A systematic review and economic evaluation of intraoperative tests (RD-100i OSNA system and Metasin test) for detecting sentinel lymph node metastases in breast cancer

Commissioned by **NIHR HTA Programme**
Project Number: 12/14/01
On behalf of National Institute of Health and Clinical Excellence
Produced by: Peninsula Technology Assessment Group (PenTAG), University of Exeter

Authors: **Nicola Huxley**, Associate Research Fellow in Modelling, PenTAG
Tracey Jones-Hughes, Research Fellow HTA, PenTAG
Helen Coelho, Research Fellow HTA, PenTAG
Tristan Snowsill, Associate Research Fellow in Modelling, PenTAG
Chris Cooper, Senior Information Specialist, PenTAG
Katie Cooper, Research Fellow in Health economics, ScHARR
Yang Meng, Research Fellow in Health economic modelling, ScHARR
Chris Hyde, Professor of Public Health, PenTAG
Ruben Mujica-Mota, Senior Lecturer in Health Economics, PenTAG

Correspondence to **Ruben Mujica-Mota**
Peninsula Technology Assessment Group (PenTAG)
University of Exeter Medical School
Veysey Building
Salmon Pool Lane
Exeter EX2 4SG
email: ruben.mujica-mota@pcmd.ac.uk

Date completed

Declared competing interests of authors: None

Description of any pecuniary relationship with sponsors, both personal and of the TAR centre. If there are none, please state 'none': None

Acknowledgements

We would like to thank:

We would also like to acknowledge the help of Sue Whiffin and Jenny Lowe for their administrative support throughout the project.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA programme. Any errors are the responsibility of the authors.

This report should be referenced as follows: Huxley N, Jones-Hughes T, Coelho H, Snowsill T, Cooper C, Meng Y, Cooper K, Hyde C, Mujica-Mota R. A systematic review and economic evaluation of intraoperative tests (RD-100i OSNA system and Metasin test) for detecting sentinel lymph node metastases in breast cancer. (2012) University of Exeter (Report).

Please refer to the International Committee of Medical Journal Editors (ICMJE) Uniform Requirements for manuscripts submitted to biomedical journals, see http://icmje.org/

Contributions of authors	
Nicola Huxley	Developed the short term model and adapted the long term model, executed the economic model and wrote the sections on the design and results of the economic model
Yang Meng	Co-authored the original long term economic model , advised on its adaptation to the analysis, and contributed to writing the report
Katie Cooper	Co-authored the original long term model, and contributed to writing the report
Ruben Mujica-Mota	Contributed to developing the protocol. Led the systematic reviews of economic evaluations, contributed to the design of the analysis and writing and editing of the report. Overall lead for the project and final report.
Tracey Jones-Hughes	Assessed abstracts and titles for inclusion, led the systematic review of clinical effectiveness, and contributed to the writing and editing of the report
Helen Coelho	Assessed abstracts and titles for inclusion, and contributed to the writing and editing of the report
Tristan Snowsill	Performed statistical analysis and contributed to the writing of the report
Chris Cooper	Designed and carried out literature searches for the systematic reviews and identification of model parameters, and contributed to

	the writing and editing of the report
Chris Hyde	Developed the protocol. Contributed to the systematic review, design of the model, and to the writing and editing of the report. Director of TAR group at PenTAG and Guarantor of the report

Contents

List of tables

List of figures

List of abbreviations

AE	adverse events
AJCC	American Joint Committee on Cancer
ALN	axillary lymph nodes
ALND	axillary lymph node dissection
CIS	carcinoma-in-situ
CEA	cost-effectiveness analysis
CEAC	cost-effectiveness acceptability curves
CHEC	Consensus on Health Economic Criteria
CI	confidence interval
CRD	NHS Centre for Reviews and Dissemination
DCIS	ductal carcinoma-in-situ
DLY	discounted life year
DTA	diagnostic test accuracy
FNAC	fine needle aspiration cytology
FN	false negative
FP	false positive
HR	hazard Ratio
HRG	Healthcare Research Group
HRQL	health related quality of life
HTA	Health Technology Assessment
HUI	health utility index
ICER	incremental cost effectiveness ratio
ITC	Isolated tumour cells
LN	lymph node
MDT	Multi-disciplinary team
MTA	multiple technology assessment
N/A	not applicable
NCCN	National Comprehensive Cancer Network
NCCN FCSI	National Comprehensive Cancer Network FACT CRC Symptom Index
NCI-CTC	National Cancer Institute Common Terminology Criteria
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence
NIHR	National Institute for Health Research
NØ	no regional lymph node invasion
NR	not reported
OS	overall survival
OSNA	one step nucleic acid amplification
PenTAG	Peninsula Technology Assessment Group

QALY	quality-adjusted life year
QoL	quality of life
qRT-PCR	quantitative reverse transcriptase – polymerase chain reaction
RCT	randomised controlled trial
RECIST	Response Evaluation Criteria In Solid Tumors
SD	standard deviation
SE	standard error
SLN	sentinel lymph node
SLNB	sentinel lymph node biopsy
SROC	summary receiver operator characteristics
T0	no evidence of tumour
T1, T2, T3, T4	stage of cancer
T_{is}	tumour in situ
TAB	tissue allocation bias
TNM	tumour node metastases
TN	true negative
TP	true positive
UK	United Kingdom
US	ultrasound

Glossary

Axillary lymph node	Lymph nodes located in the axilla or armpit
Axillary lymph node dissection (ALND)	Removal of some or all lymph nodes within the axilla
Cost-effectiveness analysis	An economic analysis that converts effects into health terms and describes the costs for additional health gain.
Crossing point	The point during quantitative PCR where the level of fluorescence is above background
Decision modelling	A theoretical construct that allows the comparison of the relationship between costs and outcomes of alternative healthcare interventions.
Distant metastases	Cancer that has spread from the original (primary) tumor to distant organs or distant lymph nodes.
False negative	Incorrect negative test result – number of diseased persons with a negative test result.
False positive	Incorrect positive test result – number of non-diseased persons with a positive test result.
Incremental cost effectiveness ratio	The difference in the mean costs of two interventions in the population of interest divided by the difference in the mean outcomes in the population of interest.
Index test	The test whose performance is being evaluated.
Meta-analysis	Statistical techniques used to combine the results of two or more studies and obtain a combined estimate of effect.
Loco-regional metastases	Metastasis (spread) of a cancer only within the region in which it arose.
Lymph node	An organ of the immune system which filters foreign particles and bacteria from lymph fluid.
Metastatic disease	The spread of cancer from one organ or body part to another organ or body part
Polymerase chain reaction (PCR)	A technology used for amplifying DNA sequences
Primary tumour	A tumour growing at the anatomical site where tumour progression began
Quality adjusted life year (QALY)	A measure of health gain, used in economic evaluations, in which survival duration is weighted or adjusted by the patient's quality of life during the survival period.
Quantitative reverse transcriptase – polymerase chain reaction (qRT-PCR)	The cloning of genes by reverse transcribing the RNA of interest into its DNA complement through the use of reverse transcriptase. Subsequently, the newly synthesized cDNA is amplified using traditional PCR and may be quantitatively measured by using fluorescent probes
Receiver Operating Characteristic (ROC) curve	A graph which illustrates the trade-offs between sensitivity and specificity which result from varying the diagnostic threshold.

Reference standard	The best currently available diagnostic test, against which the index test is compared.
Regional metastases	The spread of cancer beyond the initial site to regional lymph nodes.
Reverse Transcription Loop Mediated Isothermal Amplification (RTLAMP).	A technique to amplify mRNA directly from tissue lysates
Sensitivity	Proportion of people with the target disorder who have a positive test result.
Sentinel lymph node	The first lymph node(s) to which cancer cells are likely to spread from a primary tumour
Specificity	Proportion of people without the target disorder who have a negative test result.
Summary Receiver Operating Characteristic (ROC) plot or curve	A diagram used to plot the results of included studies in systematic review of test accuracy studies. If a meta-analysis is performed it may include either a summary point or a summary curve or both
Tissue allocation bias (TAB)	Occurs when tumour deposits exist in different portions of tissue which are separately allocated to the index and reference test.
True negative	Correct negative test result – number of non-diseases persons with a negative test result.
True positive	Correct positive test result – number of diseased persons with a positive test result.

1 Executive Summary

1.1 Background

- One of the key steps in the management of breast cancer is determining whether there is spread to the axillary lymph nodes (ALN) from the main (primary) tumour
- Sentinel lymph node biopsy (SLNB) is first done at the same time as removal of the main tumour to determine whether there are regional metastases in the SLNs, the first ALNs into which the breast drains lymph. If there are any more than isolated tumour cells in the SLNs, complete axillary lymph node dissection (ALND) is required because of the possibility that tumour cells have spread beyond the SLNs into the other ALNs
- Determining whether the SLNB is positive is usually done by histopathology, examining slides under a microscope, after the operation to remove the primary tumour, so there is a delay before the ALND is performed
- If positivity of SLNB could be established during the operation, intraoperatively, ALND could be performed without delay with potential benefits for the patient and the health service
- OSNA and Metasin are two types of test which claim to be able to accurately diagnose regional metastases in the SLNs sufficiently quickly to be used intraoperatively
- OSNA is an automated molecular test that uses one-step nucleic acid amplification technology. The test analyses and amplifies genetic material (mRNA) and detects the presence of the Cytokeratin 19 (CK19) gene. OSNA does not require the mRNA to be extracted and purified from the tissue before being analysed.
- Minimal details are available for Metasin, which detects the presence of CK19 and mammaglobin. However, it appears the RNA must be extracted from tissue, purified and quantified, prior to nucleic acid amplification and analysis.
- OSNA and Metasin could be used as a replacement for post-operative histopathology or as an adjunct to it. As a replacement all of each SLN would be used by wither OSNA or Metasin; as an adjunct half of each node would be used by either OSNA or Metasin, and half used for histopathology if the OSNA or Metasin result was negative
- The main consequences of introducing OSNA or Metasin relative to current practice are predicted to be:

- ALND will be performed as a single operation following immediately after primary tumour removal, rather than a separate second operation as in current practice
- This in turn may lead to reduced anxiety in patients who no longer have to wait to find out if they need a second operation and reduced time to adjuvant treatment, if this is required
- The reduced time to ALND may however complicate the decision making process of the multi-disciplinary team (MDT)
- The adverse effects of ALND may be less where it is performed immediately after the primary tumour removal than if it is performed later as a second operation
- One rather than two operations may also lead to reduced hospital costs
- There will be increased costs associated with OSNA or Metasin, which will be off-set by reduced histopathology costs where OSNA or Metasin are used as replacement tests
- Any potential benefits above will be off-set if OSNA or Metasin introduces diagnostic errors, indicated by either its sensitivity or specificity being less than 100%
- If false-negatives are introduced, women with macro- or micrometastases will be misidentified as SLNB negative, and they will not undergo ALND. This may compromise their outcome with respect to breast cancer
- If false-positives are introduced, women who are SLNB negative will be misidentified as having macro- or micrometastases, with any resulting side-effects, but without any benefit in outcome with respect to breast cancer

- This report sets out to test these claims
- Tissue allocation bias is a major challenge in evaluating OSNA and Metasin, particularly its accuracy. Dividing the SLNs between the test of interest (OSNA or Metasin) and the test with which this is being compared (usually histopathology) means that the tumour may only be present in the section given to one or other test. Apparent errors identifying the tumour may therefore not be the fault of the test, but a problem with sampling.

1.2 Objective

To evaluate the clinical effectiveness and cost effectiveness of OSNA and Metasin if used in the NHS in England for the intraoperative analysis of metastases in sentinel lymph nodes of breast cancer patients.

1.3 Methods

1.3.1 Clinical effectiveness systematic review

- The assessment comprises a systematic review of clinical and cost-effectiveness studies, a review and a critique of data supplied by the manufacturer, and a *de novo* economic analysis.
- A systematic review was conducted to summarise the evidence on the clinical effectiveness of RD100i (OSNA) and Metasin for the intra-operative analysis of breast cancer metastases in sentinel lymph nodes. The search strategy focused on the interventions in context of the specific area in which the tests are applied: the lymph nodes.
- The following bibliographic databases were searched in this review: Medline, Medline in process and Embase (all via OVID), Web of Science (including conference proceedings, via ISI), the Cochrane Library (all) and HEED (via the Cochrane Collaboration). The searches did not use any form of limit (e.g. date).
- The following trials registries were also searched: NIH ClinicalTrials.gov, Current Controlled Trials, WHO International Clinical Trials Registry Platform (ICTRP), EU Clinical Trials Register. Google was also searched to identify grey literature and conference publications. Items included after full-text screening were forward citation chased using Web of Science (Thompson Reuters).
- Critical appraisal was performed using the Cochrane Risk of Bias tool[1] and the QUADAS-2 tool[2]. Results were summarised in tables and text, stratified by level of data (patient or node), node analysed (sentinel or axillary) and correction for tissue allocation bias.
- Studies were only included in the meta-analysis if the numbers of true positives, true negatives, false negatives and false positives were all reported in the text or could be unambiguously inferred from other figures in the text. Meta-analysis was not performed where there were fewer than four included studies, as required for the bivariate method. Summary receiver operating characteristic (SROC) curves and

summary estimates of sensitivity and specificity, with 95% CIs were calculated where possible.

1.3.2 Cost-effectiveness systematic review

- A search of the economic evaluation literature sought to identify studies of intra-operative testing options for metastatic disease in early breast cancer. In addition to the electronic sources searched for the clinical effectiveness review, the NHS Health Economic Evaluation and EconLit databases were searched for cost –effectiveness and cost-utility studies and a supplementary manual search of the bibliography of relevant studies.

1.3.3 PenTAG cost-effectiveness analysis

- The PenTAG model was split into two sections: the diagnostic pathway and the management pathway.
- The diagnostic pathway was a decision tree built to represent the diagnosis of regional metastases in the sentinel lymph nodes. Three pathways were examined: current practice histopathology (the 'gold standard'), replacement testing of the full node by intraoperative testing, half node intraoperative testing followed by histopathology on the other half node.
- In the diagnostic pathway, patients who were diagnosed with sentinel lymph node metastases received axillary dissection (ALND). For those diagnosed intraoperatively, this occurred during the same surgery as their sentinel lymph node biopsy (SLNB) and for those diagnosed by histopathology, this occurred in a follow up surgery.
- Diagnostic accuracy was taken from the clinical effectiveness systematic review. For OSNA this was split into studies that included adjustment for tissue allocation bias (TAB) and those that did not.
- Patients incurred costs depending on their diagnostic strategy, their surgery, any additional hospital stay and adverse events. The intraoperative test costs (OSNA and Metasin) were derived from information provided by the technology sponsors. Costs for histopathology were taken from Cutress et al., 2010.[3] Other costs were derived from a variety of sources. Notably the cost of surgery was calculated using two different strategies: firstly, the same strategy as ScHARR[4] using NHS Reference costs and secondly using the YHEC costing approach from Burke and Patton 2010.[5]
- Short term disutility for patients waiting for results or undergoing a second operation was investigated

- Outcomes from the diagnostic pathway included cost-effectiveness based on cost-accuracy (cost per case accurately diagnosed, cost per node positive case detected, cost per node negative case detected) and cost-effectiveness based on short term cost-utility.
- The management pathway was concerned with lifetime results and used an updated version of a previous model.[4] This model was a discrete event simulation that followed individual patients through a series of health states calculating their accrued costs and QALYs.
- Patients who received different diagnoses in the diagnostic pathway were treated differently in the management pathway and so accrued a different set of costs and QALYs. These cost and QALYs were then attached to the short term results to give overall costs and QALYs.
- Most parameters for the management model were taken from Cooper et al., 2011 after forward searching failed to produce any new information.[4] Instead the studies that Cooper and colleagues had used were carefully checked to approve their values. Where possible, updated costs were found and where this was impossible, previous costs were updated to 2010 prices.
- Univariate, deterministic sensitivity analyses were conducted to examine the effect of sensitivity, specificity, prevalence and cost adjustments on the various outcomes.
- Sensitivity and specificity were assessed using a threshold analysis with ranges taken from the clinical effectiveness systematic review. Prevalence of sentinel lymph node metastases was adjusted from 20% (from the clinical effectiveness systematic review) to 10% and 40%. Costs of tests and surgeries and management costs were adjusted by +/-10%.
- Metasin was assessed separately from OSNA illustratively because the underlying accuracy data were not published.

1.4 Results

1.4.1 Clinical effectiveness systematic review

- Eighteen studies were included that investigated the performance, particularly test accuracy, of either OSNA or Metasin on detecting metastases in the sentinel or axillary lymph nodes of breast cancer patients. Two studies were included for Metasin, however, both were unpublished and in draft form. The remaining sixteen studies reported on OSNA, with two papers reporting the same study.

- The majority of studies were considered to be at low risk of bias, although many were unclear regarding their method of patient recruitment and lacked detail on patient characteristics.
- Reported outcomes were limited, for example no data was found for clinical outcomes, such as patient anxiety and number of repeat operations. Only one included study provided evidence for time in operating theatre.
- In accuracy studies the reference standard (i.e. histopathology), although plausible, may be performed with varying levels of analysis, and as such, may not be a true indicator of the target condition.
- The main issue within the included studies has been tissue allocation bias (TAB), which occurs when a different portion of tissue is allocated to the index and reference test and cannot then be re-used between them. Studies have dealt with this in a variety of ways, some re-analysing both histopathology and molecular samples, some choosing to re-analyse just one technology and some doing neither.
- Other quality concerns included unclear sampling methods, e.g., no evidence was given of sample replicates and reproducibility for molecular analysis. Furthermore, unless otherwise mentioned, test failures were not reported. Blinding of outcome assessors was also often not reported and papers were unclear on robustness of histopathology, e.g. whether results were checked by a second party.
- It should be noted that more than one sentinel lymph node may be removed from a patient, which means that results for a study may be presented by patient or by individual node or both. Analyses using both units of measure were considered useful.
- A potential conflict of interest also features heavily, since one of the two unpublished Metasin studies was performed at the institution in which the technology was developed and the majority of the OSNA studies were financially supported by Sysmex.
- Overall, with regard to the validity of the test accuracy results, the studies were reasonably well done and produce consistent results. However, there is a strong assumption in the accuracy studies that the reference standard is a true measure of the target disorder.
- A summary of results is presented below.

	Sample type	Number of studies	Adjustment for TAB	Sensitivity (95% CI)	Specificity (95% CI)
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pooled OSNA	Patient	5	No	84.5 (74.7–91.0)	91.8 (87.8–94.6)
Pooled OSNA	Patient	3	Yes	91.3 (83.6–95.6)	94.2 (91.2–96.2)
Pooled OSNA	SLN	4	No	79.9 (74.2–84.6)	95.5 (94.1–96.5)
Pooled OSNA	SLN	5	Yes	89.0 (82.1-93.4)	97.5 (96.6-98.2)

- As there were only two studies for Metasin, a meta-analysis was not performed. Since these data were taken from draft papers, prior to peer review, the results, [redacted], must be used with caution.
- As displayed above, some studies have adjusted for TAB, generally taking a conservative approach by excluding affected samples. This does improve the test accuracy, increasing sensitivity from 79.9% to 89.0% and increasing specificity from 95.5% to 97.5%. Although it is not possible to be definitive, we consider the adjustments to be reasonable.
- With regard to the time taken to perform OSNA, despite the lack of detail in the studies explaining which aspects of the procedure were monitored, the time ranges from less than 30 minutes to 39.6 minutes for one node. This increases by approximately 5 to 10 minutes per additional node analysed.

1.4.2 Cost effectiveness systematic review

- Two studies were identified by the search. One was a study of OSNA conducted in one centre in Spain. It provided evidence on resource use and costs, but was limited in its assessment of benefits to patients or in terms of improved diagnostic accuracy. The other study was an evaluation of a diagnostic testing option in the UK that has been withdrawn from the market (GeneSearch) and is therefore no longer relevant for the present evaluation.
- The cost effectiveness systematic review did not reveal any directly relevant studies and therefore a de novo model was justified.

1.4.3 Independent ERG assessment

- In the base case, short term cost-accuracy results show that in general OSNA half node was not cost effective, as it was either dominated by tests that were equally accurate, but less costly (i.e. OSNA full node when detecting node-negative cases) or extended dominated by histopathology, where histopathology had a smaller incremental cost-effectiveness ratio (ICER) per correct diagnosis than OSNA half node, when they were both compared to OSNA full node. In all cases where OSNA half node was dominated or extended dominated, the ICER per additional diagnostic yield of histopathology versus OSNA full node was less than £13,000. However, for node-positive case detection, using NHS Reference Costs, OSNA half node dominated histopathology (having the same detection rate but being less costly) and had an ICER of £16,123 per additional node-positive case detected compared to OSNA full node.
- When the sensitivity and specificity of OSNA were altered to use values from studies that adjusted for TAB, the cost-accuracy results remained fairly consistent, although the ICERs were increased for the greater values of sensitivity and specificity, to reflect the reduced differential in accuracy relative to histopathology. However, when values from Khaddage were used (sensitivity 100%, specificity 97.2%), OSNA half node was consistently dominated by OSNA full node, since they had the same diagnostic accuracy, but full node did not incur the additional cost of histopathology. The ICERs between histopathology and full node increased to £27,300 per patient correctly diagnosed and per node-negative case detected using NHS Reference Costs (£17,100 per patient correctly diagnosed and per node-negative case detected using YHEC costs) and OSNA full node dominated histopathology for detecting node-positive cases (because their diagnostic yield was equal, but OSNA full node was less costly).
- Short term cost-utility values showed OSNA full node consistently dominating OSNA half node and histopathology as it had higher QALY gains and was less costly. As these results only accounted for costs incurred in the diagnostic phase and did not consider any benefits relating to accuracy of the tests, they were not affected by adjusting the sensitivity and specificity of OSNA for TAB.
- Long term results revealed that OSNA half node was extended dominated by histopathology when they were both compared to OSNA full node, as histopathology had a higher QALY yield than OSNA half node and a smaller incremental cost per QALY gained compared to OSNA full node. The ICER between histopathology and OSNA full node was less than £5,000 per QALY gained, for both costing strategies.

When these results were adjusted to use TAB values, half node remained extended dominated using values provided by Frere Belda and colleagues. At the values reported by Snook and co-workers OSNA half node was extended dominated under YHEC costs, whereas with NHS Reference costs it resulted in £8,063 per QALY gained relative to the full node option. Histopathology compared to OSNA full node had ICERs under £10,000 per QALY gained using NHS Reference costs or YHEC costs. However, when values from the study by Khaddage and colleagues were used, full node OSNA dominated both half node OSNA and histopathology under both costing scenarios.

- Base case results demonstrated a link between accuracy and both short term and long term measures of cost effectiveness, which was investigated further in the sensitivity analyses. Base case long term results indicated that sensitivity may be more influential than specificity so individual univariate threshold analyses were conducted on sensitivity and specificity. The ICERs of histopathology for all measures of cost effectiveness increased with increases in sensitivity or specificity, but the long term ICERs remained below £20,000 per QALY gained for all values of sensitivity up to 95%. Short term cost-accuracy results for specificity increased from £8,945 per additional patient correctly diagnosed at 95% specificity, to £22,761 per additional patient correctly diagnosed at 100% specificity. As for long term results for specificity the maximum ICER was £8,430 per QALY gained when OSNA had 100% specificity, and at 70% specificity OSNA was dominated by histopathology. This implied that changes to specificity may have more of an impact in the short term, but that changes to sensitivity may have a much greater impact on long-term cost effectiveness.
- Sensitivity analysis was also conducted on the effect of prevalence of sentinel lymph node metastases in the patient population. When prevalence was reduced to 10% histopathology dominated OSNA half node in terms of short term cost-accuracy and long term cost effectiveness and had an ICER of £2,626 per QALY gained relative to OSNA full node. Short term cost-utility results were unaffected, with OSNA full node continuing to dominate. Increasing the prevalence to 40% most notably affected the long term ICERs, with half node OSNA dominating histopathology. The ICER of OSNA half node relative to OSNA full node was £2,208 per QALY gained.
- Altering individual costs and utility parameter values in both sections of the model had very little impact on overall cost-effectiveness results. This highlighted the importance of having the correct information for the diagnostic accuracy of OSNA as these results demonstrated that this was the most influential outcome parameter.

Results for Metasin were provided on a purely illustrative basis, as the only values available for the sensitivity and specificity of Metasin were from an unpublished, draft paper that has not yet been peer reviewed. Furthermore the only cost information available was provided on behalf of the author of the draft paper and did not reflect the cost of Metasin with a CE mark, which it has recently received, but for which there is no list price yet available (list prices became available after the analyses for this review had been completed) . Cost accuracy analysis for Metasin was more favourable than in the case of OSNA, reflecting the [redacted] lower costs of Metasin. As with OSNA, short term cost-utility results demonstrated that intraoperative testing on a full node basis would dominate all other strategies. In the long term Metasin half node had an in incremental cost per QALY gained below £13,000 relative to its full node alternative, which suggests that it is the cost-effective choice since histopathology had an ICER of £467,113, using NHS Reference Costs, and £246,089, using YHEC costs, against Metasin half node.

1.5 Conclusions

- The research study evidence-base for OSNA and Metasin is restricted to evidence on their test accuracy (sensitivity and specificity) relative to a reference standard of histopathology
- All other conclusions are based on the predictions of a health economic model in a linked-evidence approach
- OSNA [redacted] appear to be effective in reducing the number of separate second ALND operations, which leads to savings in cost and benefit to patients. However this is at the expense of diagnostic errors, both false negatives and false positives.
- Overall the cost-effectiveness evidence is inconclusive. Depending on the view about the best source of estimates on sensitivity and specificity for OSNA, the model analysis may favour favour histopathology or OSNA. The evidence on Metasin is incipient and may only be suggestive; in general the potential long term benefits of increased accuracy with histopathology more than compensate for its disadvantage in terms of expediency of test results but such balance is sensitive to how different studies address the issue of TAB (i.e. the study by Khaddage and colleagues vs. other studies)
- Therefore, uncertainty about the estimates of accuracy leads to considerable uncertainty about cost-effectiveness

- Considerable caution needs to be exercised in interpreting the results of the studies of Metasin because they are based on draft papers which have not been submitted for publication. Cost-effectiveness estimates are thus illustrative and it is not valid to conclude on the relative effectiveness and cost-effectiveness of the two tests of interest.

1.6 Suggested research priorities

The uncertainty in decision making due to lack of data and variation in available estimates on diagnostic accuracy, short-term implications of diagnostic approaches and long term patient management may be reduced by undertaking further research

Peer reviewed and formally published research on Metasin is essential.

Greater clarity on the true costs of the alternative tests, and the variation in resource utilisation at the level of the patient, would help to identify the significance and the distribution of costs and benefits of intra-operative diagnosis.

Observational studies may be useful to verify empirically confirm and quantify the reduction in the numbers of operations, the anxiety caused to patients, and its impact on their quality of life, by delayed availability of test results and second operations, and the costs implications to hospitals of the introduction of intraoperative testing in SLNB.

However improving on the accuracy of estimates for OSNA in particular, overcoming concerns about TAB and the validity of currently available reference standards may be challenging. A test-treat randomised trial may be the only way to truly resolve whether introduction of intraoperative testing in SLNB would be effective and thus cost-effective. The outcome would need to be locoregional recurrence rates or even survival in order to capture the trade-off between the potential short term gains associated with single operations to achieve ALND and the longer term disbenefits arising from false negative and false positive cases occurring with intraoperative testing.

Also, a strong assumption in the report is that ALND is the usual best treatment if micro and macrometastases, or their equivalents are identified in a SLNB. Evidence on this is evolving needs to be followed closely as it could impact on decisions about intraoperative testing in SLNB in the future.

2 Background

2.1 Nature of disease

Breast cancer affects the breast tissue of either women or men, although the latter is relatively uncommon. In it cells making up the breast begin to divide in an uncontrolled manner forming tumours. Initially they are confined within the structures of the breast where the cells would normally occur, and the tumour is referred to as ductal carcinoma-in-situ (DCIS). Later the cells become locally invasive extending into parts of the breast and surrounding tissue where they would not normally occur. The breast cancer cells also enter the lymph drainage system, which in the breast leads to breast cancer cells lodging and growing in lymph nodes (LN), particularly those in the armpit – the axillary lymph nodes (ALN). These are called regional metastases. The size of ALN containing local metastases can vary from a near normal LN size of 0.5 cm or less in length to greatly enlarged, 4cm in length.

In addition the breast cancer cells can enter the blood stream, from which they can spread to distant parts of the body such as lungs, bones and the liver. These are called distant metastases. Without treatment the proliferation of the breast cancer cells and their spread around the body leads to death. The reasons why breast cancer develops are not completely understood, although genetic predisposition has been increasingly recognised as a contributing factor in a minority of women.[6]

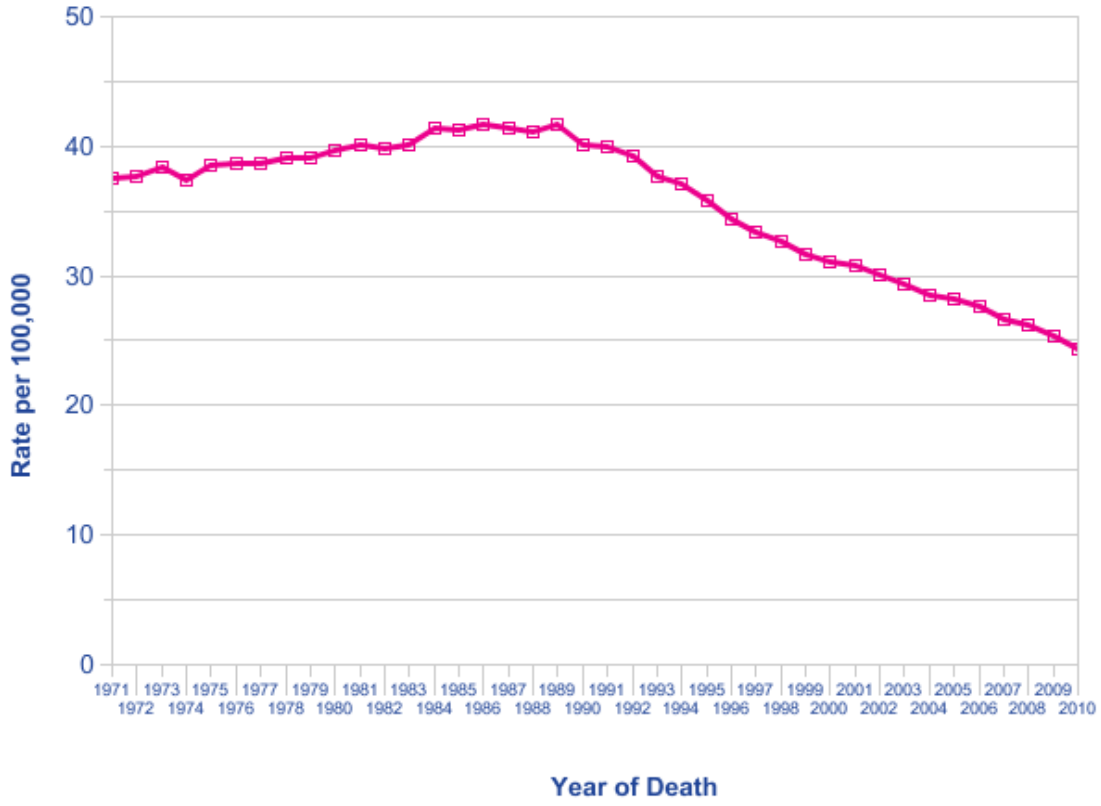
Table 1. Breast cancer incidence for England and Wales, 2010

	Age (years)											
Population and Country	10-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85+	Total
Women, Wales *	90	133	241	278	280	358	357	219	262	184	223	2625
Women, England **	1798	2528	4164	4637	4005	5618	5009	3430	3492	3073	3505	41259
Men, Wales *	0	0	0	0	4	1	2	5	0	2	1	15
Men, England **	7	11	11	19	30	37	47	59	47	48	37	353
Total	1895	2672	4416	4934	4319	6014	5415	3713	3801	3307	3766	44252

Sources:
* Welsh Cancer Intelligence and Surveillance Unit. Cancer Incidence in Wales 2006-2010.

Breast cancer is a very important challenge to health. Each year in England [7] and Wales ENREF 3 [8] approximately 40,000 individuals develop breast cancer. As is illustrated in Table 1, the majority of these are women over the age of 40. Further breast cancer is responsible for over 10,000 deaths each year in England and Wales[9], although the mortality rates have been declining from a peak in the mid-1980's (Figure 1). Breast cancer is the most commonly occurring female cancer and the second most common cause of cancer death after lung cancer accounting for 1 in 6 of all female cancer deaths. Approximately one in nine women will develop breast cancer at some stage of their life.

Figure 1. Age-standardised rates for female breast cancer mortality in the UK, 1971-2010. [9]



Source: http:www.cancerresearchuk.org/cancer-info/cancerstats/types/breast/mortality (Last accessed 4/1/13)

2.1.1 Staging of breast cancer

There is a well-defined system for recording the severity and extent of spread of breast cancer. This is based on tumour size, spread to LNs and presence of distant metastases, the TNM classification.[10, 11,12, 13] This is given in Table 2 and Table 3.

Table 2. Description 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 the 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:	Redrawn from: 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).

Table 3. 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:	Redrawn from: 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		

Increasing size of the main tumour and spread to ALNs are two of the key features denoting worsened stage. The presence of distant metastases denotes the worst stage IV, irrespective of the size of the main tumour or spread to regional LNs. The majority of patients present with Stage I or Stage II as illustrated in Table 4.[14] These figures combined with information on how tumour size inter-relates with nodal status suggests that in the UK 33% of patients present with spread to LNs.[15] This agrees with a systematic review with 13 included studies undertaken to inform the NICE clinical guideline on early and locally advanced breast cancer which reported a mean prevalence of LN positive status as 31.4% with a range of 18 to 59%.[16] Other sources have suggested a higher value of patients presenting with spread to LNs of 41%.[4]

By definition, early breast cancer is cancer which has not spread beyond the breast or the ALNs on the same side as the tumour, that is Stages I (any), II (any) or IIIA.

Table 4. Proportion of patients by stage categories for breast cancer in East of England residents (2006-2009)

Stage	N	% among all patients	% among patients with known stage
I	6788	38	41
II	7361	41	45
III	1490	8	9
IV	821	5	5
Unknown	1376	8	n/a
All (excluding unknown)	17836 (16460)		
Source	Lyratzopoulos G, Abel GA, Barbiere JM *et al*. Variation in advanced stage at diagnosis of lung and female breast cancer in an English region 2006-2009. *Br Journal of Cancer* 2012; **106:** 1068-1075.		

Recent modifications to the staging system for breast cancer also recognise the size of the local metastases in the lymph nodes. Macrometastases are defined as tumour deposits where one dimension is above 2mm. Micrometastases, deposits which are only discernible microscopically, measure >0.2 mm with no dimension being >2mm. Isolated tumour cells are also recognised as part of the N0 category and by definition no dimension of any collection of isolated tumour cells must exceed 0.2 mm.

2.1.2 Prognosis

With early identification and treatment the outlook for patients with breast cancer is good. The overall 5 year survival rate is approximately 80%.[17] These rates vary with age, with patients over 70 having survival rates below 80% as shown in Table 5.

Table 5. Five year survival rates for female breast cancer according to age in England 2005-2009.

	Age (years)						
	15-39	40-49	50-59	60-69	70-79	80-99	All ages
5 year survival rate (%)	84	89	90	90	81	69	85
Source:	http:www.cancerresearchuk.org/cancer-info/cancerstats/types/breast/survival (Last accessed 4/1/13)						

Stage also influences survival as shown in Table 6. In women diagnosed with breast cancer in the West Midlands between 1985-9 and followed up until 1999, 5 year survival was 88% for Stage I, 69% for Stage II, 43% for Stage III and 12% for Stage IV.[18] This pattern is maintained in data from the USA from the National Cancer Data Base who were diagnosed in 2001 and 2002.[19]

Table 6. Five-year survival rates according to stage of disease

	Stage of disease						
	I	II		III			IV
		IIA	IIB	IIIA	IIIB	IIIC	
UK*	88%	69%		43%			12%
USA**	88%	81%	74%	67%	41%	49%	15%
Source	* 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. **http://www.cancer.org/cancer/detailedguide/breast-cancer-survival-by-stage (Last accessed 09/11/2012)						

In addition to age, tumour size and spread, prognosis is also related to tumour grade and receptor status. In respect of the latter, positive oestrogen receptor status and positive progesterone receptor status denote better prognosis and over expression of human epidermal growth factor 2 poorer prognosis. These additional factors are incorporated into tools to estimate prognosis such as the Nottingham Prognostic Index[20] and Adjuvant! Online[21] which are used to guide treatment, particularly the use of adjuvant chemotherapy in early breast cancer. The effectiveness and cost-effectiveness of gene expression profiling and expanded immunochemistry tests to enhance the Nottingham Prognostic Index and Adjuvant! Online is currently under consideration by NICE's Diagnostic Assessment Committee.[22]

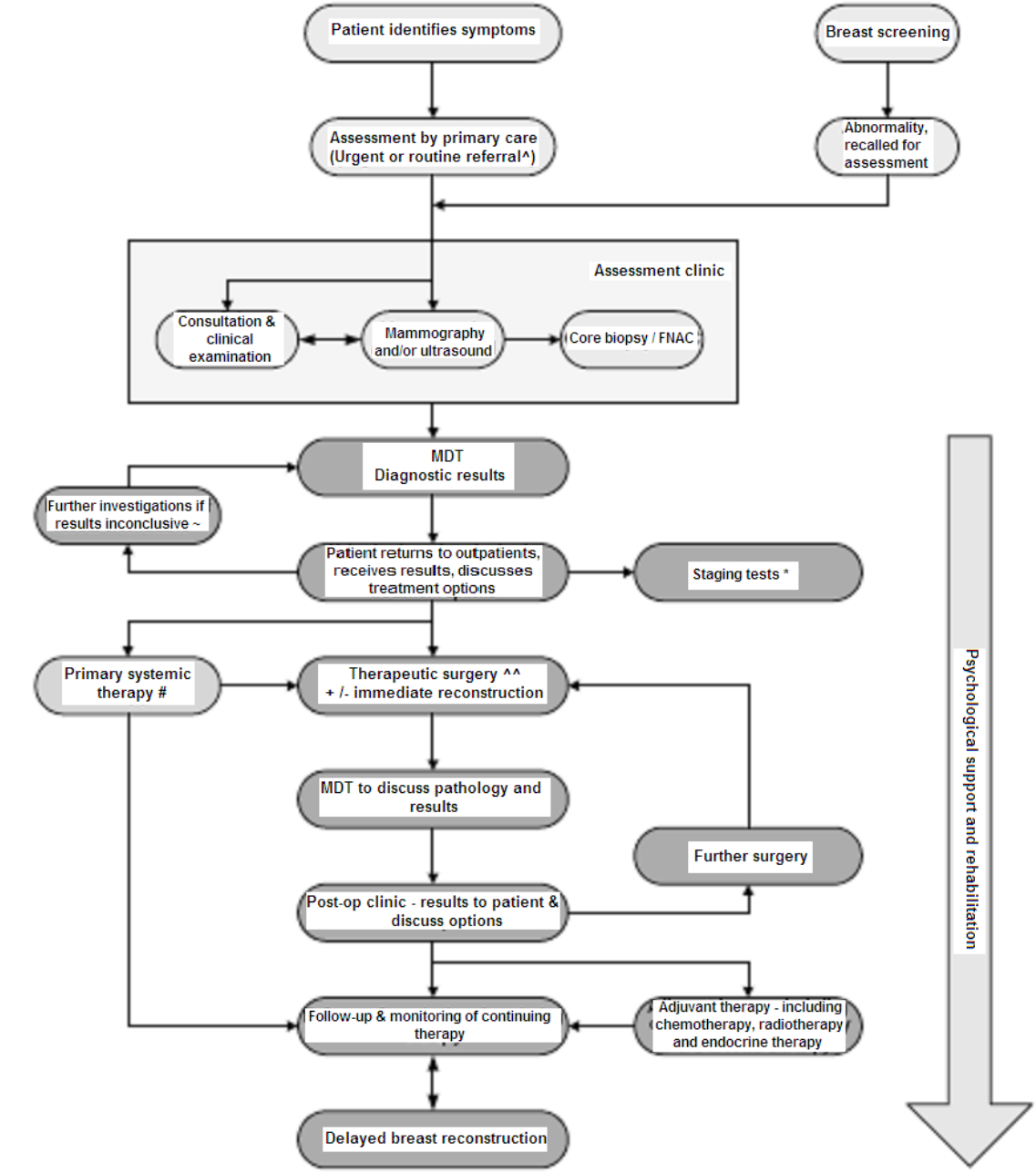
The implications for prognosis of micrometastases as compared to macrometastases are currently unclear, but they are counted as local metastasis and treated in a similar manner to macrometastases (see below). In contrast isolated tumour cells are not currently counted as local metastasis and assumed to have a similar to prognosis to no lymph node metastases (N0).[23]

2.2 Management of disease

2.2.1 General clinical pathway for suspected breast cancer

The tests and treatment advised for patients suspected with breast cancer are outlined in the NICE clinical guideline on early and locally advanced breast cancer.[16] This summarised as an algorithm reproduced in Figure 2.

Figure 2. Clinical pathway for breast cancer



Key:
^ Following the publication of the Cancer Reform Strategy (Department of Health, 2007), by December 2009 all patients presenting with breast problems referred by their GP to a specialist should be seen within two weeks, in England.
~ Include repeat core biopsy/open biopsy/MRI etc.
* Not all patients will require staging: Scottish Intercollegiate Guidelines Network (2005) Management of breast cancer in women: A national clinical guideline. SIGN Publication No. 84. Edinburgh: SIGN, 2005. ISBN: 1 899893 34 2.
# For elderly or unfit patients, surgery may not be appropriate. For locally advanced but non metastatic, primary systemic therapy precedes therapeutic surgery in order to reduce size of tumour.
^^ Could include breast conservation (WLE), mastectomy & axillary staging (SLNB, sampling or clearance).

Source: 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.

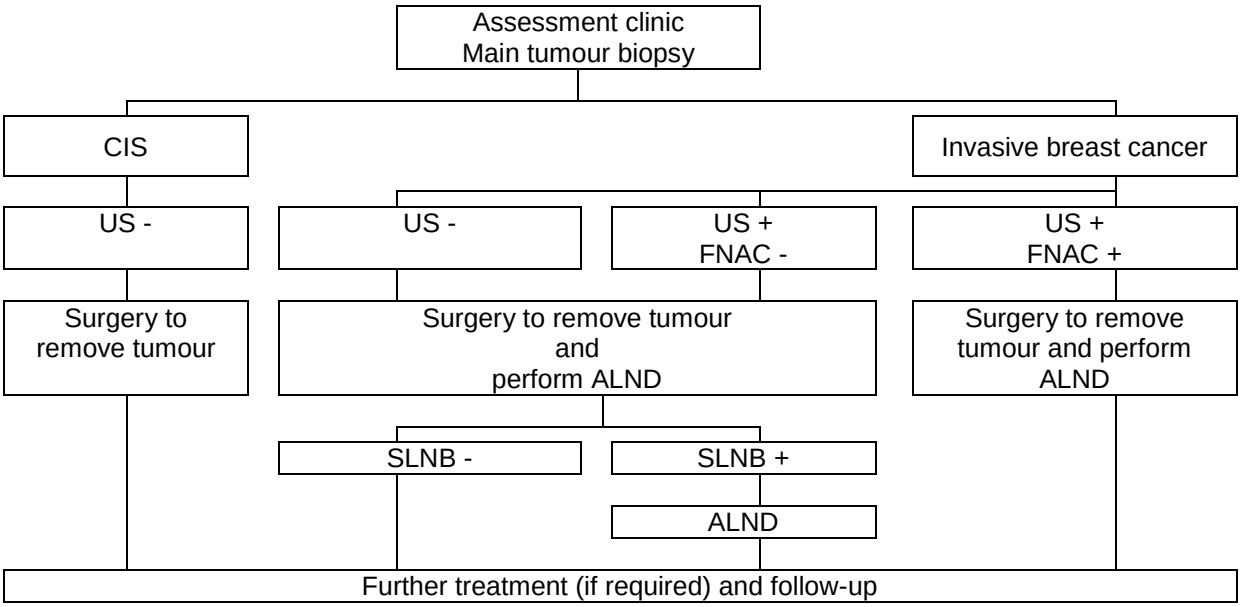
In the initial assessment anyone suspected of breast cancer whether via primary care or the breast screening programme is referred to an assessment clinic where clinical examination is repeated, mammography and ultrasound undertaken and a biopsy of the lesion(s) performed (usually a core biopsy). Ultrasound of the axilla is performed and further biopsies of suspicious LNs taken (fine needle aspirations or core). A multi-disciplinary team (MDT) meeting considers the results from these investigations and makes a definitive recommendation regarding proposed management to discuss with the patient.

If a cancer is detected, surgery is usually performed to remove the breast cancer unless neoadjuvant therapy is considered appropriate. Staging of the axilla is performed in cases of invasive breast cancer in the manner described in the next section. All findings, including the pathology of the removed tumour and the results of any staging procedures are then discussed at a further MDT meeting where decisions are then made about further surgery and whether adjuvant hormone therapy, chemotherapy, immunotherapy or radiotherapy is required.

2.2.2 Clinical pathway for staging of breast cancer and subsequent surgery to the axilla

The detailed steps for investigating whether a suspected breast cancer has spread to the axilla and the degree of that spread are also outlined in the NICE guidelines.[16] These are summarised in Figure 3. Further information on key steps in this pathway is given in following sections.

Figure 3. Clinical pathway detail for staging of breast cancer and subsequent surgery to the axilla



Four different scenarios are recognised depending on the findings resulting from the initial assessment:

a) Likely DCIS. Provided axillary ultrasound (US) reveals no abnormalities and the patient is not considered high risk, no further investigations of the axilla are undertaken

b) Likely invasive breast cancer with negative axillary US (no abnormal LNs identified). In these patients at the time of the surgical removal of the main breast tumour, a sentinel lymph node biopsy (SLNB) is undertaken. In a SLNB a weakly radioactive solution and a blue dye are injected into the breast before surgery to identify the first LN/s to which the breast drains lymph in a particular individual. The SLNs become blue and/or can be detected using a radioactivity counter. They are most frequently found in the axilla of the same side. These sentinel lymph node(s) (SLN) are then removed to see if the cancer has spread from the original site. This is done by histopathology which involves cutting very thin slices of the SLNs, staining them and then carefully examining them under a microscope by a medically qualified specialist. Histopathology takes several days and sometimes further investigations are required for a definitive diagnosis e.g. immunohistochemistry.. If no breast cancer cells or single, isolated breast cancer cells are found in the SLN/s no further action needs to be taken. However, if the breast cancer has spread to the SLNs, all the relevant lymph nodes in the axilla need to be removed in a further operation called axillary

lymph node dissection (ALND). This provides treatment by removing all the tumour cell-bearing LNs and others within a defined anatomical boundary and it provides detailed staging information by allowing the number of LNs with metastases to be precisely quantified as the LNs removed are subjected to further histopathological examination.

Occasionally SLNB cannot be undertaken. In this case a 4 node sample may be performed instead in which 4 LNs are removed and examined without specific evidence that they are the SLNs to give some further information.

c) Likely invasive breast cancer with a positive axillary US but normal US guided fine needle aspiration cytology (FNAC). Again at the time of the surgical removal of the main breast cancer, a SLNB is undertaken with the same actions as b) above if cancer is not found or confirmed to have spread to the SLNs.
d) Likely invasive breast cancer with positive axillary ultrasound, and confirmed abnormality on US guided FNAC. In this case the patient would proceed directly to ALND without a preceding SLNB.

The preference in NICE guidance for the use of a strategy employing US and FNAC to triage need for SLNB was underpinned by a model-based cost-effectiveness analysis undertaken as part of the preparation the NICE clinical guideline on early and locally advanced breast cancer.[16]

2.2.3 Accuracy of clinical examination

Clinical examination of the axilla involving palpation is subject to error as a method of detecting spread of breast cancer to the ALNs. The sensitivity of the technique has been estimated as 46% based on pooling of a number of studies.[24-29] The fact that over half of axillary metastases are not detected by palpation is the reason why additional investigations are required. The main reason why palpation is unsuccessful is often because the presence of metastases does not always lead to a change in size or texture of the axillary lymph nodes, coupled with the fact that the axillary lymph nodes are not always easy to examine.

2.2.4 Accuracy of ultrasound guided fine needle aspiration and cytology

Like palpation, ultrasound and FNAC are imperfect techniques. The accuracy was considered in detail as part the NICE clinical guideline on early and locally advanced breast cancer and the relevant section is reproduced in full in Appendix 1.[16]

The key facts identified were:

- LNs can be visualised by US in 81% of cases, although there is considerable variation
- Using LNs that were suspicious on ultrasound based on their size (> 5mm) and configuration as the diagnostic criteria, sensitivity was 69% and specificity was 75%. This was when patients with palpable and non-palpable axillary lymph nodes were combined. The accuracy was improved when just cases with palpable LNs were included and worsened with cases with non-palpable LNs
- The staging performance of US guided FNAC was sensitivity 43% and specificity 100% with an accompanying positive predictive value of 99% and a negative predictive value of 72%.

2.2.5 Accuracy and adverse effects of axillary lymph node dissection

ALND, described earlier, has been considered the 'gold standard' procedure for staging the axilla. It 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.[4]

However, ALND is associated with significant complications, including a 21% incidence of arm lymphoedema (general swelling)[30-32], a 22% incidence of seromas (pockets of fluid under the skin)[30,33] and a 14% infection rate[30,34]. In addition, insertion of a surgical drain during surgery is commonplace (79%) and usually necessitates prolongation of hospital stay.[34] Pain, limited mobility, numbness and sensory loss are also common. 80% of women are claimed to suffer some adverse event .[35] It is for these reasons that there has been a focus on performing ALND for is therapeutic effects applicable in patients with spread of breast cancer to the ALNs, rather than as a more widely applied diagnostic tool.

2.2.6 Accuracy and adverse effects of sentinel lymph node biopsy

SLNB, described earlier, although still a surgical procedure of the axilla is much simpler than ALND and associated with a much lower rate of side-effects. Thus the incidence of lymphoedema falls to 7%,[36] seroma to 7%,[30,33] surgical drain requirement to 2%[34,37] and infection incidence to 2%[30,34,37].

SLNB was considered in detail as part the NICE clinical guideline on early and locally advanced breast cancer and the relevant section is reproduced in full in Appendix 2.[16] The key features identified were:

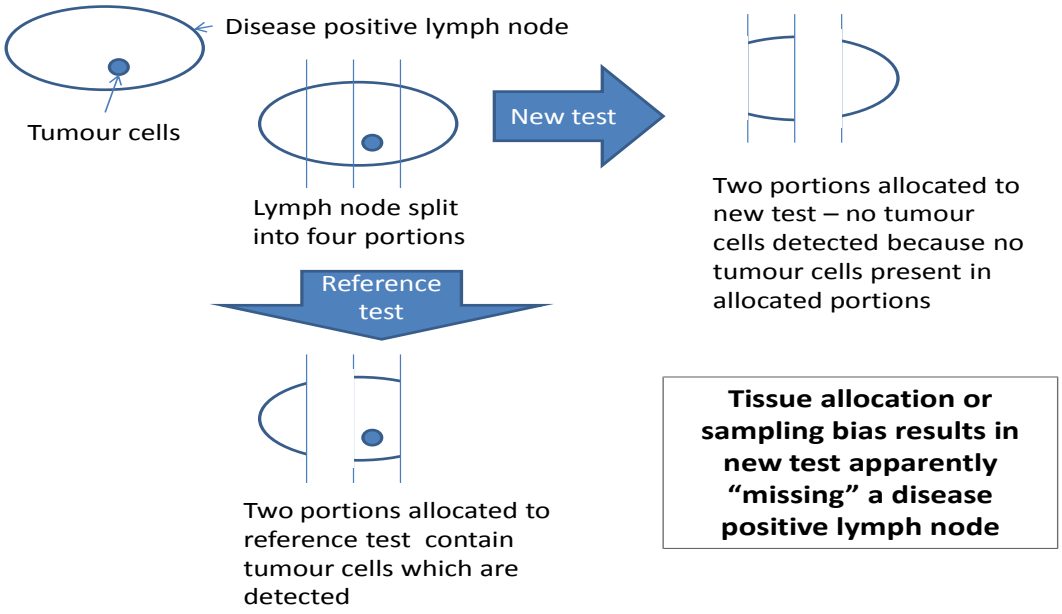
- The overall sentinel lymph node localisation rate was 96.4%
- The pooled estimate of false negative rate was 7% i.e. sensitivity is 93%
- The mean proportion of patients with positive sentinel lymph nodes was 42%
- Patients treated by SLNB do not appear to have poorer rates of disease-free survival or overall survival, or of axillary recurrence in the short term, compared to patients treated by axillary clearance employing ALND

2.2.7 Challenge to measuring accuracy – tissue allocation bias

Accuracy indicates the degree to which a test of interest correctly identifies whether a patient/sample has the target disease (its sensitivity) or correctly identifies that the patient/sample does not have the disease (its specificity), the true disease state being identified by applying a reference standard to all patients/samples.

A challenge for measuring the accuracy of tests aiming to identify whether breast cancer has spread to ALNs is that the tumour cells are not evenly distributed throughout the LN. Further the tests of interest often consume the LN so that once it has been used for one test it cannot be used by another necessitating that the sample be partitioned if multiple tests are to be done. Thus apparent errors in accuracy may be introduced not just because a new test truly fails to identify tumour cells which are present, but also because the portion of the lymph node used in the test was not the portion which contained the tumour cells. This is referred to as tissue allocation or sampling bias, which can lead to underestimation of sensitivity where the portion allocated to the new test does not contain the tumour cells. This is illustrated in the diagram below. Underestimation of specificity can also occur if the new test is allocated a portion of the LN containing tumour cells and the reference standard the portion without the tumour cells.

Figure 4. Diagram illustrating the effect of tissue allocation or sampling bias



The impact of tissue allocation bias needs careful consideration both in the context of the particular new technologies under consideration, but also because histopathology is a commonly used reference test, and it itself is affected by tissue allocation bias as the portion of the lymph node submitted to histopathology may not contain the tumour cells. This is further complicated by the fact that histopathology only examines a finite number of the slices in the portion of the lymph node allocated to it, which even though likely to be equally spaced throughout the lymph node portion, do not represent all of it.

The implications of tissue allocation bias are re-visited later in the background and throughout the report.

2.2.8 Other approaches to treating the spread of breast cancer tumour cells beyond the sentinel lymph nodes

The current management indicated above, particularly that micrometastases and macrometastases in the SLNs should be followed by ALND is based on the then consensus indicated by the NICE guideline published in 2009. There is however on-going debate about treatment of the possible spread of breast cancer tumour cells to ALNs beyond the SLNs. Pre-eminently this has been precipitated by the results of a randomized clinical trial comparing ALND with no ALND in women with invasive breast cancer and SLN spread[38] - the American College of Surgeons Oncology Group Z0011 trial. The participants were adult

women with histologically confirmed invasive breast cancer clinically 5cm or less, no palpable adenopathy,and an SLN metastatic breast cancer documented by frozen section, touch preparation, or haematoxylin-eosin staining on permanent section. They were ineligible if they had three or more positive SLNs, matted nodes, or gross extra-nodal disease, or if they had received neoadjuvant hormonal therapy or chemotherapy. 445 were randomly allocated to ALND following SLNB and 446 to no ALND following SLNB, although this number of participants was well below the target for recruitment, 1900 in total. At a median follow-up of 6.3 years, 5 year overall survival was 91.8% with ALND and 92.5% without ALND. This is equivalent to an adjusted HR of 0.87 (95% CI 0.62 to 1.23).5 year disease-free survival was 82.2% with ALND and 83.9% without ALND. This is equivalent to an adjusted HR of 0.88 (95% CI 0.65 to 1.25). The conclusion was that among patients with limited SLN metastatic breast cancer treated with breast conservation and systemic therapy, the use of SLND alone compared with ALND did not result in inferior survival. Although the lack of power of the study and limited radiotherapy quality assurance are important provisos, it does explain why there is growing caution about the use of ALND where the amount of spread to the SLNs is limited and women are also likely to receive adjuvant therapy.

In addition to possible changes in the use of ALND, there is also debate about the role of axillary irradiation as an alternative to ALND where SLNB is positive. Currently there are no prospective data to guide indications for axillary irradiation in the absence of ALND. The results of the ongoing AMAROS trial[39] comparing axillary irradiation with ALND after SLNB are awaited. In the absence of level 1 evidence, pragmatic recommendations for local-regional irradiation have been suggested.[40] It may be appropriate for example to consider axillary irradiation after SLNB for patients with low volume macrometastases (1-2 positive nodes) or high risk micrometastases. It should be noted that there is no data on the role of axillary irradiation after a positive sentinel node analysed by OSNA or equivalent technologies.

2.3 Description of technologies under assessment

2.3.1 Rationale

One of the problems with current practice with respect to investigating whether breast cancer has spread to the ALNs is the need to wait for the histopathology results from SLNB indicating whether spread has or has not occurred. Thus if spread to the ALNs has indeed occurred there is inevitably a delay before performing ALND relative to a situation where

excision of the suspected breast cancer and ALND could be done in one operation. Also the individual must be admitted to hospital, be operated on and receive an anaesthetic for a second time. The operation of ALND itself may be more difficult because of the recent prior operation and complication rates may as a result be higher than if the ALND immediately followed the SLNB. Adjuvant treatments, if required, may also be delayed.

Two new methods, similar in approach, of examining LNs removed in SLNB intraoperatively have been claimed as ways to achieve SLNB followed immediately by ALND where required, avoiding delay and a separate second operation . The new methods may also however limit the opportunity for a MDT to consider the appropriateness of ALND taking account of all the information that could potentially be available at the time of the second operation under current practice, unless each possible outcome has been discussed preoperatively.

2.3.2 One step nucleic acid amplification (OSNA)

The RD100i OSNA system (henceforth referred to as OSNA) is an automated molecular test that uses one-step nucleic acid amplification technology. The test analyses and amplifies genetic material (mRNA) from solubilised biopsy samples of SLN tissue and detects the presence of the Cytokeratin 19 (CK19) gene, a biological marker associated with breast cancer and not normally present in lymph node tissue. It is claimed that the RD 100i OSNA test will provide a result within a short time and therefore, can be used during breast surgery to determine if other lymph nodes should be removed at the same time as the initial tumour.

OSNA does not require the mRNA to be extracted and purified from the tissue before being analysed. The expression level of CK19 mRNA correlates with the size of the lymph node tumour cell foci. Since the foci may not be evenly distributed throughout the node, the system provides more accurate results if more of the node is analysed because there is less risk of tissue allocation bias (sample bias). The result is most accurate if the entire node is used, but then no follow-up histopathology is possible. The system can be used with half of the lymph node (one piece or alternate slices), allowing for the possibility of follow-up histopathology but potentially decreasing the accuracy of the results due to the increased risk of tissue allocation bias. The time to results is dependent on the number of lymph nodes analysed, but the test takes approximately 30 - 45 minutes. It includes both the time to prepare the lymph nodes, dissecting them out and trimming away fat, solubilising them and running the test. The OSNA test result is expressed both quantitatively and qualitatively; - for

lymph node negative test results, + (> 250 copies of CK19 m RNA / µl) for lymph nodes with a micrometastatic tumour burden and ++ (>5000 copies of CK19 m RNA / µl) for lymph nodes with a macrometastatic tumour burden. 250 copies of CK19 m RNA / µl is thus the threshold or cut-off level defining the tumour load in the SLNs above which further treatment with ALND is triggered.

The analyser amplifies and detects the CK19 mRNA by using 6 different primers which have been specifically designed to avoid the amplification of CK19 pseudogenes or their transcripts; amplification of these would lead to false positive results. Undesired amplification of genomic DNA is avoided by precipitation of DNA at low pH during sample preparation and the isothermal reaction temperature of 65°C.

The manufacturer estimates that 1% of breast tumours do not express CK19 mRNA and therefore, if cancer spreads to the lymph nodes from these tumours, CK19 mRNA will not be detected even though the lymph nodes are metastatic. Pre-screening of tumour biopsies for CK19 expression could be carried out before using the RD100i OSNA test to reduce the small risk of false negative results for sentinel lymph nodes with actual tumour cell foci.

A (Conformitée Européenne) CE mark has been obtained for this technology.

2.3.3 Metasin

The Metasin test is an intraoperative molecular test developed within the NHS at the Princess Alexandra Hospital in Harlow, Essex. The claims for its effect on management of patients with breast cancer are similar to those for OSNA . The test has similarities to a discontinued commercial test (Veridex Genesearch BLNA assay) and uses the technique of quantitative reverse transcriptase PCR (qRT-PCR) to detect two predictive markers of metastases, CK19 and mammaglobin. Mammaglobin is expressed mainly by breast epithelial cells and high levels of mammaglobin are associated with breast cancer. A reference gene, PBGD, is used to confirm the validity of the mRNA used in the test and two other controls, positive and negative, are also included. The test uses reagents that can be purchased from Roche and Qiagen and can be used on any platform (PCR machine). This in-house test differs from the discontinued commercial test by using distinctly different and unique primer-probe combinations to detect the CK19 and mammaglobin genes. The test is reported to take 26 minutes to results after 6-10 minutes for extracting and purifying mRNA from the tissue.

Although somewhat unclear, Metasin appears to be semi-quantitative according to Sundaresan.[41] The threshold is calculated by crossing point values (Cp) which occur when the fluorescence, from the DNA associated fluorescent probes, increases above background during amplification. The values for micrometastasis are quoted as Cp >25 and <32 for CK19 and Cp>25.9 and <32 for mammaglobin. Presumably, the values for macrometastasis are above these ranges.

Pre-screening of tumour biopsies for CK19 mRNA and mammaglobin mRNA expression could be carried out before using the Metasin test because like the CK19 biomarker, mammaglobin is not expressed in all breast tumours. The proportion of breast cancer tumours that do not express mammaglobin mRNA is not known.

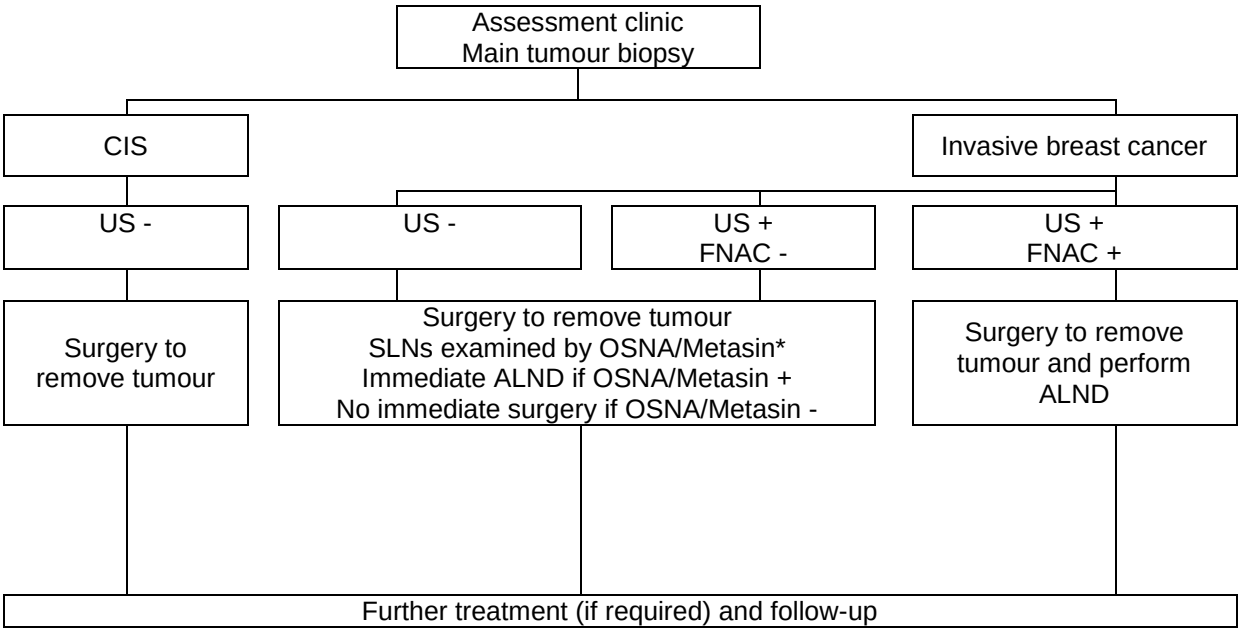
A (Conformitée Européenne) CE mark has very recently been obtained for this technology.

2.3.3.1 Proposed clinical pathway

Either OSNA or Metasin could be used as a replacement for current normal practice, in which case the SLNs are used in their entirety. However, OSNA or Metasin could also be used adjunctively where half of each lymph node (one piece or alternate slices) is used for the OSNA or Metasin intraoperative testing and the remaining half examined using standard postoperative histopathology. Where the new tests have been introduced in practice the first model is the most commonly followed as it maximises the claimed benefits of the new technology.

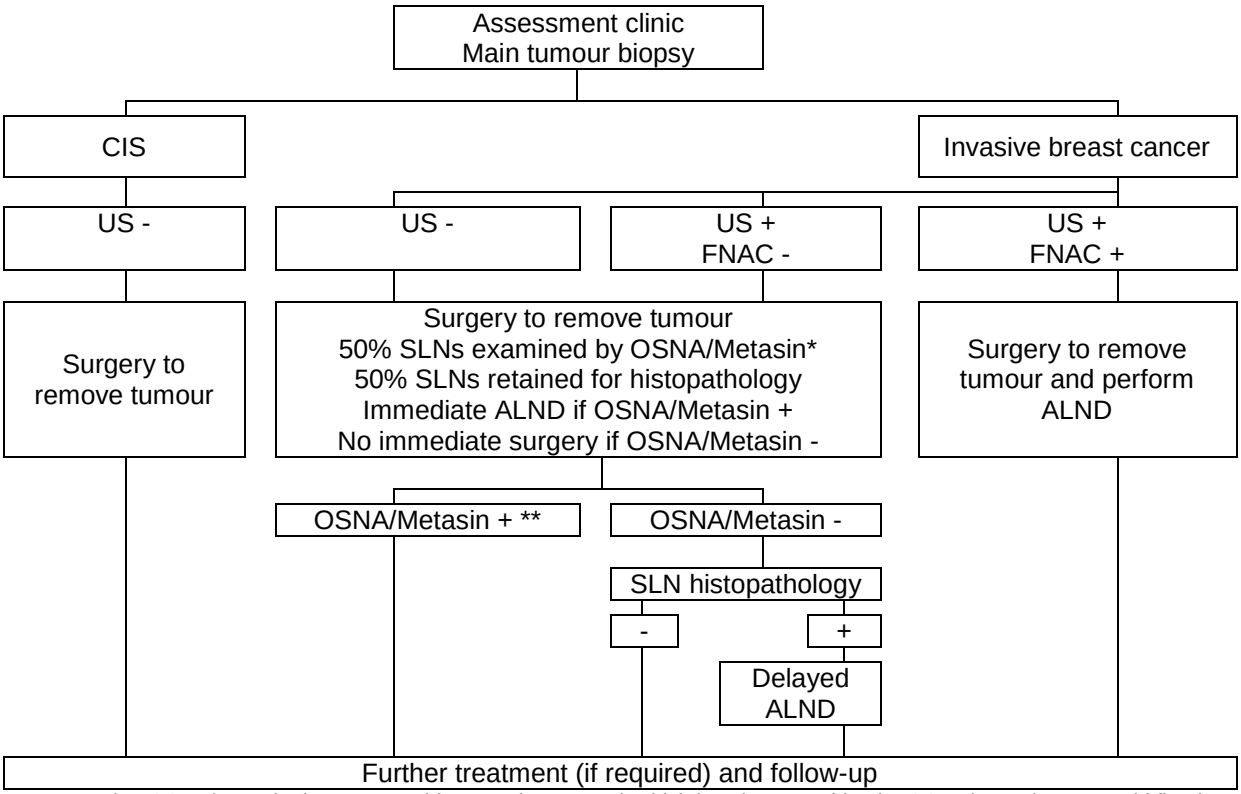
The proposed clinical pathways are illustrated in the figures below. This emphasizes that changes to the clinical pathway only occur to patients whose US is negative, or to patients whose US is positive but their FNAC negative.

Figure 5. Proposed clinical pathway detail for staging of breast cancer and subsequent surgery to the axilla – OSNA or Metasin used as a replacement for SLNB



* In case that OSNA/Metasin does not provide a result or a result which is uninterpretable, the OSNA/Metasin test would first be repeated using remaining solubilised LN. Failing this a 4 node sample could be done, followed by a delayed ALND in the case of positive LNs being identified in the sample, or an immediate ALND with the patient's prior informed consent for contingency.

Figure 6. Proposed clinical pathway detail for staging of breast cancer and subsequent surgery to the axilla – OSNA or Metasin used adjunctively



* In case that OSNA/Metasin does not provide a result or a result which is uninterpretable, the OSNA/Metasin test would first be repeated using remaining solubilised LN. Failing this histopathology will be done on the remaining 50% of the SLN.
** No histopathology done on SLN if OSNA/Metasin +. ALND sample will still be examined histopathologically as normal

2.3.4 Other technologies

There are two other pathological methods that can be used intraoperatively, frozen section and touch imprint cytology. Frozen section involves a section of the lymph node being snap-frozen, stained and sliced before being viewed by a consultant histopathologist. Touch imprint cytology involves the lymph node being sliced and the cut surface of the node imprinted on to a slide, which is then stained and viewed by a consultant histopathologist. Both intraoperative pathological methods can be used to determine if ALND needs to be performed at the same time as the first surgery and post-operative histopathology analysis is usually carried out to reduce the risk of a false negative result. However, in practice, these intraoperative methods are rarely used because they have low accuracy and pathology resources are very limited within the NHS. [42]

2.3.5 Measuring the accuracy of OSNA and Metasin

As already introduced earlier one aspect of the evaluation of new tests is measuring their accuracy, by calculating their sensitivity and specificity. This requires specification of the best available method of identifying the target condition of interest, the reference standard. In the case of the new technologies in question this is the true presence of breast cancer cells in the SLNs. The ideal reference standard is thus histopathological examination of the SLNs in their entirety. However, practically this is impossible because tissue is required by both the tests of interest and the reference standard. This leads to a compromise reference standard where LNs are split into several sections, often four, and alternate sections allocated to either the test whose accuracy is being evaluated or the reference standard. However as introduced earlier this means that tissue allocation or sampling bias will operate and so this needs to be carefully considered when interpreting the results of test accuracy studies. This can include a careful analysis of discrepant results to try to identify whether sampling is a potential explanation for apparent false negative or false positive results, recognising that there are other reasons for these. Limitations of histopathology as a reference standard given that all the lymph node sample can never be examined because of the finite number of slices that can be taken are important amongst these.

Further it is known that the histopathology of SLNs has a false negative rate relative to examining all lymph nodes removed in an ALND. Accepting use of histopathological examination of SLNs as the reference standard for OSNA and Metasin implies that these error rates of SLNB histopathology are also suffered by the new tests of interest. Whether

these errors could be avoided by the new tests of interest, could theoretically be investigated by using ALND findings as the reference standard, but the ethical considerations of exposing all OSNA/Metasin negative patients to the side-effects of ALND greatly reduce the acceptability of such an approach.

2.3.6 Implications for comparing the effectiveness and cost-effectiveness of OSNA and Metasin with current practice

The following report needs to extend its assessment beyond the accuracy of the new tests of interest, to their impact on patients and the health service. In the usual situation where there is no direct and rigorous research evidence on whether introducing OSNA or Metasin will lead to improved patient outcomes, a linked evidence approach using economic modelling is likely to be required. In this the likely consequences of errors in diagnosis are translated into outcomes. The estimates of sensitivity and specificity from accuracy studies are generally used to capture the difference in error rates between the new tests and current practice, so it is important that the new tests as used in the accuracy studies are similar to the way they will be used in practice and that the reference standard used in the accuracy studies is as close as possible to current diagnostic practice. The issues raised in the preceding section suggest that the similarity between current histopathological practice and the reference standard in the accuracy studies may require close attention. An immediate handicap however is that what constitutes average current practice with regard to histopathological examination of SLNs is difficult to define. As a consequence it may need to be accepted that although histopathological practice in research represents best practice, it is reasonable to consider that it is close enough to average current practice for the purpose of this report.

2.3.7 Main potential consequences of using OSNA and Metasin as compared with current practice

Although based on claims which require substantiation, (one of the main purposes of the report), the foregoing background suggests that main anticipated effects of introducing intraoperative OSNA or Metasin for those undergoing SLNB are:

- ALND will be performed as a single operation following immediately after primary tumour removal, rather than a separate second operation as in current practice

- This in turn may lead to reduced anxiety in patients who no longer have to wait to find out of they need a second operation and reduced time to adjuvant treatment, if this is required
- The reduced time to ALND may however complicate the decision making process of the MDT
- The adverse effects of ALND may be less where it is performed immediately after the primary tumour removal than if it is performed later as a second operation
- One rather than two operations may also lead to reduced hospital costs
- There will be increased costs associated with OSNA or Metasin, which will be off-set by reduced histopathology costs where OSNA or Metasin are used as replacement tests
- Any potential benefits above will be off-set if OSNA or Metasin introduces diagnostic errors, indicated by either its sensitivity or specificity being less than 100%
- If false-negatives are introduced, women with macro- or micrometastases will be misidentified as SLNB negative, and they will not undergo ALND. This may compromise their outcome with respect to breast cancer
- If false-positives are introduced, women who are SLNB negative will be misidentified as having macro- or micrometastases, with any resulting side-effects but without any benefit in outcome with respect to breast cancer

The economic model will need to attempt capture all of the above potential consequences.

3 Definition of the decision problem

The question addressed by this health technology assessment is as set out in the final scope published by NICE, and is reproduced here for reader convenience.

A protocol was developed a priori by the authors to address the decision problem. The aspects of this involving systematic review were registered on PROSPERO registration number CRD42012002889.

The methods used to address specific aspects of the decision problem are detailed at the beginning of each of the relevant chapters which follow.

3.1 Decision question

Are the RD100i OSNA system and any alternative technologies identified during scoping, clinically effective and cost effective if used in the NHS in England?

3.1.1 Population

Individuals with invasive breast cancer who undergo a sentinel lymph node biopsy.

3.1.2 Intervention

- The RD100i OSNA system using a whole node sample.
- The RD100i OSNA system using a half node sample with postoperative histopathology confirmation.

3.1.3 Alternative diagnostic technologies

- The Metasin test using a whole node sample (Intraoperative in-house molecular test developed at Princess Alexandra Hospital, Harlow, Essex).
- The Metasin test using a half node sample with postoperative histopathology confirmation.

3.1.4 Comparators

Post-operative standard histopathology alone.

3.1.5 Health care setting

Secondary and tertiary care settings.

3.1.6 Health outcomes

3.1.6.1 Clinical considerations

The intermediate measures for consideration include:

- Diagnostic test accuracy
- Test failure rate
- Discordant test results
- Time to test result
- Duration of anaesthesia

The clinical outcomes for consideration include:

- Patient anxiety associated with waiting time for result and not knowing extent of surgery prior to operation
- Number of repeat operations (except for re-excision of positive margins)
- Time in operating theatre
- Time to start and nature of adjuvant therapy
- Morbidity and mortality from biopsies, axillary dissections, first and second operations and treatment of cancer
- Adverse events from false test results including patient distress and sequelae.

Data on these outcomes are likely to be used along with clinical utility scores to estimate Quality-Adjusted Life Years (QALYs).

3.1.6.2 Cost considerations

The cost analysis will be based on the UK NHS setting and comprise both NHS and Personal Social Services (PSS) costs.

The costs for consideration include:

- Cost of equipment, any additional tests (pre-screening), reagents and consumables
- Staff and training of staff
- Maintenance of equipment
- Costs associated with surgeon time and the management of operating theatre time
- Medical costs arising from on-going care following test results including those associated with surgery, time spent in hospital, and treatment of cancer.

- Medical costs arising from adverse events including those associated with biopsies, surgery, cancer treatment and false test results.

The cost of the hardware for the RD100i OSNA system is approximately £70,000 (excluding VAT). The consumable cost is approximately £150 - £250 per patient (excluding VAT). This consumable cost is dependent upon the number of tests performed per theatre day and the number of patient samples tested. The maintenance cost is £6,180 per annum (excluding VAT) following the expiry of the 1 year warranty.

4 Assessment of clinical effectiveness

4.1 Methods for reviewing effectiveness

The diagnostic accuracy of the tests OSNA and Metasin was assessed by a systematic review of research evidence. The review was undertaken following the principles published by the NHS CRD.[43]

4.1.1 Identification of studies

The following bibliographic databases were searched in this review: Medline, Medline in process and Embase (all via OVID), Web of Science (including conference proceedings, via ISI), the Cochrane Library (all) and HEED (via the Cochrane Collaboration). The searches did not use any form of limit (e.g. date). See Appendix 3 for details.

The following trials registries were also searched: NIH ClinicalTrials.gov, Current Controlled Trials, WHO International Clinical Trials Registry Platform (ICTRP), EU Clinical Trials Register. Google was also searched to identify grey literature and conference publications.

Items included after full-text screening were forward citation chased using Web of Science (Thompson Reuters).

Searches were de-duplicated and managed using Endnote (X5).

Relevant studies were then identified in two stages. Titles and abstracts returned by the search strategy were examined independently by two researchers (TJH and HC) and screened for possible inclusion. Disagreements were resolved by discussion. Full texts of the identified studies were obtained. Two researchers (TJH and HC) examined these independently for inclusion or exclusion, and disagreements were again resolved by discussion.

4.1.2 Inclusion and exclusion criteria

4.1.2.1 Population

Studies of individuals with invasive breast cancer who underwent a (sentinel) lymph node biopsy during the primary operation to excise a suspected breast cancer were included.

4.1.2.2 Interventions and comparators

Studies of OSNA or Metasin as used at the thresholds recommended by the manufacturer or designer were included.

The reference standard was post-operative histopathology, performed on fresh sections of tissue.

Frozen section and touch imprint cytology were excluded as comparators as they were not felt to be sufficiently feasible for widespread implementation (or intervention).

4.1.2.3 Outcomes

No study was excluded on the basis of outcomes, provided it appeared relevant to those listed in the decision problem.

- Test failure rate
- Diagnostic test accuracy
- Discordant test results
- Time to test result
- Duration of anaesthesia/ time in operating theatre
- Number of repeat operations (except for re-excision of positive margins)
- Time to start and nature of adjuvant therapy

The clinical outcomes for consideration include:

- Patient anxiety associated with waiting time for result and not knowing the extent of surgery prior to operation
- Adverse events from false test results including patient distress and sequelae
- Morbidity and mortality from biopsies, axillary dissections, first and second operations and treatment of cancer

4.1.2.4 Study design

For the review of test accuracy, the protocol made provision for all study designs unless evidence on the intervention and outcome of interest was already available from designs less open to bias as judged with reference to standard hierarchies of evidence.

Systematic reviews were used as a source for finding further studies and to compare with our systematic review. For the purpose of this review, a systematic review was defined as one that has:

- a focused research question
- explicit search criteria that are available to review, either in the document or on application
- explicit inclusion/exclusion criteria, defining the population(s), intervention(s), comparator(s), and outcome(s) of interest
- a critical appraisal of included studies, including consideration of internal and external validity of the research
- a synthesis of the included evidence, whether narrative or quantitative

Studies were excluded if they did not match the inclusion criteria, and in particular were:

- pre-clinical and animal
- reviews, editorials and opinion pieces
- case reports
- studies with <10 participants

Beyond this, no study design was excluded unless evidence on the intervention and outcome of interest is already available from study designs less open to bias as judged with reference to standard hierarchies of evidence.[43]

4.1.3 Data extraction strategy

Data were extracted by one reviewer (TJH) using a standardised data extraction form in Microsoft Access 2010 and checked by a second reviewer (HC). Disagreements were resolved by discussion, with involvement of a third reviewer if necessary. Data extraction forms for each included study can be found in Appendix 4.

4.1.4 Critical appraisal strategy

The methodological quality of the studies was assessed, where applicable to the design of the study, according to criteria specified by the Cochrane Collaboration's Tool for Assessing Risk of Bias.[43] The QUADAS-2 was used for test accuracy studies.[2]

Quality was assessed by one reviewer and judgements were checked by a second. Any disagreement was resolved by discussion, with involvement of a third reviewer as necessary. The two instruments are summarised below. Results were tabulated and the relevant aspects described in the data extraction forms.

4.1.4.1 Internal validity

The instruments sought to assess the following considerations:

Cochrane Collaboration's Tool for Assessing Risk of Bias

- Was the allocation sequence adequately generated?
- Was allocation adequately concealed?
- Was knowledge of the allocated intervention adequately prevented during the study?
- Were incomplete outcome data adequately addressed?
- Are reports of the study free of suggestion of selective outcome reporting?
- Was the study apparently free of other problems that could put it at high risk of bias?

QUADAS-2

- Description of patient selection
- Was a consecutive or random sample of patients enrolled?
- Was a case-control design avoided?
- Did the study avoid inappropriate exclusions?
- Could the selection of patients have introduced bias?
- Are there concerns that the included patients do not match the review question
- Description of index and reference tests
- Was the index test assessor blind to the results of the reference standard and vice versa?
- Was a threshold pre-specified?
- Could the conduct or interpretation of the index test or reference standard have introduced bias?
- Are there concerns that the conduct or interpretation of the question have introduced bias for the index test or reference standard?
- Is the reference standard likely to classify the target condition?
- Description of patient flow and timing
- Did all patients receive a reference standard and was it the same test for each?
- Were all patients included in the analysis?
- Could the patient flow have introduced bias?

4.1.4.2 External validity

External validity was judged according to the ability of a reader to consider the applicability of findings to a patient group and service setting. Study findings can only be generalisable if

they provide enough information to consider whether a cohort is representative of the affected population at large. Therefore studies that appeared to be typical of the UK breast cancer population with regard to these considerations were judged to be externally valid.

4.1.5 Methods of data synthesis

Details of the extracted data and quality assessment for each individual study are presented in structured tables and as a narrative description. Any possible effects of study quality on the effectiveness data are discussed. Data on test accuracy are presented as sensitivity, specificity and concordance, where available.

4.1.5.1 Meta-analysis

Meta-analysis of diagnostic test accuracy was performed using the bivariate method[44] implemented in Stata/SE 12.1[45] using the command `metandi`[46]. Studies were only included in the meta-analysis if the numbers of true positives, true negatives, false negatives and false positives were all reported in the text or could be unambiguously inferred from other figures in the text. Meta-analysis using the full bivariate method was not performed where there were fewer than four included studies as the model cannot generally be estimated with fewer than four studies. Where the full bivariate model could not be estimated (either due to insufficient studies or other convergence errors) we reduced the complexity of the model (as done in e.g., Pennant et al. 2010)[47] by setting the correlation parameter to zero (effectively reducing the model to two independent univariate random effects analyses) and performing the analysis directly using the Stata command `xtmelogit`.

The bivariate method, when calculated using maximum likelihood estimation and without covariates, is equivalent to the hierarchical summary receiver operating characteristic (HSROC) model [48-50] and this can be used to provide a summary receiver operating characteristic (SROC) curve and prediction region as well as a summary estimate and confidence region of sensitivity and specificity.

.The SROC curve is designed to show how sensitivity and specificity are traded off against each other in different studies, through variation of the positivity threshold. If, and only if, there is reason to believe the positivity threshold might vary between studies we provide a SROC curve and prediction region.

4.1.6 Interpreting the results from the diagnostic studies

4.1.6.1 Test accuracy

In most of the studies, the accuracy of the interventions has been evaluated against the reference (gold) standard of post-operative histopathology. The results are generally reported as follows:

- Sensitivity – True Positive/(True Positive + False Negative). This is the probability of detecting the presence of metastases in someone with metastases.
- Specificity – True Negative/(False Positive + True Negative). This is the probability of not detecting metastases in someone without metastases. In this instance a high specificity is required to avoid unnecessary ALND.
- Positive predictive value (PPV) – True positive/(True Positive + False Positive). This is the probability of someone with a positive result actually having metastases.
- Negative predictive value (NPV) – True Negative/(True Negative + False Negative). This is the probability of someone with a negative test result actually not having metastases.
- Accuracy or concordance with reference standard – (True Positive +True Negative)/(True Positive +True Negative + False Positive + False Negative). This is the percentage of test results correctly identified by the test, i.e. the rate of agreement with the reference standard
- Discordance – cases of disagreement between the reference and index test

4.1.6.2 Discordant case analysis

Many studies used a single gate design, whereby in a single sample of individuals with unknown metastatic status, portions of the same node were allocated to either a molecular assay or histopathology. However, due to the spatial distribution of metastasis within a lymph node, and the use of different parts of the lymph nodes for different diagnostic tests, tissue allocation bias (TAB) may occur (i.e. discordant results between tests may be due to genuine differences between the tissue samples). Unfortunately, it is not possible to use the same tissue for both tests; the tissue used for the molecular assays cannot be used for histology, and the formalin-fixed and paraffin-embedded tissue required for permanent sections is not suitable for quantitative mRNA measurements.

Some studies have attempted to address this issue by performing extensive histopathology and molecular techniques on discordant cases, in order to ascertain whether the results are

true and therefore due to differences in allocated tissue. Results are then adjusted (discordant cases due to TAB are removed).

4.2 Results of test accuracy

The results of the assessment of clinical effectiveness will be presented as follows:

1. An overview of the quantity and quality of available evidence together with a table summarising all included trials (Table 8, page 59), a table of patient characteristics (Table 9, page 65) and a summary table of key quality indicators (Table 10, page 69)
2. A critical review of the available evidence for each of the stated research questions covering:

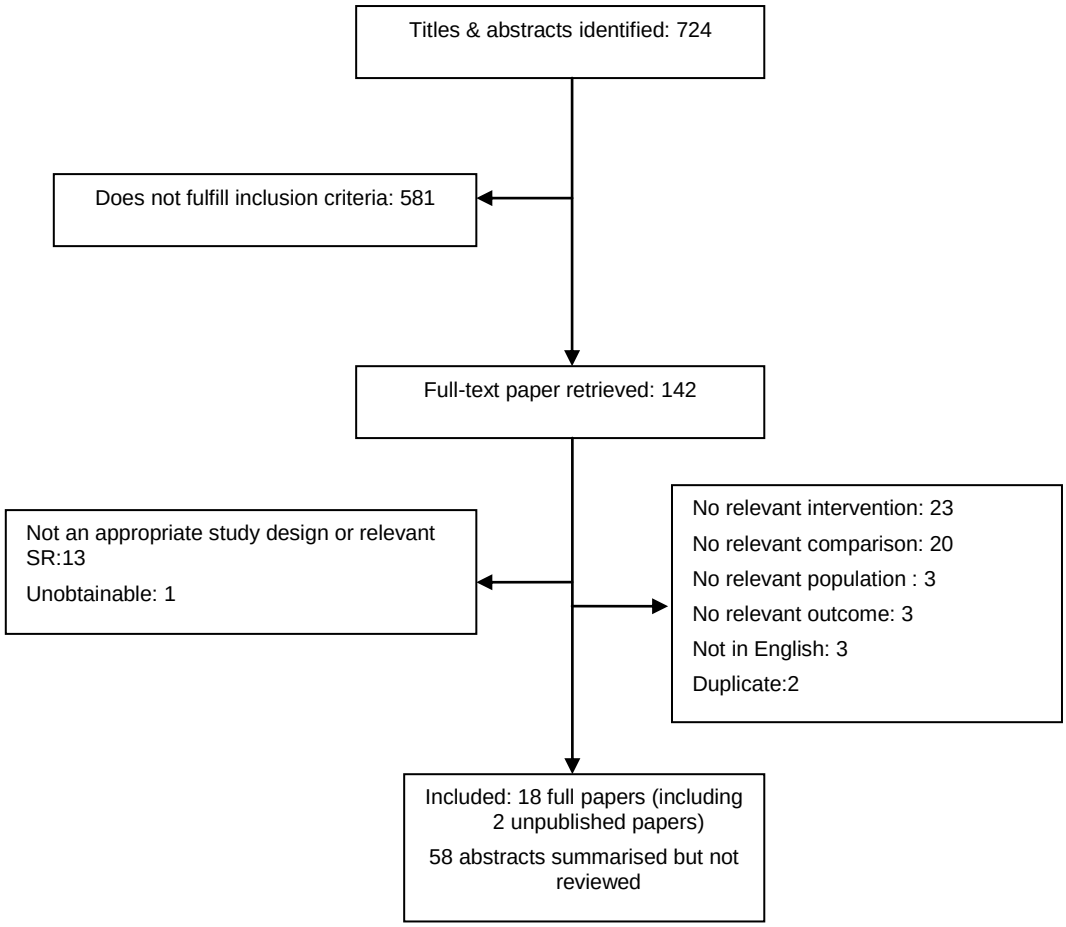
    - the quantity and quality of available evidence,
    - a summary table of the study characteristics,
    - a summary table of the population characteristics,
    - study results in terms of sensitivity, specificity and discordant case analysis presented in narrative and tabular form,
    - comparison of the results in terms of time to analysis
    - a summary table of abstracts identified, but not included in the review

4.2.1 Quantity and quality of research available

4.2.1.1 Number of studies identified.

The electronic searches retrieved a total of 665 titles and abstracts. Fifty nine additional papers were found by searching the bibliographies of included studies and by forward chasing. A total of 581 papers were excluded, based on screening title and abstract. Full text of the remaining 143 papers was requested for more in-depth screening, to give a total of 16 published and two unpublished papers included in the review. The process of study selection is shown in Figure 7.

Figure 7. Summary of study selection



4.2.1.2 Number of studies excluded

Papers were excluded for at least one of the following reasons: duplicate publications, narrative reviews, and publications (systematic reviews and individual studies) not considering relevant intervention, population, comparison or outcomes. The bibliographic details of studies retrieved as full papers and subsequently excluded, along with the reasons for their exclusion are detailed in Appendix 5.

4.2.1.3 Number and description of included studies

Seventeen test accuracy studies were included for OSNA, with two unpublished papers for Metasin which have been assessed using the STARD criteria below.[41,51]

Table 7. STARD assessment for Metasin papers

Section and Topic	STARD criteria	Sundaresan	McDowell
TITLE/ABSTRACT/ KEYWORDS	Identify the article as a study of diagnostic accuracy.	✗	[redacted]
INTRODUCTION	State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups.	✓	[redacted]
METHODS			
Participants	Describe the study population: The inclusion and exclusion criteria, setting and locations where the data were collected.	✗	[redacted]
	Describe participant recruitment: Was recruitment based on presenting symptoms, results from previous tests, or the fact that the participants had received the (evaluated) index tests or the (golden) reference standard?	✗	[redacted]
	Describe participant sampling: Was the study population a consecutive series of participants defined by the selection criteria in items 3 and 4? If not, specify how participants were further selected.	✗	[redacted]
	Describe data collection: Was data collection planned before the index test and reference standard were performed (prospective study) or after (retrospective study)?	✓	[redacted]
Test methods	Describe the reference standard and its rationale.	✓	[redacted]
	Describe technical specifications of material and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard.	✗	[redacted]
	Describe definition of and rationale for the units, cut-offs and/or categories of the results of the index tests and the reference standard.	✗	[redacted]
	Describe the number, training and expertise of the persons executing and reading the index tests and the reference standard.	✗	[redacted]
	Describe whether or not the readers of the index tests and reference standard were blind (masked) to the results of the other test and describe any other clinical information available to the readers.	✗	[redacted]
Statistical methods	Describe methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g. 95% confidence intervals).	✓	[redacted]
	Describe methods for calculating test reproducibility, if done.	✗	[redacted]
RESULTS			
Participants	Report when study was done, including beginning and ending dates of recruitment.	✗	[redacted]
	Report clinical and demographic characteristics of the study population (e.g. age, sex, spectrum of presenting symptoms, co morbidity, current treatments, recruitment centers).	✗	[redacted]
	Report the number of participants satisfying the criteria for inclusion that did or did not undergo the index tests and/or the reference standard; describe why participants failed to receive either test (a flow diagram is strongly recommended).	✗	[redacted]
Test results	Report time interval from the index tests to the reference standard, and any treatment administered between.	N/A	[redacted]
	Report distribution of severity of disease (define criteria) in those with the target condition; other diagnoses in participants without the target condition.	✗	[redacted]
	Report a cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard.	✓	[redacted]
	Report any adverse events from performing the index tests or the reference standard.	✗	[redacted]
Estimates	Report estimates of diagnostic accuracy and measures of statistical uncertainty (e.g. 95% confidence intervals).	✓ (no confidence intervals)	[redacted]
	Report how indeterminate results, missing responses and outliers of the index tests were handled.	✗	[redacted]
	Report estimates of variability of diagnostic accuracy between subgroups of participants, readers or centers, if done.	✗	[redacted]
	Report estimates of test reproducibility, if done.	✗	[redacted]
DISCUSSION	Discuss the clinical applicability of the study findings.	✓	[redacted]

The search also identified 58 abstracts, some of which repeated the data in the full papers, others provided supplementary information and others again were not associated with a full paper. The data in the abstracts have been compiled as a table in Appendix 6. However, due to the lack of accompanying details, no quality assessment has been performed. All included citations are detailed in Table 8.

Table 8. Summary information of included test accuracy studies

Author	Year	Patients (n)	SLN or ALN (n)	Centre	Design	Outcomes
			Metasin			
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Sundaresan[41]	Unpub	1265	2279 SLN	Multicentre, UK	Single gate	Test accuracy, time to analysis
			OSNA			
Bernet, L.[52]	2011	NR	55 SLN	Multicentre, Spain	Observation	Time to analysis
Bernet Vegue, L. [53]	2012	55	567 Non-SLN	Multicentre, Spain	Single gate	Test accuracy
Castellano, I.[54]	2012	279	Unclear	Turin, Italy	Cohort	Test accuracy , non-SLN involvement
Choi, Y.[55]	2010	199	284 SLN	Seoul, Korea	Single gate	Test accuracy
Feldman, S.[56]	2011	496	1044 SLN	Multicentre, USA	Single gate	Test accuracy
Godey, F.[57]	2012	723	Unclear	Rennes, France	Cohort	Test accuracy
Guillen-Paredes, M.P.[58]	2011	80	114 SLN	Murcia, Spain	Cohort	Operating time, days in hospital, costs.
Khaddage, A.[59]	2011	46	80 SLN	Saint-Etienne, France	Single gate	Test accuracy
Le Frere Belda, M.A.[60]	2011	233	503 SLN	Multicentre, France	Single gate	Time to analysis, test accuracy
Osako, T. [61]	2011	183	Non-SLN	Cancer Institute Hospital, Tokyo, Japan	Cohort	Test accuracy
Schem, C.[62]	2009	93	343 ALN	University Clinic of Schleswig-Holstein, Albertinen Hospital, Germany	Single gate	Test accuracy
Snook, K. L.[63]	2011	204	395 SLN	Multicentre, UK	Single gate	Test accuracy , time to analysis
Tamaki, Y.[64]	2009	198	674 ALN+SLN	Multicentre, Japan	Single gate	Test accuracy
Tamaki, Y.[65]	2012	417	775 SLN	Multicentre, Japan	Single gate	Test accuracy
Tsujimoto, M. [66]	2007	101	325 ALN + SLN	Multicentre, Japan	Single gate	Test accuracy
Visser, M.[67]	2008	32	346 ALN	Alkmaar and Amsterdam	Single gate	Test accuracy

SLN sentinel lymph node, ALN axillary lymph node, non-SLN non-sentinel lymph node, NR not reported, single gate – a study design in which only patients with the target condition are recruited and receive both the index test and reference standard

4.2.1.4 Study characteristics

OSNA and Metasin have standard procedures and thresholds, therefore unless otherwise stated, the included studies have complied with manufacturer's instructions. Both methods are semi-quantitative and differentiate between micro and macro metastases, although McDowell only reports positive and negative results.[51] (McDowell)

In contrast to the molecular methods, there is some heterogeneity with regard to the reference standard, particularly with the number of levels examined, for example, one level analysis will involve analysis of one section of node, whereas five-level analysis will examine five sections. As such, for one level histopathology, it is likely that macrometastases will be identified, but micrometastases may be missed. Micrometastasis is considered to be >0.2mm and <2mm and macrometastasis >2mm. Since there is no way to analyse the whole node, this method cannot be 100% sensitive.

Some studies report cases of isolated tumour cells (ITC) with histopathological analysis. These will fall below the threshold for the OSNA and Metasin and are generally considered lymph node negative, since their clinical significance is unknown.(59)

The majority of included studies comply with a single gate design, a single sample of individuals with unknown metastatic status was assessed by both the diagnostic test under scrutiny and the reference standard. No studies utilising a two-gate design, where the test under scrutiny is performed on a sample that includes individuals with known metastatic status (using the reference standard), were identified. However, three cohort studies have been included, where different patient populations were utilised for each test. The inclusion of cohort studies enabled the identification of data based upon whole node analysis, something that would not be possible for OSNA or Metasin using classic diagnostic test accuracy study designs, such as single-gate (patients with only the target condition are recruited), or two-gate (two sets of patients are recruited, one with the target condition and one without).

A general issue for all of the studies is that a portion of node tissue is allocated to the index test and a second portion is allocated to the reference standard. As such, the tests are analysing different tissue, which cannot be reused between tests. Since metastases may be distributed unevenly, tissue allocation bias (TAB) may occur i.e. metastases may exist in the tissue provided for one test, but not in the other.

It should also be noted that studies examining axillary lymph nodes as well as, or instead of, sentinel lymph nodes were included.

4.2.1.4.1 Metasin

There is currently no published evidence for the test accuracy of Metasin, however, we have received drafts of papers reporting on two single gate studies.

Sundaresan et al. report on a postoperative evaluation of Metasin using 2279 nodes from 1265 patients with breast cancer, where 608 of the 1265 cases were from patients already assessed by Genesearch-BLN.[41] Six centres contributed tissue homogenates, although 2 centres were only able to provide frozen RNA. Three level histology, i.e. examination of three sections, was performed although the authors state that all laboratories followed their own SOPs for analysis of slides. The authors also mention that there was no uniformity of the protocols followed for sentinel and axillary clearance.

[redacted]

4.2.1.4.2 OSNA

Cohort studies

An assessment of the reliability of OSNA as a single test on SLNs was reported by Castellano et al.[54] This was a single centre cohort study with 279 patients. Histology was performed on fresh SLNs sliced at 2 mm, then step sections at 100 µm. Positive rates for both cohorts were presented.

A cohort study reported by Godey et al. compares the positivity rate of OSNA performed in a routine clinical setting, to historical postoperative histology results for 723 patients.[57] Histology examination consisted of 250 µm sections of the node, until no tissue was left.

Guillen Parades et al. present a retrospective cost-benefit analysis.[58] This cohort study aimed to analyse the economic costs of the intraoperative OSNA assay compared to the conventional postoperative histological and immunohistochemical assay. Histology was performed in 4mm sections of the node. Results include operative time, days in hospital, costs and post-operative complications.

Osako et al. reports on a retrospective cohort study of 183 patients at a single centre.[61] Intra-operative OSNA was compared with one level histology (examination of one slice) performed on 5 to 7 non-SLNs from the same patient. Positive rates were described for both methods.

Single gate studies

The aim of a multicentre study presented by Bernet et al. was to compare OSNA with histology and evaluate its feasibility intra-operatively.[52] Fifty five SLNs were investigated via a single gate study design, however, the results appear to include touch imprint and frozen section analysis supported by post-operative histopathology. Therefore these results are not included in the review. Relevant outcomes which are included in the review are time of extraction, dissection, preparation and analysis.

A second paper by Bernet et al. reported on the Breast Complete Lymphadenectomy OSNA Study for Enhanced Review-I (B-CLOSER-I).[53] Eight hospitals were involved in this single gate study comparing histopathology and OSNA for the pathologic staging of ALNs after identification of positive SLNs in patients with primary breast cancer.

Fifty five patients were recruited consecutively, providing 567 non-SLNs for analysis. Both OSNA and histopathology were performed post-operatively. Tissue used for OSNA analysis was stored at -80^{o}C. Two phases are reported, both utilising a single gate design: the validation phase and routine use. For the validation phase, histopathology was five level (examination of five slices), whereas for routine use a central 1mm slice was used for histopathology (one level), with the remainder being used for OSNA. For both phases, discordant cases were investigated using additional molecular analysis, performed by quantitative reverse transcriptase – polymerase chain reaction (qRT-PCR). Discussed outcomes include discordance during the validation phase.

The sensitivity, specificity and accuracy of intra-operative OSNA, as compared to three level histology, are presented by Choi et al. where 199 patients had 284 SLNs analysed.[55] With regard to discordance, clinical information, status of non-SLNs and expression of CK19 protein in lymph node metastasis foci were evaluated on a patient basis.

A multicentre single gate study reported by Feldman et al. recruited 496 patients to give 1044 SLNs for analysis.[56] The node was sectioned into 6, with alternate slices for histopathology and OSNA. The slices for histopathology were then dissected at 200 µm intervals for H&E staining and pan-CK. The results were evaluated by blinded pathologists. For discordant case analysis, blank histopathology was checked and OSNA re-tested with Western Blot and qRT-PCR.

Khaddage et al report on a multicentre, single gate study based in France, with concordance, sensitivity and specificity as outcomes.[59] A validation phase with 46 patients and a routine phase with 197 patients was performed. Histopathology for the validation

phase was five level and the routine phase, one level. Both node and patient level analysis are presented. The results of the OSNA investigation were not known to the histopathologist and vice versa. Discordant case analysis of the validation phase was performed by qRT_PCR. (ref)

Le Frere-Belda et al. present a study to assess the intra-operative diagnostic performance of OSNA versus extensive histological evaluation.[60] This was a multicentre single gate study in 8 French clinical centres. Alternate slices of dissected SLNs were used, along with five level histology. It should be noted that 2 centres re-used frozen section samples which may impair integrity for final histology.

In cases of discordance, with positive OSNA and negative histology, 200 µm skip spaces for all slices were analysed. Samples were also sent to Sysmex for blind molecular analysis by qRT-PCR.

A study by Schem et al. considers the performance of OSNA compared to 5 level histology.[62] This two centre study was a blinded, single gate, experimental design, using 343 ALNs. OSNA samples were stored frozen at -80^{o}C. Discordant samples were analysed by further levels of histology, Western Blot and qRT-PCR. In addition, 120 histopathologically negative samples were cut into further levels for specificity. Sensitivity and discordance were also reported.

Snook et al. report on a UK single gate prospective multicentre evaluation of OSNA involving 4 centres.[63] Two hundred and four patients were recruited, providing 395 SLNs although there are no further details on recruitment. OSNA slices were snap-frozen at -80^{o}C. Five-level histology was performed and molecular analysis for discordant samples was performed by qRT-PCR. Sensitivity, specificity, accuracy, PPV, NPV and discordance were reported.

A Japanese multicentre study of two single gate trials is reported by Tamaki 2009, examining the validity of OSNA for clinical use.[64] Seven centres were involved, two in trial 1, three in trial 2 and two in both. Trial 1 compared intra-operative OSNA with detailed histology for detection of metastasis of 124 ALNs. Alternate slices of the node were allocated to each technology. For histology, the tissue was dissected in 0.2mm sections. Trial 2 was designed to replicate routine use, so only one-level histology was performed. For discordant cases, however, histology was performed as per trial 1, alongside Western Blotting analysis. Sensitivity, specificity and discordance are reported.

The same trial was reported by Tsujimoto et al., who present data from 6 centres for 101 patients and 325 ALNs.[66] Intra-operative OSNA was compared to three-level histology.

Results were examined by three third party (blinded) pathologists. The authors state that calculations for sensitivity and specificity were not appropriate because separate tissue was used for both tests. Therefore the results were evaluated as concordance with histology. Discordant samples were analysed by qRT-PCR and Western Blot.

A subsequent Japanese multicentre study by Tamaki et al. investigated clinical use of OSNA compared to one level histology.[65] This was a single gate design involving 198 patients and 674 ALNs.

Finally, Visser et al. report on a single gate study, testing OSNA on 346 ALN as compared to 5 level histology.[67] To investigate whether results were influenced by sampling bias, the histologic work up was extended to all levels in the first 120 histologically negative lymph node samples (as were paraffin blocks of discordant cases). In addition, the homogenised lymph node lysates of samples were subjected to qRT-PCR and Western Blot analysis.

4.2.1.5 Population characteristics

In general, patient characteristics were poorly reported, as were inclusion and exclusion criteria. Not all studies provided age range and often, only information on tumour staging was provided. Comparable characteristics are presented in Table 9, page 65.

4.2.1.5.1 Metasin

[redacted]Sundaresan et al. provide no details on patient characteristics.[redacted][41]

4.2.1.5.2 OSNA

Two of the cohort studies reveal reasonably homogenous patient populations with regard to age and clinical status.[54],[61] A relatively small sample size was used for each patient group in the study reported by Guillen-Parades, with a difference of ~ 7 years for mean age.[58] The histology group had proportionally more T2 than OSNA.

Across many of the single gate studies, minimal information is given regarding population characteristics. In general, the populations are heterogeneous, with studies including patients across the spectrum of tumour and nodal staging, whereas other studies have included only one or two stages. It should also be noted that for some studies, the number of recruited patients is small, whereas the number of nodes analysed is comparatively large.

Table 9. Summary of patient characteristics

	Bernet Vegue[53]	Castellano[54]		Choi[55]	Feldman[56]	Frere Belda[60]	Godey[57]		Guillen-Parades[58]		Khaddage[59]	
Intervention	O+H	O	H	O+H	O+H	O+H	O	H	O	H	O+H	O+H
No.	55	110	169	199	496	233	258	355	35	45	46	197
Median age, yrs (range)	59 (23-87)	66.7 (38-82)	61.2 (23-86)	49.9[b]	58.8[c] (28-88)	58 (30-93)	56.8[c]	56.9[c]	55.54[c]	61.89[c]		
Clinical stage (%)												
0				11 (5.5)		41 (17.7)						
I	21 (38.2)			132 (66.3)		175 (75.4)						
II	22 (40)			54 (27.1)		13 (5.6)						
III	12 (21.8)			2 (1.0)		2 (0.9)						
IV				11 (5.5)		1 (0.4)						
Tumour size (%)[a]												
<1 cm		33 (30)	41 (24)									
1.1-1.5 cm	44 (80.0)	19 (17)	45 (27)									
>1.5 cm	7 (12.7)	58 (53)	83 (49)									
Clinical tumour classification	1 (1.8)											
T0	3 (5.5)										7	1
Tis				8 (4.0)	21 (4.2)				2	3	0	21
T1	49 (89.1)			129 (64.8)	327 (65.9)				16	13	34	141
T2	6 (10.9)			56 (28.1)	124 (25)				17	29	2	30
T3	55			2 (1.0)	5 (1)						2	1
T4	59 (23-87)										1	1
Tx				4 (2.0)	19(3.8)							
Nodal status (%)												
pN0	21 (38.2)			153 (76.9)	387 (78)	225 (97.0)						
pN1	22 (40)			37 (18.6)	84 (16.9)	7 (3.0)						
pN2	12 (21.8)			5 (2.5)	14 (2.8)							
pN3				4 (2.0)	4 (0.8)							
Histopathologic type (%)												
IDC		81 (74)	109 (64)	165 (82.9)	348 (70.2)	164 (70.4)	212	313	31	37	36	148
ILC	44 (80.0)	16 (14)	29 (17)	9 (4.5)	40 (8.1)	34 (14.6)	46	42	2	5	5	16
DCIS	7 (12.7)			9 (4.5)		23 (9.9)			1	2	5	21
Other	1 (1.8)	13 (12)	31 (18)	16 (8.1)	109 (21.7)	12 (5.2)			1	1	0	12
HER2 (%)	3 (5.5)											
Negative		108 98)	144 (85)									
Positive	49 (89.1)	2 (2)	25 (15)			13 (6.3)						

[a]reported as mm, corrected, [b]42.2%distribution, [c] mean

	Osako[61]		Schem[62]	Snook[63]	Tamaki 2009[64]		Tamaki 2012[65]	Tsujimoto[66]	Vegue[52]	Visser[67]
Intervention	O	H	O+H	O+H	O+H T1	O+H T2	O+H	O+H	O+H	O+H
No.	119	64	93	204	36	185	439	101	55	32
Median age, yrs (range)	53 (27-86)	56 (39-81)			55.9[c]	54.7[c]	56.1[c] (25-90)		59 (23-87)	
Clinical stage (%)										
0					2 (6)	14 (9)		5		0
I					8 (24)	51 (31)	183 (43.9)	41	21 (38.2)	8
II					14 (41)	64 (40)	110 (26.4)	49	22 (40)	15
III					5 (15)	7 (4)	70 (16.8)	5	12 (21.8)	7
IV					0	0		1		2
Unknown					2 (6)	0	54 (12.9)	5		
Tumour size (%)[a]										
<10 mm										
1.1-1.5 cm										
>1.5 cm										
Clinical tumour classification (%)										
T0										
Tis							50 (12)			
T1			46	133			254 (60.9)			
T2			36	60			111 (26.6)			
T3			4	5			2 (0.5)			
T4			7							
Nodal status (%)										
pN0			46					60		14
pN1	115 (96.6)	62 (96.9)	27					35		10
pN2	3 (3.4)	2 (3.1)	13					2		6
pN3			7					4		2
Histopathologic type (%)										
Invasive ductal carcinoma	110 (92.4)	57 (89.1)	68	160 (78.8)	32 (94)	130 (79)	305 (73.1)	87	44 (80.0)	30
Invasive lobular carcinoma	4 (3.4)	2 (3.1)	21	22 (10.8)	1 (3)	7 (4)	24 (5.8)	4	7 (12.7)	2
Ductal carcinoma in situ					0	18 (11)	53 (12.7)	5	1 (1.8)	
Other			4	16 (7.9)	1 (3)	9 (5)	35 (8.4)	5	3 (5.5)	
HER2 (%)										
Negative	106 (89.1)	55 (85.9)					334 (87.8)		49 (89.1)	
Positive	13 (10.9)	9 (14.1)					51 (12.2)		6 (10.9)	

[a]reported as mm, corrected, [b]42.2%distribution, [c] mean

4.2.1.6 Assessment of study quality

A summary of the quality assessment of studies included in this review is shown in Table 10 followed by a narrative summary. The main issue within the included studies has been tissue allocation bias (TAB), which occurs when tissue cannot be re-used between the index and reference test. In short, due to an uneven distribution of metastases, the test results may be discordant but accurate for the tissue received. As such, many studies attempted to address this, either by more extensive histology or additional molecular analyses, such as qRT-PCR and Western blotting for CK19 protein. Results should then be re-calculated after adjustment for TAB, generally with an increase in OSNA sensitivity and specificity.

Other concerns included lack of detail on patient recruitment, minimal information on patient characteristics and unclear sampling methods, i.e., no evidence was given of sample replicates and reproducibility for molecular analysis. Furthermore, unless otherwise mentioned, test failures, such as operator or instrumental error, were not reported. It was also often not mentioned whether the outcome assessors were blinded; for example it is not inconceivable that the pathologist performing intra-operative molecular analysis also performed the more subjective post-operative histopathology. Papers were also often unclear on robustness of histopathology, e.g. whether results were checked by a second party.

A potential conflict of interest also features heavily, [redacted] and the majority of the OSNA studies were financially supported by Sysmex.

Table 10. Summary of quality assessment

QUADAS 2 Domain		OSNA																Metasin	
		First author (date)																	
		Bernet, L. (2011)[d52]	Bernet Vegue, L. (2012)[53]	Castellano, I. (2012)[54]	Choi, Y. (2010)[55]	Feldman, S. (2011)[56]	Frere Belda, M (2012)[60]	Godey, F. (2012)[57]	Guillen-Paredes, M. (2011)[58]	Khaddage, A. (2011)[59]	Osako, T. (2011)[61]	Schem, C. (2009)[62]	Snook, K. L. (2011)[63]	Tamaki, Y. (2009)[64]	Tamaki, Y. (2012)[65]	Tsujimoto, M. (2007)[66]	Visser (2008)[67]	McDowell, A (unpub)[51]	Sundaresan, V. (unpub)[41]
Patient selection	Was a consecutive or random sample of patients enrolled? (Y/N/U)	N	Y	U	U	U	U	U	U	U	Y	U	U	U	U	U	U	[redacted]	U
	Was a cohort study design avoided?[a](Y/N/U)	NA	Y	N	Y	Y	Y	N	N	Y	N	Y	Y	Y	Y	Y	Y	[redacted]	Y
	Did the study avoid inappropriate exclusions? (Y/N/U)[g]	U	Y	Y	Y	Y	Y	N	Y	Y	Y	U	Y	Y	Y	U	Y	[redacted]	U
	Could the selection of patients have introduced bias? (H/L/U)	U	L	U	U	U	U	U	U	U	L	U	U	U	U	U	U	[redacted]	U
	Concerns that the included patients do not match the review question? (H/L/U)	L	L	L	L	L	L	L	L	L	L	U	L	L	L	L	L	[redacted]	L
Index test	Were the index test results interpreted without knowledge of the results of the reference standard? (Y/N/U)	U	U	NA	U	Y	U	NA	Y	Y	U	Y	Y	Y	U	Y	U	[redacted]	U
	If a threshold was used, was it pre-specified? (Y/N/U)	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	[redacted]	Y
	Could the conduct or interpretation of the index test have introduced bias? (H/L/U)[e]	L	U	L	L	L	L	U	L	L	L	L	L	L	L	H	L	[redacted]	U
	Are there concerns that the index test, its conduct, or interpretation differ from the review question? (H/L/U)	L	L	L	L	L	L	U	L	L	L	L	L	L	L	L	L	[redacted]	L
Reference standard	Is the reference standard likely to correctly classify the target condition? (Y/N/U)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	[redacted]	Y

	Were the reference standard results interpreted without knowledge of the results of the index test? (Y/N/U)	NA	U	NA	U	Y	Y	U	NA	Y	U	U	Y	Y	U	U	U	▌	U
	Could the reference standard, its conduct, or its interpretation have introduced bias?[f] (H/L/U)	L	U	U	U	L	U	U	U	L	L	U	U	L	U	L	U	▌	U
	Are there concerns that the target condition as defined by the reference standard does not match the review question?	L	L	L	L	L	L	U	L	L	L	L	L	L	L	L	L	▌	L
Flow and timing	Did all patients receive a reference standard? (Y/N/U)	NA	Y	N	Y	Y	Y	N	NA	Y	N	Y	Y	Y	Y	Y	Y	▌	Y
	Did all patients receive the same reference standard? (Y/N/U)	NA	Y	Y	Y	Y	Y	N	NA	Y	Y	Y	Y	Y	U	Y	Y	▌	Y
	Were all samples (that should have been[b]) included in the analysis? (Y/N/U)	U	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	▌	Y
	Could the patient flow have introduced bias? (H/L/U)	U	L	U	U	U	L	U	U	L	L	U	U	L	L	U	U	▌	U
Additional items	Were samples suspected of TAB excluded from the analysis? (H/L/U)[c]	NA	N	NA	N	Y	Y	NA	NA	Y	NA	Y	Y	Y	N	N	Y	▌	N
	Are there concerns about selective reporting of outcomes? (H/L/U)	U	L	L	L	L	L	H	L	L	L	L	L	L	L	U	L	▌	L

Y:yes N:no NA:not applicable U:unclear H: high risk of bias L:low risk of bias

[a] The QUADAS 2 asks whether a *case-control* design has been avoided (i.e. a two-gate diagnostic test accuracy study). No two-gate studies were found for this review, but cohort studies (i.e. those comparing a population receiving histology with a separate population receiving OSNA/Metasin) were identified. It was decided that these studies should be included because they provide OSNA/Metasin data when the whole node is used. This question, therefore, has been adapted to flag up whether the study is a cohort study or a single-gate diagnostic accuracy study.
[b] The QUADAS 2 question "Was there an appropriate interval between index test(s) and reference standard?" has been omitted because the review protocol was designed such that studies with an inappropriate interval between intervention and reference standard (i.e. those using intraoperative histology as a reference standard) were excluded from the review.
[c]This question is designed to ensure that bias has not been introduced by excluding samples. However, discordant samples that are likely due to TAB should be excluded. A 'Yes' response here, therefore, refers to all samples other than those deemed to be subject to TAB (irrespective of whether TAB samples were excluded or not).
[d] Study only assessed on Trial 2
[e] None of the molecular tests provided evidence of reproducibility, however, this is not considered an issue with bias
[f] The reference standard is prone to observer bias
[g] DCIS was not considered inappropriate inclusion

4.2.1.6.1 Metasin

[redacted]

57

[redacted]

The second draft paper was presented by Sundaresan, who created the Metasin test.[41] This study design is experimental, where six centres contributed tissue homogenates, which were then compared to historical 3-level histology. As such, no patient characteristics are included. Two centres were only able to provide frozen homogenates, which may impact on the quality of RNA. Technical details of Metasin are not provided. No cases were excluded due to suspected TAB and no retrospective discordant case analysis was performed. However, where possible, RNA was re-analysed by an independent panel of markers. It is not clear whether initial analysis was performed by blinded pathologists. Failed assays were reported (1.2%). According to the authors this was possibly due to insufficient RNA in the sample.

4.2.1.6.2 OSNA

Cohort studies

Castellano provides a comprehensive flowchart.[54] Patients were not selected randomly, although patient groups were compared, and no significant differences were found for the characteristics reported. Histology involved 2mm slices of the whole node which were then sectioned at 100µm.

Godey et al. compare a historical cohort using a 1mm central slice sectioned every 250 um.[57] A flow diagram documents the pathway of patients through the study. However, reporting of information on patient characteristics for the whole study is minimal. The authors mention that for three patients there were technical problems with OSNA, but this was not elaborated on.

Guillen-Paredes and colleagues provide a retrospective cost benefit analysis using two relatively small patient groups.[58] Inclusion criteria are detailed, but patient characteristics

presented are minimal. Histology involved preparing 4mm sections of the lymph node, subsectioned into 15 x 4µm slices for examination.

The final cohort study is reported by Osako et al. where OSNA is compared to historical single section histology.[61] Detailed exclusion criteria and comprehensive patient characteristics are included, with no significant heterogeneity being found between the groups. Some of the nodes were frozen at -80^{o}C prior to OSNA, which may or may not affect mRNA. Laboratory consumables were funded by Sysmex.

Single gate studies

Bernet et al. report two trials within one paper.[52] For Trial 1, it was unclear whether fresh or frozen sections were used as the comparator, therefore the results are not included in this review. Trial 2, investigates the timing of procedural steps for OSNA for 55 cases, and was, therefore, included in the review.

A second paper by Bernet reports on a single gate study using only one level histopathology, with the remaining tissue used for OSNA, which is likely to result in substantial TAB.[53] The authors claim this approach was chosen to replicate routine use. Patients were recruited consecutively across 8 hospitals, although the achieved sample size is relatively small at 55 patients. Reasonably comprehensive reporting of patient characteristics was given. No details were given on blinding of pathologists and there was no further investigation into discordant cases, although further details such as copies/µL were given. There may be a conflict of interest since the study was funded by Sysmex España S.L.

The study reported by Choi et al. gives no information on recruitment and no diagram of patient flow. However, analysis on all appropriate samples appears to have been performed and clinicopathological characteristics of patients are given.[55] Of note, macrometastasis or micrometastasis was confirmed by both or either intraoperative histopathological examination of frozen section specimens and postoperative histopathological (three-level) examination with permanent tissue specimens. Although it appears from this that some of the reference standard data may, therefore, have been based on frozen section alone, closer examination of the entirety of the data has led us to believe that this was not the case. In particular, data are presented on discordant frozen section and final histopathology results, and these suggest that there were no cases where frozen section results were positive and final histology negative. As such, it appears that a positive histology result was based on two scenarios: first, both the frozen section and the final histology were positive, and second, the frozen section results were negative and the final histology positive. Therefore, all positive histology cases would have had a positive final histology result.

The study is unclear on whether the pathologists were blinded and provides no detail on test failure or replicates for OSNA. Discordance was evaluated using clinical information, status of non-SLNs and CK19 protein expression in metastatic foci of lymph nodes. It is unclear how the CK19 was determined. This study was also supported by Sysmex (Kobe, Japan).

Feldman et al. report a study funded by Sysmex America, Inc. comparing OSNA with apparently extensive histology of 200 µm intervals of approximately 3 x 1 mm slices.[56] Patient recruitment is not mentioned, although patient characteristics are comprehensively reported. No patient flow diagram is included and the number of SLNs after discordance (1018) does not comply with the numbers before discordance (1044) minus the resolved cases (28). The population is somewhat heterogeneous, including all classifications of tumour and lymph node status. Histopathologists were blinded and discordant case analysis was performed using Western blotting and qRT-PCR, although no details are given of this.

A single gate study reported by Le Frere Belda et al. details a flow chart of the SLN samples as well as patient characteristics.[68] It investigates a reasonably heterogeneous population, although the majority had PN0 node status. No information was given on recruitment, other than inclusion criteria. Duplicate samples were used for OSNA. Five-level histology was performed, although, it should be noted, in five centres frozen sections were re-used for this, which may have degraded the sample. Discordant case analysis was performed by extensive histopathology and SLN homogenates shipped to Sysmex (Norderstedt, Germany) for blind molecular analysis. It is unclear whether original analysis was blinded. One sample was excluded due to a manipulation error. Laboratory consumables were funded by the Sysmex Corporation.

The study reported by Khaddage et al. compares OSNA to five level histology for a clinical evaluation and against one level histology for routine use.[59] Although replicates are not mentioned, positive and negative controls are confirmed. No patient flow diagram is displayed and there are no details on recruitment other than minimal inclusion/exclusion criteria. However clinical characteristics of the patient population are given and all samples appear to have been analysed. Pathologists were blinded to both test results. Discordant case analysis was performed only on the clinical study using only qRT-PCR (2 cases), where TAB is likely to be less of an issue than for the routine study (17 cases). No details of test failures, other than one case of a false positive due to an invalid control, were given. The study was funded by the Sysmex, Kobe, Japan.

The study reported by Schem et al. provides minimal information on patient characteristics or recruitment.[62] Five-level histology was used as the reference standard and the outcome

assessors were blinded. The samples for OSNA were frozen at -80^{o}C. Specificity was calculated on the first 120 negative nodes (of 343), undergoing extended histology. Histology was also extended on discordant cases with Western blotting and qRT-PCR performed on the homogenates. Comprehensive details are provided on all these techniques other than replicates and test failures. This study was also supported by Sysmex.

Snook et al. reports a prospective study comparing OSNA with five-level histology.[63] No patient flow diagram is included, although numbers appear accurate, and minimal characteristics are detailed. It is unclear whether recruitment was consecutive. Outcome assessors were blinded to OSNA results. Discordant case analysis was performed by Western blotting and qRT-PCR. No financial contribution was received from any organisation; however, support in the form of training and advice was provided by Sysmex Life Science.

Tamaki et al. present one of two papers on a Japanese multicentre study.[64] A validity and a routine use trial are reported. For the validity trial, histopathology was extensive with node sections taken at 0.2mm intervals. For routine use, a more standard 3-level method was used. Brief patient characteristics are presented for each trial with no details on recruitment. No patient flow diagram is produced, but the number of nodes extracted and analysed comply. Outcome assessors were blinded and discordant cases were analysed by Western blotting and qRT-PCR, although technical details are not provided. This study was supported by Sysmex.

Tsujimoto et al.[66] appears to report on the same trial as Tamaki et al. [64]where OSNA is compared to 3-level histology. Minimal details are provided on patient characteristics and with no description of recruitment. Lymph nodes were stored at -80^{o}C until used. Samples were assayed in duplicate. It is unclear from this report whether pathologists were blinded, although if this is in fact the same trial as described in Tamaki et al. then it reported that they were. Histopathological samples were examined by three third party pathologists. As in Tamaki et al., it is reported that discordant case analysis was performed by Western blotting and qRT-PCR, although it is reported here that these tests were not performed on all samples. Sysmex are acknowledged but there is no explicit mention of financial support.

A second paper by Tamaki et al. reports another multicentre study based in Japan.[65] Comprehensive clinical patient characteristics are given with minimal details on recruitment. Description of the OSNA assay is sparse and only one-level histology is employed. No discordant case analysis was performed.

The final study by Visser et al. employed three-level histology on ALNs, which, if negative, underwent a further four-level investigation.[67] The patient number was relatively small at 32, although the total number of nodes analysed was 346. No details were provided on patient recruitment and patient characteristics were minimally reported. There was no evidence of replicates for OSNA, but for discordant case analysis, RT-PCR was performed in duplicate. Frozen samples were used for OSNA (first 120 histologically negative samples). The pathologists do not appear to have been blinded.

4.2.2 Assessment of test accuracy

The assessment of test accuracy for both OSNA and Metasin is hindered by TAB, and by comparison with an inconsistent reference standard. Where discordant samples are further investigated (usually using either more extensive histopathology or molecular analysis by QT-PCR or Western blotting) and attributed to TAB, analyses can be adjusted by excluding such cases. However, adjustment cannot be made for heterogeneity between studies with regards the reference standard; histopathology may be performed extensively across five levels of a lymph node or more perfunctorily using only one level of a central slice.

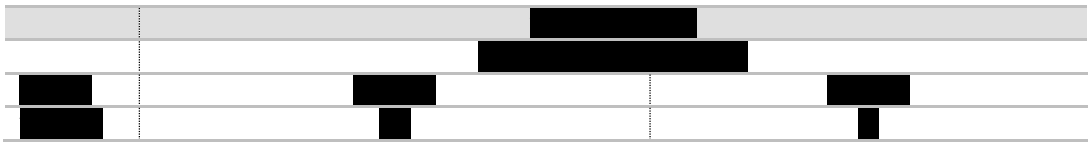
It should be noted that isolated tumour cells (ITCs), which are found by immunohistochemical staining and not detected by OSNA are considered lymph node negative since their clinical significance is unknown. [67]

Individual results are described alongside a narrative description below. Summarised results, which have been stratified, are presented in Table 27 to Table 32 (pages 97-99) at the end of this section.

4.2.2.1 Metasin

[redacted]

11



[redacted]		

The second paper, by Sundaresan, which like the first, is unpublished and therefore has not undergone formal peer review, report a sensitivity of 92%, a specificity of 97% and an accuracy of 96% after using a more experimental, postoperative study design for OSNA.[41] [redacted] used three level histology as the reference standard. Twenty cases (1.6%) were reported as assay fails, which the authors consider may have been due to insufficient RNA in the material submitted for molecular analysis. The authors report fifty six patients to be discordant and, although the authors do not present evidence, they consider TAB to be largely responsible.

Table 12. Correlation between Metasin and histopathology for Sundaresan[41]

	n = 1265 patients	
	Three level histopathology	
Metasin	Positive	Negative
Positive	249	36
Negative	20	940

4.2.2.2 OSNA

Cohort studies

The Castellano et al. study indicates the rate of negative cases determined by standard histology (73%) was highly comparable with the rate of the negative SLNs determined by OSNA (71%), with no adjustment for TAB.[54] The percentage of micrometastases detected by OSNA was significantly higher (18%) than that determined by the standard procedure (8%; P<0.01), whereas the rate of macrometastases detected by OSNA was lower (11% v 20%). The authors mention that this may be a reflection of different patient populations rather than the different techniques, however, the negative rates were similar.

The authors hypothesise that the morphological evaluation of the diameter of metastases does not perfectly correlate with the tumour load as evaluated by mRNA copies of CK19 and that, here, there could be an overestimation of macro metastasis (>2mm) versus micrometastasis by histology (>0.2mm and <2mm) or an underestimation of macro versus micro by OSNA.

The study presented by Godey et al.[57], report a positive rate of 26.4%, which is comparable to the rate reported by Castellano et al. (29%).[54] The time required for the final results of the OSNA assay (including transport to the laboratory) depended on the number of SLNs

studied. The mean time was 32.9 min (standard deviation 4.9, n=94) for 1 SLN (n = 94), 36.4 min (standard deviation 4.5 min, n=144) for 2 SLNs, 41.6 min (standard deviation 5.2 min, n=87) for 3 SLNs and 48.5 min (standard deviation 8.7, n=39) for 4 SLNs.

Guillen-Parades present a cost/benefit analysis.[58] They report a shorter total operative time for the OSNA patients, where the mean total time difference for histopathology (mean: 78 min; standard deviation:48.02) compared to OSNA (mean: 62.14 min; standard deviation: 21.93) was statistically significant (P < .005). However, when only considering the time of the first operation, the time for histology is shorter (histology 57.11 min v OSNA 62.14 min), although this difference was not statistically significant (p =0.15).

Mean hospital stay after the first operation in the histology group 1.8 days (range: 1–13; standard deviation: 2.04), while hospital stay for the second operation was 2.41 days (range: 1–6; standard deviation: 1.29, over 12 patients), resulting in an overall calculation of 2.44 days, as compared to a mean of 1.54 days for OSNA (range: 1–4; standard deviation: 0.78), which was a statistically significant difference (P<.001). The analysis of complications (minor, major and no complications) of the histology group 1 (57, considering the 1st and 2nd operations) compared to those of the OSNA group (35), showed statistically significant differences (P=.015).

The final cohort study is presented by Osako et al.[61] The entire node was used for OSNA group, whereas only a single section of the node was used for the histology group. No significant difference was seen between positive rates of histological macrometastases and OSNA (++) (18.8% 95%CI; 10.5-30.8 vs 25.2%, 95% CI 17.9-34.2; P=0.420). However, a significant difference was displayed between histological micrometastases and OSNA (+) (1.6%, 95%CI; 0.1-9.5 vs 30.3% 95%CI; 22.3-39.5; P<0.001).

Table 13. Positive rates for metastases reported in cohort studies

Positive rate (%)	Castellano[54]	Godey[57]	Osako[61]
Histology	28	24.8	20.3 (11.7-32.6)
OSNA	29	24.4	55.5 (46.1-64.5)

Single gate studies

The first study presented by Bernet has only been included in the review for the time to for OSNA analysis (Section 4.2.3, page 108).[52]

The second study reported by Bernet (Vegue) et al. only investigated a 1mm central node slice by histopathology, with the remainder allocated to OSNA (therefore a high rate of discordance was predicted).[53] Values for sensitivity and specificity were not given in the paper. A statistically significant overall discrepancy was shown between OSNA and histopathology (P<0.001).

Table 14. Correlation between OSNA and histopathology for Bernet Vegue et al.[53]

(n=567 non-SLN)			
One level histopathology			
OSNA	Macrometastasis	Micrometastasis	Negative
++	1	4	14
+	0	1	25
+i/low expression	0	0	8
-	0	0	514

Discordant results were observed in 51 (9%) nodes, although the discordant case analysis table only displays results for 47 nodes. In the histopathology negative nodes, OSNA was identified ++ in 2.5%, + in 4.5% and low expression in 1.4%. The authors suggest OSNA is of particular relevance for the identification of low volume metastases.[53]

When individual nodes were considered, there were no false negatives for OSNA. In contrast, over 80% of discordant nodes found to be positive for metastases were not identified by conventional, one-level histology. The authors state that one level histology was chosen to reflect the minimum standard used in many laboratories.[53]

The study reported by Choi gives a sensitivity of 77.8% (95%CI, 0.60-0.90), which is comparatively low.[55] The authors indicate this is likely to be due to TAB, however the dissection of the node is similar to other studies. Choi et al. report a specificity of of 96.3% (95%CI 0.92-0.99) and accuracy of 93% (95%CI, 0.88-0.96).

Table 15. Correlation between OSNA and histopathology for Choi et al.[55]

n=199 pts				
Three level histopathology				
OSNA	Macrometastasis	Micrometastasis	ITC	Negative
++	19	2	1	1
+	3	3	0	4
+i	1	0	0	0
-	4	4	3	154

Fourteen cases of discordance occurred. In one false positive case (OSNA+, histology -), the histopathological result of ITC detection as well as the existence of metastasis in other SLNs was attributed to the localization of metastasis foci in the lymph node. For 3 more

discordant cases OSNA was close to cut-off, indicating the existence of weak positivity in the sample. For another, even though OSNA assay provided (++) judgment, the authors were unable to find any other findings in this evaluation to support this judgment. They hypothesised that this was due to the localization of metastasis foci in the lymph node as the probability of OSNA false positives is considered to be very low based on the results of previous studies. For another O+H- case, OSNA assay provided positive results for each of 2 lymph nodes from this patient. This was also considered to be due to localization of metastasis foci in a lymph node.

The false negative cases (OSNA-, histology+) were investigated immunohistochemically to confirm the protein level of CK19 in tumours. Four cases showed expression of CK19 protein in less than 10% of the tumour. The cause of these discordances may have been the inability of OSNA assay to detect metastasis of breast cancers with low CK19 expression. In two other cases, metastases were histopathologically found by the postoperative examination of permanent specimens. CK19 IHC of 2 cases was not performed due to insufficient tissue and remains inconclusive.

Following adjustment for TAB, the study by Feldman et al. reports the sensitivity as 82.7%, the specificity as 97.7% and the accuracy as 95.8%.[56]

Table 16. Correlation between OSNA and histopathology for Feldman et al.[56]

n=1044 SLN				
Three level histopathology				
OSNA	Macrometastasis	Micrometastasis	ITC	Negative
++	77	9	1	8
+	9	12	0	29
-	9	22	14	854

Seventy one cases were considered discordant. Although OSNA failed to detect 9 macrometastases and 22 micrometastases that were identified by reference pathology, it detected 9 macrometastases and 29 micrometastases that were identified as negative or as isolated tumor cells by reference pathology. All of those 9 macrometastases were identified as true misses by reference pathology, and recalculation of the assay performance for macrometastasis yielded a PPV of 100% for an OSNA++ result.

For macrometastases, all 9 discordant OSNA++ results were identified as true-positive, and 9 of 29 discordant OSNA+ results also were identified as true-positive upon discordant case analysis. Although 20 of the 29 discordant OSNA+ results could not be confirmed, the median CK19 copy numbers in these samples were close to the assay cut-off, suggesting that they were very small micrometastases, and this did not rule out the possibility of tissue allocation bias.

The sensitivity and specificity of OSNA were lower than the values reported in previous smaller studies, possibly due to: evaluation solely of SLNs; slicing at 1-mm intervals rather than at 2-mm intervals, and large numbers of micrometastases.

Le Frere Belda et al. report the following, sensitivity and specificity for OSNA compared to 5 level histology was 91.1 (95% CI, 80.3–97.1) and 97.2 (95% CI, 95.1–98.6), respectively and subsequent to TAB adjustment.[60]

Table 17. Correlation between OSNA and histopathology for Le Frere Belda[60]

n=503 SLN				
Five level histopathology				
OSNA	Macrometastasis	Micrometastasis	ITC	Negative
++	37	6	0	3
+	5	3	1	23
-	3	9	27	386
n=233 patients				
Five level histopathology				
OSNA	Macrometastasis	Micrometastasis	ITC	Negative
++	22	6	0	3
+	2	3	3	17
-	2	7	17	151

The authors report 39 cases of discordance and mention that since different parts of the node were used for each method because each technique required different tissue preparation, discrepancies between OSNA and histological results were expected.[60] Twenty-seven cases were histology negative/OSNA positive. All 27 samples remained histologically negative after further examination of histological slices, whereas 15 samples had OSNA-positive lysates after further molecular analysis by quantitative reverse transcriptase-polymerase chain reaction (qRT-PCR), suggested by the authors to be TAB. Twelve histology positive/OSNA negative samples were found. One was not investigated by further molecular analysis and seven were negative

For the study presented by Godey and colleagues[57], a positive rate of 24.4% is reported for OSNA and 24.8 for histology, which is comparable to the rate reported by Castellano and colleagues (29%).[54]

Khaddage et al. reported on OSNA compared to both five- level histology and a routine one-level histology. For the extensive histology after adjustment for TAB, presented below, the sensitivity, specificity and accuracy was 100%. 98.4% and 98.7%.

Table 18. Correlation between OSNA and histopathology for Khaddage[59]

n=80 SLN (validation study)				
5 Level Histopathology				
OSNA	Macrometastasis	Micrometastasis	ITC	Negative

++	11	2	-	0
+	2	0	-	1
-	0	0 (2)	2	60
n=46 patients (validation study)				
5 Level Histopathology				
OSNA	Macrometastasis	Micrometastasis	ITC	Negative
++	6	2	0	0
+	0	0	0	1
-	0	0 (2)	2	33
n=197 patients (routine use)				
1 Level Histopathology				
OSNA	Macrometastasis	Micrometastasis	ITC	Negative
++	9	1	1	2
+	8	7	1	13
-	0	0	1	154

Values in parentheses indicate case numbers before discordant case investigation

Only 2 results appear to be discordant for five-level histology, where micrometastases were detected by histology, but the OSNA sample was negative. Further molecular analysis by qRT-PCR for CK19 and two breast cancer specific markers displayed negative results, therefore TAB was assumed to have occurred. Seventeen cases were discordant for routine use, all H-O+, three of which had SNLN involvement and fourteen indicating micrometastases and the likelihood of TAB.

The study reported by Schem et al. 2009 indicates a sensitivity of 98.1% and a specificity of 91.7%.[62] After adjustment for TAB, sensitivity increased to 100% and specificity to 96.5%. Twenty eight cases were considered discordant.

Table 19. Correlation between OSNA and histopathology for Schem et al.[62]

n=343 ALN				
Five level histopathology				
OSNA	Macrometastasis	Micrometastasis	ITC	Negative
++	90	7	0	9
+	7	-	1	16
-	0	2	2	209

RNA and proteins were extracted from the lysates of the 28 discordant cases, followed by qRT-PCR and Western Blotting for CK19. If the data obtained by additional analyses were consistent with the results obtained by OSNA, TAB was considered likely and these samples were excluded from the sample cohort.

In both of the OSNA negative/histology positive and 11 out of 26 OSNA positive/histology negative samples, discordant case analysis revealed equivalent results to the ones seen in the OSNA assay. As such, it cannot be fully excluded that even a higher proportion of discordant results were due to TAB because the homogenates were exposed to long storage and transport conditions which might have lowered the concentration and quality of RNA and

proteins. This is especially true for OSNA samples with copy numbers close to the cut-off level as qRT-PCR, and Western blot investigation is then also likely to be close to the detection limit.

Eleven of the 26 OSNA positive/histology negative samples had CK19 mRNA copy numbers/µL below 1,000. The authors suggest that with 250 copies/µL as the cut-off level, these positive OSNA results very likely indicate a low tumour burden in the lymph nodes which was probably absent in the tissue sections used for histological investigation.[62] For samples close to cut-off level, long storage and travel can affect concentration and quality of RNA and proteins.

Snook et al. compare OSNA to five level histology.[63] OSNA was reported to have a sensitivity of 91.7%, a specificity of 96.9% and an accuracy of 96.0%, after adjustment for TAB.

Table 20. Correlation between OSNA and histopathology for Snook et al.[63]

n=395 SLN				
Five level histopathology				
OSNA	Macrometastasis	Micrometastasis	ITC	Negative
++	48	1	0	0
+	8	9	0	10
-	4	2	20	293
n=194 patients				
Five level histopathology				
OSNA	Macrometastasis	Micrometastasis	ITC	Negative
++	33	1	0	0
+	5	5	1	7
-	4	1	11	126

A total of 33 cases were considered discordant, with 17 due to TAB following qRT-PCR and Western blotting. Sixteen samples were reported as truly discordant, 10 false positive and 6 false negative. The OSNA false positives with a low copy number were attributed to the possibility of small micrometastases not apparent in histology. Although the authors mention there is the possibility of contamination since CK-19 is expressed on the surface of cells of epithelial origin, benign or malignant, such as breast, colonic or gastric cells.[69] Ideally, the cut-off value reduces the likelihood of this.

Since histopathology is prone to sampling bias, i.e., the number of levels of histopathology can influence the result, the authors also question whether histopathology is appropriate as a gold standard. Even employing an intensive regimen of five levels per slice plus dual IHC, as used in the reported study, the remaining step sections were not examined. With regard

to further molecular analysis, the authors consider the possibility of a reduction of RNA concentration in the frozen lysate of SLN sample due to the freeze–thaw effect, which could account for some of the discordant cases, despite repeat molecular testing.

Tamaki et al., present results for 2 trials[64], which is also reported by Tsujimoto et al.[66] Trial 1 compared the specificity of OSNA with detailed histology (0.2mm sections), to confirm the validity of the index test. Trial 2 was intended to replicate routine use where only one-level histology took place.

Table 21. Correlation between OSNA and histopathology for Tamaki et al.[64]

n=124 ALN Trial 1				
0.2 mm Section histopathology				
OSNA	Macrometastasis	Micrometastasis	ITC	Negative
Positive	16	3		3
Negative	0	1		101
n=551 ALN Trial 2				
Three-level histopathology				
OSNA	Macrometastasis	Micrometastasis	ITC	Negative
Positive	64	6		22
Negative	4	6		348

Sensitivity and specificity for trial 1 was reported as 95% and 97%, respectively, whereas for trial 2, sensitivity and specificity were 87.5% and 94% before adjustment for TAB and 87.7% and 94.3% after adjustment.

In trial 2, thirty two nodes displayed discordant results (7.1%). For eight of the ten false-negative nodes, not all of the serial sections contained metastasized foci and therefore may not have been in the OSNA samples. Two nodes from different patients with the false-negative cases displayed a very weak expression of CK19 mRNA. The primary tumours of the patients also showed negative staining for CK19 as confirmed by immunohistochemistry.

ITCs were found in the remaining specimens of 5 of the 22 false-positive nodes after additional sectioning, which had not been detected by routine pathologic examination using 2-mm interval sections. In another eight false-positive nodes, no tumour cells were found in the pieces remaining after the pathologic examination, but lymphatic vascular invasions were observed in the main tumours of these nodes. In addition, the lysate of two of them preserved for the OSNA assay contained a significant amount of CK19 protein.

As for pathologically positive lymph nodes (diameter of metastasis, >0.2 mm), 87.7% could be detected by the OSNA assay, whereas 94.1% of macrometastases (>2 mm) were assessed as positive.

The second paper by Tamaki et al. 2012 compare OSNA with one-level histology.

Table 22. Correlation between OSNA and histopathology for Tamaki et al.[65]

n=417 patients (SLN)		
One-level histopathology		
OSNA	Positive	Negative
Positive	58	36
Negative	8	315

Forty four cases were discordant, considered to be inevitable by the authors due to the method of tissue sampling. Eight cases were OSNA-negative, in which postoperative pathologic examination identified metastasis. Of these, 7 patients who had micrometastasis identified, discordance was suspected to be due to uneven allocation of minuscule metastases in an SLN. However, in the remaining 1 patient with macrometastasis, low expression of the CK19 protein in the main tumour was confirmed, leading to a negative OSNA result. Although, the authors state it is not necessarily the case that low expression of CK19 protein is reflected with low expression of CK19 mRNA.[65]

There were also 36 OSNA positive histology negative discordant cases, including 2 with isolated tumour cells in the SLNs assessed by pathology. Of these, 7 patients had non-SLN metastases subsequently identified, therefore OSNA had made an accurate assessment. Microinvasion was suspected in a core-needle biopsy specimen from 1 patient. Two patients had widespread DCIS that measured >6 cm, and another had multiple lesions. The remaining 2 patients had high-grade DCIS.

The authors suggest that, despite the cut-off, OSNA can detect metastases with high sensitivity even in tumours diagnosed pathologically as DCIS, and that such findings may result in an upgrade of the clinical stage of such tumours.[65]

Tsujimoto et al.[66] report on the same trial as Tamaki et al.[64] OSNA sensitivity, compared with three-level histology for the trial with 325 nodes (both ALN and SLN) was reported as 91.1% with an accuracy of 98.2%.

Table 23. Correlation between OSNA and histopathology for Tsujimoto et al.[66]

n= 325 SLN and ALN				
Three level histopathology				
OSNA	Macrometastasis	Micrometastasis	ITC	Negative
++	34	0	0	0
+	6	3	0	4
-	0	2	13	263

n= 81 SLN				
Three level histopathology				
OSNA	Macrometastasis	Micrometastasis	ITC	Negative
++	11	0	0	0
+	1	2	0	1
-	0	2	3	61
n= 144 SLN from pN0 pts				
0.2mm Interval histopathology				
OSNA	Macrometastasis	Micrometastasis	ITC	Negative
++	0	0	0	0
+	0	0	0	0
-	0	0	3	141

Histology based on three level IHC only, as H&E was considered by authors to beless sensitive

Six discordant cases were observed (for the trial with 325 nodes) between the OSNA assay and histopathology. Western blot analysis of the 2 discordant cases showed the presence of an amount of CK19 protein equivalent to micrometastasis. The authors considered that although the possible presence of benign epithelial cells such as glandular inclusions in the lymph nodes cannot be eliminated, the results may be better explained by the presence of metastatic foci in the lymph nodes in light of the results of the specificity study and the amount of CK19 protein expression.[66] This is further supported by the lack of false positives reported for the pN0 patients. Two other cases were negative according to the OSNA assay, but were judged positive for micrometastasis according to three-level histopathology.

The final single gate study included in this review was presented by Visser et al., comparing OSNA to five level histology.[67] After adjustment for TAB, a sensitivity of 95.3%, specificity of 97.1% and an accuracy of 96.8%, was revealed.

Table 24. Correlation between OSNA and histopathology for Visser et al.[67]

n=346 ALN				
Five level histopathology				
OSNA	Macrometastasis	Micrometastasis	ITC	Negative
++	50	4	0	2
+	2	5	0	13
-	1	2	3	264

In order to establish the level of discordance due to TAB, the first 120 histologically negative lymph nodes, as determined by five-level histology, were cut into further levels at intervals of 250µm.

Seven cases were considered to be due to tissue allocation bias. A further 18 unresolved cases were reported. In 8 cases, only Western Blot analysis for CK19 protein could be obtained because poor quality RNA did not allow qRT-PCR. One of 3 histology positive/OSNA- negative samples yielded negative results for all 3 markers. In the lysates of the other 2 samples, CK19 protein levels, were slightly above the cut-off level suggesting the

presence of small tumour deposits. In 11 histologically negative/OSNA positive samples low CK19 mRNA copy numbers (250–750/µL) were found with OSNA. Six of these could be further analyzed by qRT-PCR, whereas the remaining 5 samples suffered from poor RNA quality. The same was true for 1 sample with high CK19 mRNA copy number.

4.2.2.3 Discordance

Studies show a range of discordance from 3.7% to 8.6% (Table 25), with further details in Table 26. The greatest discordance was generally for OSNA false positives. This is a concern since this would potentially lead to unnecessary, complete axillary dissection in patients. This may be due to epithelial contamination although, for OSNA, the CK19 copy-number cut-off corresponds to the presence of 5000 tumour cells. Therefore, it is unlikely that positive results might occur because of epithelial displacement or illegitimate transcription, which involve only 500-1000 non-tumour cells.[67] By contrast, false negative tests may result in a lack of treatment in the early stages of cancer, leading to a worse prognosis. The patient may also be required to undergo a second operation for ALND.

Although the true number of discordant cases and those due to TAB cannot be fully differentiated, the studies which have adjusted their results display increased specificity and sensitivity for OSNA. Interestingly, the studies which analyse a full node with extensive histology showed no significant difference in positive rates between OSNA and histology[54,57], whereas the study using one level histology demonstrated a significantly higher positive rate for OSNA.[61]

Table 25. Cases of discordance

First author	Total cases (%)*	Attributed to TAB (%)	H+/O- (%)	H-/O+ (%)
Bernet Vegue[53]	8.3	NR	0	8.3
Choi[55]	7*	NR	4	3
Feldman[56]	6.8	2.7*	3.0	3.6
Frere Belda[60]	7.7	4.3	2.3	5.4
Khaddage (study phase)[59]	3.7	0	2.5	1.2
Khaddage (routine use)[59]	8.6	NR	0	8.6
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Schem[62]	8.2	3.8	0.6	7.6
Snook[63]	8.3	4.3	3.0	5.3
Sunderasan[41]	4.2	NR	1.5	2.6
Tamaki, Trial 2[64]	7.1	NR	2.2	4.9
Tamaki[65]	5.7	NR	1.0	4.6
Tsujimoto[66]	1.8	NR	0.6	1.2
Visser[67]	5.2	2	0.87	3.2

Table 26. Details of discordance analysis for individual studies

Study	N	Discordant analysis	Unadjusted		Adjusted		Numbers reclassified		Comments on nature and impact of
Bernet Vegue									
Axillary node level	567	No further analysis performed	TP	6	TP	-	FP to TN	-	Results not adjusted.
			FP	39	FP	-	FP to TP	-	
			FN	0	FN	-	FP to FN	-	
			TN	522	TN	-	FP excluded	-	
							FN to TN	-	
							FN to TP	-	17 FP had <1000 copies/ul, relatively close to cut-off for micrometastases.
							FN to FP	-	
							FN excluded	-	
							Other	-	
Choi									
Patient level	199	In discordant cases, clinical information, status of non-SLNs and expression of CK19 protein in lymph node metastasis foci were evaluated on a patient basis	TP	27	TP	-	FP to TN	-	Results not adjusted.
			FP	6	FP	-	FP to TP	-	
			FN	8	FN	-	FP to FN	-	
			TN	157	TN	-	FP excluded	-	
							FN to TN	-	1 FP and 1 FN – discordance attributed to location of metastases.
							FN to TP	-	
							FN to FP	-	
							FN excluded	-	
							Other	-	3 FP – low copy number, weak positivity
Feldman									
SLN level	1044	Blank tissue sections were stained with CK19-specific antibody; back-up samples were retested with OSNA to check for operator errors; Western blot analysis of CK19 and RT-PCR.	TP	107	TP	125	FP to TN	-	2 pathology assessment reversals after TAB. Unclear whether this affects TP or TN result.
			FP	38	FP	20	FP to TP	-	
			FN	31	FN	-	FP to FN	-	
			TN	868	TN	-	FP excluded	-	
							FN to TN	-	
							FN to TP	-	
							FN to FP	-	28 discordant results resolved, but specific results not given other than 18 FP found to
							FN excluded	-	

						Other	-	be TP by histology, assumed TAB, and
Khaddage								
Patient level (study phase)	46	Discordant case analysis consisted of qRT-PCR.	TP 8 FP 1 FN 2 TN 35		TP 8 FP 1 FN 0 TN 35	FP to TN 0 FP to TP 0 FP to FN 0 FP excluded 0 FN to TN 0 FN to TP 0 FN to FP 0 FN excluded 2 Other 0		2 FN – one case confined to two of 5 level histology, in the other case, one slice was metastases free for histology. Further molecular analysis was negative indicating TAB.
SLN level (study phase)	80		TP 15 FP 1 FN 2 TN 62		TP 15 FP 1 FN 0 TN 62	FP to TN 0 FP to TP 0 FP to FN 0 FP excluded 0 FN to TN 0 FN to TP 0 FN to FP 0 FN excluded 2 Other 0		1 FP – close to cut-off level indicating very small tumour deposits.
Patient level (clinical phase)	197		TP 25 FP 17 FN 0 TN 155		TP - FP - FN - TN -	FP to TN 0 FP to TP 0 FP to FN 0 FP excluded 0 FN to TN 0 FN to TP 0 FN to FP 0 FN excluded 0 Other 0		
Le Frere Belda								
Patient level	233	False positive (O+H-) : Slides made for all the portions	TP 33		TP 32	FP to TN 0		Study assumes that TAB is the only reason

		(b & d)of the half node allocated to histology which were not used initially. False positive (O+H-) and false negative (O + H -):	FP FN TN	23 9 168	FP FN TN	12 3 168	FP to TP FP to FN FP excluded FN to TN FN to TP FN to FP FN excluded Other	0 0 11 0 0 0 6 1 TP	why there may be discordance. If initial histology was confirmed or initial OSNA + or – results were confirmed, TAB was assumed to have occurred and the patient affected excluded from the analysis
SLN level	503	Remaining homogenate from the portions (a & c) of the ½ allocated to OSNA shipped to Sysmex for blind molecular analysis	TP FP FN TN	51 27 12 413	TP FP FN TN	51 12 5 413	FP to TN FP to TP FP to FN FP excluded FN to TN FN to TP FN to FP FN excluded Other	0 0 0 15 0 0 0 7 0	
McDowell									
[redacted]									
[redacted]									

Schem										
ALN level	343	For discordant results, histological work-up was extended until no tissue remained in the paraffin blocks. The homogenates were also analysed by Western blot and qRT-PCR.	TP FP FN TN	104 26 2 211	TP FP FN TN	104 15 0 211	FP to TN FP to TP FP to FN FP excluded FN to TN FN to TP FN to FP FN excluded Other	0 0 0 11 0 0 0 2 0	Provided that the supplemental analyses gave the same result as OSNA (either histology or molecular), the samples were excluded as an uneven distribution of the metastases (TAB) was likely to be the cause. The 2 FN also gave negative results for
Snook										
Patient level	194	Homogenates were analysed using qRT-PCR and Western blot analysis.	TP FP FN TN	- - - -	TP FP FN TN	44 8 5 137	FP to TN FP to TP FP to FN FP excluded FN to TN FN to TP FN to FP FN excluded Other	- - - - - - - - -	After discordant case analysis and the exclusion of samples affected by TAB, 9 patients were excluded from final analysis. No further details given.
SLN level	395		TP FP FN TN	- - - -	TP FP FN TN	66 10 6 313	FP to TN FP to TP FP to FN FP excluded FN to TN FN to TP	0 0 0 11 0 0	33 discordant nodes. 17 considered affected by TAB (11 FP, 6

							FN to FP	0	FN) were excluded from final analysis.
							FN excluded	6	
							Other		
Sundaresan									
Patient level	1265	Not described – deeper level histology	TP	249	TP	251	FP to TN	0	1 FN changed to ITCs. Deeper levels of histology revealed metastases in 2 cases.
			FP	36	FP	34	FP to TP	2	
			FN	20	FN	19	FP to FN	0	
			TN	940	TN	941	FP excluded	0	
							FN to TN	1	
							FN to TP	0	
							FN to FP	0	
							FN excluded	0	
							Other	0	
SLN level	2279	Not described	TP	341	TP	343	FP to TN	0	
			FP	60	FP	58	FP to TP	2	
			FN	35	FN	34	FP to FN	0	
			TN	1770	TN	1771	FP excluded	0	
							FN to TN	1	
							FN to TP	0	
							FN to FP	0	
							FN excluded	0	
							Other	0	
Tamaki (2009) Trial 2									
SLN level	450	For discordant nodes, the remaining pathologic specimen blocks were sectioned at 0.2mm intervals and examined with H&E and IHC for CK19. The lysate was examined for CK19 protein expression	TP	70	TP	71	FP to TN	0	For 8 FN, uneven localization of tumour cells found in remnants of nodes. 2 other FN showed faint expression of CK19 by IHC.
			FP	22	FP	21	FP to TP	1	
			FN	10	FN	10	FP to FN	0	
			TN	348	TN	348	FP excluded	0	
							FN to TN	0	
							FN to TP	0	
							FN to FP	0	
							FN excluded	0	

		by Western blotting					Other	0	
Tamaki (2012)									
Patient level	417	No additional analysis mentioned in method.	TP FP FN TN	58 36 8 315	TP FP FN TN	- - - -	FP to TN FP to TP FP to FN FP excluded FN to TN FN to TP FN to FP FN excluded Other	- - - - - - - - -	No adjustments made, however, 7 FN were, micrometastases in histopathology . Macrometastases in one FN, but IHC revealed low CK19 protein expression. ITC in SLN from 2 FP and non-SLN metastases in 7 FP patients. Therefore 9 FP harboured cancer cells in ALN
Tsujimoto									
Patient level (ALN and SLN)	325	qRT-PCR and CK19 Western blot analysis of the lysates were carried out.	TP FP FN TN	43 4 2 280	TP FP FN TN	- - - -	FP to TN FP to TP FP to FN FP excluded FN to TN FN to TP FN to FP FN excluded Other	- - - - - - - - -	Adjustment of results not performed. 6 discordant cases. In 2 FP, micrometastasis was observed in the Western blot analysis. For the other 2 FP cases, Western blot was not performed. Explanation of discordance for FN unclear.
SLN level	81	Histopathology was repeated, examined and evaluated by third party pathologists.	TP FP FN TN	14 1 2 64	TP FP FN TN	- - - -	FP to TN FP to TP FP to FN FP excluded FN to TN FN to TP FN to FP FN excluded Other	- - - - - - - - -	
Visser									

ALN level	346	For discordant samples, histology was extended to all levels. The lysates were analysed with qRT-PCR and Western blot analysis.	TP FP FN TN	61 15 3 267	TP FP FN TN	61 8 3 267	FP to TN FP to TP FP to FN FP excluded FN to TN FN to TP FN to FP FN excluded Other	0 0 0 7 0 0 0 0 0	If further analysis yielded a result compatible with a positive OSNA result, these samples were excluded from the final analysis due to the indication for sampling bias. 7 FP were considered to be due to sampling bias, leaving 11 unresolved. In 8

4.2.2.4 Summary of test accuracy

The results are summarised and stratified according to SLN, ALN, patients, before TAB and after TAB in Table 27 to Table 32(pages 97 to 99). An overall summary is provided in Table 33, page 100,

With regard to Metasin, the results must be used with caution, since they have been taken from unpublished, non-peer reviewed papers. These papers are also very much in draft form and therefore lacking in detail. That said, with regards to quality, many of the issues (such as lack of information about replicate measurement and therefore no estimate of the reproducibility and the robustness of the test) apply to both papers reporting OSNA and [redacted] reporting Metasin.

OSNA detects only CK19 protein expression. It has been reported that 98.2% of breast cancer tumours express CK19 protein, leaving 1.8% patients for whom this technique may be invalid.[62] However, Tamaki et al. suggest it is not necessarily the case that low expression of CK19 protein is reflected with low expression of CK19 mRNA.[64] In contrast, Metasin, which also identifies mammaglobin, produces an increased sensitivity of 92% as compared to OSNA ranging from 77.8 – 80%.[41]

As the reference standard, IHC is capable of detecting ITCs, unlike OSNA (and presumably Metasin). However, since the American Society of Clinical Oncology (ASCO)[70] and NICE clinical guideline for early breast cancer[16] indicates they have an unknown clinical significance, and there are insufficient data to recommend appropriate treatment, ITCs were considered histologically negative throughout this review.

Although the aim of OSNA and Metasin is to perform intraoperative molecular analysis, some studies employed a more experimental approach and used frozen samples of RNA, either initially or in discordant case analysis. Both storage and the freeze thaw effect have been known to adversely affect RNA, effectively leading to a reduction in concentration.

Histopathology will always be limited by sampling bias, since only a certain number of slices are taken. This can be increased or reduced according to requirements, but will have cost implications and there will always be unanalysed tissue. Furthermore, there is the possibility of observer subjectivity. In contrast, for OSNA or Metasin, the whole node can be used and the semi-quantitative test is objective.

There is also the issue of TAB. Unfortunately, due to the nature of the OSNA/Metasin, and of histopathology, it is not possible to use the same tissue for both tests, and thus TAB is likely to occur. There is no diagnostic test accuracy (DTA) study design that can resolve this issue,

and whilst cohort studies can provide us with some data on whole node analysis, the tissue being analysed by each test is still, of course, different. Studies have attempted to mitigate the issue of TAB by investigating discordant cases. However, there is a lack of consistency between studies with regards the methods used to investigate discordant cases, and there will always be uncertainty attached to decisions about whether TAB has occurred.

With regard to overall results, the range of specificity is lower for results presented by patient, rather than node, before and after adjustment for TAB as shown in Table 27 to Table 34. The results produced by meta-analysis agree with this. Sensitivity before TAB is similar for both patient and node, whereas after TAB, the sensitivity is slightly greater for nodes. The changes may be due to that fact that some studies use a larger number of nodes from a small number of patients.

With the exception of Feldman, the range of sensitivity and specificity for ALN and SLN appear to be similar.

Table 27. Results for Patients before TAB

First author	Patient (n)	Histology	H+/O+	H-/O-	H+/O-	H-/O+	% Sensitivity (95% CI)	% Specificity (95% CI)
Metasin								
[redacted]	[redacted]	NR	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Sundaresan[41]	1265	3 level	249	940	20	36	92 (89-95)[a]	97 (95-97)[a]
OSNA								
Choi (SLN)[55]	199	3 level	27	157	9	6	77.8 (60-90)	96.3 (92-99)
Frere Belda (SLN)[60]	233	5 level	33	168	9	23	78.6 (63.1–89.7)	88.0 (82.4–92.3)
Khaddage (SLN)[59]	46	5 level	8	35	2	1	80.0 (44.4-97.5)[a]	97.2 (85.5-99.9)[a]
Khaddage (SLN)[59]	197	1 level	25	155	0	17	100 (88.7-100)[b]	90.1 (84.6-94.1)[b]
Tamaki (SLN) [65]	417	1 level	58	315	8	36	87.9 (77.5-94.6)[b]	89.7 (86.1-92.7)[b]
Bernet Vegue (Non SLN)[53]	55	1 level	6	26	0	23	100 (41.4-100)[b]	53.1 (38.3-67.5)[b]

NR – not reported. For clarity O+ and O- refer to positive and negative results for OSNA or Metasin, [a]CI calculated by PenTAG . [b] CI, sensitivity and specificity calculated by PenTAG

Table 28. Patients after TAB (SLN only)

First author	Patient (n)	Histology	H+/O+	H-/O-	H+/O-	H-/O+	% Sensitivity (95% CI)	% Specificity (95% CI)
OSNA								
Frere Belda [60]	215	5 level	32	168	3	12	91.4 (76.9–98.2)	93.3 (88.6–96.6)
Khaddage[59]	46	5 level	NR	NR	NR	NR	100	97.2
Snook[63]	194	5 level	44	137	5	8	89.8 (77.8-96.6)[a]	94.5 (89.4-97.6)[a]

NR – not reported. For clarity O+ and O- refer to positive and negative results for OSNA or Metasin. [a]CI calculated by PenTAG . [b] CI, sensitivity and specificity calculated by PenTAG

Table 29. Results for SLN before TAB

First author	Sample no	Histology	H+/O+	H-/O-	H+/O-	H-/O+	% Sensitivity (95% CI)	% Specificity (95% CI)
Metasin								
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Sundaresan[41]	2279	3 level	341	1770	35	60	93 (87.3-93.4)[c]	97 (95.8-97.5)[a]
OSNA								
Feldman[56]	1044	3 level	107	868	31	38	77.5 (69.7-84.2)	95.8 (94.3-97.0)
Frere Belda[60]	503	5 level	51	413	12	27	80.9 (69.0–89.8)	93.9 (91.2–96.0)
Khaddage[59]	80	5 level	15	62	2	1	88.2 (63.6-98.5)[a]	98.4 (91.5-100)[a]
Tsujimoto[66]	81	3 level	14	64	2	1	87.5 (61.7-98.4)[b]	98.5 (91.7-100)[b]

NR – not reported. For clarity O+ and O- refer to positive and negative results for OSNA or Metasin. [a]CI calculated by PenTAG . [b] CI, sensitivity and specificity calculated by PenTAG. [c] Specificity calculated by PenTAG shown to be 91%.

Table 30. Results for SLN after TAB

First author	Sample no	Histology level	H+/O+	H-/O-	H+/O-	H-/O+	% Sensitivity (95% CI)	% Specificity (95% CI)
OSNA								
Feldman	1018	3 level	NR	NR	NR	NR	82.7	97.7
Frere Belda	481	5 level	51	413	5	12	91.1 (80.3–97.1)	97.2 (95.1–98.6)
Khaddage	78	5 level	NR	NR	NR	NR	100	98.4
Snook	395	5 level	66	313	6	10	91.7 (82.7-96.9)[a]	96.9 (94.4-98.5)[a]

NR – not reported. For clarity O+ and O- refer to positive and negative results for OSNA or Metasin. [a]CI calculated by PenTAG . [b] CI, sensitivity and specificity calculated by PenTAG.

Table 31. Results for ALN before TAB

Reference	Sample no	Histology level	H+/O+	H-/O-	H+/O-	H-/O+	% Sensitivity (95% CI)	% Specificity (95% CI)
OSNA								
Schem	343 ALN	5 level	104	211	2	26	98.1 (93.4-99.8)[a]	91.7 (84.3-92.7)[c]
Tamaki 2009 a	124 ALN	Sectioned at 0.2mm intervals	19	101	1	3	95 (75.1-99.9)	97.1 (91.8-99.4)
Tamaki 2009 b	450 ALN	3 level	70	348	10	22	87.5 (78.2-93.8)	94.1 (91.0-96.3)
Tsujimoto 2007	325 ALN/SLN	3 level	43	276	2	4	95.6 (84.9-99.5)[b]	98.6 (96.4-99.6)[b]
Vegue	567 ALN	1 level	6	522	0	39	100 (60.7-100)[b]	93.0 (90.6-95.0)[b]
Visser	346 ALN	3 level	61	267	3	15	95.3 (84.9 – 99.5)[a]	94.7 (96.4-99.6)[a]

For clarity O+ and O- refer to positive and negative results for OSNA or Metasin. [a]CI calculated by PenTAG . [b] CI, sensitivity and specificity calculated by PenTAG. Specificity calculated by PenTAG shown to be 89%.

Table 32. Results for ALN after TAB

Reference	Sample no	Histology level	H+/O+	H-/O-	H+/O-	H-/O+	% Sensitivity (95% CI)	% Specificity (95% CI)
OSNA								
Schem	330 ALN	5 level	NR	NR	NR	NR	100	95.6
Tamaki 2009	450 ALN	Sectioned at 0.2mm intervals	71	348	10	21	87.7 (78.5-93.9)	94.3% (95.3 – 98.8)
Visser	339 ALN	3 level	NR	NR	NR	NR	95.3	97.1

NR – not reported. For clarity O+ and O- refer to positive and negative results for OSNA or Metasin

Table 33. Overall range of central estimates for sensitivity and specificity

	Sensitivity (%)	Specificity (%)
Patients before TAB (Metasin)	[redacted]	[redacted]
Patients before TAB (OSNA - SLN)	77.8 - 80.0	88.0 - 97.2
Patients after TAB (OSNA-SLN)	89.8 - 100	93.3 - 97.2
SLN before TAB (OSNA)	77.5 - 88.2	93.9 - 98.4
SLN after TAB (OSNA)	82.7 - 100	96.9 - 98.4
ALN before TAB (OSNA)	87.5 - 98.1	91.7 - 97.1
ALN after TAB (OSNA)	87.7 - 100	94.3 – 97.1

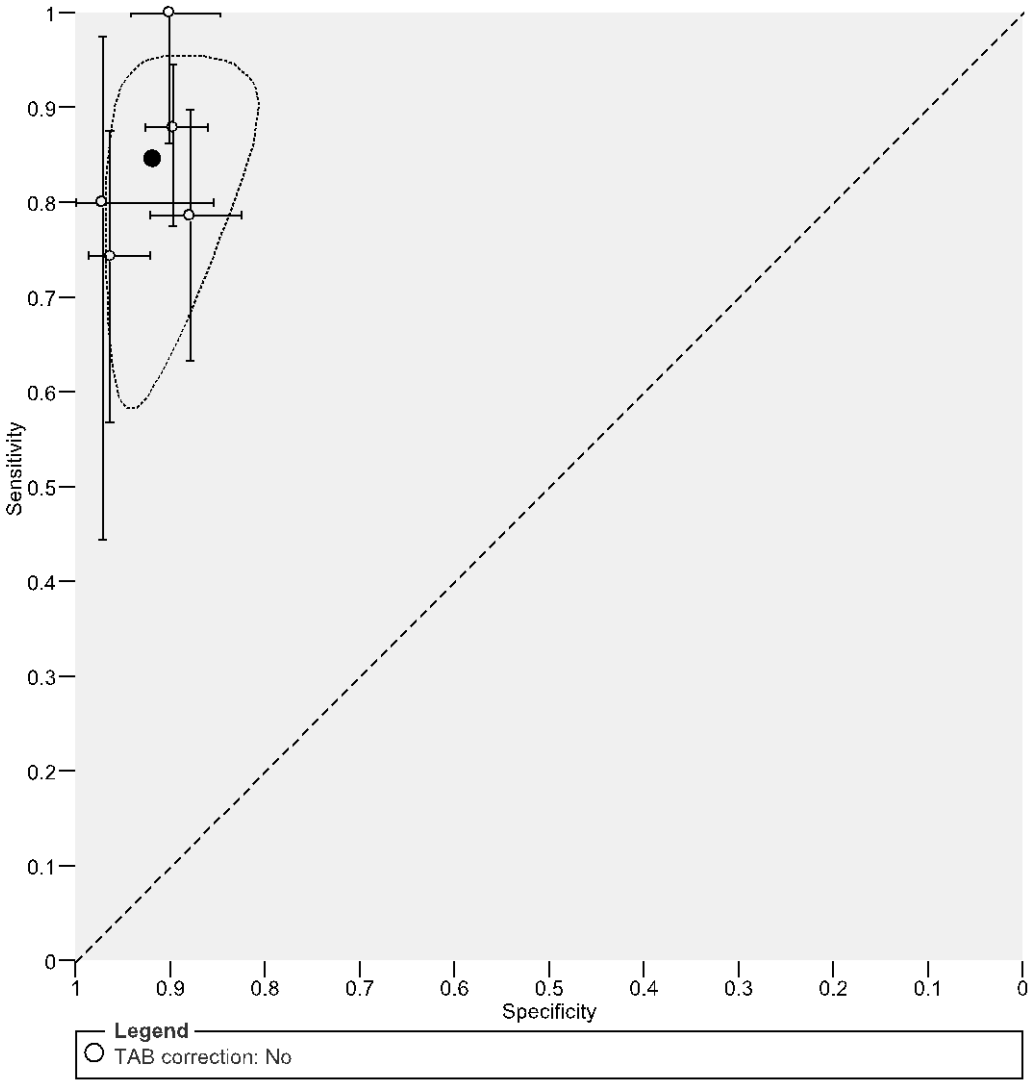
The results of the meta-analysis for OSNA are shown in Table 34.

Table 34. Meta-analyses of OSNA test accuracy

Sample type	Adjustment for TAB	Number of studies	Sensitivity (95% CI)	Specificity (95% CI)
Patient	No	5	0.845 (0.747–0.910)	0.918 (0.878–0.946)
Patient	Yes	3	0.913 (0.836–0.956)	0.942 (0.912–0.962)
SLN	No	4	0.799 (0.742–0.846)	0.955 (0.941–0.965)
SLN	Yes	5	0.890 (0.821–0.934)	0.975 (0.966–0.982)
ALN	No	6	0.951 (0.900–0.976)	0.949 (0.912–0.969)
ALN	Yes	4	0.965 (0.873–0.991)	0.962 (0.934–0.978)

For the meta-analysis of test accuracy based on analysis of SLNs without adjustment for TAB we were not aware of a compelling reason to believe the positivity threshold might vary between studies and so for consistency with meta-analyses for the other subgroups (see below) we do not include an HSROC curve or prediction region. Figure 8 shows the results of this meta-analysis.

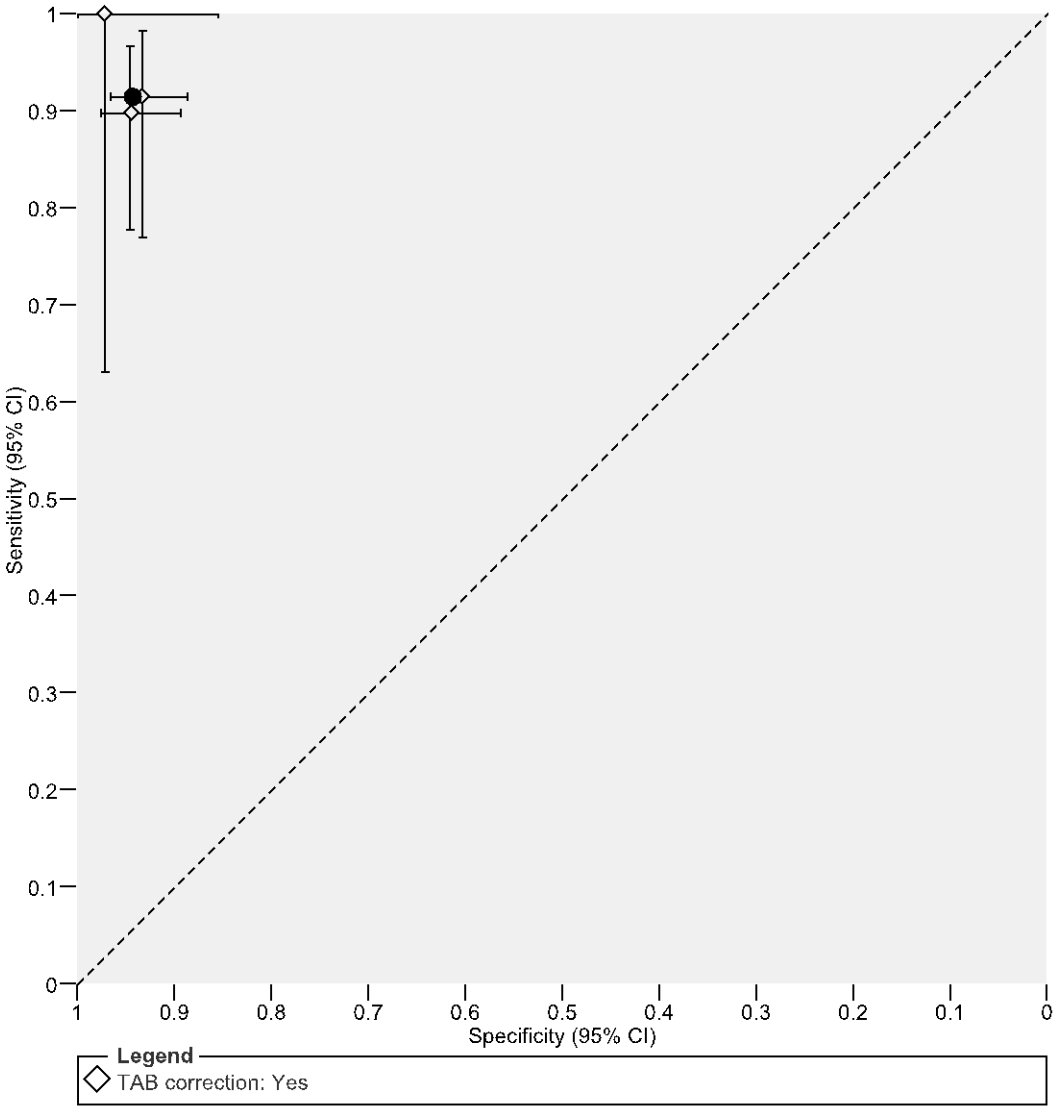
Figure 8. Meta-analysis and forest plot of test accuracy of OSNA for patients (based on analysis of sentinel lymph nodes) without adjustment for tissue allocation bias



Study	TP	FP	FN	TN	TAB correction	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Choi 2011	26	6	9	157	No	0.74 [0.57, 0.88]	0.96 [0.92, 0.99]		
Khaddage 2011	25	17	0	155	No	1.00 [0.86, 1.00]	0.90 [0.85, 0.94]		
Khaddage 2011a	8	1	2	35	No	0.80 [0.44, 0.97]	0.97 [0.85, 1.00]		
Le Frere-Belda 2012	33	23	9	168	No	0.79 [0.63, 0.90]	0.88 [0.82, 0.92]		
Tamaki 2012	58	36	8	315	No	0.88 [0.78, 0.95]	0.90 [0.86, 0.93]		

As there were only three studies to estimate the test accuracy of OSNA for patients with adjustment for TAB we set the correlation parameter to zero as described in Section 4.1.5.1, page 53. The results are displayed in Figure 9.

Figure 9. Meta-analysis and Forest plot of test accuracy of OSNA for patients (based on analysis of sentinel lymph nodes) with adjustment for tissue allocation bias

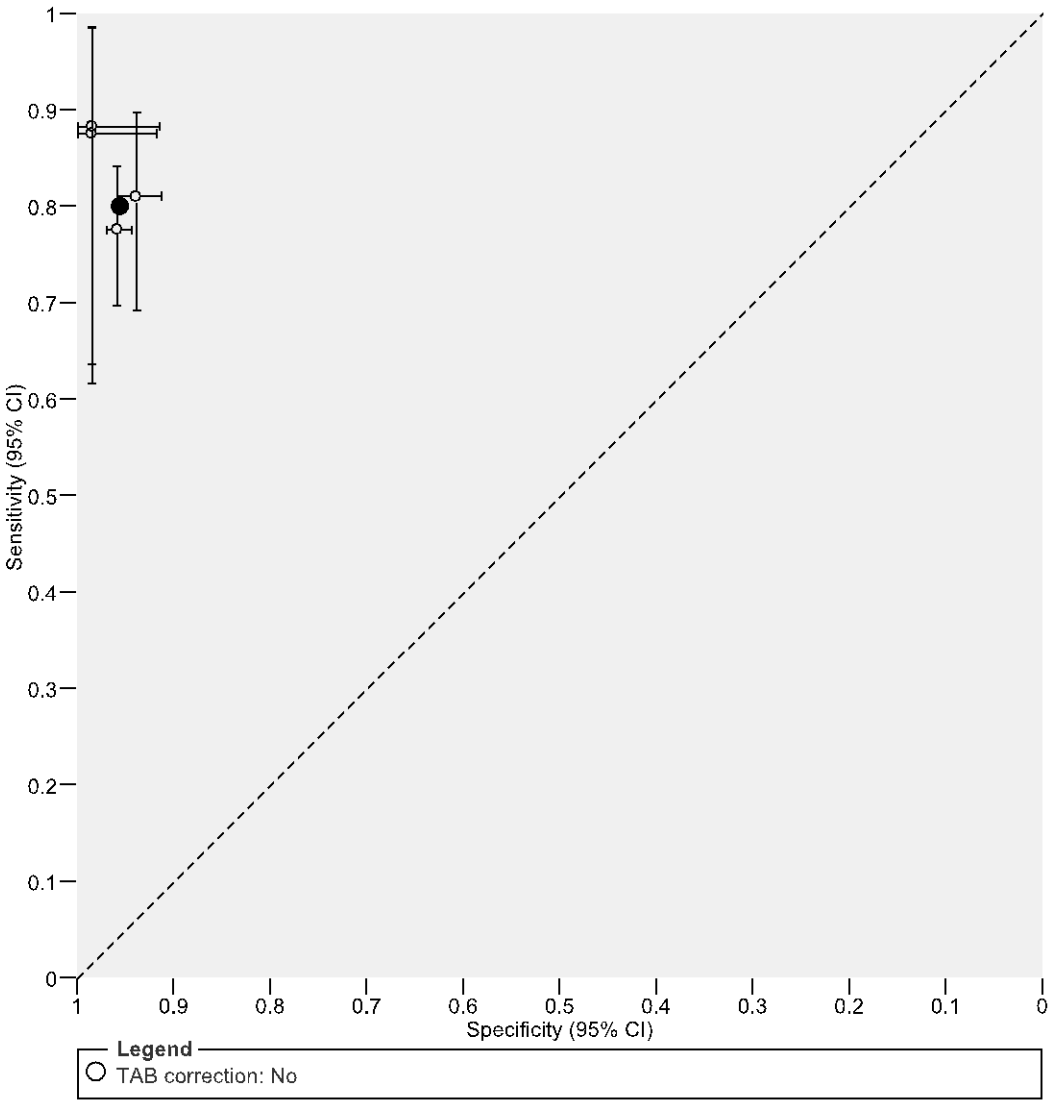


Study	TP	FP	FN	TN	TAB correction	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Khaddage 2011b	8	1	0	35	Yes	1.00 [0.63, 1.00]	0.97 [0.85, 1.00]		
Le Frere-Belda 2012a	32	12	3	168	Yes	0.91 [0.77, 0.98]	0.93 [0.89, 0.97]		
Snook 2011	44	8	5	137	Yes	0.90 [0.78, 0.97]	0.94 [0.89, 0.98]		

For the meta-analyses of test accuracy of OSNA for SLNs and ALNs we do not present SROC curves or prediction regions because the studies report the same copies/µL thresholds of 250 and 5,000 for OSNA results (+; micrometastases) and (++; macrometastases), respectively.

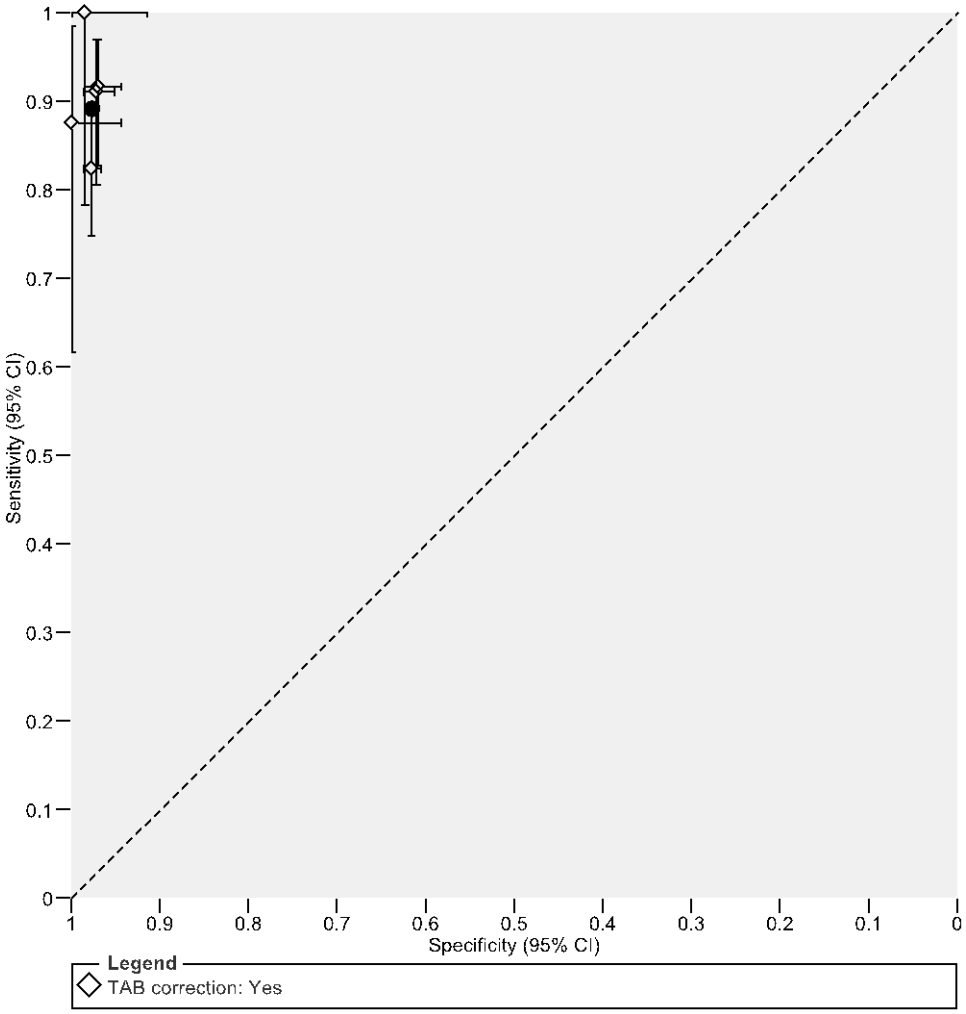
We attempted to perform a meta-analysis of test accuracy of OSNA for SLNs with and without adjustment for TAB using the full bivariate method. In both of these meta-analyses we encountered atypical results in the output such that the between-study correlation parameter ($\sigma_{AB}/\sqrt{\sigma_A^2\sigma_B^2}$ in Reitsma et al. 2005) is estimated as -1 with no estimate of standard error or confidence intervals.[71] This phenomenon has been recognised before and Riley et al. conclude it is likely to occur when there are few studies or the within-study variation is large (i.e., the studies themselves are small) and is due to sensible bounds being placed on the model to avoid the maximum likelihood estimation including a correlation outside the range $[-1,1]$.[72] Riley et al. also conclude that there is no systematic bias in the summary estimates of sensitivity and specificity introduced by this phenomenon and that confidence intervals of the summary point are conservative. However, as described in Section 4.1.5.1, page 53 we set the correlation parameter to zero and repeated the analysis, the results of which are shown in Figure 9 and Figure 10. Computed summary points and confidence intervals differed only at the third decimal point.

Figure 10. Meta-analysis and Forest plot of test accuracy of OSNA in sentinel lymph nodes without adjustment for tissue allocation bias



Study	TP	FP	FN	TN	TAB correction	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Feldman 2011	107	38	31	868	No	0.78 [0.70, 0.84]	0.96 [0.94, 0.97]		
Khaddage 2011	15	1	2	62	No	0.88 [0.64, 0.99]	0.98 [0.91, 1.00]		
Le Frere-Belda 2012	51	27	12	413	No	0.81 [0.69, 0.90]	0.94 [0.91, 0.96]		
Tsujimoto 2007	14	1	2	64	No	0.88 [0.62, 0.98]	0.98 [0.92, 1.00]		

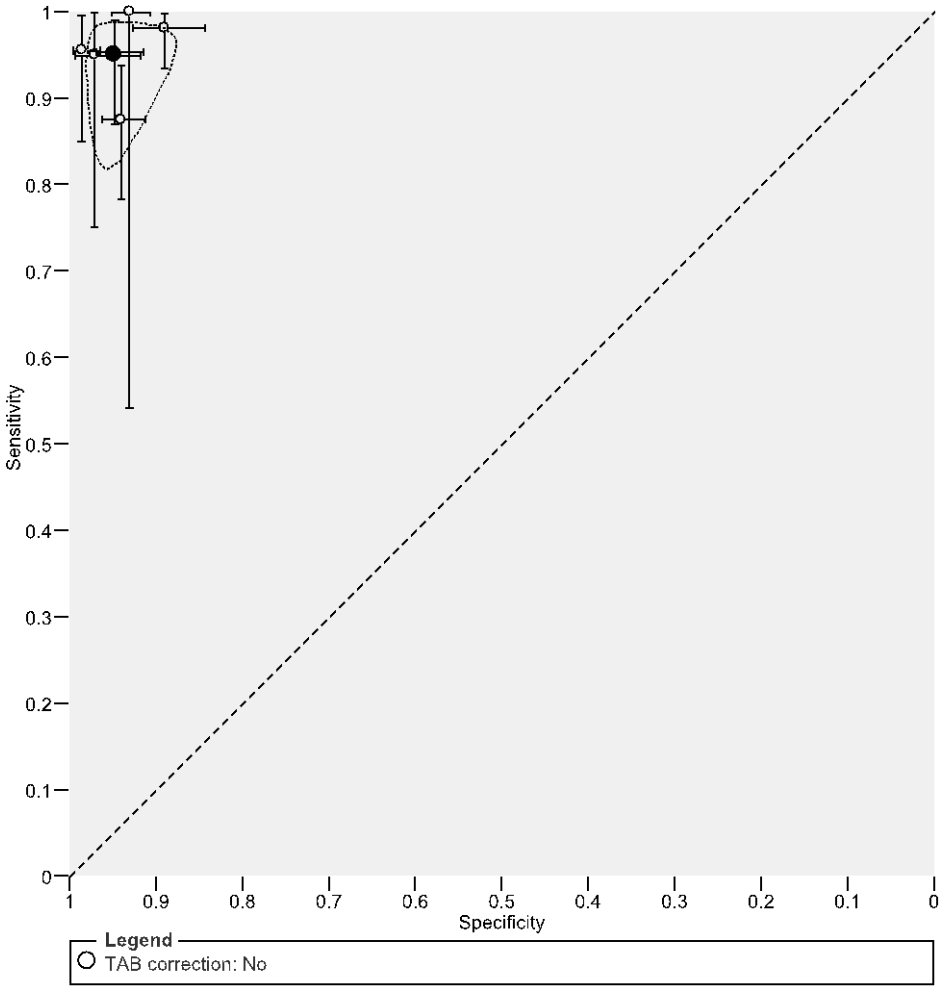
Figure 11. Meta-analysis of test accuracy of OSNA in sentinel lymph nodes with adjustment for tissue allocation bias



Study	TP	FP	FN	TN	TAB correction	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Feldman 2011a	107	20	23	868	Yes	0.82 [0.75, 0.88]	0.98 [0.97, 0.99]		
Khaddage 2011a	15	1	0	62	Yes	1.00 [0.78, 1.00]	0.98 [0.91, 1.00]		
Le Frere-Belda 2012a	51	12	5	413	Yes	0.91 [0.80, 0.97]	0.97 [0.95, 0.99]		
Snook 2011	66	10	6	313	Yes	0.92 [0.83, 0.97]	0.97 [0.94, 0.99]		
Tsujimoto 2007a	14	0	2	64	Yes	0.88 [0.62, 0.98]	1.00 [0.94, 1.00]		

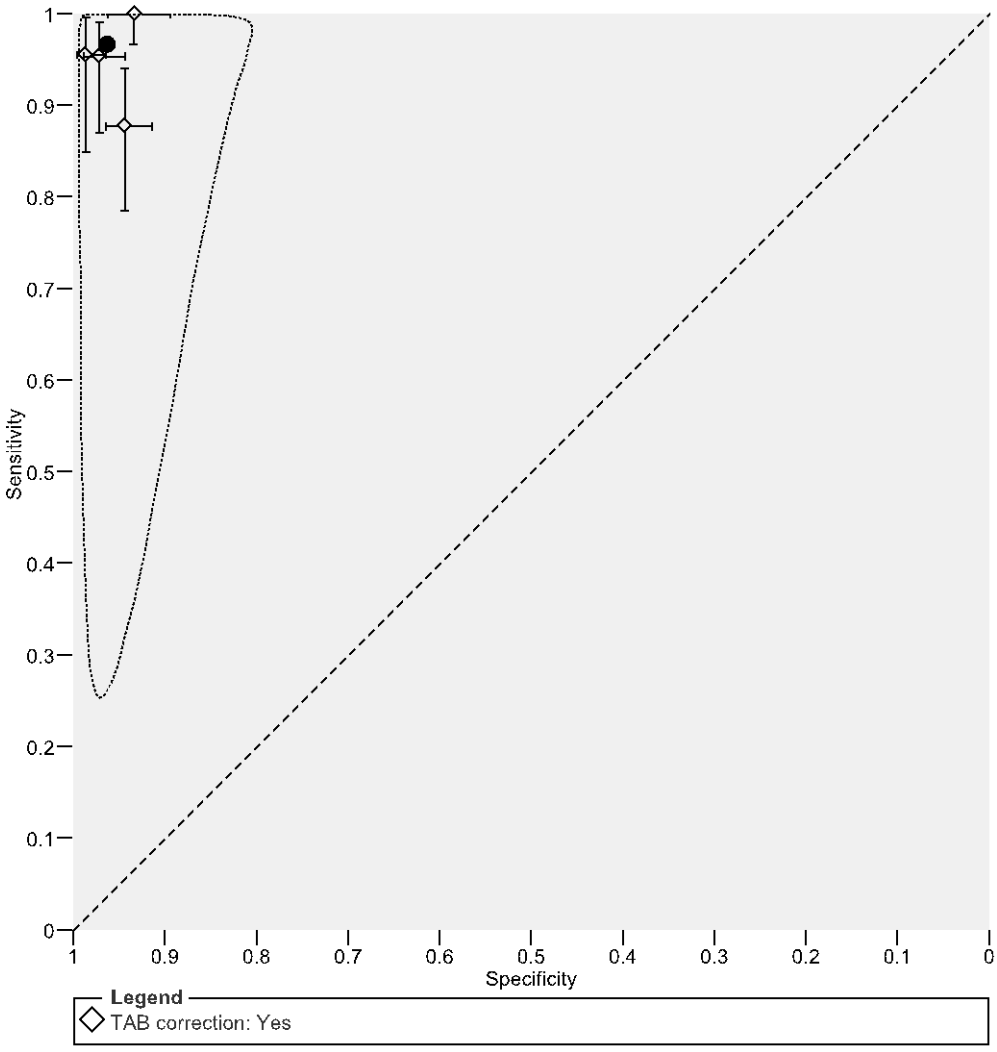
Figure 12 and Figure 13 show the results of meta-analysis of test accuracy of OSNA for ALNs respectively before and after adjustment for TAB. The bivariate method converged with no abnormal results. Note that there were six studies to inform the meta-analysis before adjustment for TAB but only four studies to inform the meta-analysis after adjustment and we believe this accounts for the significantly larger confidence region shown in Figure 13 as removing two studies from the subgroup without adjustment for TAB gives a confidence region of a comparable size.

Figure 12. Meta-analysis and Forest plot of test accuracy of OSNA for axillary lymph nodes without adjustment for tissue allocation bias



Study	TP	FP	FN	TN	TAB correction	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Schem 2009	104	26	2	211	No	0.98 [0.93, 1.00]	0.89 [0.84, 0.93]		
Tamaki 2009	19	3	1	101	No	0.95 [0.75, 1.00]	0.97 [0.92, 0.99]		
Tamaki 2009a	70	22	10	348	No	0.88 [0.78, 0.94]	0.94 [0.91, 0.96]		
Tsujimoto 2007	43	4	2	276	No	0.96 [0.85, 0.99]	0.99 [0.96, 1.00]		
Vegué 2012	6	39	0	522	No	1.00 [0.54, 1.00]	0.93 [0.91, 0.95]		
Visser 2008	61	15	3	267	No	0.95 [0.87, 0.99]	0.95 [0.91, 0.97]		

Figure 13. Meta-analysis and Forest plot of test accuracy of OSNA for axillary lymph nodes with adjustment for tissue allocation bias



Study	TP	FP	FN	TN	TAB correction	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Schem 2009a	104	15	0	211	Yes	1.00 [0.97, 1.00]	0.93 [0.89, 0.96]		
Tamaki 2009b	71	21	10	348	Yes	0.88 [0.78, 0.94]	0.94 [0.91, 0.96]		
Tsujimoto 2007a	43	4	2	276	Yes	0.96 [0.85, 0.99]	0.99 [0.96, 1.00]		
Visser 2008a	61	8	3	267	Yes	0.95 [0.87, 0.99]	0.97 [0.94, 0.99]		

4.2.3 Assessment of analysis time

The time taken for analysis of the lymph node across the studies is displayed in Table 35. It is unclear if the timings are directly comparable since the description of exactly which procedures were being monitored was sometimes ambiguous. That said, the results seem fairly consistent for both methods, between <30 to 39.6 min for 1 node increasing by approximately 5 to 10 minutes per additional node analysed.

It was also noted by Bernet et al. that the longest and most variable time period corresponded to the stage in which the node was transported from operating room to pathology department.[52] The time of macroscopic processing of the samples could also fluctuate significantly depending on the training level of the pathologist involved. The least variable time-period corresponded to the homogenisation of tissue, preparation of the diluted sample and amplification in the amplification equipment.

Table 35. Time to analysis

Nodes (n)	Median time to analysis, min								
	OSNA								Metasin
	Bernet [a,b] (range)	Choi [a,c]	Feldman[d]	Frere Belda[e]	Godey[a,f] (std)	Khadd -age[g]	Snook[h] (range)	Tsujimoto[i]	Sunda-resan
1	39.6 (26-70)	35.2	33.0	33	32.9 (4.9)		32 (22-97)	<30 min	36
2		44.8	39.6	40	36.4 (4.5)	37	42 (30-73)		42
3		50.4	45.2	48	41.6 (5.2)		51 (38-73)		46
4		50.0		54	48.5 (8.7)		62 (46-90)		

[a] mean, [b] time from receipt of node to report, [c]turnaround time, [d] interquartile mean from homogenisation to analyser output, [e] time needed for OSNA assay, [f] time required for results, [g] time from receipt of node to the result, [h] from node preparation to end of analysis, [i] all OSNA assays completed in under 30 min

The Guillen-Paredes study did not provide a time to analysis, but compared on operative time, days in hospital and complications, between histology and OSNA, as shown below in Table 36 and Table 37.[58]

Table 36. Effect of OSNA on operative time

	Mean Intervention Time, mins (sd)			Mean Days in Hospital (sd)		
	1st Operation	2nd	Total	1st	2nd	Total
Histology	57.11 (23.93)	78.33 (NR)	78 (48.02)	1.8 (2.04)	2.41 (1.09)	2.44(0.78)
OSNA	62.14 (48.02)	NA	62.14(21.93)	1.54(0.78)	NA	1.54(0.78)
	Absolute no. Intervention Time (mins)			Absolute no. Days in Hospital		

	1st Operation	2nd	Total	1st	2nd	Total
Histology	2570	940	3510	81	29	110
OSNA	2175	NA	2175	54	NA	54

Table 37. Number of complications per group

	Complications in 1st intervention			Complications in 2nd intervention		
	None	Minor	Major	None	Minor	Major
Histology	28	17	0	4	8	0
OSNA	24	10	1	N/A	N/A	N/A

Table 37 indicates that overall, patients undergoing OSNA received fewer complications, although the one major complication occurred in the OSNA group.

5 Assessment of cost-effectiveness: systematic review

5.1 Systematic review of existing cost-effectiveness evidence

5.1.1 Search strategy

The main objective of this review was to identify and evaluate published studies that examine the cost effectiveness of intraoperative tests in the diagnosis of axillary lymph node metastases. As a secondary objective, we also used the studies to inform the design of the model used in our independent economic assessment.

We reviewed published economic evaluations of intraoperative molecular assessment for metastasis in early breast cancer to identify evidence relevant to current NHS practice. In addition to electronic databases searched in the effectiveness section, the NHS Health Economic Evaluation Database (NHS HEED) and EconLit were searched for cost, cost-effectiveness and cost-utility studies. Forward citations of identified studies were searched for any relevant publications published after the initial search.

Relevant studies were then identified in two stages. Titles and abstracts returned by the search strategy were examined independently by two researchers (NH and RM) and screened for possible inclusion. Disagreements were resolved by discussion. Full texts of the identified studies were obtained. Two researchers (NH and RM) examined these independently for inclusion or exclusion, and disagreements were resolved by discussion.

5.1.2 Description of included studies

The initial search identified a total of 13 abstracts, 7 of which were of conference presentations, and the remaining constituted 4 individual studies. One study measured costs of intraoperative options that are not relevant to the UK (touch imprint and frozen section; Classe et al., 2012 [7,73]), one study analysed the diagnostic pathway leading to sentinel lymph node biopsy or ALND. [74,4] The remaining two studies were identified as relevant to our review. See Appendix 7 and Appendix 8 for details on excluded studies.

The two included studies were Cutress et al., 2010[3] and Guillen-Paredes et al., 2011.[58] A second report for the first of these studies was available from an NTAC report by Burke and Patton 2010, and was used to complement the information from its associated published paper.[5] The characteristics of the studies are summarised in Table 38, page 112. Both were single centre observational studies, comparing an intraoperative test with histopathology as the gold standard for assessing SLNB. The Cutress study was set in UK, and assessed the GeneSearch assay

while the Guillen-Paredes paper was conducted in Spain, and evaluated the OSNA assay. Both found their respective intraoperative test to be cost-effective compared to histopathology, with both assays being cost-saving whilst reducing theatre time and length of hospital stay. The UK study also considered a strategy where axillary clearance was performed on all patients instead of assessing them using SLNB, but this practice is no longer recommended by NICE [16] and the study did not find it to be cost-effective. Guillen-Paredes et al. also measured benefits through the number of minor and major complications during surgery. They found the complication incidences to be significantly fewer in the OSNA group, although this group was also the only one to have a major complication (no details were given). Neither study looked at outcomes beyond the diagnostic phase.

It must be noted that the study by Cutress and colleagues while providing a unique source of evidence on resource use and costs of intra-operative molecular testing in the UK, provides limited if any evidence on the economic outcomes of a specific intra-operative test as it refers to a test that has been withdrawn from the market (Genesearch produced by Veridex). Metasin, one of the intra-operative testing technologies being evaluated in this assessment, uses the same markers as Genesearch, CK19n and Mammaglobin, but different primer-probe combinations, and is therefore expected to perform differently in routine practice to Genesearch.

Table 38. Cost effectiveness study characteristics

Author, year published	Setting, perspective	Population	Study purpose	Study approach	Diagnostic comparators	Outcomes measured	Base results
Cutress 2010	UK healthcare setting, hospital trusts	Patients diagnosed with breast cancer eligible for SLNB	Cost-benefit analysis of the use of the intraoperative test GeneSearch to diagnose sentinel lymph node metastases	Prospective, single centre, observational study, costs estimated retrospectively	Model 1: No SLNB, axillary clearance performed Model 2: SLNB analysed by histopathology. Node negative, no further treatment. Node positive, axillary clearance in second operation Model 3: Half node SLNB analysed with GeneSearch. Node negative, no further treatment. Node positive, axillary clearance in same operation. Rest of the node analysed with histopathology	Costs *Theatre time* *Length of hospital stay*	Cost effectiveness improves with intraoperative testing- 28% avoid a second operation and cost per case reduced from £2891 to £2833
Guillen-Paredes 2011	Spain, hospital setting	Patients with early stage breast cancer, who are ultrasound negative and underwent SLNB	Cost-benefit analysis of the use of intraoperative OSNA assay to diagnose sentinel lymph node metastases	Retrospective, single centre study	Group1: Histopathology. Metastases, axillary clearance in second Group2: OSNA. Metastases, axillary clearance in first op	Costs *Theatre time* *Length of hospital stay*	Overall Group 2 had a cost saving of 319.99 euros per patient per intervention, shorter theatre time and length of hospital stay and fewer complications.

5.1.3 Quality appraisal

A quality appraisal was carried out on the two studies, using the Drummond checklist [75]. A summary of the results are provided in Table 39.

Table 39. Quality assessment of studies, using Drummond 1996[75]

Criteria	Cutress 2010	Guillen-Paredes 2011
Study design		
The research question is stated	✓	✓
The economic importance of the research question is stated	✓	✓
The viewpoint(s) of the analysis are clearly stated and justified		
The rationale for choosing alternative programmes or interventions compared is stated	partial	✓
The alternatives being compared are clearly described	✓	✓
The form of economic evaluation used is stated	✓	✓
The choice of form of economic evaluation is justified in relation to the question addressed	✗	✗
Data collection		
The source(s) of effectiveness estimates used are stated	✓	✓
Details of the design and results of effectiveness study are given (if based on a single study)	✓	✓
Details of the methods of synthesis or meta-analysis of estimates are given (if based on a synthesis of a number of effectiveness studies)	n/a	n/a
The primary outcome measure(s) for the economic evaluation are clearly stated	partial	✓
Methods to value benefits are stated	✗	✗
Details of the subjects from whom valuations were obtained were given	n/a	n/a
Productivity changes (if included) are reported separately	✗	✗
The relevance of productivity changes to the study question is discussed	n/a	n/a
Quantities of resource use are reported separately from their unit costs	✓	✓
Methods for the estimation of quantities and unit costs are described		
Currency and price date are recorded	✓	partial
Details of currency of price adjustments for inflation or currency conversion are given	✗	✗
Details of any model used are given	n/a	n/a
The choice of model used and the key parameters on which it is based are justified	n/a	n/a
Analysis and interpretation of results		
Time horizon of costs and benefits is stated	✓	✓
The discount rate(s) is stated	n/a	n/a
The choice of discount rate(s) is justified	n/a	n/a

5.1.3.1 Study design

Both studies were observational and therefore open to bias and confounding. They both stated their research question and the approach to economic evaluation, but no justification was given by either study for the economic evaluation study design used in relation to the research question. The viewpoint of both analyses was implicitly justified by the public health systems in which the studies were conducted and the local practice at the respective centres, and both studies acknowledged the limited generalisability of their findings due to the availability of different intraoperative testing technologies in other centres.

5.1.3.2 Data

Details of methods of patient recruitment were given. While both studies were based on single-gate designs and thus subject to TAB, none adjusted for it in the analysis of accuracy. Both studies reported methods of collecting health care resource quantity data and applying unit costs to them, but only the Spanish study explicitly stated the primary outcome measures of their evaluation in terms of the health benefits of intra-operative diagnosis. Both studies reported unit costs and quantities separately, but only provided explanations as to the estimation of unit costs, not the quantities. The Spanish study did not state the date of the unit costs used and neither study provided details on whether any price and currency conversion adjustments were made. Neither study valued health benefits nor examined changes in productivity or its associated costs.

5.1.3.3 Analysis and interpretation of results

Neither study analysed outcomes beyond the end of the diagnostic phase and therefore did not require the use of discount rate. Since no sensitivity analyses were provided, the degree to which cost differences were true differences as opposed to the results of chance alone, or estimated precisely cannot be established.

In summary, only one study was found for OSNA, and none for Metasin. One study was found on GeneSearch, which served as the basis for the development of Metasin (see Section 2.3.3, page 40) but may not produce the same outcomes. While the OSNA study is likely to reflect the study centre's practice in Spain, the GeneSearch study is no longer relevant as the technology has been withdrawn from the market. The validity of results in both studies is uncertain by their lack of information on the way accuracy was measured. Further, the degree of uncertainty in the estimates, and the extent to which the observed differences may be explained by chance, cannot be established. Nor can it be known to what extent the results may be generalisable across countries or indeed jurisdictions within the UK. This suggests that the existing evidence on

economic outcomes is unlikely to serve to inform medical decision making in the context studied here.

5.2 Submissions from sponsoring companies

No economic studies of Metasin were submitted by its sponsor. Only one conference abstract was found that included any information on costs; it only reported the cost per assay (£35.00; a different estimate of this, based on more detailed information obtained from the sponsor was used in the model; see Section 5.3.3.2.1). Only a file with illustrations of spreadsheets from a decision model of the costs and accuracy of OSNA developed for Belgium by Liven Annemanns, was made available, as unpublished and academically confidential work [76].

[redacted]

The model considered only the costs differences between OSNA and the status quo for diagnostic testing of metastatic early breast cancer, namely intraoperative testing by touch imprint cytology or frozen section with postoperative histopathology in negative cases in 90% of cases and only postoperative histopathology in the rest. The analysis assumed a SLN positive rate of 39%. In addition to the costs of the diagnostic tests, the analysis included the costs of further investigations after a negative intra-operative test result, the initial surgery without lymph nodes dissection, the extended surgery with intra-operative testing and ALND, and the cost of the second surgery including the hospital stay. The model analysed the costs accruing over a one year period following primary surgery for breast cancer. The authors' rationale for this choice was that 'it is anticipated that the systematic application of OSNA will lead to equivalent test results and long term patient management as current postoperative testing'

Using values of 95.6% sensitivity and 96.7% specificity, the mean per patient costs were estimated to be €771 (£603; at GDP PPP, OECD 2012[78]) with the status quo and €604 (£473) with OSNA, resulting in mean per patient cost savings with the latter of €167 (£131). The authors argue that their analysis 'would be equally applicable to the UK but should show a greater economic benefit given that most sentinel lymph node analysis in breast cancer patients in the UK market is performed post-operatively so savings in second surgery would be more pronounced'.

The results of this analysis are of limited value due to the lack of information made available to the ERG on the methods behind the analysis. The analysis was based on aggregate measures of costs, so that quantities of resource use were not discernible. The analysis claims to account for costs up to one year after the initial operation but from the available detail it is unclear whether any cost other than those associated with the diagnostic pathway were included. The analysis

did not attempt to account for the costs of adverse events nor the effects in terms of health benefits and quality of life. No account was made for the degree of uncertainty in model estimates, thus preventing an assessment of the likelihood that chance alone may explain the reported cost effects of OSNA.

5.3 Independent ERG assessment

5.3.1 Objective of analysis

The main analysis compares OSNA and Metasin with histopathology from the perspective of the NHS. The evaluation is presented for the outcomes occurring up to the staging the axilla. In addition, a separate analysis evaluated the long term outcomes of intra-operative testing options in terms of QALYs and costs. The long term analysis is intended as an illustration of the relative size of benefits of intra-operative diagnosis and the effect of uncertainty on its expected lasting impact.

5.3.2 Description of model

The model is split into two separate sections (diagnostic and management) to encompass both immediate and long-term outcomes. As the technology is expected to have a larger impact on the short term outcomes, it is this section that we particularly focus on.

5.3.2.1 Diagnostic pathway

This section of the model is used to both inform the rest of the model and provide the intermediate outcomes described in the definition of the decision problem (Section 3.1, page 46).

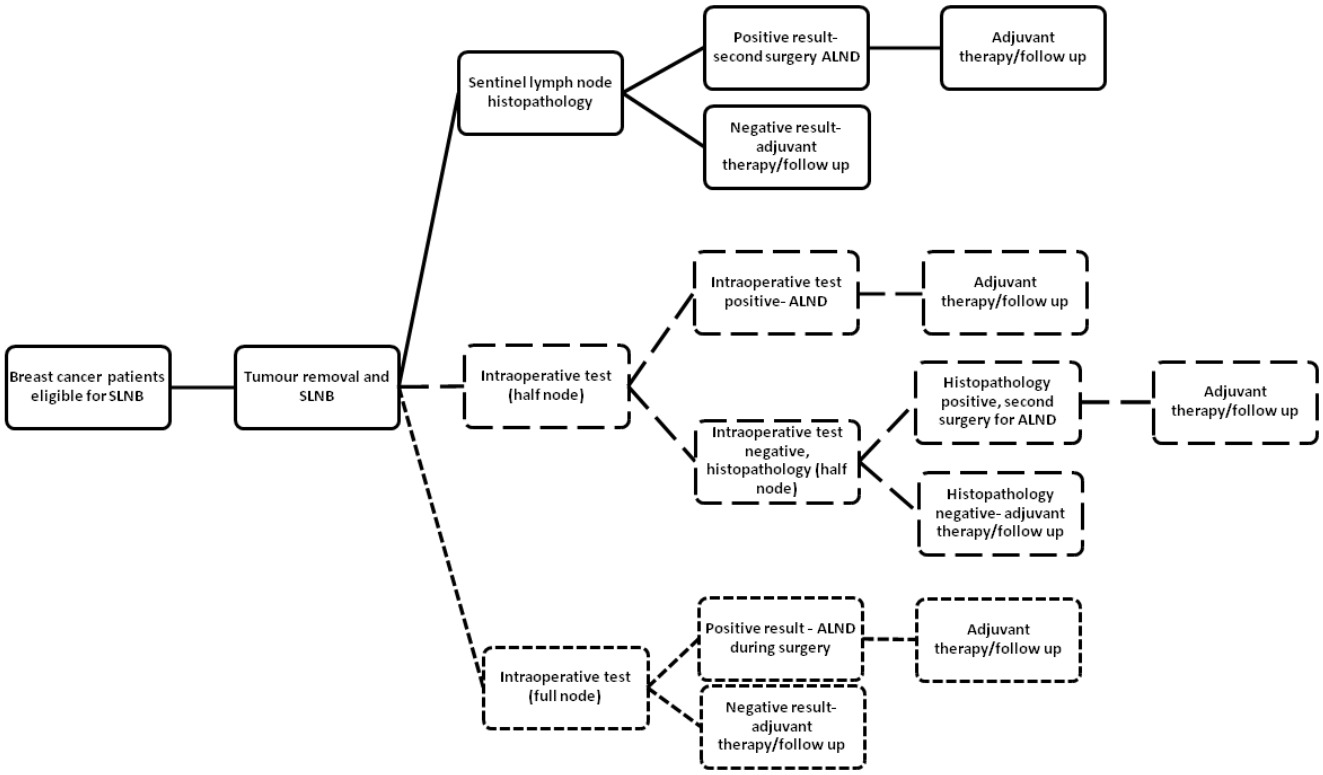
5.3.2.2 The ERG diagnostic pathway

Patients enter the model as those who have SLNB performed during their initial tumour removal. The model then splits into three different strategies, to encompass each of the possible combinations of diagnostic tests: intraoperative only (OSNA or Metasin, not in combination), histopathology only, or a combination of an intraoperative assay and follow up histopathology. These three strategies are shown in Figure 14, page 118. In the following descriptions, positive test results refer to results that indicate metastases in the sentinel lymph node. The three modelled pathways are therefore:

- Current practice: SLNB is analysed using histopathology of the full node. If positive for metastases a second surgery is performed where axillary lymph node dissection (ALND) occurs.
- Add in strategy: half of the sentinel lymph node from SLNB is analysed during the tumour removal operation using one of the intraoperative tests (OSNA or Metasin). Those with a positive result receive ALND during that surgery. For those with a negative result, the other half of the sentinel lymph node is kept to be analysed with histopathology. Patients where metastases were not detected at the intraoperative stage, but whose histopathology is positive receive ALND as a second operation.
- Replacement strategy: the full sentinel lymph node is assessed by the intraoperative test, with no histopathology. Those with a positive result will receive ALND during their tumour removal.

At this stage we calculate intermediate costs and patient outcomes. This analysis was developed in Excel.

Figure 14: Diagnostic pathway to test for axillary metastases



5.3.2.2.1 Assumptions

In this section of the model the following assumptions are made:

- Compliance to all procedures offered to the patient is 100%. This includes compliance to ALND and therefore all patients who test positive for sentinel lymph node metastases receive ALND. This assumption was adopted on the basis of the opinion elicited from clinical experts advising the ERG on this topic (see Appendix 9 for list of clinical experts).
- The failure rate of SLNB discussed in the Background (Section 2.2.6, page 36) is treated as 0. In reality, SLNB has a failure rate of <5% (the consensus from clinical experts advising the ERG on this topic). As this is relatively small and there is no evidence on the difference in impact when comparing the use of SLNB in intraoperative testing to SLNB in histopathology, we have to assume that the impact is the same and treat it as 0.
- It is unclear from the clinical effectiveness systematic review what a failed intraoperative test would be, as it is not something discussed in the studies. The number of reports that provide numbers of failed tests is small and suggests the rates are small; and our experts were not aware of failure rates, suggesting the concepts of a failed intraoperative test and the rate of failure are not yet well established. As the understanding of failure for OSNA is unclear, it is impossible to model successfully the impact of knowing when a test has

failed in the model. Therefore for the purposes of the model we assume the known failure rate to be 0. Instances where the failure status of intraoperative testing is unknown directly impact the sensitivity and specificity of the test and therefore can be implicitly incorporated into the model using a sensitivity analysis on the sensitivity and specificity parameter values.

- In the absence of relevant empirical data from studies using full node analysis, half node analysis data is used for all test accuracy data for all pathways. We expect this to underestimate the accuracy of these tests in a full node situation and thus serve as a conservative analysis of their benefits.
- Similarly, based on the data, histopathology is assumed to have sensitivity and specificity of 1 as it is used as the 'gold standard' in the evaluative studies. This may also slightly underestimate the accuracy of the intraoperative tests, as the accuracy of histopathology will not actually be 1.

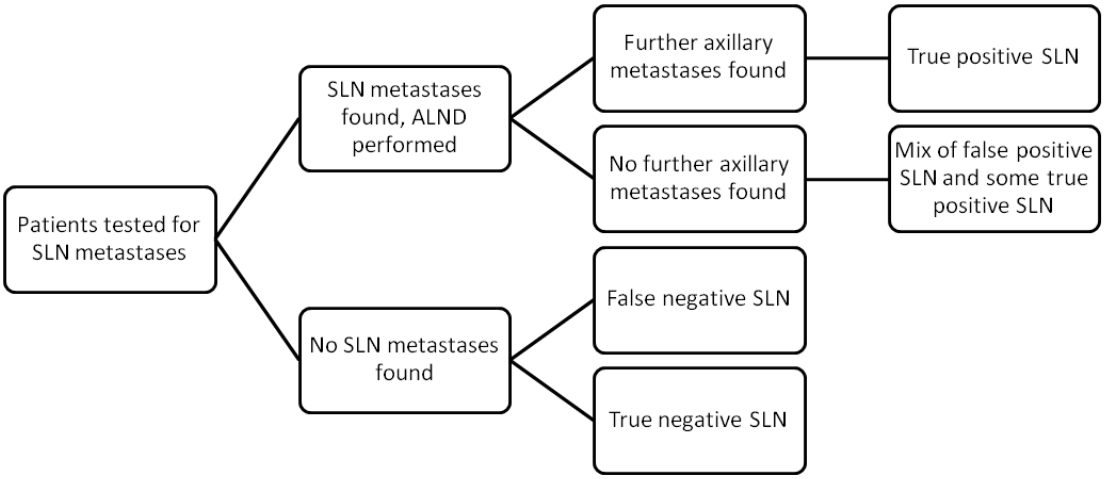
5.3.2.2.2 Post diagnosis (management pathway)

After the diagnostic pathway, our cohort splits into various management subgroups:

- Patients with sentinel lymph node metastases who test positive for metastases and who receive ALND (in either their first surgery or as a separate second surgery for ALND), *AND* who also test positive for additional axillary metastases when the material from their ALND is tested. The proportion of patients assumed to have additional axillary metastases after a false positive sentinel lymph node metastases diagnosis is assumed to be small enough that it can be modelled as 0%. The proportion of patients assumed to have additional axillary metastases after a true positive sentinel lymph node metastases diagnosis is taken from the Z0011 trial[38] and is set to 27.3%.
- Patients who test positive for metastases and who receive ALND (in either their first surgery or as a separate second surgery for ALND), *BUT* who do not test positive for additional axillary metastases when the material from their ALND is tested. These patients include those who were correctly identified with sentinel lymph node metastases and those who were incorrectly identified as having sentinel lymph node metastases.
- Patients without sentinel lymph node metastases who test negative for metastases and do not receive ALND.
- Patients with sentinel lymph node metastases who test negative for metastases and do not receive ALND.

A visual representation of these groups is given in Figure 15.

Figure 15: Post-diagnosis subgroups.



This diagram shows the pathway of patients immediately post-diagnosis and their true sentinel lymph node metastases status.

Once these subgroups are established, the model moves into its second section: the management pathway. This section of the model calculates the long term outcomes, most notably the costs and QALYs. We model each of the groups through from this point using the discrete event simulation model previously utilised by the ScHARR-TAG team[4,74] as a basis with updated parameters.

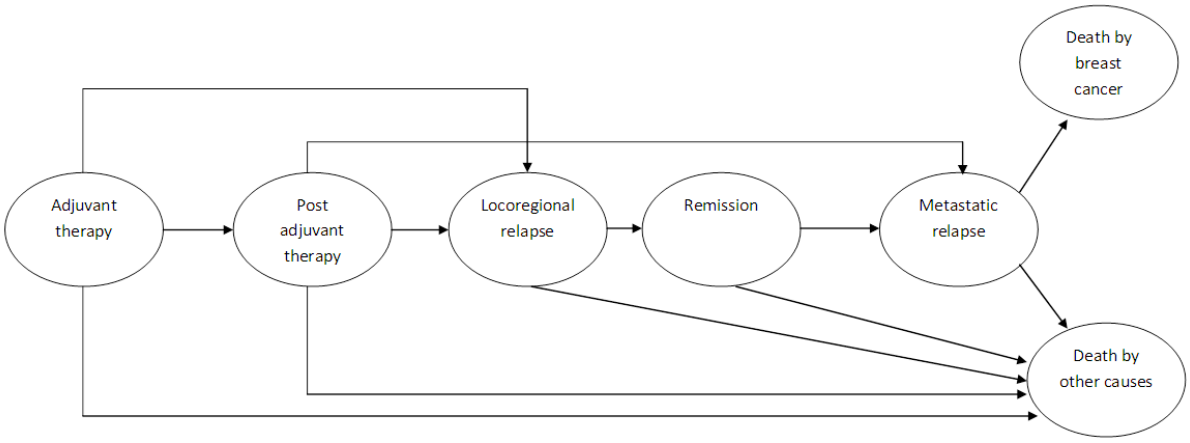
5.3.2.3 The ScHARR model structure

ScHARR uses a discrete event simulation model built in Simul8 for their economic evaluation of diagnostic imaging leading to SLNB or ALND without SLNB (without intraoperative testing) for the diagnosis of metastasis in early breast cancer. A discrete-event simulation (DES) models individual patients with individual attributes, such as the patient's diagnostic and disease history. Unlike Markov state transition models, where the transition probabilities are evaluated during fixed time intervals, the transit of patients through states in a discrete-event simulation may occur continuously according to a statistical distribution function. If a patient has the possibility of making a transition from the current to multiple states then the model samples the time to each possible destination and compares them. The destination state with the shortest time is the one to which the simulated patient moves to.[4] DES models are more flexible than Markov models, but as they are usually more complex; they take longer to build and run and, more importantly, are more demanding in terms of evidence with which to populate their parameters [79]. As the existing ScHARR model had already been subject to peer review [74], and because of the high element of uncertainty involved in extrapolating outcomes beyond the diagnostic phase in this evaluation, the benefits of building a de novo model did not justify its costs.

A basic layout of the ScHARR model is given in Figure 16. Once the process of testing for sentinel lymph node metastases, surgery for tumour removal and possible ALND have been completed, patients enter a state of adjuvant therapy. This involves chemotherapy and hormonal therapy (where appropriate) for those diagnosed with metastases and hormonal therapy alone (where appropriate) in the patients diagnosed without sentinel lymph node metastases. After adjuvant therapy the patients may move into a disease free state (post adjuvant therapy state) and potentially stay there for the rest of their lifetime. During or after their adjuvant therapy, some patients may have a locoregional relapse (see Glossary). Patients in the post adjuvant therapy state may experience locoregional or metastatic relapse. Patients in the locoregional relapse state may go into remission or advance to metastatic relapse. Patients in remission may remain in that state until death, or enter a metastatic relapse state (for the purposes of the model there cannot be a further locoregional relapse). The model does not consider metastatic disease curable and therefore patients in this state may move to a death state, either from breast cancer or other causes. Patients in all states can die from other causes.

The authors of the ScHARR model claimed that there were problems in deciding whether axillary lymph node metastases identified in follow up are due to recurrence or previous misdiagnosis (a previous false negative). As this problem was beyond the scope of their project, they did not explicitly model it. This issue is similarly beyond the scope of our assessment and therefore we adopt their approach as reasonable for our purpose.

Figure 16. ScHARR model for post diagnosis of axillary metastases



The following are the ScHARR model assumptions:

- Lymphoedema is the only long-term adverse event considered and is classified as either mild/moderate or severe, based on the research data available. It affects both costs and quality of life for the rest of the patient's life. We also only model lymphoedema as no relevant papers on other long term adverse events were found. To incorporate the effect

of lymphoedema on patients we populated the respective parameters using evidence from studies of patient reported outcomes for lymphoedema.

- Adjuvant therapy occurs for a maximum 5-year period, consistent with the recommended follow up period stated in the NHS Guidelines for early breast cancer[16], but the model allows for patients to move to another state before the end of this 5 year period. In this state, patients who test positive for additional axillary metastases, confirmed on ALND, receive chemotherapy for half a year, followed by hormonal therapy for 4.5 years. Patients who test negative for sentinel lymph node metastases or only test positive for sentinel lymph node metastases and do not test positive for further axillary metastases on ALND, receive hormonal therapy for 5 years.
- After a locoregional relapse further locoregional relapse cannot occur; only metastatic relapse may. This is a simplification of clinical practice, but remained so that the model could be kept simple and because this was not the main focus of our assessment.
- 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 they may overestimate the risk of dying due to non-breast cancer causes. In our model we use England mortality statistics, as this was the population stated in the objective in our protocol. As with ScHARR, we only use the statistics for women as a very small proportion of breast cancer cases (<1%) are in men.[17]

Since we relied on the treatment phase of the ScHARR model with limited modification, we quality assessed the model and found that it met the standard quality criteria required for health economic evaluation models. [75] The ScHARR model depicts a very similar problem to that analysed herein, but does not consider intraoperative testing approaches. Although this is presumably because intraoperative tests were outside of its remit, the authors do not state this. However, this has limited impact as it only affects the diagnostic stage of their model, which we do not use. The model does not value productivity changes, which are beyond our remit too. Overall, the ScHARR model was well reported and transparent; and we therefore felt comfortable adapting it to our purposes. During the course of our review process we have worked with two of the ScHARR model authors in adapting and updating their model for our purposes, and they have become co-authors in this report (YM, KC).

5.3.3 Source of model parameter values

5.3.3.1 Diagnostic pathway

5.3.3.1.1 Test accuracy event and probabilities

In our analyses histopathology is assumed the 'gold standard' and given an accuracy of 1, as this generally is how the sensitivities and specificities of the intraoperative tests have been assessed in the literature (see Section 4.2.1.4, page 60).

Table 40. Test accuracies in the base case

Test	Sensitivity	Specificity	Source
Histopathology	1	1	Assumed (used as 'gold standard' in studies)
OSNA	84.5%	91.8%	No adjustment for TAB: Meta-analysis results Section 4.2.2, page 75
OSNA	i)91.4% ii) 100% iii) 89.8%	i) 93.3% ii) 97.2% iii) 94.5%	Adjustment for TAB i)Frere Belda[60] ii) Khaddage[59] iii)Snook[63]
Metasin	92.6%	96.3%	No adjustment for TAB Sundaresan[80]

Accuracies for Metasin have been provided by authors of unpublished papers[41,51,80] and accuracies for OSNA were provided by the patient level results of the clinical systematic review. Sufficient studies on OSNA that did not adjust for TAB were found to allow us to do a meta-analysis for the sensitivity and specificity of the test. As there were too few studies on OSNA unadjusted for TAB and only two studies on Metasin, we used the meta-analysed accuracy values for OSNA without TAB and the Sundaresan paper (which did not adjust for TAB) for our base case analysis and used the other studies[60,59,63] individually in our sensitivity analyses. In neither histopathology nor intraoperative testing have we accounted for the false negative rate associated with SLNB. This rate is normally less than 10%[81] (for instance, Cooper et al., 2011 model it as 7%[4]) and happens either as a result of metastases being in a node other than the SLNB being examined or due to the preparation procedure for histopathology. As there is no data on what proportion of the failure rate would specifically affect histopathology (or intraoperative testing), we have not modelled it. It is therefore important to note that the model may underestimate the accuracy of the intraoperative tests.

The node positive prevalence is set at 20% in the base case which is consistent with the values found in the systematic review of accuracy studies.

5.3.3.2 Model parameter values

We found no new relevant studies in the population of interest that served to update the probabilities of adverse events in the ScHARR model. For the short term adverse event probabilities, we used the original model's values.[4] We examined the original studies that ScHARR used to populate both long and short term outcomes to verify that the probability values for events in their long term model applied to our intermediate outcomes. For lymphoedema we confirmed ScHARR's values by returning to the same original sources as ScHARR and similarly splitting patients into those with mild/moderate lymphoedema and those with severe lymphoedema. In the studies patients were divided into these groups either on the basis of their responses (Crane-Okada (2008)[31]), or on the swelling measurement (over 5cm was severe in the McLaughlin study [32]). Patients were also split by the treatment to the lymph nodes; either SLNB or ALND. In all the studies ALND followed SLNB and therefore the reported probabilities included the risk level of SLNB. The costs, disutility and probabilities for lymphoedema are incorporated in the treatment phase of the model, since they are adverse events with long term implications.

Table 41. Probabilities associated with adverse events

Event	Probability	Source
Short term adverse events		
Infection		
SLNB	0.072	Cooper et al. (2011)[4]
ALND	0.221	
Seroma		
SLNB	0.02	Cooper et al. (2011)[4]
ALND	0.792	
Surgical drain		
SLNB	0.021	Cooper et al. (2011)[4]
ALND	0.142	
Long term adverse events		
SLNB	6.8%	Crane-Okada et al. (2008)[31]
Mild/moderate lymphoedema	4.5%	Adjusted using Mak et al. (2009)[82]
Severe lymphoedema	2.3%	
ALND	21.4%	Crane-Okada et al. (2008)[31] Mclaughlin et al. (2008)[32] Blanchard et al. (2003)[30]
Mild/moderate lymphoedema	14.2%	Adjusted using Mak et al. (2009)[82]
Severe lymphoedema	7.2%	Adjusted using Mak et al., (2009)[82]

5.3.3.2.1 Costs

Costs for tests were taken directly from manufacturer/academic submissions. Where only a range of costs was provided (for instance the costs for OSNA were given as £300-£400[83], we

took the midpoint of that range as our base cost (in the case of OSNA £350). The costs of Metasin provided by the sponsor of the technology were used. After the analysis had been completed the sponsor provided an indicative update of the cost (£60-£80 for kit reagents, up to 4 nodes, and reagents and consumables costs of £20 per node), which was not used to revised the analysis due the limited time to do that and because results were not likely to be significantly altered.(However with a private company having taken over the commercial development of the test, the cost-effectiveness analysis may be updated if and when a finalised list price for the C-E marked Metasin test is received.) Histopathology costs were based on data provided by NTAC, which in turn was based on evidence from a micro-costing study conducted at the Queen Alexandra Hospital (Cutress et al., 2010[3], reflated to 2010 prices), reported in technical detail by researchers at the York Health economics consortium (YHEC; Burke et al., 2011 [5]) and includes the cost of confirmation by immunohistochemistry.

For surgery costs we relied mainly on NHS reference costs. An overall average value for breast surgery was calculated using costs for lumpectomy and mastectomy (HRG codes JA07D, JA07E, JA07F, JA07G, JA07H for mastectomy; HRG codes JA09E, JA09F,JA09G, JA09H for lumpectomy, chosen on advice from ScHARR)[84]. These costs were weighted by 2/3 for lumpectomy and 1/3 for mastectomy to represent the split between lumpectomy and mastectomy cases according to the opinion from our specialist advisors. The cost of a single surgery for ALND used the reference cost for procedures on the lymphatic system (HRG code WA24Z). There were no reference costs that included extra costs for SLNB or ALND procedures undertaken during a breast surgery. We therefore used the same procedure as employed by Cooper et al., 2011 to incorporate these additional costs. Using the results reported by Pandharipande et al. [85], we calculated the ratio between the cost of a breast surgery with SLNB compared to a breast surgery alone ($7,537/$5,264) and the ratio between breast surgery with SLNB and ALNB compared to breast surgery with SLNB alone ($11,244/$7,537). The base cost of breast surgery (£1804.90 from the NHS reference costs) is then multiplied by these ratios to give costs for breast surgery with SLNB and breast surgery with SLNB and ALND.

Costs of short term adverse events were taken from the same source as ScHARR (Jeruss et al. 2006) as no other papers were identified. No papers were identified for updating the annual lymphoedema costs and as these were originally provided for ScHARR by the Sheffield Lymphoedema Service, we considered this an appropriate source and updated the costs to 2010.

Table 42. Unit costs of diagnostic and primary surgery services

Test or surgery	Costs	Source
Histopathology	£472	Cutress et al., (2010)[3]
OSNA	£350	Manufacturer Information Submitted to NICE[76]
Metasin	£74	Manufacturer Personal Communication 19/11/2012
Breast surgery with SLNB	£2584	NHS Reference costs (2010-2011)[84] updated using same method as ScHARR using Pandharipande et al. (2008)[85]
With intraoperative test	£12	Additional time cost calculated: 3 minutes for 54% patients, Ng et al. (2010)[86] Extra cost updated to 2010, Burke and Patton (2010)[5]
With ALND	£3855	Breast surgery adjusted using Cooper et al, (2011) technique[4]
Secondary operation for ALND	£3569	NHS Reference costs (2010-2011)[84]
Additional hospital stay for surgery with ALND	£106	Burke and Patton (2010), updated to 2010 costs[5]
Additional hospital stay for second surgery for ALND	£512	Burke and Patton (2010), updated to 2010 costs[5]
Short term adverse event	£333	Jeruss et al., (2006), updated to 2010 prices[87]
Mild lymphoedema	£71	ScHARR updated to 2010 prices[4]
Moderate/severe lymphoedema	£1269	ScHARR updated to 2010 prices[4]

Most studies only report the time it takes to run an intraoperative test and not the impact this has on the length of surgery. The exception is a recent abstract by Ng et al. [86] which stated that in 54% of their cases breast surgery was complete before the results for OSNA were received. The median time over this surgery was 3 minutes. Applying this to an entire cohort gives an average waiting time of 1.62 minutes. To cost this, we applied this time delay to costs from the micro-costing study from YHEC [5] which reports unit costs for various surgical procedures. These costs were updated to 2010 costs using the cost convertor developed by Shemilt 2010[88]. Using the costs for a surgeon (£159.74 per hour), theatre staff (£2.60 per minute) and anaesthetist staff (£124.24 per hour) we calculated that the extra cost of waiting for intraoperative results was £11.88. We used this value for all intraoperative tests, even though it was only calculated using OSNA data, as there was no equivalent data for Metasin and none of the reported Metasin processing times were contrary to any of the assumptions made.

A consequence of an ALND procedure, either in the first breast surgery or a separate surgery, is extra days spent in hospital by the patient. This extended stay was calculated using the results of the YHEC micro-costing study[5], subtracting the length of stay for a standard surgery (2.1 days) from those that include ALND (2.7 days) or the combined stay of first and second surgeries (5 days) for those that have ALND as a separate surgery. These were multiplied by the updated 2010 ward costs (£176.44 per day[5]) to give the cost of this extra hospital stay. The cost of short term adverse events was adjusted to 2010 prices, since, as reported previously, no more recent reports were found.[4,87]

Univariate sensitivity analyses of unit costs of tests of primary surgery were conducted using the value ranges on cost drivers. The recently published paper by YHEC provides unit costs and resource use and is widely known. We therefore examined costs based on their information. To be sure it would directly compare with our approach using NHS reference costs, we examined the length of stay per operation both in our reference case and for YHEC. In our base case, the average length of stay for a single breast surgery was roughly 2 days compared with YHEC's 2.1 (Table 42, page 126). The average length of stay for surgery on the lymphatic system (our ALND alone procedure) was roughly 2.65 days, which is not dissimilar from the YHEC average of 2.9.

We therefore calculated surgery costs based on the YHEC information, with unit costs (Table 43) updated to 2010 costs. As in our base analysis, we assumed that the ward stay would be the same for all patients without axillary clearance, regardless of how the SLNB was analysed. YHEC reported intraoperative analysis times that were shorter than the surgery times and so no additional time was factored in for this case. On the advice of our experts, the time taken by the MDT is the same for all diagnostic testing procedures and therefore the second operation for ALND does not incur an additional MDT cost as they do in the YHEC report.

Table 43. Resource use

Resources	Surgery with SLNB	Second surgery	Surgery with axillary clearance
Procedure time (minutes)	54	60	
Procedure time - biopsy and analysis (minutes)			53
Procedure time - operation (minutes)			40
Anaesthetic Prep/Recovery (minutes)	20	20	20
Theatre turnaround time (minutes)	20	20	20
Length of stay (in days)	2.1	2.9	2.7
Hospital sterilisation & disinfection unit (trays)	1	1	1
Physiotherapist (minutes)		10	10
Multi-disciplinary Team Appointment (minutes)	10		10
Multi-disciplinary Team Meeting (minutes)	3		3

Source: Burke M, Patton T. The Cost Impact of Implementing Intra-Operative Testing for the Diagnosis of Patients with Metastatic Breast Cancer in England. York Health Economics Consortium NHS Technology Adoption Centre (2010)

Total surgery costs, given in Table 44, are calculated by multiplying the resource use by the relevant unit cost. For a basic surgery this includes the theatre times for each member of the surgical team, theatre stock cost, ward costs for length of stay and the cost of the MDT meetings. As additional ward costs for surgery with ALND and ALND alone are already calculated separately as part of the base case analysis, these are not included in those particular operation costs. However, these two surgeries do incur the cost of a physiotherapist, due to the ALND.

Table 44. Unit costs updated from 2008 to 2010

Resources	Unit Costs (£)
Direct costs	
Surgeon (per hour)	£159.74
Ward costs (per day)	£176.44
Indirect costs	
Theatre staff (per minute)	£2.60
Anaesthetist staff (per hour)	£124.24
Theatre stock consumables (by case)	£64.73
Anaesthetic administration & junior doctors (by case)	£68.91
Hospital sterilisation & disinfection unit (by case)	£17.75
Anaesthetic gases (by case)	£25.06
Physiotherapist (per hour)	£25.06
Multidisciplinary team appointment (per hour)	£81.43
Multidisciplinary team meeting (per hour)	£644.16

Source: Burke M, Patton T. The Cost Impact of Implementing Intra-Operative Testing for the Diagnosis of Patients with Metastatic Breast Cancer in England. York Health Economics Consortium NHS Technology Adoption Centre (2010)

Table 45. Surgery costs calculated using Burke and Patton (2010)[5]

Surgery	Costs
Cost of operation with SLNB (with ward costs, MDT)	£1,282
Operation with SLNB and ALND	£1,580
ALND only	£1,284

Source: Burke M, Patton T. The Cost Impact of Implementing Intra-Operative Testing for the Diagnosis of Patients with Metastatic Breast Cancer in England. York Health Economics Consortium NHS Technology Adoption Centre (2010)

As Burke and Patton did not present values for the uncertainty, for sensitivity analysis, we used a separate report by Burke and Setters[89] for alternative values to identify resource use that might affect the costs. The ones that directly applied were a shortening in the surgery with ALND by 10 minutes; and changes in the meeting length of the MDT (range 10-20 minutes). By varying the surgery costs by +/- 10% we were able to incorporate the variations these changes caused.

5.3.3.2.2 Utilities

Patients who undergo testing by histopathology have to wait around two weeks for these results.[76] We surmise that this would incur some level of disutility to patients due to the associated anxiety of waiting, and imputed it using the health state valuation equation provided by Dolan.[90] Dolan estimated the health state values of the EQ-5D classification system, which measures health related quality of life in terms of five dimensions: mobility; self-care; ability to perform usual activities; pain/discomfort; and anxiety/depression. In turn, each of these dimensions is measured on three levels: no problem, moderate problems, and severe problems.

The utility weights for these levels were estimated from responses to relative valuation questions in a survey of the UK general public. For our purposes, only the utility decrement due to anxiety/depression was relevant and adopted the one specific to the severe level. The equation for this takes the form

$$Y = \alpha + 2\text{AD} + \text{A2} + \text{N3}$$

Where *α* is the constant (0.081) applied to any level of disutility, AD is the constant (0.071) applied for any level of disutility associated with anxiety or depression, A2 is the constant (0.094) applied to severe levels of anxiety or depression and N3 is the constant (0.269) applied when any of the five dimensions of the EQ-5D are recorded as severe. Our patients were assumed to already have utility less than one (meaning we did not need to apply the *α* value); to be moving from a state of no anxiety/depression to severe anxiety/depression; and that this anxiety/depression would be the only dimension of the EQ-5D they had that was severe. This gave us a decrement of (-0.236-0.269= -0.505) Once we had used the formula to calculate a value for *Y*, we adjusted it to apply for only two weeks. This gave us a reduction in undiscounted QALYs of 0.019 for the two weeks spent waiting for histopathology results.

We also assumed that patients who undergo a second operation would have a disutility. To account for this, we used utilities from the literature.[91] None were specific, but there was a value of 0.62 for the 2 months following breast cancer surgery. We subtracted this from the highest utility in the adjuvant therapy states (0.82) and adjusted for the 2 months to give a disutility of 0.03.

5.3.3.3 Treatment phase

5.3.3.3.1 Health state transition probabilities

No studies were found to allow us to update all the health state transition probability values appropriately and therefore we returned to the papers used by the ScHARR model and kept all the parameters and standard errors they used.[4] The papers used included previous cost-effective models [92, 85, 93] and a prospective cohort study.[94] The cost effectiveness models used data from the Early Breast Cancer Trialists' Collaborative Group[92], Adjuvant! Online[85]; and a retrospective US study by Chang et al. 2003[93].

Table 46. Health state transitions probabilities

Transition	Probability	Standard error	Distribution	Source
Annual probability of locoregional recurrence (no sentinel lymph node metastases/no additional axillary metastases)	0.03	0.0017	Beta	Cooper et al. (2011)[4,92]
Annual probability of locoregional recurrence (sentinel lymph node metastases and additional axillary metastases)	0.09	0.0052	Beta	Cooper et al., (2011)[4,92]
Annual probability of locoregional recurrence (sentinel lymph node metastases, tested negative)	0.14	0.0082	Beta	Cooper et al., (2011)[4,92]
Annual probability of metastatic recurrence(True negative, false positive)	0.0023	0.00014	Beta	Cooper et al., (2011)[4,85]
Annual probability of metastatic recurrence(True positive)	0.0052	0.00030	Beta	Cooper et al., (2011)[4.85]
Annual probability of metastatic recurrence (false negative)	0.0094	0.00054	Beta	Cooper et al., (2011)[4,85]
Annual probability of metastatic relapse from locoregional recurrence	0.18	0.010	Beta	Cooper et al., (2011)[4,94]
Annual probability of death from locoregional recurrence	0.3	0.017	Beta	Cooper et al., (2011)[4.92]
Annual probability of metastatic relapse from remission	0.13	0.0075	Beta	Cooper et al, (2011)[4,94]
Annual probability of death from metastatic relapse	0.37	0.021	Beta	Cooper et al., (2011)[4.93]

When patients enter the model, their maximum life expectancy is set using the Life Tables. It is possible for a patient to die before this time in the model, depending on their transitions through the states. Each time a patient enters a state, the time to each of the next plausible states is sampled from an exponential distribution using the annual transition probability. The state with the shortest time delay is the state that the patient then moves to, following that delay and the process is repeated. This analysis simulates the individual experience of a cohort of 5000 patients, to calculate patient level results (first order Monte Carlo) and additionally randomly samples from the independent model parameter distributions (second order Monte Carlo simulation). There is therefore a probabilistic sensitivity analysis built into the model.

5.3.3.3.2 Health state costs

Relevant studies and NHS costs were used to calculate the health state costs. The cost of adjuvant therapy was calculated using methods reported elsewhere.[4] This assumes that patients with test negative sentinel lymph node status (true or false) and patients who test positive for sentinel lymph node metastases, but have no further metastases found on ALND will receive

hormonal therapy for 5 years where appropriate. Patients who were found to have sentinel lymph node metastases and additional axillary metastases receive chemotherapy for half a year at regimens and prices of treatments commonly used in the UK.[4] They also then receive hormonal therapy for 4.5 years. To calculate an average cost per patient, assumptions are made about the hormonal therapy provided. It is therefore assumed that 81% of patients are oestrogen receptor positive and will therefore respond to hormonal therapy. [93] As in the ScHARR model, we assume that of these 90% will receive some form of aromatase inhibitor and the other 10% tamoxifen.[4] Regardless of hormonal therapy status, each patient will also have a yearly outpatient follow up appointment and a mammogram. The combined average annual cost per patient for hormonal adjuvant therapy and follow up is then calculated to be £1086.75. The cost of the post adjuvant therapy state is assumed to be £0 and the costs of locoregional and metastatic relapse and remission and death are taken from Karnon et al., 2008,[95] which was a UK cost-utility analysis that estimated these costs based on a cost study of 199 women in Edinburgh.

All costs are updated to 2010 costs as required and as we had no data on their uncertainty, values were sample from a uniform distribution of +/- 10% around the mean estimate.

Table 47. Health state costs

State	Cost	Parameter	Source
Cost adjuvant therapy (TN,FP, FN)	£1087	Annual cost	NHS reference cost outpatient follow up 10/11 [84] + mammogram NHS reference cost 02/03 [96] updated + Ward et al., (2007) hormone therapy updated [93]
Cost adjuvant therapy TP	£9447first year	Cost for 6 months	Cooper et al., (2011) cost updated[4]
	£1087 After 1st year	Annual cost after first 6 months	NHS reference cost outpatient follow up 10/11 [84] + mammogram NHS reference cost 02/03 [96] updated + Ward et al., 2007 hormone therapy updated [93]
Cost post adjuvant therapy	0	Annual cost	Cooper et al., (2011) assumption[4]
Cost of locoregional recurrence	£13745	Annual cost	Karnon et al., (2008)[95]; 2005 prices reflated to 2010 using the Hospital and Community Health Services Index, Curtis et al., (2011) [97]
Cost of remission	£108	Annual cost	Karnon et al., (2008)[95]; 2005 prices reflated to 2010 using the Hospital and Community Health Services Index, Curtis et al., (2011) [97]
Cost of metastatic relapse	£10443	Annual cost	Karnon et al., (2008)[95]; 2005 prices reflated to 2010 using the Hospital and Community Health Services Index, Curtis et al., (2011) [97]
Cost of death	£5713	Cost of event	Karnon et al., (2008)[95]; 2005 prices reflated to 2010 using the Hospital and Community Health Services Index, Curtis et al., (2011) [97]

5.3.3.3.3 Health state utilities

All health state utilities were derived from the Tengs and Wallace study[91], using the same parameters and standard errors as ScHARR. The Tengs and Wallace study was a comprehensive study and no new data were available for these particular variables. Each parameter was given a Beta distribution.

We also accounted for the general population change in utility that occurs with age, using the same formula as in the original ScHARR analysis, which adjusted a formula previously reported [98] . The utilities of the health states are multiplied by the following formula at the end of each year in the model:

$$U_{GP} = 0.9584588 - 0.0001728 \times age - 0.000034 \times age^2$$

In our model, the starting age is 56, the mode across studies in our systematic review (see Section 4.2.1.5, page 64). Although the formula was not specifically estimated for women, analysis of Canadian longitudinal population survey data suggests that the evolution of utility with age does not significantly differ by gender.[99]

Other utility values including lifetime utility decrements for lymphoedema into the treatment phase of 9.9% for mild/moderate lymphoedema and 12.3% for severe lymphoedema, were obtained from the same sources[82] used by the ScHARR model report.[4] These were estimated based on quality of life data assessed by the FACT-B + 4 (Functional Assessment of Cancer Therapy for Breast Cancer, adding a four-item arm subscale) for breast cancer patients who suffer from different degrees of lymphoedema and therefore do not necessarily represent the true utility decrements for lymphoedema. However, when ScHARR performed a sensitivity analysis on this utility it was found to not significantly affect the results[4].

Table 48. Health state utilities

State	Utility	Standard error	Distribution	Source
Adjuvant therapy TN FP FN	0.82	0.18	Beta	Tengs and Wallace (2000)[91] (adjuvant therapy)
Adjuvant therapy TP	0.74	0.26	Beta	Tengs and Wallace (2000)[91] (chemotherapy)
Post therapy	0.94	0.11	Beta	Tengs and Wallace (2000)[91]
Locoregional recurrence	0.7	0.19	Beta	Tengs and Wallace (2000)[91]
Remission	0.85	0.19	Beta	Tengs and Wallace (2000)[91] (after first recurrence)
Metastatic relapse	0.4	0.19	Beta	Tengs and Wallace (2000)[91]
Death	0			By definition

5.3.4 Implementation of sensitivity analyses

In the short term analyses we focused mainly on the cost drivers and accuracy as part of sensitivity analyses. We looked at the impact of altering the cost of the tests and surgeries, particularly the cost of a second surgery, by setting ranges appropriate to the evidence we had obtained in the literature. Where there was no additional information on the variation in costs, we looked at the cost ranges of +/- 10%. As well as looking at the changes in accuracy between the values produced from studies that did not account for TAB and those that did, we performed a threshold analysis on sensitivity, where the specificity is held and the sensitivity increased in steps of 5% from a minimum of 70% (the lower end of the CI produced in the meta-analysis). This process was repeated for specificity, holding sensitivity constant. The impact of accuracy and costs was assessed in both the short term and long term results.

For the purposes of assessing the effect of uncertainty in long term outcomes on the results, we ran supplementary univariate sensitivity analysis on the costs of adjuvant therapy, adjusting the mean +/- 10% in two different simulations. Adjuvant therapy is the cost parameter that varies the most in direct relation to the outcome of the diagnostic tests and therefore most likely to impact on model results.

We also considered the impact of the disutility associated with anxiety waiting for histopathology results and a second operation for ALND. For the disutility associated with anxiety, we examined the scenarios where the disutility was zero, some disutility (0.006 using Dolan's moderate anxiety value of 0.071, adjusted for 2 weeks), and extreme disutility (0.023), which used Dolan's formula, but assumed there was nothing else wrong with the patient. For the disutility of a second operation, we looked at the scenarios where the disutility was adjusted by +/- 20% and where the disutility was zero.

5.3.4.1.1 Discount rate

Costs and utilities are discounted at a rate of 3.5% in accordance with NICE guidance.

5.3.5 Measure used to synthesise cost and benefits

5.3.5.1 Intermediate results

The following measures were used to synthesis costs and health benefits in the diagnostic phase:

- Incremental cost per patient
- Incremental cost per case correctly diagnosed

- Incremental cost per additional positive case detected
- Incremental cost per additional negative case detected

Although we had set out present, in addition to these measures, estimates of cost per second operation avoided, this measure turned out to be of no additional informational value and the corresponding results are therefore not presented.

5.3.5.2 Long term analysis

The following measures were used to synthesis costs and health benefits over the lifetime of patients:

- Incremental costs
- Incremental QALYs
- Incremental cost per QALY gained

Since the status quo, histopathology, was assumed the gold standard in the clinical effectiveness studies, it tended to be the option with the highest benefit (only one study reported evidence that diagnostic accuracy with OSNA was sufficiently close to histopathology to generate total QALYs that were larger than those with histopathology). In order to avoid the complication of interpretation that arises with the ICER of a new technology, i.e. intraoperative testing, that has lower costs than the status quo, results are presented by ranking the three options, namely intraoperative diagnosis with half node, diagnosis with full node and histopathology in order of increased total QALYs.

5.3.6 Results

5.3.7 Base Case

5.3.7.1 Diagnostic phase

The studies used to inform the economic analyses are presented in Table 49 (see Section 4.2.1.6, page 68 for discussion on the strengths and weakness of selected studies). For the purpose of the economic analyses they were grouped by test option and whether they controlled for TAB.

Table 49. Short term costs-accuracy analysis comparing histopathology to full and half node intraoperative analysis

	Mean estimates			Incremental results	
Measure	**Histopathology**	**OSNA[1]**		**Difference OSNA half vs. full node**	**Difference Histopathology vs. OSNA half node**
		half node	**full node**		
Accuracy[a]	1.0000	0.9344	0.9034	0.0310	0.0656
Sensitivity*Prevalence[1]	0.2000	0.2000	0.1690	0.0310	0
Specificity*(1-Prevalence[1])	0.8000	0.7344	0.7344	0	0.0656
	NHS reference costs of ALND				
Costs per patient	£3,987	£3,897	£3,397	£500	£90
Incremental cost per additional patient correctly diagnosed				OSNA half node extended dominated	£6,108*
Incremental cost per additional node-positive case detected				£16,123	Histopathology dominated
Incremental cost per additional node-negative case detected				OSNA half node dominated	£8994*
	Analysis using costs based on YHEC model				
Costs per patient	£2,228	£2,284	£1,855	£429	-£56
Incremental cost per additional patient correctly diagnosed				OSNA half node extended dominated	£3,866*
Incremental cost per additional node-positive case detected				OSNA half node extended dominated	£12,046*
Incremental cost per additional node-negative case detected				OSNA half node dominated	£5,693*

1 Node positive prevalence fixed at 20%. [a] Accuracy refers to the cost per case correctly identified * Comparison is Histopathology relative to full node OSNA due to the half node OSNA option being dominated (i.e. having higher costs and lower diagnostic yield than another diagnostic option) or extended dominated (i.e. the incremental cost per additional diagnostic yield being higher than that of another, more accurate option). In the case of node-negative case detection OSNA half node was dominated by full node, as they had the same detection rate, but OSNA full node was less expensive. Similarly, for node-positive case detection histopathology and OSNA half node had the same detection rate; so their dominance was dependent purely on which was less expensive. For NHS Reference costs, this was OSNA half node, so histopathology was dominated; for YHEC costs this was histopathology, so OSNA half node was dominated. OSNA half node is stended dominated in NHS Reference costs due to having a higher incremental cost per additional diagnostic yield than histopathology compared to full node OSNA.

Ranking the diagnostic strategies by diagnostic accuracy and comparing the incremental costs and yields with increasingly accurate diagnostic strategies, allowed the calculation of incremental costs per additional patient correctly diagnosed ratios and the identification of dominated options, i.e. those that had higher costs and lower yield than the alternative. Table 49 summarises the cost-accuracy analysis for accuracy estimates unadjusted for TAB. Under the NHS Reference Costs costing system, OSNA half node dominated histopathology for each additional node positive case detected; but histopathology dominated OSNA half node for each additional node negative case detected and had an ICER of £8,994 when compared to OSNA full node, which may be seen as cost effective depending upon threshold. In the case where YHEC cost values are used, histopathology dominated OSNA half node analysis for all measures of accuracy and consistently had an ICER under £13,000 when compared to OSNA full node.

The discrepancy between the YHEC cost results and the NHS Reference cost results occurred for a couple of reasons. Firstly, YHEC costs for surgery were significantly lower than the NHS Reference costs, giving greater wait to other costs, such as adverse event costs; but there was also a difference in ratios between the types of surgery (breast surgery with SLNB alone, breast surgery with SLNB and ALND, breast surgery for ALND alone). In particular, the cost ratio between histopathological and intraoperative test positive patients in YHEC was significantly lower than when using the NHS reference costs. This meant the cost impact of a second surgery for patients diagnosed by histopathology was greatly reduced; and costs such as the additional cost of histopathology in OSNA test negative patients and short term adverse events in the half node strategy had a greater influence. This made the half node strategy more expensive than histopathology alone. This gave rise to the seemingly inconsistent results between the NHS Reference costs where half node OSNA was not dominated by histopathology and the YHEC costs where it was.

This analysis was repeated for the three individual OSNA studies that adjusted for TAB and the incremental results are reported in Table 50. Similar results were found in the case of the Frere-Belda and Snook values, but not in the best case scenario provided by Khaddage where sensitivity was 100% and specificity was 97.2%. Here full node analysis had the same accuracy as half node, so histopathology was compared directly with full node. OSNA full node dominated histopathology for node positive diagnosis, and any additional node negative patients detected by histopathology had an ICER £27,300, under the NHS Reference Costs. This reduced to £17,100 under YHEC costing.

Table 50: Short term cost-accuracy results comparing histopathology to OSNA with sensitivity and specificity adjusted for TAB

	Incremental results Frere-Belda		Incremental results Snook		6 Incremental Results Khaddage	
Measure	**Difference OSNA half vs. full node**	**Difference Histopath-ology vs. OSNA half node**	**Difference OSNA half vs. full node**	**Difference Histopath-ology vs. OSNA half node**	**Difference OSNA half vs. full node**	**Difference Histopath-ology vs. OSNA half node**
Accuracy[a]	0.0172	0.0536	0.0204	0.0440	0	0.0224
Sensitivity*Prevalence[1]	0.0172	0	0.0204	0	0	0
Specificity*(1-Prevalence[1])	0	0.0536	0	0.0440	0	0.0224
	NHS reference costs of ALND					
Costs per patient	£437	£150	£458	£152	£367	£244
Incremental cost per additional patient correctly diagnosed	OSNA half node extended dominated	£8,289*	OSNA half node extended dominated	£9,463*	OSNA half node dominated	£27,300*
Incremental cost per additional node-positive case detected	£25,426	Histopathology dominated	£22,435	Histopathology dominated	OSNA half node dominated	Histopathology dominated *
Incremental cost per additional node-negative case detected	OSNA half node dominated	£10,949*	OSNA half node dominated	£13,850*	OSNA half node dominated	£27,300*
	Analysis using costs based on YHEC model					
Costs per patient	£398	-£26	£411	-£29	£367	£16
Incremental cost per additional patient correctly diagnosed	OSNA half node extended dominated	£5,254*	OSNA half node extended dominated	£5933*	OSNA half node dominated	£17,100*
Incremental cost per additional node-positive case detected	OSNA half node extended dominated	£21,629*	OSNA half node extended dominated	£18,731*	OSNA half node dominated	Histopathology dominated *
Incremental cost per additional node-negative case detected	OSNA half node dominated	£6,941*	OSNA half node dominated	£8,684*	OSNA half node dominated	£17,100*

1 Node positive prevalence fixed at 20%. [a] Accuracy refers to the cost per case correctly identified * Comparison is Histopathology relative to full node OSNA due to the half node OSNA option being dominated (i.e. having higher costs and lower diagnostic yield than another diagnostic option) or extended dominated (i.e. the incremental cost per additional diagnostic yield being higher than that of another, more accurate option). In the case of node-negative case detection OSNA half node was dominated by full node, as they had the same detection rate, but OSNA full node was less expensive. Similarly, for node-positive case detection histopathology and OSNA half node had the same detection rate; so their dominance was dependent purely on which was less expensive. For NHS Reference costs, this was OSNA half node, so histopathology was dominated; for YHEC costs this was histopathology, so OSNA half node was dominated. OSNA half node is extended dominated in NHS Reference costs due to having a higher incremental cost per additional diagnostic yield than histopathology compared to full node OSNA

Part of the cost effective assessment of the short term outcomes examined the disutility of second operations and waiting for histopathology results. The diagnostic cost and utility per

patient are reported in Table 51. For strategies that do not involve histopathology, this utility was 1 as there was neither the need to wait for test results, nor a second surgery directly resulting from the intraoperative test. Using the NHS Reference costing scenario, OSNA full node dominated half node OSNA and histopathology as its increased QALY gains were less costly. Under the YHEC costing strategy, half node OSNA provided a small QALY gain over histopathology, with an ICER of £4,832 per QALY gained, but full node OSNA continued to dominate half node analysis and histopathology. It is estimated that 4.1% of the 76.5% of half-node patients who have to wait for histopathology result would end up with a positive diagnosis as opposed to the 20% expected positivity rate with patients under histopathology.

According to these short term costs and benefits, full node OSNA is the most cost effective test for short term utility. Using the meta-analysis accuracy values for OSNA (sensitivity 84.5%, specificity 91.8%), histopathology appears to be cost effective in terms of short term accuracy, though this begins to look questionable when the effect of TAB is accounted for and OSNA has a higher sensitivity and specificity. In all scenarios OSNA is less costly than histopathology.

Table 51. Short term disutility

	Mean estimates			**Incremental results**	
Measure	**Histopathology**	**OSNA**		**Difference OSNA half node vs. histopathology**	**Difference OSNA full node vs. OSNA half node**
		half node	**full node**		
	NHS reference costs of ALND				
Costs per patient	£3,987	£3,897	£3,397	-£90	-£500
Utility	0.9739	0.9854	1	0.0115	0.0146
Incremental cost per QALY gained				OSNA half node dominated	Histopathology dominated*
	Analysis using costs based on YHEC model				
Costs per patient	£2,228	£2,284	£1,855	£56	-£429
Incremental cost per QALY gained				OSNA half node extended dominated	Histopathology dominated*

*Comparison is Histopathology relative to full node OSNA due to the half node OSNA option being dominated (i.e. having higher costs and lower diagnostic yield than another diagnostic option) or extended dominated (i.e. the incremental cost per additional diagnostic yield being higher than that of another, more accurate option)

6.1.1.1 Long term analysis

Thus far, the analysis has only considered short term costs and consequences of diagnostic options, effectively assuming no utility benefits accrue from increased diagnostic accuracy. In the long term scenario we examined costs and QALYs, which are presented in Table 52 and account for all costs and the benefits of accurate diagnosis through improved patient management net of the disutility of anxiety due to waiting for postoperative diagnosis and undergoing a second operation. The strategies were then ordered by increasing numbers of QALYs, with OSNA full node having the least QALYs (9.222) and histopathology having the most (9.321). As this demonstrates, the QALY difference was less than 0.10. This QALY difference occurs as a direct consequence of the accuracies of the test, as higher accuracies lead to more correct diagnoses. As the purpose of diagnoses is to inform management strategies, an increased number of correct diagnoses lead to an increase in the correct management and therefore patients gain the most QALYs possible. The QALY gain from the higher accuracy of histopathology is small, but may be seen as cost effective: under the NHS Reference costing And ignoring OSNA half node, as this is a strategy is extended dominated and unlikely to continue once OSNA has been validated, histopathology had an ICER of £4,324 per QALY gained, compared to full node OSNA. Under YHEC costing, histopathology dominated OSNA half node, i.e. it had higher costs and fewer QALYs. In this scenario the ICER comparing histopathology to OSNA full node was £2,150 per QALY gained.

Using the TAB adjusted values (presented in Table 53) half node OSNA remained extended dominated using the Frere Belda values. In this case, histopathology compared to OSNA full node had ICERs of £9,493 per QALY gained using NHS Reference costs and £5,215 using YHEC costs. Using the values for Snook, OSNA half node was extended dominated using YHEC costs (and histopathology had an ICER compared to full node OSNA of less than £4,850 per QALY gained), but under NHS Reference costs, OSNA half node was no longer extended dominated and had an ICER compared to OSNA full node of £8,063 per QALY gained. Comparing histopathology to OSNA half node, the ICER became £14,967 per QALY gained. More importantly, Khaddage suggested a much better result for OSNA, with full node OSNA dominating half node and histopathology (full node OSNA had a higher QALY gain and lower costs) . This occurred due to the influence of the short term disutility for anxiety and second operation applied (where applicable) to patients undergoing histopathology testing in the histopathology and OSNA half node arms. Without this disutility the results using the Khaddage values for sensitivity and specificity altered slightly so that, although in the long term OSNA half node was still dominated by OSNA full node, this was now because they had the same QALY gain, but the full node was less expensive. The ICER between histopathology and OSNA full

node was £68,432 per QALY gained using NHS Reference Costs and £41,619 per QALY gained using the YHEC costs, suggesting histopathology was not that cost effective compared to OSNA full node.

Table 52. Long term outcomes comparing histopathology to intraoperative analysis

Measure	Mean estimates			Incremental results	
	Histopathology	OSNA half node	OSNA full node		
				Difference OSNA half node vs. OSNA full node	Difference Histopathology vs. OSNA half node
	NHS reference costs of ALND				
Cost per patient (discounted)	£20,530	£20,523	£20,099	£424	£7
QALYs (discounted)	9.321	9.307	9.222	0.085	0.015
Incremental cost per QALY gained				OSNA half node extended dominated	£4,324*
	Analysis using costs based on YHEC model				
Costs per patient (discounted)	£18,771	£18,910	£18,556	£353	-£139
Incremental cost per QALY gained				OSNA half node extended dominated	£2,150*

* Comparison is Histopathology relative to full node OSNA due to the half node OSNA option being dominated or extended dominated. Costs and QALYs discounted at a rate of 3.5%

Table 53. Long term incremental outcomes comparing histopathology to intraoperative analysis (TAB adjusted)

Measure	Incremental results					
	Frere Belda[60]		Snook[63]		Khaddage[59]	
	Difference OSNA half node vs. OSNA full node	Difference Histopathology vs. OSNA half node	Difference OSNA half node vs. OSNA full node	Difference Histopathology vs. OSNA half node	Difference OSNA half node vs. Histopathology*[2]	Difference OSNA full node vs. OSNA half node*[2]
		NHS reference costs of ALND				
Cost per patient (discounted)	£395	£82	£408	£96	-£216	-£367
QALYs (discounted)	0.041	0.010	0.051	0.006	0.0025	0.0151
Incremental cost per QALY gained	OSNA half node extended dominated	£9,493*	£8,063	£14,967	Histopathology dominated	OSNA half node dominated
		Analysis using costs based on YHEC model				
Costs per patient (discounted)	£356	-£94	£361	-£85	-£367	-£13
Incremental cost per QALY gained	OSNA half node extended dominated	£5,215*	OSNA half node extended dominated	£4,850*	Histopathology dominated	OSNA half node extended dominated

* Comparison is Histopathology relative to full node OSNA due to the half node OSNA option being dominated or extended dominated .*[2]OSNA strategies have a greater number of QALYs, therefore the order of comparison is switched. In both costing strategies, OSNA full node dominates OSNA half node i.e. has lower costs and greater benefits.

6.1.2 Sensitivity Analysis

For the purposes of sensitivity analysis we chose to only report the findings of the NHS Reference costing strategy.

6.1.2.1 Accuracy

As the TAB results already indicated, test accuracy has a direct impact on the cost-effectiveness of the tests. A threshold analysis was therefore conducted to investigate sensitivity and specificity separately. In the case of threshold analysis for sensitivity, specificity was held constant and sensitivity increased by steps of 5% over a range of 70-100%. The opposite was then performed for specificity. This sensitivity analysis was conducted on the full node OSNA results, where the change in accuracy would be most notable (the overall sensitivity for OSNA half node was fixed at 100%). Short term utility results are not reported as the utility for OSNA was not affected by the accuracy of the test.

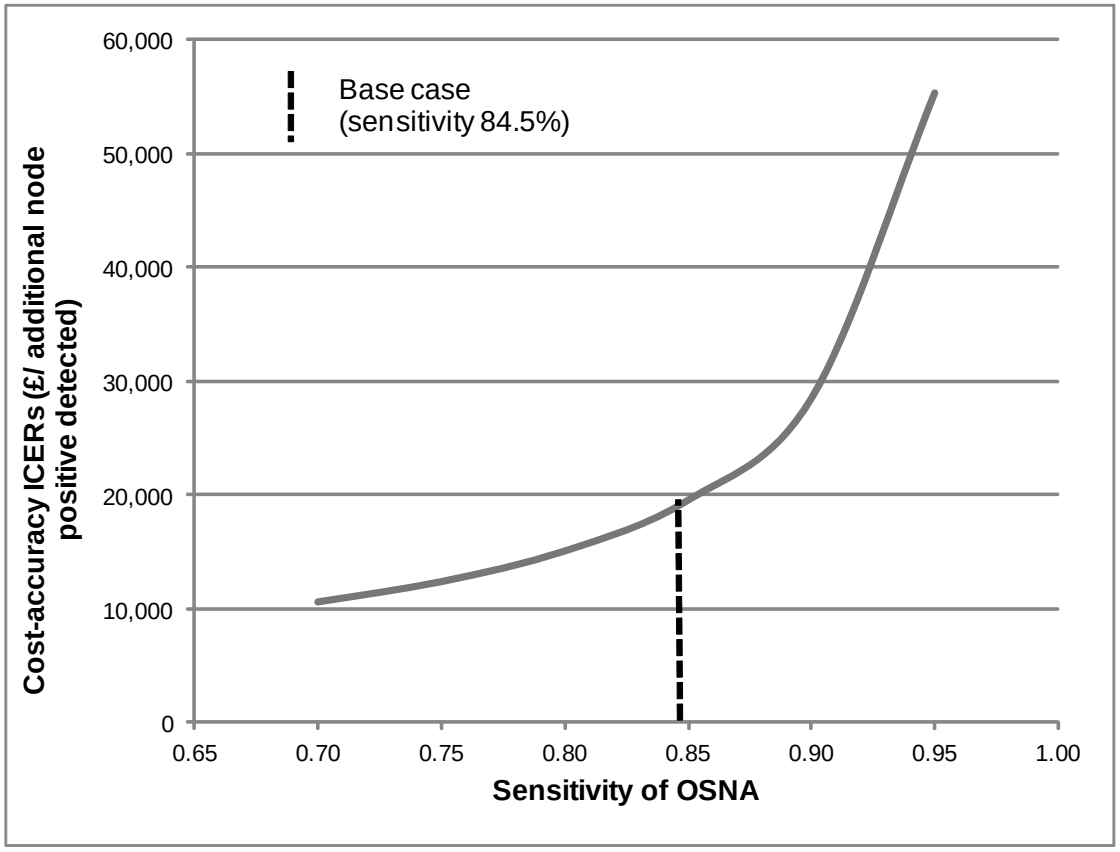
The results of the threshold analysis for sensitivity are presented in Table 52 and Table 53. Table 54 demonstrates that when specificity is held at 91.8%, the proportion of node positive cases detected increased by 1% each time the sensitivity increased by 5%. The short term costs of OSNA also increased by roughly £18 (therefore decreasing the difference in costs between OSNA and histopathology) and the combination of both increased the cost-accuracy ICERs between histopathology and OSNA. The ICER for cost per case correctly identified increased from £6,108 per additional case correctly identified by histopathology when OSNA had 70% sensitivity, to £8,162 per additional case correctly identified by histopathology when OSNA had 100% sensitivity. The ICERs for cost per node positive case detected had consistently larger ICERs, from £10,685 per additional node positive case detected for OSNA sensitivity 70% to OSNA dominating histopathology when OSNA sensitivity was 100%. However, in this case, the ICER increased much faster as the sensitivity of OSNA neared 100%, as Figure 17 shows.

Table 54: Accuracy results for threshold analysis for sensitivity

	Increase in accuracy							
Measure	**Histopathology vs.OSNA[1] with sensitivity**							
	Base Case: 84.5%	**70%**	**75%**	**80%**	**85%**	**90%**	**95%**	**100%**
Accuracy[a]	0.0966	0.1256	0.1156	0.1056	0.0956	0.0856	0.0756	0.0656
Sensitivity*Prevalence [1]	0.031	0.06	0.05	0.04	0.03	0.02	0.01	0
		NHS reference costs of ALND						
Costs per patient	£590	£641	£623	£606	£588	£571	£553	£535
Incremental cost per additional patient correctly diagnosed	£6108	£5,104	£5,394	£5,737	£6,153	£6,666	£7,315	£8,162
Incremental cost per additional node-positive case detected	£19,033	£10,685	£12,470	£15,147	£19,609	£28,532	£55,303	Histopath-ology dominated

[1] Node positive prevalence fixed at 20%. [a]Accuracy refers to the cost per case correctly identified. Histopathology was dominated when it had the same diagnostic yield as OSNA, but was more expensive.

Figure 17. Comparison of cost-accuracy results for threshold analysis of OSNA sensitivity.



In this figure, the ICER of Histopathology relative to full node OSNA increases with OSNA sensitivity.

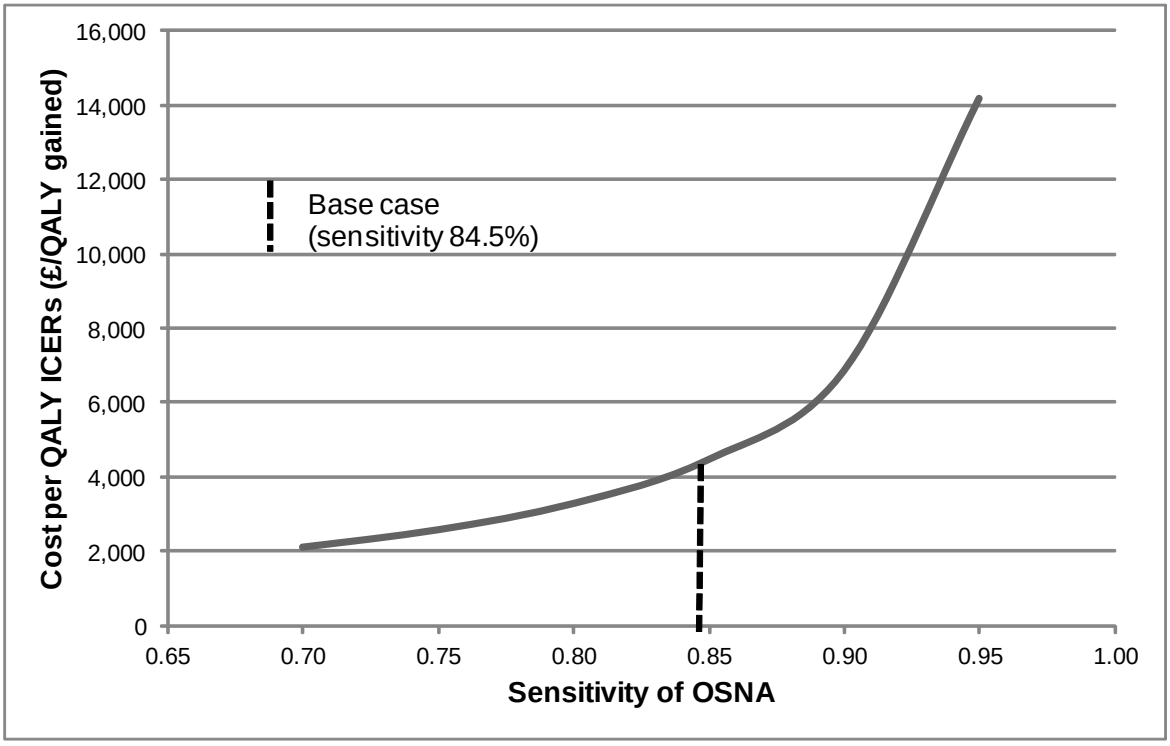
Long term results. presented in Table 55., show a similar finding to the accuracy results: as the sensitivity of OSNA increased, so did the ICER for cost per QALY gained by histopathology. These ranged from £2,119 per QALY gained when OSNA had sensitivity 70% to £14,193 per QALY gained when OSNA had 95% sensitivity. At 100% sensitivity OSNA dominated histopathology, having more QALYs and fewer costs. Unlike the accuracy results, the cost difference between histopathology and OSNA also increased each time the sensitivity increased. Again, as the sensitivity of OSNA neared 100% the ICERs began to increase much faster, as demonstrated in Table 55.

Table 55. Long term results for threshold analysis for sensitivity

	Increase in accuracy							
Measure	**Histopathology vs.OSNA[1] with sensitivity**							
	Base Case: 84.5%	**70%**	**75%**	**80%**	**85%**	**90%**	**95%**	**100%**
QALYs (discounted)	0.0997	0.1939	0.1614	0.1289	0.0964	0.0639	0.0314	-0.0011
		NHS reference costs of ALND						
Costs per patient (discounted)	£431	£411	£418	£425	£431	£438	£445	£452
Incremental cost per QALY	£4,324	£2,119	£2,588	£3,294	£4,476	£6,862	£14,193	Histopathology dominated

Costs and QALYs discounted at a rate of 3.5%

Figure 18. Comparison of long term cost effectiveness for threshold analysis of OSNA sensitivity



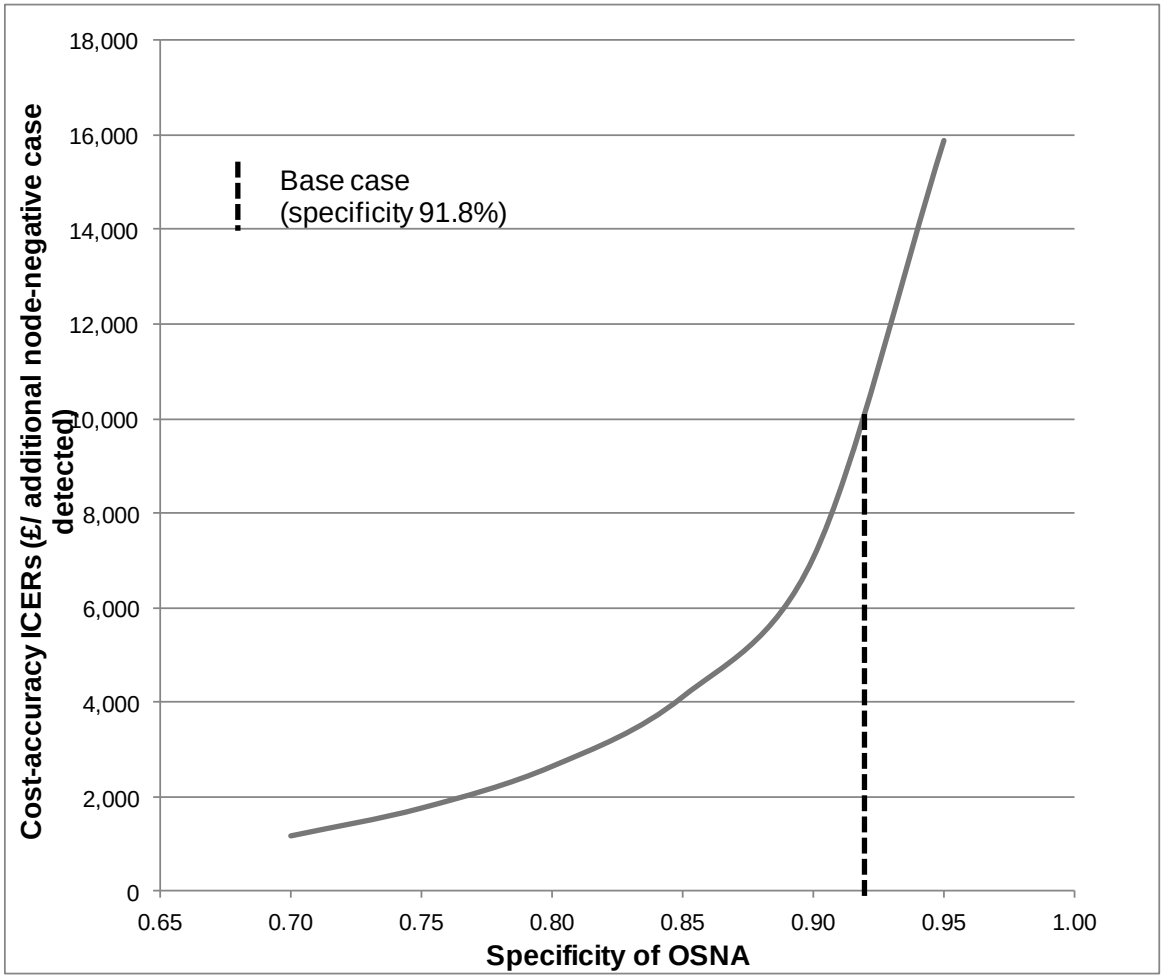
The results of the threshold analysis for specificity are presented in Table 56 and Table 57. These demonstrate that when sensitivity is held at 84.5%, the proportion of node negative cases detected by OSNA increases by 4% each time the specificity increased by 5% and the costs decreased by roughly £70. As with the threshold analysis for sensitivity, the accuracy ICERs for histopathology increased with the specificity of OSNA. In terms of the cases correctly identified, the ICERs ranged from £1,043 per additional case correctly diagnosed by histopathology when OSNA had specificity 70% to £22,761 per additional case correctly diagnosed when OSNA had specificity 100%. For the node negative cases, the ICERs ranged from £2,671 per node negative case detected when OSNA had specificity 70% to OSNA dominating histopathology when OSNA had specificity 100%. This time both cost-accuracy measures had ICERs that increased faster as the specificity of OSNA approached 100%, as demonstrated in Figure 19.

Table 56: Accuracy results for threshold analysis for specificity

		Increase in accuracy						
Measure		**Histopathology vs.OSNA[1] with specificity**						
	Base case: 91.8%	**70%**	**75%**	**80%**	**85%**	**90%**	**95%**	**100%**
Accuracy[a]	0.0966	0.2710	0.2310	0.1910	0.1510	0.1110	0.0710	0.0310
Specificity*(1-Prevalence)[1]	0.0656	0.24	0.20	0.16	0.12	0.08	0.04	0
		NHS reference costs of ALND						
Costs per patient	£590	£283	£353	£424	£494	£565	£635	£706
Incremental cost per additional patient correctly diagnosed	£6,108	£1,043	£1,529	£2,218	£3,272	£5,087	£8,945	£22,761
Incremental cost per additional node-negative case detected	£8,994	£1,178	£1,766	£2,648	£4,118	£7,058	£15,878	Histo-pathology dominated

[1] Node positive prevalence fixed at 20%. [a] Accuracy refers to the cost per case correctly identified Here strategies that are dominated have the same detection rate, but are more expensive.

Figure 19. Comparison of cost-accuracy results for threshold analysis of OSNA specificity.



In this figure, the ICER for histopathology increases as the specificity of OSNA increases. The node-negative cases curve appears truncated because at the maximum specificity (100%) OSNA dominates histopathology (same diagnostic yield, lower costs), which is not represented on the graph.

In the long term cost effectiveness results, presented in Table 57 and Figure 20, the long term costs of OSNA decreased as the specificity and QALY gain increased. This meant that for the lowest specificity of 70%, OSNA was dominated by histopathology as it was both more expensive and had fewer QALYs. The largest ICER for histopathology, when OSNA had 100% specificity was £8,430 per QALY gained. As with the previous results, the ICER increase was more pronounced as the cost difference between the histopathology and OSNA increased (i.e. as the specificity of OSNA increased), but this increase was not as severe as it was for the short term outcomes for the specificity threshold analysis, nor the long term cost effectiveness outcomes for the sensitivity threshold analysis.

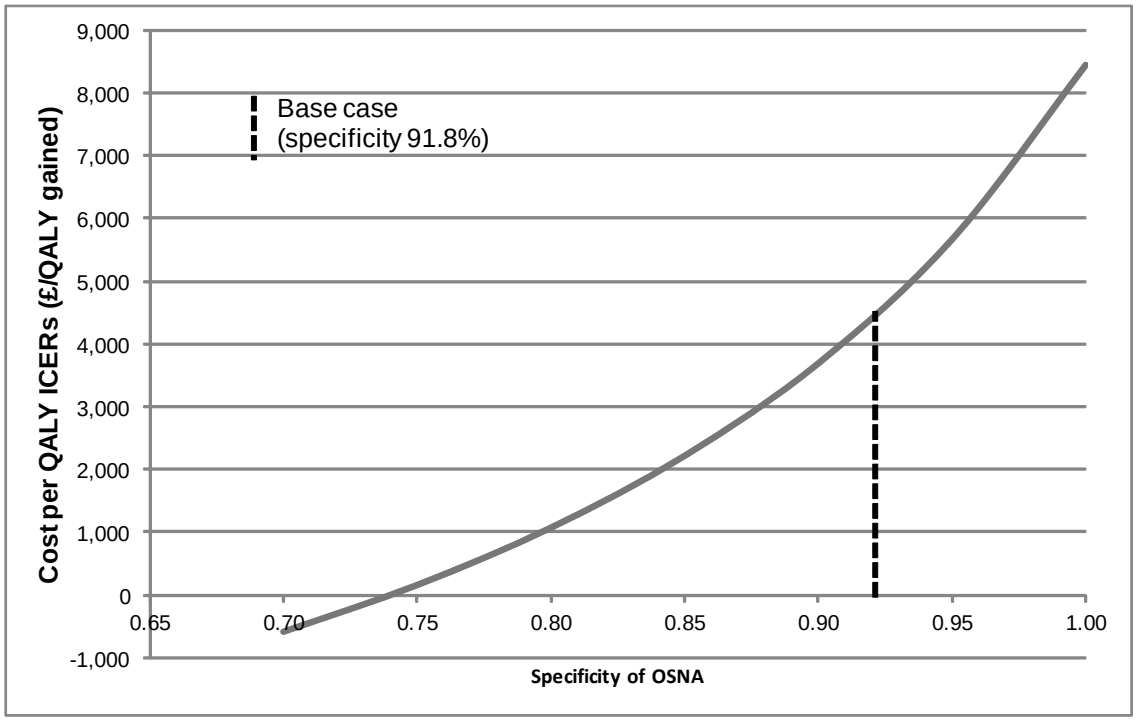
Table 57. Long term results for threshold analysis for specificity

	Increase in accuracy
Measure	**Histopathology vs.OSNA[1] with specificity**

	Base case: 91.8%	70%	75%	80%	85%	90%	95%	100%
QALYs (discounted)	0.0097	0.1660	0.1508	0.1356	0.1203	0.1051	0.0899	0.0747
	NHS reference costs of ALND							
Costs per patient (discounted)	£431	-£98	£24	£145	£266	£387	£508	£630
Incremental cost per QALY	£4,324	OSNA dominated	£156	£1,068	£2,210	£3,683	£5,655	£8,430

OSNA was dominated at specificity 70% as it had fewer QALYs and higher costs than histopathology. Costs and QALYs discounted by a rate of 3.5%.

Figure 20. Comparison of long term cost effectiveness for threshold analysis of OSNA



These threshold analyses demonstrate no obvious significant difference to the base case results, though all ICERs increased as either the sensitivity or specificity increased. Staring with about 90% the changes in sensitivity or specificity resulted in exponential increases in the ICERs. Individually, the findings from the base case are repeated, with specificity seemingly having a larger impact on the short term outcome of cost-accuracy and sensitivity on the long term cost-effectiveness.

Overall these threshold analyses suggest that if the true values of sensitivity and specificity for OSNA lie within the range of 90-100%, close to those of the sensitivity and specificity of histopathology, the cost effectiveness of OSNA may increase greatly. Furthermore, sensitivity

and specificity are not independent and therefore altering one would alter the other. For simplicity and because the relationship between the two is not yet properly understood, this analysis does not analyse the relationship between specificity and sensitivity, and therefore it may not accurately represent what the true results would be. However, as the TAB results from Khaddage demonstrate, when both specificity and sensitivity are sufficiently high, OSNA becomes much more cost effective.

6.1.2.2 Prevalence

The prevalence of lymph node metastases in the population directly affects the accuracy of tests and their outcomes. To investigate the effect fully, we considered the effect of prevalence when our base case value of 20% was halved to 10% and doubled to 40%.

When prevalence was 10%, histopathology dominated OSNA half node and had ICERs under £20,000 per additional case correctly identified for histopathology versus OSNA full node, both overall and when split into node negative and node positive patients. Short term cost utility results remained the same, with full node OSNA dominating the rest. In the long term, histopathology dominated half node OSNA, and the ICER for histopathology compared to full node OSNA was £2,626 per QALY gained.

When the prevalence was increased to 40%, accuracy results were similar to those of the base case, with histopathology being cost effective for additional cases correctly identified, OSNA half node dominating histopathology for detecting additional node positive patients and OSNA full node dominating half node for additional node negative patients detected. The ICER comparing histopathology to OSNA full node for detecting node negatives cases increased to £24,680 per additional case detected. OSNA once again dominated the short term utility analyses. In the long term half node OSNA dominated histopathology, as it now had slightly higher QALY gains (0.0001). OSNA half node compared to full node had an ICER of £2,208 per QALY gained.

6.1.2.3 Short term costs

In the short term the costs of the tests were altered by +/- 10%. Altering the cost of histopathology did not affect the overall outcomes, simply increasing or reducing individual ICERs accordingly. Decreasing the cost of OSNA to £315 similarly did not greatly affect the results, but increasing it to £385 caused histopathology to dominate half node OSNA in the long term, with an ICER of £3,972 when comparing histopathology to full node OSNA.

The cost of a separate ALND surgery was altered using values provided in the NHS Reference Costs. Low second surgery costs (£1,608) meant that short term histopathology dominated half node OSNA for all accuracy ICERs and comparing to full node, all ICERs were below £7,000 per additional case detected. For short term utility full node OSNA still dominated half node and histopathology. In the long term histopathology dominated half node OSNA and when comparing to full node, histopathology had an ICER of £388 per QALY gained. High second surgery costs (£4,871) had no significant effect on short term outcomes. Long term ICERs increased to £15,384 per QALY gained when comparing histopathology to OSNA half node and £5,469 per QALY gained when comparing half node OSNA to full node.

6.1.2.4 Long term costs

Adjuvant therapy costs were altered by +/-10%. This only affected the long term results and did not greatly influence their results. High costs for patients undergoing hormonal adjuvant therapy (£1195) increased the ICERs slightly, with histopathology compared to full node OSNA having an ICER of £4,353. Low costs for patients undergoing chemotherapy (£8,502) had the smallest ICERs, reducing the ICER between histopathology and full node OSNA to £4,237 per QALY gained.

Tables of these sensitivity analyses can be found in Table 58 and Table 59. These include sensitivity analyses performed on the disutility associated with anxiety from waiting for results and a second operation and the costs of the first breast surgery (with or without ALND), but there were no significant differences in results. Overall, we found that the most influential parameters were the sensitivity and specificity of OSNA, which makes good data on these values very important to have.

Table 58. Sensitivity analyses results for cost-accuracy

Parameter	**Base Case**	Sensitivity Analysis	Cost Accuracy ICERs using NHS Reference costs					
			Incremental cost per additional patient correctly diagnosed		Incremental cost per additional node-positive case detected		Incremental cost per additional node-negative case detected	
			Difference OSNA half node vs. OSNA full node	**Difference Histopath-ology vs. OSNA half node**	**Difference OSNA half node vs. OSNA full node**	**Difference Histopath-ology vs. OSNA half node**	**Difference OSNA half node vs. OSNA full node**	**Difference Histopath-ology vs. OSNA half node**
Base case	N/A	N/A	OSNA half node extended dominated	£6,108*	£16,123	Histopathology dominated	OSNA half node dominated	£8,994*
Prevalence	20%	10%	OSNA half node extended dominated	£3,112*	OSNA half node extended dominated	£17,930*	OSNA half node dominated	£3,766*
		40%	OSNA half node	£10,919*	£9,132	Histopathology	OSNA half node	£24,680*

			extended dominated			dominated	dominated	
Cost of histopath-ology	£472	£425	OSNA half node extended dominated	£5,619*	£14,957	Histopathology dominated	OSNA half node dominated	£8,274*
		£519	OSNA half node extended dominated	£6,597*	£17,289	Histopathology dominated	OSNA half node dominated	£9,714*
Cost of OSNA	£350	£315	OSNA half node extended dominated	£6,470*	£16,123	Histopathology dominated	OSNA half node dominated	£9,528*
		£385	OSNA half node extended dominated	£5,746*	£16,123	Histopathology dominated	OSNA half node dominated	£8,461*
Cost of second surgery	£3,569	£1,608	OSNA half node extended dominated	£2,048*	OSNA half node extended dominated	£6,383*	OSNA half node dominated	£3,016*
		£4,871	OSNA half node extended dominated	£8,805*	£17,425	Histopathology dominated	OSNA half node dominated	£12,966*

* Comparison is Histopathology relative to full node OSNA due to the half node OSNA option being dominated or extended dominated. OSNA half node dominated when diagnostic yield is same as OSNA full node, but full node is less expensive. OSNA half node is extended dominated when histopathology has a smaller ICER for additional diagnostic yield compared to OSNA full node, than that of OSNA half node. Histopathology is dominated when OSNA half node has same diagnostic yield, but is less expensive.

Table 59. Sensitivity analyses for short term and long term cost-effectiveness

Parameter	**Base Case**	Sensitivity Analysis	Short term and long term cost-effectiveness ICERs using NHS Reference Costs.			
			Short term incremental cost per QALY gained		Long term incremental cost per QALY gained (cost and QALYs discounted)	
			Difference OSNA half node vs. OSNA histopathology	**Difference OSNA full node vs. OSNA half node**	**Difference OSNA half node vs. OSNA full node**	**Difference Histopathology vs. OSNA half node**
Base case	N/A	N/A	OSNA half node extended dominated	Histopathology dominated*	OSNA half node extended dominated	£4,324*
Prevalence	20%	10%	OSNA half node extended dominated	Histopathology dominated*	OSNA half node extended dominated	£2,626*
		40%	OSNA half node extended dominated	Histopathology dominated*	£2,208	Histopathology dominated
Cost of histopathology	£472	£425	OSNA half node extended dominated	Histopathology dominated*	OSNA half node extended dominated	£3,850*
		£519	OSNA half	Histopathology	OSNA half	£4,797*

			node extended dominated	dominated*	node extended dominated	
Cost of OSNA	£350	£315	OSNA half node extended dominated	Histopathology dominated*	OSNA half node extended dominated	£4,675*
		£385	OSNA half node extended dominated	Histopathology dominated*	OSNA half node extended dominated	£3,972*
Cost of second surgery	£3,569	£1,608	OSNA half node extended dominated	Histopathology dominated*	OSNA half node extended dominated	£388*
		£4,871	OSNA half node extended dominated	Histopathology dominated*	£5,469	£15,384
Annual cost of adjuvant therapy (hormone therapy and follow up)	£1,087	£978	N/A	N/A	OSNA half node extended dominated	£4,311*
		£1,195	N/A	N/A	OSNA half node extended dominated	£4,353*
6 month cost of adjuvant therapy (chemotherapy and follow up)	£9,447	£8,502	N/A	N/A	OSNA half node extended dominated	£4,237*
		£10,392	N/A	N/A	OSNA half node extended dominated	£4,402*
Disutility of anxiety waiting for histopathology results	0.019	None	OSNA half node extended dominated	Histopathology dominated*	OSNA half node extended dominated	£3,618*
		0.006	OSNA half node extended dominated	Histopathology dominated*	OSNA half node extended dominated	£3,805*
		0.023	OSNA half node extended dominated	Histopathology dominated*	OSNA half node extended dominated	£4,463*
Disutility of second surgery for ALND	0.033	None	OSNA half node extended dominated	Histopathology dominated*	OSNA half node extended dominated	£4,052*
		0.027	OSNA half node extended dominated	Histopathology dominated*	OSNA half node extended dominated	£4,266*
		0.04	OSNA half node extended dominated	Histopathology dominated*	OSNA half node extended dominated	£4,382*

* Comparison is Histopathology relative to full node OSNA due to the half node OSNA option being dominated or extended dominated. When stated, costs and QALYs discounted at a rate of 3.5%. Note that for short term utility the order is reversed, as OSNA full node has the highest QALY yield and histopathology the lowest.

6.1.3 Metasin Results

As an alternative index test, we compared Metasin to histopathology. This was performed as an illustrative exercise, as the accuracy values for Metasin are from a draft paper that had not yet been peer reviewed at the time of this review. Furthermore, the only cost information for Metasin was provided from personal correspondence with The Princess Alexandra Hospital on behalf of the author of the draft paper. This value only reflected the cost of the non-CE marked Metasin test, since no list price for the test is yet available though Metasin has received a CE mark (*The cost effectiveness analyses may be updated if and when a finalised list price for the CE-marked Metasin test is received.)* Updated cost estimates were made available after modelling analyses were completed, but as these were provided at a later date and the changes to cost did not seem significantly different from those used in the analysis, these were not implemented.

Cost-accuracy results are reported in Table 60 and report that the ICER between histopathology and Metasin half node was £15,695 per additional case detected under NHS Reference costs and £8,869 under YHEC costing. As with OSNA, Metasin's half node intraoperative analysis dominated histopathology for additional node positive cases, as it had the same accuracy, but was less expensive than histopathology. Furthermore, full node Metasin dominated half node for detecting node-negative cases. The ICER for node-negative cases comparing histopathology to full node Metasin was £30,453 per additional node negative case detected with NHS Reference costs and £22,484 with YHEC costs. The cost effectiveness of histopathology compared to Metasin therefore depends on the costing strategy and threshold used.

Table 60: Short term costs-accuracy analysis comparing histopathology to full and half node intraoperative Metasin analysis

	Mean estimates			**Incremental results**	
Measure	**Histopathology**	**Metasin[1]**		**Difference Metasin half vs. full node**	**Difference Histopathology vs. Metasin half node**
		half node	**full node**		
Accuracy[a]	1.0000	0.9704	0.9556	0.0148	0.0296
Sensitivity*Prevalence[1]	0.2000	0.2000	0.1852	0.0148	0
Specificity*(1-Prevalence[1])	0.8000	0.7704	0.7704	0	0.0296
	NHS reference costs of ALND				
Costs per patient	£3,987	£3,523	£3,086	£437	£465
Incremental cost per additional patient correctly diagnosed				Metasin half node extended dominated	£20,302*
Incremental cost per additional node-positive case detected				£29,515	Histopathology dominated
Incremental cost per additional node-negative case				Metasin half node dominated	£30,453*

detected					
	Analysis using costs based on YHEC model				
Costs per patient	£2,228	£1,966	£1,562	£403	£263
Incremental cost per additional patient correctly diagnosed				Metasin half node extended dominated	£14,990*
Incremental cost per additional node-positive case detected				£27,230	Histopathology dominated
Incremental cost per additional node-negative case detected				Metasin half node dominated	£22,484*

1 Node positive prevalence fixed at 20%. [a] Accuracy refers to the proportion of cases correctly identified * Comparison is Histopathology relative to full node OSNA due to the half node Metasin option being dominated (same diagnostic accuracy as full node, more expensive) or extended dominated (histopathology had a smaller ICER when compared to full node Metasin than half node Metasin did). Here strategies that are dominated have the same detection rate, but are more expensive.

Short term utility results for Metasin are reported in Table 61. Regardless of costing strategy Metasin dominated histopathology, having fewer costs and higher utilities. It is estimated that of the 78.5% of half node patients with negative results, who have to wait for histopathology test results, 1.9% would go on to have a positive diagnosis, in contrast with the 20% of expected positive diagnosis with post-operative histopathology. In the long term, histopathology compared to half node Metasin has an ICER of over £460,000 per QALY gained for NHS Reference costs and an ICER over £240,000 per QALY gained using YHEC costs. The ICER between half node and full node Metasin remained under £13,000 for both costing strategies. These figures are reported in Table 62.

Table 61: Short term utility for Metasin

	Mean estimates			**Incremental results**	
Measure	**Histopathology**	**Metasin**		**Difference Metasin half node vs. histopathology**	**Difference Metasin full node vs. Metasin half node**
		half node	**full node**		
	NHS reference costs of ALND				
Costs per patient	£3,987	£3,523	£3,086	-£465	-£437
Utility	0.9739	0.9849	1	0.0110	0.0151
Incremental cost per QALY gained				Metasin half node extended dominated	Histopathology dominated*
	Analysis using costs based on YHEC model				
Costs per patient	£2,228	£1,966	£1,562	-£263	-£403
Incremental cost per		-	-	Metasin half node extended dominated	Histopathology dominated*

QALY gained					

Histopathology and Metasin half node were both dominated by Metasin full node (and Metasin full node had a smaller ICER than Metasin half node compared with histopathology), which had a higher QALY gain and lower costs.

Table 62. Long term costs and QALYs for Metasin

	Mean estimates			Incremental results	
Measure	**Histopathology**	**Metasin half node**	**Metasin full node**		
				Difference Metasin half node vs. Metasin full node	**Difference Histopathology vs. Metasin half node**
	NHS reference costs of ALND				
Cost per patient	£20,530	£20,103	£19,702	£401	£427
QALYs	9.321	9.320	9.288	0.032	0.001
Incremental cost per QALY gained				£12,374	£467,113
	Analysis using costs based on YHEC model				
Costs per patient	£18,771	£18,546	£18,179	£367	£225
Incremental cost per QALY gained				£11,329	£246,089

7 Discussion

7.1 Statement of principal findings

7.1.1 Clinical effectiveness

Eighteen studies were included that investigated the performance of either OSNA or Metasin on detecting metastases in the sentinel or axillary lymph nodes of breast cancer patients. Two studies were included for Metasin; both were unpublished and in draft form. The remaining sixteen studies reported on OSNA, with two papers reporting the same study.

The studies were assessed against the Cochrane Risk of Bias tool[1] and the Quadas-2.[2] The majority of studies were considered to be at low risk of bias, although many were unclear regarding their method of patient recruitment and lacked detail on patient characteristics. Often, no evidence was given of sample replicates and reproducibility for molecular analysis. Furthermore, test failures were not reported for most studies. Reported outcomes were also limited, for example no data were found for clinical outcomes, such as patient anxiety and number of repeat operations. Only one included study provided evidence for time in operating theatre.[58] A potential conflict of interest also features heavily, since one of the two unpublished Metasin studies was performed at the institution in which the technology was developed and the majority of the OSNA studies were financially supported by Sysmex.

With regard to test accuracy, there are two main issues. The first is the strong assumption that the reference standard is the most accurate measure of the target disorder. In this case, the reference standard (i.e. histopathology), although plausible, has been performed with varying levels of analysis, and as such, may not be a true indicator of the target condition. The second concern was tissue allocation bias (TAB), which occurs when a different portion of tissue is allocated to the index and reference test and cannot then be re-used between them. Studies have dealt with this in a variety of ways, some re-analysing both histopathology and molecular samples, some choosing to re-analyse just one technology and some doing neither. The majority of studies which have adjusted for TAB, have taken a conservative approach by excluding affected samples, which we consider to be a reasonable practice.

A summary of results is presented in Table 63. As there were only two studies for Metasin, a meta-analysis was not performed. The displayed data for this test were taken from draft papers, prior to peer review. Therefore, the results, [redacted], must be used with caution.

Table 63. Summary of pooled results

	Sample type	Adjustment for TAB	No. of studies	Sensitivity (95% CI)	Specificity (95% CI)
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pooled OSNA	Patient	No	5	84.5 (74.7–91.0)	91.8 (87.8–94.6)
Pooled OSNA	Patient	Yes	3	91.3 (83.6–95.6)	94.2 (91.2–96.2)
Pooled OSNA	SLN	No	4	79.9 (74.2–84.6)	95.5 (94.1–96.5)
Pooled OSNA	SLN	Yes	5	89.0 (82.1-93.4)	97.5 (96.6-98.2)
Pooled OSNA	ALN	No	6	95.1 (90.0–97.6)	94.9 (91.2–96.9)
Pooled OSNA	ALN	Yes	4	96.5 (87.3–99.1)	96.2 (93.4–97.8)

It should be noted that more than one sentinel lymph node may be removed from a patient, which, as shown in Table 63, may have implications for sensitivity and specificity data when grouped either by patient or by node. Adjustment for TAB, clearly improves the test accuracy, increasing sensitivity from 79.9% to 89.0% and increasing specificity from 95.5% to 97.5% for OSNA at the node level, with similar increases at patient level.

Studies which analysed ALNs were also included in this review. Increased sensitivity (ALN 95.1% vs. SLN 79.9%) and similar specificity (ALN 94.9% vs SLN 95.5%) to SLN results were seen before adjustment for TAB. Subsequent TAB adjustment increases the SLN sensitivity, but has little effect on ALN data.

With regard to the time taken to perform OSNA, despite the lack of detail in the studies explaining which aspects of the procedure were monitored, the time ranges from less than 30 minutes to 39.6 minutes for one node. This increases by approximately 5 to 10 minutes per additional node analysed.

Overall and despite the concerns with the reference standard and TAB, we feel the studies were well performed and produced consistent results.

7.1.2 Cost effectiveness

The short term results, up to the diagnostic phase, suggest that OSNA is less accurate but saves costs and quality of life losses to patients due to anxiety from delayed diagnosis and second operation relative to the referent standard of histopathology. In Metasin the evidence base is currently only indicative and requires validation, as it is limited to two unpublished studies, and may therefore be neither of the same quality nor as generalisable as the respective evidence on OSNA which has accumulated over several available studies from a wider variety of settings. It is possible that there is more confidence in the accuracy of OSNA and this should be taken into account when assessing the technologies, as OSNA still has the potential for cost savings compared with current practice.

Limiting our analysis to the evidence on OSNA versus histopathology, revealed that there may be a trade-off between histopathology and intra-operative testing but that this depends on estimate of costs adopted. Valuing resource use at the current national reference costs, and judged on the ability to identify node-positive patients, histopathology was found to be inferior to OSNA. Whether the higher costs of using OSNA adjunctively are justified by the additional sensitivity of the test strategy, however, depends on whether decision makers are prepared to pay £16,000 per additional case detected. On the other hand, when estimates derived from a microcosting study developed by the York Health Economics Consortium in a study commissioned by NTAC are used to value resource use, histopathology may be cost effective as long as the cost-effectiveness threshold is above £12,000 per node positive case detected. In contrast, when the disutility of anxiety and a second operation is taken into account and the diagnostic options are compared in terms of the costs and such disutility, OSNA emerges as clearly preferred option with lower costs and less harm (disutility) inflicted on patients.

The above results, would be sufficient for medical decision making if one entertained the idea that more accurate diagnosis does not have significant health benefits for patients (i.e. QALY gains). Despite the inherent uncertainty of extrapolating benefits from the diagnostic phase into the remaining lifetime of patients undergoing SLNB, the present analysis has adapted an existing model of early breast cancer patient management developed by the School of Health and Related Research (ScHARR), to gain insight about the potential significance of benefits due to diagnostic accuracy gains in this clinical area. After accounting for the disutility of anxiety due to waiting for post-operative diagnostic test results and a second operation, the long term analysis using the adopted ScHARR model suggest that histopathology may produce large enough benefits from improved patient management over intra-operative testing approaches to offset its short term disadvantage due to the costs and disutility associated with delayed diagnosis and second operations. Indeed results suggest that as long as the NHS is willing to pay more than

£4,000 per QALY gained histopathology should be retained as the optimal diagnostic approach for metastatic diagnosis in early breast cancer.

The base-case result that histopathology is cost-effective is however not robust to adjustment for tissue allocation bias. Of the three studies that provided good quality of evidence on sensitivity and specificity, two supported histopathology as the cost-effective option at the conventional £20,000 per QALY threshold, whereas the third, a study by Khaddage and colleagues results in OSNA full node being unambiguously dominant.

A comment is also warranted regarding the relevant cost-effectiveness threshold for the present analyses. Typically in cost-effectiveness analyses, the status quo is less effective and less costly than the new technologies against which it is being compared. In the present case the reverse applies, since histopathology, the current standard is both more effective (i.e. has higher QALYs) and more costly that the new, intraoperative approaches. To facilitate interpretation we have presented results of the three arm comparison (intra-operative test with half vs. full node vs. histopathology) by ranking the three options by increasing level of effectiveness and calculating incremental costs, benefits and ICERs between adjacent options. That way we can retain the standard interpretation of ICERs for decision making. However, it is argued by some economists that the threshold for cost-effectiveness should be higher for new technologies that involve foregoing some benefit in exchange for cost savings, as in the case of OSNA or Metasin relative to histopathology in our present analysis. If this is the case, then histopathology may well have an ICER relative OSNA larger than the conventional £20,000 per QALY and still be considered cost-effective. The problem is that no conventional thresholds for new technologies with lower costs and lower benefits exist.

7.2 Strength and limitations of assessment

7.2.1 Clinical effectiveness

The strengths of this systematic review are that it was conducted by an independent research team using the latest evidence in line with a pre-specified protocol. The search strategy did not restrict by study design and also included forward chasing. The studies were independently screened by two reviewers, with data extraction and quality appraisal performed by one reviewer and checked by a second. Any disagreements were resolved by consensus. A large number of

abstracts were identified which were not included in the review, since a quality appraisal could not be performed. However, a table of results has been compiled in the Appendix for interest.

The assessment of Metasin should be treated with caution. Only two studies were included, both of which were unpublished and therefore had not received peer review.

For all studies, there was a wide variance in histopathology performed, e.g., one-level, three-level or five-level. It is questionable whether the use of one-level histopathology is appropriate as a reference standard. Use of discordant case analysis was inconsistent and results sometimes attributed to TAB, which may or may not be appropriate, although generally, a conservative approach of sample exclusion was used. Some studies reported patient level data, node data, ALNs or SLNs. Often for the ALN analysis, the recruited population was relatively small, but a large number of nodes analysed. On some occasions it was unclear whether fresh or frozen samples had been used for histopathology.

7.2.2 Cost effectiveness

The scope of our analysis was limited by some assumptions adopted to address data limitations. We have not accounted for the possibility of second operations due to involved or close margins of the primary tumour. Indeed, a second operation rate for OSNA of 9% vs. 39% for histopathology, including 5% and 13% for margins insufficiency without ALND resection, respectively, has been reported.[100] No previous analysis has looked at this complex issue in terms of its costs, let alone its long term health implications. Our current analysis does not look at how the cost and health benefit balance of intra-operative testing relative to histopathology and between intra-operative testing options may be affected by the current controversy about the effectiveness of ALND, particularly in relation to micro-metastatic disease.[101] If part of the claimed advantages by OSNA, the detection of micro-metastatic involvement, has no associated survival benefits its long term cost-effectiveness may be overestimated in our present analysis. Undertaking objective analysis of this issue is complicated however, since the threshold definitions of micrometastasis of histopathology and OSNA differ and because of the current uncertainty associated with the emerging evidence on the benefits of ALND in small tumours found in the sentinel lymph node.

Another limitation that originated from the lack of evidence to inform our analysis pertained to the potential loss of valuable information for patient management that full node use by intra-operative testing incurs relative to histopathology.[102] The extent to which this may affect our results is open to highly speculative judgement.

We have not looked at the cost-effectiveness implications of possible options that have been considered for increasing the sensitivity of diagnostic strategies considered here. For example, OSNA reliance on the CK19 marker detecting metastasis may lead to missing the small proportion of patients with metastatic breast cancer whose tumour does not express the marker. It has been suggested that pre-operative testing for the primary tumour may identify the subgroup of patients without CK19 expression and who require an alternative approach for metastatic diagnosis. The same could be adopted for Metasin, to exclude the possibility that those testing negative may be false negative cases due to lack of expression of CK19 and Mammaglobin, the markers used by the test to identify positive cases.

7.3 Uncertainties

7.3.1 Clinical effectiveness

The primary concern for the studies where a portion of the node was provided for molecular analysis and a portion provided for histopathology, was TAB. Many studies attempted to address this, however, whether the adjusted results were truly due to TAB or whether a proportion of cases had been missed, is unclear.

The histopathology methods were sometimes ambiguously reported, and the variety compared is likely to be reflected in the heterogenous data.

A significant proportion of the studies were funded by the manufacturers, therefore there may be a potential conflict of interest.

7.3.2 Cost effectiveness

Throughout our analyses we have assumed that histopathology is the gold standard. There are however indications that histopathologists may miss low volume metastases.[103,104,105] Due to the small frequency of this omission relative to the occurrence of tissue allocation bias, our analysis may have adopted a slightly conservative stance.

We present our results under two costing approaches that have been used by other assessment groups previously. They have significant implications for the decision since, as discussed for the analysis of the evidence on OSNA versus histopathology, the choice of one or the other approach in the present analysis led to status quo, histopathology, being unambiguously inferior to intra-operative testing as opposed to an option whose cost-effectiveness depended on a judgement about the value of its additional accuracy relative to its costs.

8 Conclusions

8.1 Implications for service provision

OSNA and Metasin may provide some advantages over histopathology. They are automated, standardised and objective, with the capacity to be used intraoperatively. Furthermore, the whole node can be analysed, removing the concern of TAB. However, information gained by histology, such as the location of metastatic foci and morphological evaluation are not possible. OSNA displays a higher rate of false positives, possibly due to contamination from epithelial cells. In contrast, false negatives may occur if the breast cancer does not express CK19 protein.

[redacted]

Adoption of OSNA or Metasin may bring significant benefits to patients and cost savings to the NHS. However there remain important questions regarding whether ultimately those benefits are worth the costs of less accurate diagnosis and potentially inferior long term health outcomes for patients.

8.2 Suggested research priorities

The uncertainty in decision making due to lack of data and variation in available estimates on diagnostic accuracy, short-term implications of diagnostic approaches and long term patient management may be reduced by undertaking further research. Priority should be placed on documenting the performance of intra-operative tests under different protocols. There is limited information about the performance of such tests under routine established practice and many of the existing studies are likely to apply to an early experience with those technologies. Peer reviewed and formally published research on Metasin is also essential as the presumed low cost of this test makes it a potentially attractive option for the NHS.

Greater clarity on the true costs of the alternative tests, and the variation in resource utilisation at the level of the patient, would help to identify the significance and the distribution of costs and benefits of intra-operative diagnosis. Observational studies may be useful to verify empirically confirm and quantify the reduction in the numbers of operations, the anxiety caused to patients, and its impact on their quality of life, by delayed availability of test results and second operations, and the costs implications to hospitals of the introduction of intraoperative testing in SLNB. While

much uncertainty is likely to remain about the magnitude of long term benefits of increased diagnostic accuracy, a natural target for great returns on investment in research may be found in documenting what the impact on patient reported outcomes is of uncertainty associated with delayed diagnosis and second operation. Indeed we may expect greater prominence of this aspect of the decision making problem in populations with greater prevalence of node positive SLN.

However improving on the accuracy of estimates for OSNA in particular, overcoming concerns about TAB and the validity of currently available reference standards may be challenging. A test-treat randomised trial may be the only way to truly resolve whether introduction of intraoperative testing in SLNB would be effective and thus cost-effective. The outcome would need to be locoregional recurrence rates or even survival in order to capture the trade-off between the potential short term gains associated with single operations to achieve ALND and the longer term disbenefits arising from false negative and false positive cases occurring with intraoperative testing.

Also, a strong assumption in this report is that ALND is the usual best treatment if micro and macrometastases, or their equivalents, are identified in a SLNB. Evidence on this is evolving and needs to be followed closely as it could impact on decisions about intraoperative testing in SLNB in the future.

9 References

1. Collaboration TC. *Cochrane Collaboration's Tool for Assessing Risk of Bias.* (last accessed
2. Whiting PF, Westwood ME, Rutjes AW *et al*. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. *BMC Medical Research Methodology* 2006; **6:** 9.
3. Cutress RI, McDowell A, Gabriel FG *et al*. Observational and cost analysis of the implementation of breast cancer sentinel node intraoperative molecular diagnosis. *J Clin Pathol* 2010; **63:** 522-529.
4. Cooper KL, Meng Y, Harnan S *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 Asses* 2011; **15:** 1-+.
5. Burke MP, T. *The Cost Impact of Implementing Intra-Operative Testing for the Diagnosis of Patients with Metastatic Breast Cancer in England*. 2010.
6. Dumitrescu RG, Cotarla I. Understanding breast cancer risk – where do we stand in 2005? . *J Cell Mol Med* 2005; **9:** 208-221.
7. (WCISU). WCISU. Cancer incidence in Wales 2003–2007 [cited 23 Jan 2009]; Available from: www.wales.nhs.uk/sites3/Documents/242/Cancer%20Incidence%20in%20

Wales%202003–2007.pdf
8. Statistics OfN. Registrations Series MB1. . 2010.
9. UK CR. Breast cancer mortality statistics. [cited 04/01/2013]; Available from: http:www.cancerresearchuk.org/cancer-info/cancerstats/types/breast/mortality/#sex
10. Singletary SE, Connolly JL. Breast cancer staging: working with the sixth edition of the AJCC Cancer Staging Manual. . *CA Cancer J Clin* 2006; **56:** 37-47.
11. UK. CR. TNM breast cancer staging. [cited 29 Nov 2010]; Available from: www.cancerhelp.org.uk/help/default.
12. Society AC. How is breast cancer staged? [cited 29 Nov 2010]; Available from: www.cancer.org/docroot/CRI/content/CRI_2_4_3X_How_is_breast_cancer_staged_5.asp
13. UK. CR. Number stages of breast cancer. [cited 29 Nov 2010]; Available from: www.cancerhelp.org.uk/help/default.asp?page=3315.
14. Lyratzopoulos G, Abel GA, Barbiere JM *et al*. Variation in advanced stage at diagnosis of lung and female breast cancer in an English region 2006-2009. *Br Journal of Cancer* 2012; **106:** 1068-1075.
15. Carter C, Allen C, Henson DE. Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases. *Cancer* 1989; **63:** 181-187.
16. National Institute for HC, Excellence. Early and locally advanced breast cancer: diagnosis and treatment. London: NICE; 2009.
17. Statistics OfN. Survival Rates in England, patients diagnosed 2001–2006 followed up to 2007. [cited; Available from: www.statistics.gov.uk/StatBase/Product.asp?vlnk=14007. 2007.
18. Unit WMCI. 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. . In: Unit WMCI, editor. Birmingham; 2002.
19. cancer.org. [cited 9 Nov 2012]; Available from: http://www.cancer.org/cancer/detailedguide/breast-cancer-survival-by-stage (
20. Galea MH, Blamey RW, Elston CE, Ellis IO. The Nottingham Prognostic Index in primary breast cancer. . *Breast Cancer Res Treat* 1992; **22:** 207-219.
21. Online. AOA. Decision making tools for health care professionals. [cited 29 Nov 2010]; Available from: www.adjuvantonline.com/index.jsp
22. NICE.org. [cited 9 Nov 2012]; Available from: http://guidance.nice.org.uk/DT/4
23. Pazaiti A FI. Which patients need an axillary clearance after sentinel node biopsy. *International Journal of Breast Cancer* 2011**:** 9.
24. Kvistad KA, Rydland J, Smethurst HB *et al*. Axillary lymph node metastases in breast cancer: preoperative detection with dynamic contrast-enhanced MRI. . *Eur Radiol* 2000; **10:** 1464-1471.

25. Mumtaz H, Hall-Craggs MA, Davidson T *et al*. Staging of symptomatic primary breast cancer with MR imaging. *AJR* 1997; **169:** 417-427.
26. Smith IC, Welch AE, Chilcott F *et al*. Gamma emission imaging in the management of breast disorders. *Eur J Surg Oncol* 1998; **24:** 320-329.
27. Wahl RL, Siegel BA, Coleman RE, Gatsonis CG, Group PS. 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:** 227-285.
28. Wallace IW, Champton HR. Axillary nodes in breast cancer. *Lancet* 1972; **1:** 217-218.
29. 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:** 765-772.
30. 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-Chicago* 2003; **138:** 482-487.
31. 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:** 1996-2005.
32. McLaughlin SA, Wright MJ, Morris KT *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:** 5213-5219.
33. Purushotham AD, Upponi S, Klevesath MB *et al*. Morbidity after sentinel lymph node biopsy in primary breast cancer: results from a randomised controlled trial. *J Clin Oncol* 2005; **23:** 4312-4321.
34. Mansel RE, Fallowfield L, Kissin M *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:** 566-609.
35. D'Angelo-Donovan D, Dickson-Witmer D, Petrelli NJ. Sentinel lymph node biopsy in breast cancer: A history and current clinical recommendations. *Surgical Oncology* 2012**:** 1-5.
36. 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:** 1560-1568.
37. Wilke LG, McCall LM, Posther KE *et al*. Surgical complications associated with sentinel lymph node biopsy: results from a prospective international cooperative group trial. *Ann Surg Oncol* 2006; **13:** 491-500.
38. Giuliano AE HK, Ballman KV, Beitsch PD, Whitworth PW, Blumencranz PW et al. Axillary dissection vs no axillary dissection in women with invasive breast cancer and sentinel node metastasis. A randomized clinical trial. . *JAMA* 2011; **305:** 569-575.
39. Straver ME Meijnen P VTGea. Role of axillary axillary clearance after a tumor-positive sentinel node in the administration of adjuvant therapy in early breast cancer. . *J Clin Oncol* 2010; **28:** 731-737.
40. Haffty BG HK, Harris JR, Buccholz TA. . Positive sentinel nodes without axillary dssiection: implications for the radiation oncologist. . *J Clin Oncol* 2011; **29:** 4479-4481.
41. Sundaresan V. Metasin-BLNA, the NHS solution for the intra-operative assessment of sentinel lymph nodes from breast cancer patients: The multi-centre validation of 1265 cases.; Unpublished.
42. Programme NDA. Intraoperative tests (RD-100i OSNA system and Metasin test) for detecting sentinel lymph node metastases in breast cancer.
43. NHS Centre for Reviews and Dissemination. *Undertaking systematic reviews of research on effectiveness. ,* 2 edn. York: NHS Centre for Reviews and Dissemination, 2001.
44. Reitsma JB, Glas AS, Rutjes AWS *et al*. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. . *Journal of Clinical Epidemiology* 2005; **58:** 982-990.
45. StataCorp. Stata Statistical Software: Release 12. College Station, TX: StataCorp LP.; 2011.
46. Harbord RM, Whiting PF. Metandi: Meta-analysis of diagnostic accuracy using hierarchical logistic regression. . *The Stata Journal* 2009; **9:** 211-229.

47. Pennant M TY, Pennant L, Davenport C, Fry-Smith A, et al. A systematic review of positron emission tomography (PET) and positron emission tomography/computed tomography (PET/CT) for the diagnosis of breast cancer recurrence. *Health Technol Asses* 2010; **14**.
48. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. . *Statistics in Medicine* 2001; **20:** 2865-2884.
49. Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JAC. A unification of models for meta-analysis of diagnostic accuracy studies. . *Biostatistics* 2007; **8:** 239-251.
50. Arends LR, Hamza TH, van Houwelingen JC *et al*. Bivariate random effects meta-analysis of ROC curves. . *Medical Decision Making* 2008; **28:** 621-638.
51. McDowell A, Yiangou C, Wise M *et al*. Intra-operative sentinel node assessment by QRT-PCR: experience from a single centre.; Unpublished.
52. Bernet L, Cano R, Martinez M *et al*. Diagnosis of the sentinel lymph node in breast cancer: a reproducible molecular method: a multicentric Spanish study. *Histopathology* 2011; **58:** 863-869.
53. Bernet Vegue L, Rojo F, Hardisson D *et al*. Comparison of molecular analysis and histopathology for axillary lymph node staging in primary breast cancer: Results of the B-CLOSER-I study. *Diagnostic Molecular Pathology* 2012; **21:** 69-76.
54. Castellano I, Macri L, Deambrogio C *et al*. Reliability of Whole Sentinel Lymph Node Analysis by One-Step Nucleic Acid Amplification for Intraoperative Diagnosis of Breast Cancer Metastases. *Annals of Surgery* 2012; **255:** 334-342.
55. Choi YL, Ahn SK, Bae YK *et al*. One-step Nucleic Acid Amplification (OSNA): Intraoperative Rapid Molecular Diagnostic Method for the Detection of Sentinel Lymph Node Metastases in Breast Cancer Patients in Korean Cohort. *Journal of Breast Cancer* 2010; **13:** 366-374.
56. Feldman S, Krishnamurthy S, Gillanders W *et al*. A Novel Automated Assay for the Rapid Identification of Metastatic Breast Carcinoma in Sentinel Lymph Nodes. *Cancer* 2011; **117:** 2599-2607.
57. Godey F, Leveque J, Tas P *et al*. Sentinel lymph node analysis in breast cancer: contribution of one-step nucleic acid amplification (OSNA). *Breast Cancer Research & Treatment* 2012; **131:** 509-516.
58. Guillen-Paredes MP, Carrasco-Gonzalez L, Chaves-Benito A *et al*. One-step nucleic acid amplification (OSNA) assay for sentinel lymph node metastases as an alternative to conventional postoperative histology in breast cancer: A cost-benefit analysis. *Cirugia Espanola* 2011; **89:** 456-462.
59. Khaddage A, Berremila SA, Forest F *et al*. Implementation of molecular intra-operative assessment of sentinel lymph node in breast cancer. *Anticancer Research* 2011; **31:** 585-590.
60. Le Frere-Belda MA, Bats AS, Gillaizeau F *et al*. Diagnostic performance of one-step nucleic acid amplification for intraoperative sentinel node metastasis detection in breast cancer patients. *International Journal of Cancer* 2377; **130:** 2377-2386.
61. Osako T, Iwase T, Kimura K *et al*. Accurate staging of axillary lymph nodes from breast cancer patients using a novel molecular method. *British Journal of Cancer* 1197; **105:** 1197-1202.
62. Schem C, Maass N, Bauerschlag DO *et al*. One-step nucleic acid amplification-a molecular method for the detection of lymph node metastases in breast cancer patients; results of the German study group. *Virchows Archiv* 2009; **454:** 203-210.
63. Snook KL, Layer GT, Jackson PA *et al*. Multicentre evaluation of intraoperative molecular analysis of sentinel lymph nodes in breast carcinoma. *British Journal of Surgery* 2011; **98:** 527-535.
64. Tamaki Y, Akiyama F, Iwase T *et al*. Molecular Detection of Lymph Node Metastases in Breast Cancer Patients: Results of a Multicenter Trial Using the One-Step Nucleic Acid Amplification Assay. *Clinical Cancer Research* 2009; **15:** 2879-2884.
65. Tamaki Y, Sato N, Homma K *et al*. Routine clinical use of the one-step nucleic acid amplification assay for detection of sentinel lymph node metastases in breast cancer patients Results of a Multicenter Study in Japan. *Cancer* 2012; **118:** 3477-3483.
66. Tsujimoto M, Nakabayashi K, Yoshidome K *et al*. One-step nucleic acid amplification for intraoperative detection of lymph node metastasis in breast cancer patients. *Clinical Cancer Research* 2007; **13:** 4807-4816.

67. Visser M, Jiwa M, Horstman A *et al*. Intra-operative rapid diagnostic method based on CK19 mRNA expression for the detection of lymph node metastases in breast cancer. *International Journal of Cancer* 2008; **122:** 2562-2567.
68. Belda M, Barra CC, Crouet H *et al*. Intra-operative sentinel lymph node metastasis detection in breast cancer by "One-step Nucleic Acid Amplification (OSNA)" - results of the French multicentre prospective study. *Ejc Supplements* 2008; **6:** 54-54.
69. Chu PG, Weiss LM. Keratin expression in human tissues and neoplasms. *Clin Cancer Res* 2002; **13:** 4807-4816.
70. Lyman GH AG, Somerfield MR, Benson III AB, Bodurka HJ, Burstein H et al. American Society of Clinical Oncology Guideline Recommendations for Sentinel Lymph Node Biopsy in Early-Stage Breast Cancer. *J Clin Oncol* 2005; **23:** 7703-7720.
71. Reitsma JB GA, Rutjes AWS, Scholten RJPM, Bossuyt PM, Zwinderman AH. . Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. . *Journal of Clinical Epidemiology* 2005; **58:** 982-990.
72. Riley RD, Abrams KR, Sutton AJ, Lambert PC, Thompson JR. Bivariate random-effects meta-analysis and the estimation of between-study correlation. *BMC Medical Research Methodology* 2007; **7**.
73. Classe JM, Baffert S, Sigal-Zafrani B *et al*. Cost comparison of axillary sentinel lymph node detection and axillary lymphadenectomy in early breast cancer. A national study based on a prospective multi-institutional series of 985 patients 'on behalf of the Group of Surgeons from the French Unicancer Federation'. *Ann Oncol* 2012; **23:** 1170-1177.
74. Meng Y, Ward S, Cooper K, Harnan S, Wyld L. Cost-effectiveness of MRI and PET imaging for the evaluation of axillary lymph node metastases in early stage breast cancer. *Ejso-Eur J Surg Onc* 2011; **37:** 40-46.
75. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. *Brit Med J* 1996; **313:** 275-283.
76. Byron S. Personal communication, 2012.
77. Dauplat MM, Penault-Llorca F. Sentinel lymph node biopsy intraoperative evaluation: State of art and emerging molecular assays. *Medecine Nucleaire-Imagerie Fonctionnelle Et Metabolique* 2010; **34:** 23-26.
78. Organisation for Economic Co-operation and Development O. PPPs and exchange rates. [cited 2012 1/11]; Available from: http://stats.oecd.org/Index.aspx?datasetcode=SNA_TABLE4
79. Karnon J. Alternative decision modelling techniques for the evaluation of health care technologies: Markov processes versus discrete event simulation. *Health Econ* 2003; **12:** 837-848.
80. Sundaresan V. Personal communication, 2012.
81. Cody HS. Clinicopathologic factors associated with false-negative sentinel lymph-node biopsy in breast cancer. *Annals of Surgery* 2006; **244:** 324-324.
82. Mak SS, Mo KF, Suen JJS *et al*. Lymphedema and quality of life in Chinese women after treatment for breast cancer. *Eur J Oncol Nurs* 2009; **13:** 110-115.
83. (NICE) NIfHaCE. RD100 i OSNA Brief. 2012.
84. Health Do. NHS Reference Costs 2010-2011. 2010-2011 [cited 2012 5/11]; Available from: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_131140
85. Pandharipande PV, Harisinghani MG, Ozanne EM *et al*. Staging MR Lymphangiography of the Axilla for Early Breast Cancer: Cost-Effectiveness Analysis. *Am J Roentgenol* 2008; **191:** 1308-1319.
86. Ng VV, Charlton FI, Cunnick GH. The use of intra-operative one step nucleic acid amplification (OSNA) to analyse sentinel lymph nodes in patients with breast cancer-does it impact on operating times? *British Journal of Surgery* 2011; **98:** 106-106.
87. Jeruss JS, Hunt KK, Xing Y *et al*. Is intraoperative touch imprint cytology of sentinel lymph nodes in patients with breast cancer cost effective? *Cancer* 2006; **107:** 2328-2336.
88. Shemilt I, Thomas J, Morciano M. A web-based tool for adjusting costs to a specific target currency and price year. *Evid Policy* 2010; **6:** 51-59.
89. Burke M, Setters J. *Breast Lymph Node Assay*. 2010.

90. Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997; **35:** 1095-1108.
91. Tengs TO, Wallace A. One thousand health-related quality-of-life estimates. *Med Care* 2000; **38:** 583-637.
92. 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:** 568-576.
93. Ward S, Rees A, Wilkinson A *et al*. Taxanes for the adjuvant treatment of early breast cancer: Systematic review and economic evaluation. *Health Technol Asses* 2007; **11:** 1-+.
94. 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 Tr* 1997; **45:** 181-192.
95. Karnon J, Delea T, Barghout V. Cost utility analysis of early adjuvant letrozole or anastrozole versus tamoxifen in postmenopausal women with early invasive breast cancer: the UK perspective. *Eur J Health Econ* 2008; **9:** 171-183.
96. Health Do. NHS Reference Costs 2002-2003. 2002-2003 [cited 2012 12/11]; Available from: http://collections.europarchive.org/tna/20100509080731/http://dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4070195
97. Curtis L. *Unit Costs of Health & Social Care 2011*. Unit PSSR, 2011.
98. Ara R, Brazier JE. Populating an Economic Model with Health State Utility Values: Moving toward Better Practice. *Value Health* 2010; **13:** 509-518.
99. Asakawa K, Senthilselvan A, Feeny D, Johnson J, Rolfson D. Trajectories of health-related quality of life differ by age among adults: Results from an eight-year longitudinal study. *J Health Econ* 2012; **31:** 207-218.
100. Klingler S, Marchal F, Rauch P *et al*. Intraoperative detection of lymph node metastasis using one-step nucleic acid amplification (OSNA) in breast cancer patients: Effect on second surgery rate and delay for adjuvant therapy. *J Clin Oncol* 2012; **30:** abstr 10517.
101. Giuliano AE, McCall L, Beitsch P *et al*. Locoregional Recurrence After Sentinel Lymph Node Dissection With or Without Axillary Dissection in Patients With Sentinel Lymph Node Metastases. *Annals of Surgery* 2010; **252:** 426-433.
102. Cserni G. Intraoperative analysis of sentinel lymph nodes in breast cancer by one-step nucleic acid amplification. *J Clin Pathol* 2012; **65:** 193-199.
103. Mesker WE, Torrenga H, Sloos WCR *et al*. Supervised automated microscopy increases sensitivity and efficiency of detection of sentinel node micrometastases in patients with breast cancer. *J Clin Pathol* 2004; **57:** 960-964.
104. Weaver DL, Krag DN, Manna EA *et al*. Detection of occult sentinel lymph node micrometastases by immunohistochemistry in breast cancer - An NSABP protocol B-32 quality assurance study. *Cancer* 2006; **107:** 661-667.
105. Cserni G, Bianchi S, Vezzosi V *et al*. The value of cytokeratin immunohistochemistry in the evaluation of axillary sentinel lymph nodes in patients with lobular breast carcinoma. *J Clin Pathol* 2006; **59:** 518-522.