for any abnormal images	100% ALND (some SLNB in addition)	At least level II. Nodes fixed in formalin, embedded in paraffin and stained with H&E. Depending on size, each lymph node sectioned into two or three parts, and one or more sections prepared from each part
Barranger 2003[87]	PET only	Fast 6 hours prior. Blood glucose required to be < 7 mmol/l. 4 MBq/kg FDG injected, scan 1 hour later. Whole-body scan and abdominal tomoscintigraphy in supine position, and thoracic tomoscintigraphy in prone position with arms extended, using mammoscintigraphy table. Reconstructed by iterative algorithm without attenuation correction. Two experienced nuclear medicine physicians interpreted scans. Visual analysis; consensus regarding presence or absence of abnormal uptake in axilla (and degree of uptake)	100% ALND (all SLNB in addition)	*SLNB*: (all with blue dye, 28 also with lymphoscintigraphy). 3 mm sections, H&E staining and immunohistochemistry with anti-cytokeratin antibody cocktail. *ALND*: level I and II. Sections stained with H&E

Study	PET or PET/CT	PET details	Reference standard (summary)	Reference standard (detail)
Guller 2002[88]	PET only	Fast ≥ 12 hours prior. 5 MBq/kg FDG injected, scan 90 minutes later. Whole-body scan. Prone position. Not reported who analysed. Considered positive if detectable lymph node metastases present	ALND and/or SLNB	SLN identified by combination of blue dye and radioisotope (99 m-Tc-colloid). Histopathologic analysis of SLN (H&E staining, immunohistochemistry with CK22 and Lu-5, step sections of lymph nodes)
Nakamoto 2002[89]	PET only	Fast ≥ 4 hours prior. 370 MBq FDG injected. Arms at side of body during scan. Transmission scans obtained using 68-Ge rod source for purpose of attenuation correction. Dynamic images at breast level, then static images of axillary area, obtained shortly after injection of FDG. Images reconstructed using filtered back projection algorithm and Hanning filter. Images generated with and without attenuation correction. Two experienced nuclear medicine physicians independently reviewed images. Visual analysis. Considered positive if uptake was probably or definitely abnormal on consensus between analysts. Attenuation-corrected and non-attenuation-corrected images analysed separately. SUV also measured	100% ALND	No details reported
Rieber 2002[90]	PET only	Eight-hour fast. Average 370 MBq FDG injected. Scan carried out 45–60 minutes after injection. Prone position, elevated arms. Emission scans at three bed positions. No attenuation correction. Clearly visible, focally increased tracer uptake = positive scan. One experienced observer	Histology (no further details)	Unclear – simply states histology
van der Hoeven 2002[91]	PET only	Six-hour fast, serum glucose measured. 370 MBq FDG in contralateral arm 60 minutes before scan. Emission scans of chest (two-dimensional mode, two bed positions) with patient in supine position. Images reconstructed with filtered back-projection using a Hanning filter, resulting in spatial resolution of ~7 mm full-width at half maximum. Three experienced clinicians interpreted images. Final diagnosis was taken as the result agreed upon by at least two of the three. Considered positive if PET scan showed at least moderately enhanced uptake	ALND and/or SLNB (patients were stratified for ALND or SLNB largely depending on tumour size)	ALND nodes sliced at one or two levels (< 1 cm and > 1 cm, respectively) and stained with CAM5.2 before examination SLNB used combined blue dye/radiocolloid approach. Sentinel nodes were sliced at four-to-six other levels if the first slice was negative. H&E and CAM5.2 staining. (CAM5.2 is a immunohistochemical stain)
Greco 2001[92]	PET only	Fast ≥ 5 hours prior. Required to have normal fasting blood glucose levels. 400 MBq (11 mCi) FDG injected. Most patients scanned in supine position with arms raised. Two transmission scans in bed position before FDG injection for attenuation correction. Two static emission scans 45–60 minutes after FDG injection. Images attenuation-corrected with filtered back-projection using a Hanning filter. Three nuclear medicine physicians interpreted. Visual analysis; considered positive if axillary uptake higher than surrounding tissue	100% ALND	ALND (three levels). Lymph nodes formalin fixed, paraffin embedded, sectioned, stained with H&E
Noh 1998[93]	PET only	Injection of 370 MBq (10 mCi) FDG. Scan 60 minutes later. Whole-body emission scan and regional transmission scan for attenuation correction. SUV measured. No details of who analysed or what was considered positive	Histology (no further details)	No details reported

DOI: 10.3310/hta15040

Study	PET or PET/CT	PET details	Reference standard (summary)	Reference standard (detail)
Smith 1998[94]	PET only	Six-hour fast, injection of 185 MBq of FDG, contralateral arm, scan 40 minutes later. Prone position, arms by side. Imaging at two bed positions. Two experienced independent observers who reached consensus. Visual analysis; considered positive if abnormal uptake in axilla	45/50 (90%) ALND; 5/50 (10%) FNAC (large/locally advanced)	4-NS or level III ALND nodes embedded in paraffin blocks, sections of each node cut at two-to-four levels, examined by a single pathologist using standard preparation and histochemical techniques
Adler 1997[95]	PET only	Four-hour fast, injection of 740 MBq (20 mCi) FDG (note high dose of FDG). Supine position, arms at sides. One- or two-level transmission scans of breast and axilla before FDG injection. One- or two-level emission scans 1 hour after FDG injection. Reconstructed using filtered back projection with Hann filters. Interpreted by two independent observers, discrepancies resolved by consensus. Visual analysis for increased FDG uptake compared with background. Within this review, considered positive if probably or definitely abnormal, i.e. score of 4–5 on 1–5 scale	100% ALND	*ALND*: at least level II. Formalin fixed, paraffin embedded, single 3 µm planes, stained with H&E
Avril 1996[96]	PET only	Four-hour fast, injection of 270–390 MBq FDG, scan 60–80 minutes later. Prone position, arms at sides. Whole-body PET scanner. Emission scans, then transmission scans for attenuation correction. Attenuation-corrected images reconstructed using filtered back projection with Hanning filter. Normalised for dose of FDG and body weight. Interpreted by two nuclear medicine physicians, consensus reached. Visual analysis; considered positive if focally increased FDG uptake in axilla	37/41 (90%) ALND; 4/41 (10%, locally advanced) no reference standard; assumed positive due to clinical evidence	ALND (levels I and II and possibly III)
Utech 1996[97]	PET only	At least 4 hours fast, serum glucose monitored. 10 mCi FDG given intravenously 1 hour before emission scan. Whole-body scan, four bed positions. Transmission scan before FDG injection. Emission scan 1 hour after FDG injection. Images reconstructed via filtered back projection. Three experienced nuclear radiologists reviewed scan, and final reading by experienced nuclear medicine physician. Visual analysis; considered positive if discrete focal uptake in axilla greater than background	55/124 (44%) ALND; 69/124 (56%) modified radical mastectomy	ALND level not reported

ADD, additional non-sentinal lymph node biopsy; i.v., intravenous; mCi, millicurie; MBq, megabecquerel; SUV, standardised uptake value.

Magnetic resonance imaging studies

Study	Type of MRI	MRI details	Reference standard (summary)	Reference standard (detail)
Kimura 2009[98]	USPIO-enhanced	USPIO administered intravenously 24-hours prior to MRI through 5 µm microfilter (2.6 mg/kg iron diluted in 100 ml of 0.9% saline solution) by drip infusion for 90 minutes. MRI on 1.5T MRI scanner. 3-inch surface coil placed over SLN area identified by RI lymphoscintigraphic images. Arms elevated, prone position. T1W axial and coronal MR images and T2*W axial and coronal MR images obtained. Other technical details given. Images evaluated by one radiologist and one surgeon; reached consensus. Nodes classified as (A) low signal intensity on T2*W images, (B1) partial high signal intensity on T2*W, isointense to the fatty tissue on T1W images, (B2) partial high signal intensity on T2*W images, not identified as fat on T1W images and (C) high signal intensity on T2*W images. Types C and B2 were classified as metastatic, therefore criteria for positivity was USPIO uptake	ALND and/or SLNB	*SLNB*: blue dye and radioactive tracer-guided procedure in all patients. SLNs bisected or sectioned into 2 mm slices and embedded. H&E staining used. Immunohistochemical-keratin staining if H&E negative. *ALND*: if SLNB positive. Details on level not given. Non SLN H&E only
Harada 2007[99]	USPIO-enhanced	Images both pre-contrast and 24–36 hours post-administration of USPIO, which was given as infusion over 90 minutes (2.6 mg/kg iron diluted in 100 ml 0.9% saline). Transverse and coronal T1-weighted and T2-weighted images, patients in supine position, cardiac surface coil used. (Other technical details of MRI given.) Analysed independently by one radiologist and one surgeon; consensus via discussion. Various alternative criteria were used to define a positive result, as follows. The main analysis (giving the highest sensitivity and specificity) used: (a) contrast uptake on USPIO-enhanced images (homogenous USPIO intake = benign; heterogeneous or no uptake = metastatic); (b) comparison of contrast uptake on pre-contrast and USPIO-enhanced images (homogenous USPIO intake = benign; heterogeneous or no uptake = metastatic); (c) size of short axis > 5 mm (pre-contrast images); (d) size of short axis > 10 mm (pre-contrast images); (e) abnormal morphology (replacement of central fatty hilum and/or irregular margins) (pre-contrast images)	100% ALND	*ALND (levels I and II)*: nodes formalin fixed, sectioned, stained with H&E and studied under microscope
Memarsadeghi 2006[100]	USPIO-enhanced	Images both pre-contrast, and 24–36 hours post-administration of USPIO, which was given as infusion (2.6 mg/kg iron diluted in 100 ml saline at rate of 4 ml/minute). T1-weighted images (in transverse plane) and T2-weighted images (in transverse and sagittal planes). Patients in supine position, arms by sides and then elevated (sagittal plane arms elevated only), surface coil used. (Other technical details of MRI given.) Pre-contrast and USPIO-enhanced images interpreted independently by two experienced radiologists (consensus via discussion, third radiologist if consensus could not be reached). Criteria for positive node: (a) for non-enhanced images, (i) *size* of short axis > 10 mm, and/or (ii) *shape* round rather than oval (ratio of long-to-short axis < 1.5); and (b) for USPIO-enhanced images, (i) visual analysis comparing contrast uptake on enhanced and non-enhanced images (homogenous decrease in uptake on enhanced vs non-enhanced image = benign; no decrease or partial decrease in uptake or uniformly high uptake on enhanced image = metastatic) and/or (ii) quantitative analysis showing signal-to-noise ratio decreased by < 30% on enhanced vs non-enhanced image	100% ALND	*ALND (levels I and II)*: histopathological analysis, regarded as metastatic if tumour cells present at light microscopy

DOI: 10.3310/hta15040

Study	Type of MRI	MRI details	Reference standard (summary)	Reference standard (detail)
Stadnik 2006[101]	USPIO-enhanced (also used gadolinium for one of the MRI sequences)	USPIO on slow-drip infusion for 24–36 hours prior to imaging (2.6 mg/kg iron diluted in 100 ml 0.9% saline), 3 ml/minute. 1.5T magnet, surface coil on axilla, patient supine. Sequences in transverse plane (section thickness 3 mm for all): T2-weighted 2D spin-echo images; T1-weighted 3D gradient echo; T2*-weighted 2D gradient echo. After injection of gadolinium (0.1 mmol/kg Gd), T1-weighted 3D gradient echo with spectral fat saturation. Images analysed in consensus by two experienced radiologists. Criteria for positivity based on pattern of nodal enhancement after USPIO administration (considered benign if normal morphology with uniform or central signal drop; considered metastatic if normal morphology or focal or global volume increase, with or without partial signal drop). Stated that USPIO-enhanced T1 gradient echo after gadolinium injection and fat saturation was a useful method	100% ALND	Nodes considered positive if tumour cells present at light microscopy
Michel 2002[102]	USPIO-enhanced	USPIO on slow drip infusion for 24–36 hours prior to imaging of axilla (2.6 mg/kg iron body weight diluted in 100 ml 0.9% saline soln). 2 ml/minute for first 10 minutes, to test tolerance, then 4 ml/minute. Infusion for 30 minutes overall. Technical details of MRI given. MRI of the primary breast tumour also performed using two contrast agents (USPIO and 0.1 mmol/kg Gd), but only USPIO used for imaging axilla. Criteria for metastatic node: size (short-axis diameter >10 mm), shape (round rather than oval, i.e. ratio of long-to-short axes <1.5), USPIO uptake (heterogeneous uptake or lack of uptake = malignant; homogenous uptake = benign; rim enhancement could be classified as malignant if on round node with other imaging properties). Not clear whether nodes were considered positive if they met just one of these criteria or some/all these criteria. Images interpreted by radiologist and an experienced resident	100% ALND	Block resection of levels I and II of the ipsilateral side, though later mentions level III clearance in results section. Nodes positive if cancer cells positive at light microscopy, independent of results of immunohistochemical staining
Murray 2002[103]	Dynamic gadolinium-enhanced	Contrast agent: Gd injected at 0.1 mmol/kg. Anatomical imaging prior to Gd enhanced imaging at oblique sagittal plane to optimise axillary imaging. More detail about how to interpret images given. Uses technique of comparing node uptake to uptake in adjacent fat tissue. Single experienced analyst interpreted images. Criteria for metastatic node: (a) size (cross-sectional area >4 mm^2) AND signal intensity change $>21\%$ (compared with adjacent fat tissue); (b) signal intensity change only; and (c) size only	100% ALND	Level II, bisection of smaller nodes, 2 mm macrosection of lager nodes, H&E staining, then embedded in paraffin, and 5 µm slices mounted on slides. In women with primary lobular carcinoma, immunocytochemical staining for cytokeratin with CAM5.2 also performed

DOI: 10.3310/hta15040

Study	Type of MRI	MRI details	Reference standard (summary)	Reference standard (detail)
Kvistad 2000[41]	Dynamic gadolinium-enhanced	Prone position, coil covering breast and axilla. Pre-contrast axial T1-weighted spin-echo images obtained. Then, injection of 0.1 mmol/kg gadodiamide followed by 20 ml isotonic saline. Following contrast injection, six T1-weighted dynamic contrast-enhanced 3D images acquired. Both pre-contrast images and dynamic contrast-enhanced images analysed independently by two radiologists (discrepancies resolved by consensus). Time-vs-signal intensity curves obtained. Various alternative criteria were used to define a positive result, as follows. The main analysis (giving the highest sensitivity and specificity) used: (a) positive if signal intensity increased by >100% during the first post-contrast image (after contrast injection) compared with the pre-contrast image; (b) decrease in signal intensity between second and sixth post-contrast images (known as 'positive wash-out sign' and previously shown to be sign of malignancy); (c) size of short axis >5 mm; and (d) abnormal morphology (replacement of central fatty hilus and/or irregular margins)	100% ALND	*ALND (at least levels I and II)*: nodes paraffin embedded, sectioned and stained with H&E
Mumtaz 1997[42]	Gadolinium-enhanced	Enhancement with Gd 0.1 mmol/kg. Extensive technical details given. Criteria for metastatic node: size >5 mm, higher than soft-tissue intensity and enhancement with Gd. Not clear whether nodes were considered positive if they met just one of these criteria or some/all these criteria. Assessed by single radiologist	100% ALND	None given
Yeung 2002[64]	In vivo proton MR spectroscopy	MRI performed first to identify enhancing lesions, using Gd 0.2 mmol/kg body weight. Then areas of interest (including the largest visible axillary lymph node) imaged by in vivo proton MR spectroscopy. Point-resolved spectroscopic (PRESS) sequence used. Three water-suppressed spectra acquired for each volume of interest. Spectra analysed for choline level; signal-to-noise ratio ≥2 indicated that choline present	100% ALND	Few technical details given

Gd, gadopentetate dimeglumine; RI, radioisotope; SLN, sentinel lymph node.

Health Technology Assessment programme

HTA Commissioning Board *(continued)*

Professor James Raftery,
Chair of NETSCC and Director of the Wessex Institute, University of Southampton

Professor Barney Reeves,
Professorial Research Fellow in Health Services Research, Department of Clinical Science, University of Bristol

Professor Martin Underwood,
Warwick Medical School, University of Warwick

Professor Marion Walker,
Professor in Stroke Rehabilitation, Associate Director UK Stroke Research Network, University of Nottingham

Dr Duncan Young,
Senior Clinical Lecturer and Consultant, Nuffield Department of Anaesthetics, University of Oxford

Observers

Dr Morven Roberts,
Clinical Trials Manager, Health Services and Public Health Services Board, Medical Research Council

HTA Clinical Evaluation and Trials Board

Chair,
Professor Sallie Lamb,
Director,
Warwick Clinical Trials Unit,
Warwick Medical School,
University of Warwick and Professor of Rehabilitation,
Nuffield Department of Orthopaedic, Rheumatology and Musculoskeletal Sciences,
University of Oxford

Deputy Chair,
Professor Jenny Hewison,
Professor of the Psychology of Health Care,
Leeds Institute of Health Sciences,
University of Leeds

Programme Director,
Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Members

Professor Keith Abrams,
Professor of Medical Statistics, Department of Health Sciences, University of Leicester

Professor Martin Bland,
Professor of Health Statistics, Department of Health Sciences, University of York

Professor Jane Blazeby,
Professor of Surgery and Consultant Upper GI Surgeon, Department of Social Medicine, University of Bristol

Professor Julia M Brown,
Director, Clinical Trials Research Unit, University of Leeds

Professor Alistair Burns,
Professor of Old Age Psychiatry, Psychiatry Research Group, School of Community-Based Medicine, The University of Manchester & National Clinical Director for Dementia, Department of Health

Dr Jennifer Burr,
Director, Centre for Healthcare Randomised trials (CHART), University of Aberdeen

Professor Linda Davies,
Professor of Health Economics, Health Sciences Research Group, University of Manchester

Professor Simon Gilbody,
Prof of Psych Medicine and Health Services Research, Department of Health Sciences, University of York

Professor Steven Goodacre,
Professor and Consultant in Emergency Medicine, School of Health and Related Research, University of Sheffield

Professor Dyfrig Hughes,
Professor of Pharmacoeconomics, Centre for Economics and Policy in Health, Institute of Medical and Social Care Research, Bangor University

Professor Paul Jones,
Professor of Respiratory Medicine, Department of Cardiac and Vascular Science, St George's Hospital Medical School, University of London

Professor Khalid Khan,
Professor of Women's Health and Clinical Epidemiology, Barts and the London School of Medicine, Queen Mary, University of London

Professor Richard J McManus,
Professor of Primary Care Cardiovascular Research, Primary Care Clinical Sciences Building, University of Birmingham

Professor Helen Rodgers,
Professor of Stroke Care, Institute for Ageing and Health, Newcastle University

Professor Ken Stein,
Professor of Public Health, Peninsula Technology Assessment Group, Peninsula College of Medicine and Dentistry, Universities of Exeter and Plymouth

Professor Jonathan Sterne,
Professor of Medical Statistics and Epidemiology, Department of Social Medicine, University of Bristol

Mr Andy Vail,
Senior Lecturer, Health Sciences Research Group, University of Manchester

Professor Clare Wilkinson,
Professor of General Practice and Director of Research North Wales Clinical School, Department of Primary Care and Public Health, Cardiff University

Dr Ian B Wilkinson,
Senior Lecturer and Honorary Consultant, Clinical Pharmacology Unit, Department of Medicine, University of Cambridge

Observers

Ms Kate Law,
Director of Clinical Trials,
Cancer Research UK

Dr Morven Roberts,
Clinical Trials Manager, Health Services and Public Health Services Board, Medical Research Council

Diagnostic Technologies and Screening Panel

Members

Chair,
Professor Lindsay Wilson Turnbull,
Scientific Director of the Centre for Magnetic Resonance Investigations and YCR Professor of Radiology, Hull Royal Infirmary

Professor Judith E Adams,
Consultant Radiologist, Manchester Royal Infirmary, Central Manchester & Manchester Children's University Hospitals NHS Trust, and Professor of Diagnostic Radiology, University of Manchester

Mr Angus S Arunkalaivanan,
Honorary Senior Lecturer, University of Birmingham and Consultant Urogynaecologist and Obstetrician, City Hospital, Birmingham

Dr Stephanie Dancer,
Consultant Microbiologist, Hairmyres Hospital, East Kilbride

Dr Diane Eccles,
Professor of Cancer Genetics, Wessex Clinical Genetics Service, Princess Anne Hospital

Dr Trevor Friedman,
Consultant Liason Psychiatrist, Brandon Unit, Leicester General Hospital

Dr Ron Gray,
Consultant, National Perinatal Epidemiology Unit, Institute of Health Sciences, University of Oxford

Professor Paul D Griffiths,
Professor of Radiology, Academic Unit of Radiology, University of Sheffield

Mr Martin Hooper,
Service User Representative

Professor Anthony Robert Kendrick,
Associate Dean for Clinical Research and Professor of Primary Medical Care, University of Southampton

Dr Anne Mackie,
Director of Programmes, UK National Screening Committee, London

Mr David Mathew,
Service User Representative

Dr Michael Millar,
Consultant Senior Lecturer in Microbiology, Department of Pathology & Microbiology, Barts and The London NHS Trust, Royal London Hospital

Mrs Una Rennard,
Service User Representative

Dr Stuart Smellie,
Consultant in Clinical Pathology, Bishop Auckland General Hospital

Ms Jane Smith,
Consultant Ultrasound Practitioner, Leeds Teaching Hospital NHS Trust, Leeds

Dr Allison Streetly,
Programme Director, NHS Sickle Cell and Thalassaemia Screening Programme, King's College School of Medicine

Dr Alan J Williams,
Consultant Physician, General and Respiratory Medicine, The Royal Bournemouth Hospital

Observers

Dr Tim Elliott,
Team Leader, Cancer Screening, Department of Health

Dr Catherine Moody,
Programme Manager, Medical Research Council

Professor Julietta Patrick,
Director, NHS Cancer Screening Programme, Sheffield

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy Research Programme, Department of Health

Disease Prevention Panel

Members

Chair,
Professor Margaret Thorogood,
Professor of Epidemiology, University of Warwick Medical School, Coventry

Dr Robert Cook,
Clinical Programmes Director, Bazian Ltd, London

Dr Colin Greaves,
Senior Research Fellow, Peninsula Medical School (Primary Care)

Mr Michael Head,
Service User Representative

Professor Cathy Jackson,
Professor of Primary Care Medicine, Bute Medical School, University of St Andrews

Dr Russell Jago,
Senior Lecturer in Exercise, Nutrition and Health, Centre for Sport, Exercise and Health, University of Bristol

Dr Julie Mytton,
Consultant in Child Public Health, NHS Bristol

Professor Irwin Nazareth,
Professor of Primary Care and Director, Department of Primary Care and Population Sciences, University College London

Dr Richard Richards,
Assistant Director of Public Health, Derbyshire Country Primary Care Trust

Professor Ian Roberts,
Professor of Epidemiology and Public Health, London School of Hygiene & Tropical Medicine

Dr Kenneth Robertson,
Consultant Paediatrician, Royal Hospital for Sick Children, Glasgow

Dr Catherine Swann,
Associate Director, Centre for Public Health Excellence, NICE

Professor Carol Tannahill,
Glasgow Centre for Population Health

Mrs Jean Thurston,
Service User Representative

Professor David Weller,
Head, School of Clinical Science and Community Health, University of Edinburgh

Observers

Ms Christine McGuire,
Research & Development, Department of Health

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

External Devices and Physical Therapies Panel

Members

Chair,
Dr John Pounsford,
Consultant Physician North Bristol NHS Trust

Deputy Chair,
Professor E Andrea Nelson,
Reader in Wound Healing and Director of Research, University of Leeds

Professor Bipin Bhakta,
Charterhouse Professor in Rehabilitation Medicine, University of Leeds

Mrs Penny Calder,
Service User Representative

Dr Dawn Carnes,
Senior Research Fellow, Barts and the London School of Medicine and Dentistry

Dr Emma Clark,
Clinician Scientist Fellow & Cons. Rheumatologist, University of Bristol

Mrs Anthea De Barton-Watson,
Service User Representative

Professor Nadine Foster,
Professor of Musculoskeletal Health in Primary Care Arthritis Research, Keele University

Dr Shaheen Hamdy,
Clinical Senior Lecturer and Consultant Physician, University of Manchester

Professor Christine Norton,
Professor of Clinical Nursing Innovation, Bucks New University and Imperial College Healthcare NHS Trust

Dr Lorraine Pinnigton,
Associate Professor in Rehabilitation, University of Nottingham

Dr Kate Radford,
Senior Lecturer (Research), University of Central Lancashire

Mr Jim Reece,
Service User Representative

Professor Maria Stokes,
Professor of Neuromusculoskeletal Rehabilitation, University of Southampton

Dr Pippa Tyrrell,
Senior Lecturer/Consultant, Salford Royal Foundation Hospitals' Trust and University of Manchester

Dr Sarah Tyson,
Senior Research Fellow & Associate Head of School, University of Salford

Dr Nefyn Williams,
Clinical Senior Lecturer, Cardiff University

Observers

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy Research Programme, Department of Health

Interventional Procedures Panel

Members

Chair,
Professor Jonathan Michaels,
Professor of Vascular Surgery, University of Sheffield

Deputy Chair,
Mr Michael Thomas,
Consultant Colorectal Surgeon, Bristol Royal Infirmary

Mrs Isabel Boyer,
Service User Representative

Mr David P Britt,
Service User Representative

Mr Sankaran Chandra Sekharan,
Consultant Surgeon, Breast Surgery, Colchester Hospital University NHS Foundation Trust

Professor Nicholas Clarke,
Consultant Orthopaedic Surgeon, Southampton University Hospitals NHS Trust

Ms Leonie Cooke,
Service User Representative

Mr Seumas Eckford,
Consultant in Obstetrics & Gynaecology, North Devon District Hospital

Professor Sam Eljamel,
Consultant Neurosurgeon, Ninewells Hospital and Medical School, Dundee

Dr Adele Fielding,
Senior Lecturer and Honorary Consultant in Haematology, University College London Medical School

Dr Matthew Hatton,
Consultant in Clinical Oncology, Sheffield Teaching Hospital Foundation Trust

Dr John Holden,
General Practitioner, Garswood Surgery, Wigan

Professor Nicholas James,
Professor of Clinical Oncology, School of Cancer Sciences, University of Birmingham

Dr Fiona Lecky,
Senior Lecturer/Honorary Consultant in Emergency Medicine, University of Manchester/Salford Royal Hospitals NHS Foundation Trust

Dr Nadim Malik,
Consultant Cardiologist/Honorary Lecturer, University of Manchester

Mr Hisham Mehanna,
Consultant & Honorary Associate Professor, University Hospitals Coventry & Warwickshire NHS Trust

Dr Jane Montgomery,
Consultant in Anaesthetics and Critical Care, South Devon Healthcare NHS Foundation Trust

Professor Jon Moss,
Consultant Interventional Radiologist, North Glasgow Hospitals University NHS Trust

Dr Simon Padley,
Consultant Radiologist, Chelsea & Westminster Hospital

Dr Ashish Paul,
Medical Director, Bedfordshire PCT

Dr Sarah Purdy,
Consultant Senior Lecturer, University of Bristol

Professor Yit Chiun Yang,
Consultant Ophthalmologist, Royal Wolverhampton Hospitals NHS Trust

Observers

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Dr Morven Roberts,
Clinical Trials Manager, Health Services and Public Health Services Board, Medical Research Council

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy Research Programme, Department of Health

Pharmaceuticals Panel

Members

Chair,
Professor Imti Choonara,
Professor in Child Health, University of Nottingham

Deputy Chair,
Dr Yoon K Loke,
Senior Lecturer in Clinical Pharmacology, University of East Anglia

Dr Martin Ashton-Key,
Medical Advisor, National Commissioning Group, NHS London

Mr John Chapman,
Service User Representative

Dr Peter Elton,
Director of Public Health, Bury Primary Care Trust

Dr Ben Goldacre,
Research Fellow, Division of Psychological Medicine and Psychiatry, King's College London

Dr James Gray,
Consultant Microbiologist, Department of Microbiology, Birmingham Children's Hospital NHS Foundation Trust

Ms Kylie Gyertson,
Oncology and Haematology Clinical Trials Manager, Guy's and St Thomas' NHS Foundation Trust London

Dr Jurjees Hasan,
Consultant in Medical Oncology, The Christie, Manchester

Dr Carl Heneghan,
Deputy Director Centre for Evidence-Based Medicine and Clinical Lecturer, Department of Primary Health Care, University of Oxford

Dr Dyfrig Hughes,
Reader in Pharmacoeconomics and Deputy Director, Centre for Economics and Policy in Health, IMSCaR, Bangor University

Dr Maria Kouimtzi,
Pharmacy and Informatics Director, Global Clinical Solutions, Wiley-Blackwell

Professor Femi Oyebode,
Consultant Psychiatrist and Head of Department, University of Birmingham

Dr Andrew Prentice,
Senior Lecturer and Consultant Obstetrician and Gynaecologist, The Rosie Hospital, University of Cambridge

Ms Amanda Roberts,
Service User Representative

Dr Martin Shelly,
General Practitioner, Silver Lane Surgery, Leeds

Dr Gillian Shepherd,
Director, Health and Clinical Excellence, Merck Serono Ltd

Mrs Katrina Simister,
Assistant Director New Medicines, National Prescribing Centre, Liverpool

Professor Donald Singer,
Professor of Clinical Pharmacology and Therapeutics, Clinical Sciences Research Institute, CSB, University of Warwick Medical School

Mr David Symes,
Service User Representative

Dr Arnold Zermansky,
General Practitioner, Senior Research Fellow, Pharmacy Practice and Medicines Management Group, Leeds University

Observers

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Mr Simon Reeve,
Head of Clinical and Cost-Effectiveness, Medicines, Pharmacy and Industry Group, Department of Health

Dr Heike Weber,
Programme Manager, Medical Research Council

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy Research Programme, Department of Health

Psychological and Community Therapies Panel

Members

Chair,
Professor Scott Weich,
Professor of Psychiatry, University of Warwick, Coventry

Deputy Chair,
Dr Howard Ring,
Consultant & University Lecturer in Psychiatry, University of Cambridge

Professor Jane Barlow,
Professor of Public Health in the Early Years, Health Sciences Research Institute, Warwick Medical School

Dr Sabyasachi Bhaumik,
Consultant Psychiatrist, Leicestershire Partnership NHS Trust

Mrs Val Carlill,
Service User Representative

Dr Steve Cunningham,
Consultant Respiratory Paediatrician, Lothian Health Board

Dr Anne Hesketh,
Senior Clinical Lecturer in Speech and Language Therapy, University of Manchester

Dr Peter Langdon,
Senior Clinical Lecturer, School of Medicine, Health Policy and Practice, University of East Anglia

Dr Yann Lefeuvre,
GP Partner, Burrage Road Surgery, London

Dr Jeremy J Murphy,
Consultant Physician and Cardiologist, County Durham and Darlington Foundation Trust

Dr Richard Neal,
Clinical Senior Lecturer in General Practice, Cardiff University

Mr John Needham,
Service User Representative

Ms Mary Nettle,
Mental Health User Consultant

Professor John Potter,
Professor of Ageing and Stroke Medicine, University of East Anglia

Dr Greta Rait,
Senior Clinical Lecturer and General Practitioner, University College London

Dr Paul Ramchandani,
Senior Research Fellow/Cons. Child Psychiatrist, University of Oxford

Dr Karen Roberts,
Nurse/Consultant, Dunston Hill Hospital, Tyne and Wear

Dr Karim Saad,
Consultant in Old Age Psychiatry, Coventry and Warwickshire Partnership Trust

Dr Lesley Stockton,
Lecturer, School of Health Sciences, University of Liverpool

Dr Simon Wright,
GP Partner, Walkden Medical Centre, Manchester

Observers

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Dr Morven Roberts,
Clinical Trials Manager, Health Services and Public Health Services Board, Medical Research Council

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy Research Programme, Department of Health

Expert Advisory Network

Members

Professor Douglas Altman,
Professor of Statistics in Medicine, Centre for Statistics in Medicine, University of Oxford

Professor John Bond,
Professor of Social Gerontology & Health Services Research, University of Newcastle upon Tyne

Professor Andrew Bradbury,
Professor of Vascular Surgery, Solihull Hospital, Birmingham

Mr Shaun Brogan,
Chief Executive, Ridgeway Primary Care Group, Aylesbury

Mrs Stella Burnside OBE,
Chief Executive, Regulation and Improvement Authority, Belfast

Ms Tracy Bury,
Project Manager, World Confederation of Physical Therapy, London

Professor Iain T Cameron,
Professor of Obstetrics and Gynaecology and Head of the School of Medicine, University of Southampton

Professor Bruce Campbell,
Consultant Vascular & General Surgeon, Royal Devon & Exeter Hospital, Wonford

Dr Christine Clark,
Medical Writer and Consultant Pharmacist, Rossendale

Professor Collette Clifford,
Professor of Nursing and Head of Research, The Medical School, University of Birmingham

Professor Barry Cookson,
Director, Laboratory of Hospital Infection, Public Health Laboratory Service, London

Dr Carl Counsell,
Clinical Senior Lecturer in Neurology, University of Aberdeen

Professor Howard Cuckle,
Professor of Reproductive Epidemiology, Department of Paediatrics, Obstetrics & Gynaecology, University of Leeds

Professor Carol Dezateux,
Professor of Paediatric Epidemiology, Institute of Child Health, London

Mr John Dunning,
Consultant Cardiothoracic Surgeon, Papworth Hospital NHS Trust, Cambridge

Mr Jonothan Earnshaw,
Consultant Vascular Surgeon, Gloucestershire Royal Hospital, Gloucester

Professor Martin Eccles,
Professor of Clinical Effectiveness, Centre for Health Services Research, University of Newcastle upon Tyne

Professor Pam Enderby,
Dean of Faculty of Medicine, Institute of General Practice and Primary Care, University of Sheffield

Professor Gene Feder,
Professor of Primary Care Research & Development, Centre for Health Sciences, Barts and The London School of Medicine and Dentistry

Mr Leonard R Fenwick,
Chief Executive, Freeman Hospital, Newcastle upon Tyne

Mrs Gillian Fletcher,
Antenatal Teacher and Tutor and President, National Childbirth Trust, Henfield

Professor Jayne Franklyn,
Professor of Medicine, University of Birmingham

Mr Tam Fry,
Honorary Chairman, Child Growth Foundation, London

Professor Fiona Gilbert,
Consultant Radiologist and NCRN Member, University of Aberdeen

Professor Paul Gregg,
Professor of Orthopaedic Surgical Science, South Tees Hospital NHS Trust

Bec Hanley,
Co-director, TwoCan Associates, West Sussex

Dr Maryann L Hardy,
Senior Lecturer, University of Bradford

Mrs Sharon Hart,
Healthcare Management Consultant, Reading

Professor Robert E Hawkins,
CRC Professor and Director of Medical Oncology, Christie CRC Research Centre, Christie Hospital NHS Trust, Manchester

Professor Richard Hobbs,
Head of Department of Primary Care & General Practice, University of Birmingham

Professor Alan Horwich,
Dean and Section Chairman, The Institute of Cancer Research, London

Professor Allen Hutchinson,
Director of Public Health and Deputy Dean of ScHARR, University of Sheffield

Professor Peter Jones,
Professor of Psychiatry, University of Cambridge, Cambridge

Professor Stan Kaye,
Cancer Research UK Professor of Medical Oncology, Royal Marsden Hospital and Institute of Cancer Research, Surrey

Dr Duncan Keeley,
General Practitioner (Dr Burch & Ptnrs), The Health Centre, Thame

Dr Donna Lamping,
Research Degrees Programme Director and Reader in Psychology, Health Services Research Unit, London School of Hygiene and Tropical Medicine, London

Professor James Lindesay,
Professor of Psychiatry for the Elderly, University of Leicester

Professor Julian Little,
Professor of Human Genome Epidemiology, University of Ottawa

Professor Alistaire McGuire,
Professor of Health Economics, London School of Economics

Professor Neill McIntosh,
Edward Clark Professor of Child Life and Health, University of Edinburgh

Professor Rajan Madhok,
Consultant in Public Health, South Manchester Primary Care Trust

Professor Sir Alexander Markham,
Director, Molecular Medicine Unit, St James's University Hospital, Leeds

Dr Peter Moore,
Freelance Science Writer, Ashtead

Dr Andrew Mortimore,
Public Health Director, Southampton City Primary Care Trust

Dr Sue Moss,
Associate Director, Cancer Screening Evaluation Unit, Institute of Cancer Research, Sutton

Professor Miranda Mugford,
Professor of Health Economics and Group Co-ordinator, University of East Anglia

Professor Jim Neilson,
Head of School of Reproductive & Developmental Medicine and Professor of Obstetrics and Gynaecology, University of Liverpool

Mrs Julietta Patnick,
Director, NHS Cancer Screening Programmes, Sheffield

Professor Robert Peveler,
Professor of Liaison Psychiatry, Royal South Hants Hospital, Southampton

Professor Chris Price,
Director of Clinical Research, Bayer Diagnostics Europe, Stoke Poges

Professor William Rosenberg,
Professor of Hepatology and Consultant Physician, University of Southampton

Professor Peter Sandercock,
Professor of Medical Neurology, Department of Clinical Neurosciences, University of Edinburgh

Dr Philip Shackley,
Senior Lecturer in Health Economics, Sheffield Vascular Institute, University of Sheffield

Dr Eamonn Sheridan,
Consultant in Clinical Genetics, St James's University Hospital, Leeds

Dr Margaret Somerville,
Director of Public Health Learning, Peninsula Medical School, University of Plymouth

Professor Sarah Stewart-Brown,
Professor of Public Health, Division of Health in the Community, University of Warwick, Coventry

Dr Nick Summerton,
GP Appraiser and Codirector, Research Network, Yorkshire Clinical Consultant, Primary Care and Public Health, University of Oxford

Professor Ala Szczepura,
Professor of Health Service Research, Centre for Health Services Studies, University of Warwick, Coventry

Dr Ross Taylor,
Senior Lecturer, University of Aberdeen

Dr Richard Tiner,
Medical Director, Medical Department, Association of the British Pharmaceutical Industry

Mrs Joan Webster,
Consumer Member, Southern Derbyshire Community Health Council

Professor Martin Whittle,
Clinical Co-director, National Co-ordinating Centre for Women's and Children's Health, Lymington

Current and past membership details of all HTA programme 'committees' are available from the HTA website (www.hta.ac.uk)

Feedback

The HTA programme and the authors would like to know your views about this report.

The Correspondence Page on the HTA website (www.hta.ac.uk) is a convenient way to publish your comments. If you prefer, you can send your comments to the address below, telling us whether you would like us to transfer them to the website.

We look forward to hearing from you.

NETSCC, Health Technology Assessment
Alpha House
University of Southampton Science Park
Southampton SO16 7NS, UK
Email: hta@hta.ac.uk
www.hta.ac.uk

ISSN 1366-5278