

Surveillance mammography for detecting ipsilateral breast tumour recurrence and metachronous contralateral breast cancer: a systematic review

Clare Robertson · Senthil Kumar Arcot Ragupathy ·
Charles Boachie · Cynthia Fraser · Steve D Heys ·
Graeme MacLennan · Graham Mowatt ·
Ruth E Thomas · Fiona J Gilbert ·
and the Mammographic Surveillance Health Technology
Assessment Group

Received: 14 April 2011 / Revised: 20 June 2011 / Accepted: 26 June 2011 / Published online: 11 August 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract

Objectives To determine the diagnostic accuracy of surveillance mammography for detecting ipsilateral breast tumour recurrence and metachronous contralateral breast cancer in women previously treated for primary breast cancer.

Methods A systematic review of surveillance mammography compared with ultrasound, magnetic resonance imaging (MRI), specialist-led clinical examination or unstructured primary care follow-up, using histopathological assessment for test positives and follow-up for test negatives as the reference standard.

Results Nine studies met our inclusion criteria. Variations in study comparisons precluded meta-analysis. For routine ipsilateral breast tumour detection, surveillance mammography sensitivity ranged from 64–67% and specificity ranged from 85–97%. For MRI, sensitivity ranged from 86–100% and specificity was 93%. For non-routine ipsilateral breast tumour detection, sensitivity and specificity for surveillance mammography ranged from 50–83% and 57–75% and for MRI 93–100% and 88–96%. For routine metachronous contralateral breast cancer detection, one study reported sensitivity of 67% and specificity of 50% for both surveillance mammography and MRI.

Conclusion Although mammography is associated with high sensitivity and specificity, MRI is the most accurate test for detecting ipsilateral breast tumour recurrence and metachronous contralateral breast cancer in women previously treated for primary breast cancer. Results should be interpreted with caution because of the limited evidence base.

Key Points

- Surveillance mammography is associated with high sensitivity and specificity
- Findings suggest that MRI is the most accurate test for detecting further breast cancer
- Robust conclusions cannot be made due to the limited evidence base
- Further research comparing surveillance mammography and other diagnostic tests is required

C. Robertson (✉) · C. Boachie · C. Fraser · G. MacLennan ·
G. Mowatt · R. E. Thomas
Health Services Research Unit, University of Aberdeen,
3rd Floor, Health Sciences Building Foresterhill,
Aberdeen AB25 2ZD, UK
e-mail: c.robertson@abdn.ac.uk

S. K. A. Ragupathy
Radiology Department, Aberdeen Royal Infirmary,
NHS Grampian,
Aberdeen, UK

S. D. Heys
Division of Applied Medicine, School of Medicine and Dentistry,
University of Aberdeen and Aberdeen Royal Infirmary,
NHS Grampian,
Aberdeen, UK

F. J. Gilbert
Aberdeen Biomedical Imaging Centre, University of Aberdeen,
and Aberdeen Royal Infirmary, NHS Grampian,
Aberdeen, UK

Keywords Mammography · Surveillance · Diagnostic accuracy · Neoplasm recurrence, local · Neoplasm, second primary

Abbreviations

| | |
|--------|---|
| MRI | magnetic resonance imaging |
| NIHR | National Institute for Health Research |
| HTA | Health Technology Assessment |
| CDSR | Cochrane Database of Systematic Reviews |
| DARE | Database of Abstracts of Reviews of Effects |
| WHO | World Health Organisation |
| NCI | National Cancer Institute |
| NRR | National Research Register |
| QUADAS | quality assessment of diagnostic accuracy studies |
| LR | likelihood ratio |
| DOR | diagnostic odds ratio |
| IC | incalculable |

Introduction

In women previously treated for breast cancer, surveillance mammography is useful for early detection of tumour recurrence, or for confirming the absence of recurrent cancer, and for the early detection of contralateral cancers. Although published figures vary, it has been estimated that approximately 50% of local recurrences in the breast following breast conservation surgery will be detected by mammography, with the remainder being detected by clinical examination or reported by the patient [1–4]. Recurrent tumours detected by mammography are generally smaller and less invasive than those found on clinical examination [2, 4]. Lu and colleagues [5] recently conducted a systematic review to determine the impact of early detection of isolated loco-regional and contralateral recurrence on survival. The authors reported better overall survival for recurrences detected by mammography or in asymptomatic patients, with an absolute reduction in mortality of 17–28% if all breast cancer recurrences are detected early.

While tumour recurrence displays similar mammographic features to the primary lesion [4], interpretation of the surveillance mammogram is made more difficult by surgical scarring and changes to breast density caused by primary treatment. For example, following surgery and/or radiotherapy, detectable abnormalities on mammography may include the presence of old haematoma, scar formation, fat necrosis, skin thickening, increased soft tissue density in the breast and microcalcifications. Approximately 10% of palpable tumours are not clearly visible on mammography and require additional imaging techniques for their demonstration. Surveillance mammography is therefore also associated with the possibility of false-positive results, which then require further investigations that are unnecessary and have a negative impact on a woman's quality of life.

We conducted a National Institute for Health Research Health Technology Assessment programme funded project (NIHR HTA Project 07/47/01) to examine the clinical effectiveness and cost-effectiveness of different surveillance mammography regimens after the treatment of primary breast cancer in the UK in primary and secondary care settings. The work comprised: a survey of UK breast surgeons and radiologists, a series of systematic reviews (test performance review presented here), and statistical and economic modelling to determine the effectiveness, cost consequences and cost utility of differing surveillance regimens.

The primary objective of this systematic review was to determine the test performance of surveillance mammography, alone or in combination with other tests, in detecting ipsilateral breast tumour recurrence and/or metachronous contralateral breast cancer in women undergoing routine surveillance. Our secondary objective was to compare surveillance mammography performance with alternative tests, alone or in combination, in women with a previous diagnostic test result indicating suspected ipsilateral breast tumour recurrence and/or metachronous contralateral breast cancer (referred to subsequently as non-routine surveillance).

Materials and methods

We developed and followed a structured protocol. We considered randomised controlled trials of surveillance mammography and diagnostic consecutive cohort studies of surveillance mammography or other comparator tests, involving women previously treated for primary breast cancer without detectable metastatic disease at the time of presentation for their initial treatment. We also considered indirect (between-study) comparisons by comparing cohort studies analysing results of at least 100 women who received surveillance mammography, or a comparator test, or a combination of tests, with the reference standard test in the same population. We excluded case reports and studies investigating technical aspects of a test. Comparator tests included ultrasound, magnetic resonance imaging (MRI), specialist-led clinical examination and unstructured primary care follow-up (defined as absence of formal routine secondary care follow-up, which may or may not involve mammography). The reference standard was histopathological assessment for test positives and a period of follow-up for test negatives.

We chose to include studies assessing test performance for routine and non-routine surveillance patients. Adjunct tests are part of breast cancer surveillance management and the performance of diagnostic tests used for this purpose is relevant to our population of interest. The accuracy of non-

routine adjunct imaging tests may differ from the accuracy of first-line surveillance tests as the test operator is primed to evaluate a suspicious finding in the non-routine surveillance patient. It is unclear what effect this has on test accuracy but it is likely to focus attention on a particular area of the breast and may conceivably increase the diagnostic test sensitivity. Consequently, we have not attempted to mix or compare the performance of tests used for these different purposes. Similarly, because of anatomical differences between a “treated” and an “untreated” breast (due to treatment effects) it was inappropriate to combine data on test performance for the detection of ipsilateral breast tumour recurrence and metachronous contralateral breast cancer.

The following types of outcome were considered:

- Test performance in diagnosing ipsilateral breast tumour recurrence in women undergoing routine surveillance
- Test performance in diagnosing ipsilateral breast tumour recurrence in women undergoing non-routine surveillance
- Test performance in diagnosing metachronous contralateral breast cancer in women undergoing routine surveillance
- Test performance in diagnosing metachronous contralateral breast cancer in women undergoing non-routine surveillance

To be considered for inclusion, the studies had to report the absolute numbers of true-positives, false-positives, false-negatives and true-negatives, or provide information allowing their calculation, and report a per-patient analysis.

In studies reporting the above outcomes, we planned to record the following additional outcomes, if stated:

- Adverse effects (defined as physical harms) of mammography and other tests
- Acceptability of the tests
- Reliability of the tests
- Radiological/operator expertise (who conducts the test and previous experience)
- Interpretability/readability of the tests

Major electronic databases were searched using sensitive search strategies to identify diagnostic studies of surveillance mammography, MRI, ultrasound or clinical follow-up. Searches were conducted from 1990 to March 2009 and were restricted to the English language. Conference abstracts were not included. The following databases were searched for primary studies: Medline, Medline In process, Embase, Biosis, Science Citation Index, Cancerlit, while Medion, the Cochrane Database of Systematic Reviews (CDSR), Database of Reviews of Effects (DARE) and the HTA Database were searched for reports of evidence syntheses. Reports of ongoing and recently completed trials

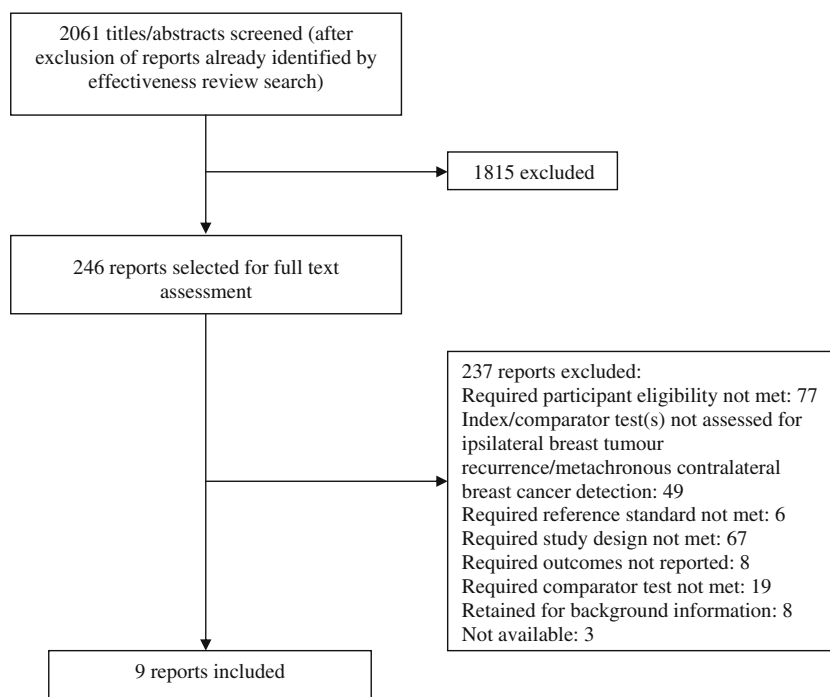
were sought from Current Controlled Trials, Clinical Trials, WHO International Clinical Trials Registry Platform, NCI Clinical Trials Database, NRR Archive and NIHR Portfolio Database. In addition, relevant websites were searched and the reference lists of all included studies were scanned for additional reports. Full details of the search strategies used are available from the authors or the full study report, currently in press (“The clinical effectiveness and cost-effectiveness of different surveillance mammography regimens after the treatment of primary breast cancer.” by Robertson et al. accepted for publication in *Health Technol Assess* 2011).

From an initial first screening round of titles and abstracts we were able to exclude reports that were clearly irrelevant to the review (e.g. did not include any of our considered diagnostic tests). We then assessed the full text versions of the remaining reports against our eligibility criteria using a screening tool comprising a checklist of our inclusion eligibility criteria, which we developed specifically for this review. One reviewer independently carried out data extraction. A second reviewer independently validated the data extraction. We calculated sensitivity, specificity, positive and negative likelihood ratios and diagnostic odds ratio for each included study.

We evaluated the quality of studies using an adapted version of the Quality Assessment of Diagnostic Accuracy Studies QUADAS tool [6]. Higher quality studies were defined as those considering a representative patient spectrum and judged to have successfully avoided partial verification bias (whether the whole or random sample of the population received reference standard verification), differential verification bias (whether patients received the same reference standard) and test review bias (whether index and reference standard test results were interpreted independently). Disagreement or uncertainty regarding data extraction or quality assessment was resolved by discussion or arbitration by a third reviewer.

Results

Figure 1 shows the flow of studies through the review. Nine studies met our inclusion criteria. Variation across the included studies precluded formal meta-analysis. We therefore present a narrative synthesis of the results. Overall, the nine studies enrolled 4002 participants. After exclusions, due to eligibility or participant drop-out, the studies included 3724 participants in their analyses. The earliest study took place in 1995 [7] and the latest in 2009 [8]. The earliest participant enrolment date given was 1992 [7] and the latest was 2003 [8]. Four studies did not give any indication of the enrolment time period [9–12]. One study took place in Sweden [7], two in the UK [10, 11], two in

Fig. 1 Flow of studies through the review process

Germany [12, 13], two in South Korea [8, 14], one in Italy [9] and one in France [15]. Across studies the ages of the participants ranged from 22–82 years [8]. Most participants were in their fifties. The median age was 53 years (interquartile range 50 to 56 years). Reported follow-up of test negatives ranged from 5 to 32 months. Table 1 provides details of the characteristics of the included studies.

Assessment of test performance

Test performance in diagnosing ipsilateral breast tumour recurrence

Table 2 shows test performance in detecting ipsilateral breast tumour recurrence in patients undergoing routine surveillance. The studies by Boné and colleagues [7] and Drew and colleagues [10] involved a total of 188 patients and reported the performance of surveillance mammography, MRI and clinical examination in routine surveillance patients. These studies reported sensitivities of 64% and 67%, and specificities of 97% and 85%, respectively, for surveillance mammography. For MRI, the studies reported sensitivities of 86% and 100% respectively, and for clinical examination 50% and 89%. Boné and colleagues [7] did not report specificity for MRI or clinical examination. The highest reported sensitivity was for MRI, and surveillance mammography combined with clinical examination (both 100%) while the highest specificity was for surveillance mammography (97%). Similarly, a high specificity of 93% was reported for MRI. The lowest reported sensitivity was

for clinical examination (50%) and the lowest specificity was for surveillance mammography combined with clinical examination (67%).

Table 3 shows test performance in detecting ipsilateral breast tumour recurrence in patients undergoing non-routine surveillance, as reported by Belli and colleagues [9], Mumtaz and colleagues [11], Rieber and colleagues [12], and Ternier and colleagues [15]. The studies involved a total of 156 patients. Across these studies, for surveillance mammography the median (range) sensitivity was 71% (50% to 83%) and specificity was 63% (57% to 75%). For MRI, the studies by Belli and colleagues [9], Mumtaz and colleagues [11] and Rieber and colleagues [12], involving a total of 193 patients, reported sensitivity of 93% and 100% (two studies) and a median (range) specificity of 94% (88% to 96%). Belli and colleagues [9] and Ternier and colleagues [15] reported the test performance of ultrasound, with sensitivities of 43% and 87%, and specificities of 31% and 73% respectively, and for clinical examination, sensitivities of 43% and 62%, and specificities of 56% and 49% respectively. The highest reported sensitivity (100%) and specificity (96%) was for MRI. The lowest reported sensitivities were (43%) for both ultrasound and clinical examination, while the lowest specificity was for ultrasound (31%).

Test performance in diagnosing metachronous contralateral breast cancer

Table 4 shows test performance in detecting metachronous contralateral breast cancer in routine surveillance patients.

Table 1 Summary of characteristics of the individual diagnostic accuracy studies

| Study ID | Study Design | Type of surveillance and primary surgery | Index tests | Comparator test | Follow-up time for verifying test negative results |
|-------------------|----------------------------|---|--|--|--|
| Belli 2002 [9] | Direct head-to-head cohort | Non-routine surveillance breast conservation patients | MRI for local recurrence | Surveillance mammography, clinical examination, ultrasound for local recurrence | MRI and clinical examination follow-up performed at 3 months. All MRI test negatives underwent cytological examination |
| Boné 1995 [7] | Direct head-to-head cohort | Routine surveillance mastectomy patients, all with breast reconstruction and implants | Surveillance mammography for local and contralateral recurrence | MRI, clinical examination for local and contralateral recurrence | Median 10 months (range 5–18 months) |
| Drew 1998 [10] | Direct head-to-head cohort | Routine surveillance breast conservation patients | MRI for local recurrence | Surveillance mammography, clinical examination, surveillance mammography + clinical examination for local recurrence | Median 341 days (range 168–451 days) |
| Kim 2009 [8] | Direct head-to-head cohort | Routine surveillance breast conservation and mastectomy patients | Adjunct ultrasound (surveillance mammography+ ultrasound) for contralateral recurrence | None | 1-2 years |
| Mumtaz 1997 [11] | Direct head-to-head cohort | Non-routine surveillance breast conservation patients | Surveillance mammography for local recurrence | MRI for local recurrence | Median 12 months (range 6–15 months) |
| Rieber 1997 [12] | Cohort | Non-routine surveillance breast conservation patients | MRI for local recurrence | None | clinical examination, ultrasound performed at 6 months follow-up. Surveillance mammography performed at 12 months' follow-up. In 22 patients a control MRI was performed at intervals of 2–16 months (mean 7.2 months) |
| Shin 2005 [14] | Cohort | Routine surveillance patients (primary surgery type not reported) | Ultrasound for local and contralateral recurrence | None | 6 months |
| Ternier 2006 [15] | Direct head-to-head cohort | Non-routine surveillance breast conservation patients | Surveillance mammography for local recurrence ² | Clinical examination, ultrasound for local recurrence | 6 months |
| Viehweg 2004 [13] | Direct head-to-head cohort | Routine surveillance breast conservation patients | MRI for contralateral recurrence | Conventional methods (surveillance mammography+clinical examination+ultrasound); MRI+Conventional methods for contralateral recurrence | 12 months |

1. *MRI* magnetic resonance imaging

2. Study authors considered computed tomography as the index test in this study but this test was not considered as an included comparator in this review

The studies by Boné and colleagues [7] and Viehweg and colleagues [13] involving a total of 202 patients, reported 67% and 91% sensitivity and 50% and 90% specificity, respectively, for MRI. Only individual studies reported the test performance of surveillance mammography, clinical

examination, and combinations of tests involving surveillance mammography. The highest reported sensitivity (100%) was for combined surveillance mammography, clinical examination, ultrasound and MRI [13], while the highest reported specificity (99%) was for combined

Table 2 Sensitivity, specificity, likelihood and diagnostic odds ratios for detecting ipsilateral breast tumour recurrence in routine surveillance patients

| Test | Study ID | Primary surgical treatment | Reported sensitivity% | Reported specificity% | LR+ | LR– | DOR (95% confidence interval) |
|--|---------------|----------------------------|-----------------------|-----------------------|------|-----------|-------------------------------|
| Surveillance mammography | Boné 1995[7] | Mastectomy | 64 | 97 | 22.2 | 0.4 | 60.3 (10.2–358.1) |
| | Drew 1998[10] | Breast conservation | 67 | 85 | 4.6 | 0.4 | 11.7 (2.6–52.4) |
| MRI | Boné 1995[7] | Mastectomy | 86 | Not reported | | | |
| | Drew 1998[10] | Breast conservation | 100 | 93 | 14.3 | <i>IC</i> | <i>IC</i> |
| Clinical examination | Boné 1995[7] | Mastectomy | 50 | Not reported | | | |
| | Drew 1998[10] | Breast conservation | 89 | 76 | 3.7 | 0.2 | 25.4 (3.0–213.9) |
| Combined surveillance mammography & clinical examination | Drew 1998[10] | Breast conservation | 100 | 67 | 3.0 | <i>IC</i> | <i>IC</i> |

MRI magnetic resonance imaging, *LR+* Likelihood ratio of a positive test result, *LR* Likelihood ratio of a negative test result, *DOR* Diagnostic odds ratio, *IC* Incalculable

surveillance mammography and ultrasound [8]. The lowest reported sensitivity (0%) was for clinical examination and the lowest specificity was for surveillance mammography, MRI and clinical examination (all 50%) [7].

None of the studies reported diagnostic accuracy of the included tests for diagnosing metachronous contralateral breast cancer in non-routine surveillance patients with a previous suspicious test result.

Test performance in diagnosing ipsilateral breast tumour recurrence and metachronous contralateral breast cancer

The study conducted by Shin and colleagues [14] was the sole study reporting overall test performance for diagnosing ipsilateral breast tumour recurrence and metachronous contralateral breast cancer. Shin and colleagues evaluated ultrasound in routine surveillance patients, reporting a sensitivity of 71% and a specificity of 98%, *LR+* 41.4, *LR* – 0.3, *OR* 138.25 (95% *CI* 61.26 to 312.04).

Quality assessment

None of the studies met all of our criteria specified for higher quality studies, although in five [8, 9, 11, 12, 15] this was due to lack of clarity as to whether reference standard results were interpreted without knowledge of index test results only. It was unclear in all but one study [14] whether the time interval between a positive test result and the histopathological reference standard was short enough to avoid improvement or progression of the condition occurring in the intervening period (disease progression bias). We are therefore uncertain of the effects of this type of bias for positive test results in these studies. All studies were judged to have appropriate follow-up time intervals for confirming negative test results and were therefore considered to be at low risk of disease progression bias for negative test results. It was unclear in the study conducted by Shin and colleagues [14], however, whether all patients with negative test results

Table 3 Sensitivity, specificity, likelihood ratio and diagnostic odds ratio for detecting ipsilateral breast tumour recurrence in non-routine surveillance patients

| Test | Study ID | Primary surgical treatment | Reported sensitivity% | Reported specificity% | LR + | LR – | DOR (95% confidence interval) |
|--------------------------|------------------|----------------------------|-----------------------|-----------------------|------|-----------|-------------------------------|
| Surveillance mammography | Belli 2002[9] | Breast conservation | 71 | 63 | 1.9 | 0.5 | 4.2 (2.6–52.4) |
| | Mumtaz 1997[11] | Breast conservation | 50 | 75 | 2.0 | 0.7 | 3 (0.6–14.0) |
| | Ternier 2006[15] | Breast conservation | 83 | 57 | 1.9 | 0.3 | 6.3 (2.5–15.6) |
| Ultrasound | Belli 2002[9] | Breast conservation | 43 | 31 | 0.6 | 1.8 | 0.3 (0.1–2.1) |
| | Ternier 2006[15] | Breast conservation | 87 | 73 | 3.2 | 0.2 | 17 (6.2–46.5) |
| MRI | Belli 2002[9] | Breast conservation | 100 | 94 | 16.0 | <i>IC</i> | <i>IC</i> |
| | Mumtaz 1997[11] | Breast conservation | 93 | 88 | 7.4 | 0.1 | 91 (7.4–1126.9) |
| | Rieber 1997[12] | Breast conservation | 100 | 96 | 24.2 | <i>IC</i> | <i>IC</i> |
| Clinical examination | Belli 2002[9] | Breast conservation | 43 | 56 | 1.0 | 1.0 | 1.0 (0.2–5.8) |
| | Ternier 2006[15] | Breast conservation | 62 | 49 | 1.2 | 0.8 | 1.5 (0.7–3.4) |

1. *MRI* magnetic resonance imaging, *LR+* Likelihood ratio of a positive test result, *LR* Likelihood ratio of a negative test result, *DOR* Diagnostic odds ratio, *IC* Incalculable

Table 4 Test performance as measured by sensitivity, specificity, likelihood ratio and diagnostic odds ratio for detecting metachronous contralateral breast cancer in routine surveillance patients

| Test | Study ID | Primary surgical treatment | Reported Sensitivity% | Reported Specificity% | LR + | LR – | DOR (95% confidence interval) |
|---|------------------|----------------------------|-----------------------|-----------------------|------|------|-------------------------------|
| Surveillance mammography | Boné 1995[7] | Mastectomy | 67 | 50 | 1.3 | 0.7 | 2.0 (0.1–78.2) |
| MRI | Boné 1995[7] | Mastectomy | 67 | 50 | 1.3 | 0.7 | 2.0 (0.1–78.2) |
| | Viehweg 2004[13] | Breast conservation | 91 | 90 | 9.4 | 0.1 | 93.1 (11.0–786.2) |
| Clinical examination | Boné 1995[7] | Mastectomy | 0 | 50 | | | |
| Combined surveillance mammography & ultrasound | Kim 2009[8] | Mastectomy | 95 | 99 | 61.5 | 0.05 | 1149.2 (148.0–8937.8) |
| | | Breast conservation | | | | | |
| Combined surveillance mammography, clinical examination & ultrasound | Viehweg 2004[13] | Breast conservation | 64 | 84 | 3.9 | 0.4 | 8.9 (2.4–33.0) |
| Combined surveillance mammography, clinical examination, ultrasound & MRI | Viehweg 2004[13] | Breast conservation | 100 | 89 | 8.9 | IC | IC |

MRI magnetic resonance imaging, LR+ Likelihood ratio of a positive test result, LR Likelihood ratio of a negative test result, DOR Diagnostic odds ratio, CI Confidence interval, IC Incalculable

received follow-up and so there is a possible risk of partial verification bias for this study. The study by Drew and colleagues [10] was considered vulnerable to partial verification bias as only those participants testing positively on MRI received reference standard verification.

Additional outcomes

None of the included studies reported data concerning adverse effects, acceptability, reliability, radiological/operator expertise and interpretability/readability of the tests. We found no discernible pattern for the histology of cancers detected and not detected both within and between diagnostic tests.

Discussion

Our findings should be interpreted with caution, as they are based on only nine studies involving a total of 3724 participants. Furthermore, the study conducted by Boné and colleagues [7] included only mastectomy patients who underwent breast reconstruction using implants. Surveillance of the chest wall and/or the reconstructed breast in patients receiving either mastectomy alone, or mastectomy with breast reconstruction and implants, varies according to different health care systems and local protocols [16–18]. These comprise an increasingly relevant sub-group of women, because of increasing rates of breast reconstruction procedures, who might receive routine surveillance mammography in the future [19]. Results from this study should be treated as distinct from the others owing to the highly selected patient population who, whilst representing a subset of our considered population, differ greatly from the wider spectrum of women who receive surveillance in practice.

Systematic reviews of diagnostic accuracy are highly complex and methodology in this area continues to evolve. In terms of strengths, we believe that the methods adopted for this review are scientifically rigorous and compatible with current guidance in this area. A limitation was that non-English language studies were excluded, potentially limiting the evidence base. Of the studies included here, few evaluated the performance of the considered tests for similar purposes. Furthermore, even where data were available it was not clinically appropriate to combine them, for example, because of differences between a “treated” and an “untreated” breast. Similarly, it was inappropriate to combine data from routine and non-routine surveillance patients. Furthermore, no data were reported by the included studies on other aspects such as adverse effects or acceptability of the tests.

Results for the index and comparator tests evaluated in this review were ascertained by subjective operator interpretation, either by visual inspection of an image of the breast (surveillance mammography, ultrasound and MRI) or by clinical examination of the breast. Data on the level of operator expertise or intra/inter-rater reliability were not reported. It is therefore unclear whether these factors had any influence on reported test accuracy within, and between, studies and therefore whether any potential test operator bias exists.

None of the studies met all of our criteria specified for higher quality studies, although they were judged to have reasonable internal validity. All but one study [8] were considered to include a representative sample and therefore have good external validity.

Conclusion

Our findings suggest that MRI can be considered to have higher diagnostic value than surveillance mammography in women previously treated for primary breast cancer. Of the

test combinations reported, surveillance mammography combined with breast ultrasound could be considered the most accurate combination of tests for detecting metachronous contralateral breast cancer. However these results should be interpreted with caution owing to the paucity of data for all diagnostic tests available for breast cancer surveillance. Further evidence on surveillance mammography and other diagnostic tests in this group of women is required in order to make a robust and informed judgement on their relative performance. Ideally a definitive randomised controlled trial should be undertaken focusing on those women at higher risk of ipsilateral breast tumour recurrence or metachronous contralateral breast cancer. Such a trial might also compare more sophisticated surveillance regimens that vary not only in terms of the frequency of mammography or other diagnostic tests but also in terms of the frequency and setting of clinical follow-up. Alternatively, high-quality, direct head-to-head studies could be undertaken comparing the diagnostic accuracy of tests used in the surveillance population.

Acknowledgement The authors would like to acknowledge Jennifer Hislop and Shihua Zhu for their assistance with screening of full text papers, data extraction and quality assessment of the studies included in the review.

The Mammographic Surveillance Health Technology Assessment Group includes: Jennifer Burr, Mike Dixon, Wilma Jack, Gill Kerr, Gill Lawrence, Graeme MacLennan, Sarah Pinder, Laura Ternent, Luke Våle and Robin Wilson.

This project is funded by the NHS R&D Programme Health Technology Assessment Programme (HTA Project 07/47/01). The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the Department of Health.

The Health Services Research Unit is funded by the Chief Scientist Office of the Scottish Government Health Directorates. The authors accept full responsibility for this article.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Cancer Research UK (2009) Breast cancer—survival statistics. Cancer Research UK, London. Available via <http://info.cancerresearchuk.org/cancerstats/types/breast/survival/>. Accessed December 2010
2. Fowble B, Solin LJ, Schultz DL et al (1990) Breast recurrence following conservative surgery and radiation: patterns of failure, prognosis, and pathologic findings from mastectomy specimens with implications for treatment. *Int J Radiat Oncol Biol Phys* 19:833–842
3. Joseph E, Hyacinthe M, Lyman GH et al (1998) Evaluation of an intensive strategy for follow-up and surveillance of primary breast cancer. *Ann Surg Oncol* 5:522–528
4. Orel SG, Troupin RH, Patterson EA et al (1992) Breast cancer recurrence after lumpectomy and irradiation: role of mammography in detection. *Radiology* 183:201–206
5. Lu WL (2009) Impact on survival of early detection of isolated breast recurrences after the primary treatment for breast cancer: A meta-analysis. *Breast Cancer Res Treat* 114:403–412
6. Whiting P, Rutjes AW, Reitsma JB et al (2003) The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 3:25
7. Bone B, Aspelin P, Isberg B et al (1995) Contrast-enhanced MR imaging of the breast in patients with breast implants after cancer surgery. *Acta Radiol* 36:111–116
8. Kim MJ, Kim EK, Kwak JY et al (2009) Sonographic surveillance for the detection of contralateral metachronous breast cancer in an Asian population. *AJR Am J Roentgenol* 192:221–228
9. Belli P, Pastore G, Romani M et al (2002) Role of magnetic resonance imaging in the diagnosis of recurrence after breast conserving therapy. *Rays* 27:241–257
10. Drew PJ, Kerin MJ, Turnbull LW et al (1998) Routine screening for local recurrence following breast-conserving therapy for cancer with dynamic contrast-enhanced magnetic resonance imaging of the breast. *Ann Surg Oncol* 5:265–270
11. Mumtaz H, Davidson T, Hall-Craggs MA et al (1997) Comparison of magnetic resonance imaging and conventional triple assessment in locally recurrent breast cancer. *Br J Surg* 84:1147–1151
12. Rieber A, Merkle E, Zeitler H et al (1997) Value of MR mammography in the detection and exclusion of recurrent breast carcinoma. *J Comput Assist Tomogr* 21:780–784
13. Viehweg P, Rotter K, Laniado M et al (2004) MR imaging of the contralateral breast in patients after breast-conserving therapy. *Eur Radiol* 14:402–408
14. Shin JH, Han BK, Choe YH et al (2005) Ultrasonographic detection of occult cancer in patients after surgical therapy for breast cancer. *J Ultrasound Med* 24:643–649
15. Ternier F, Houvenaeghel G, Lecrivain F et al (2006) Computed tomography in suspected local breast cancer recurrence. *Breast Cancer Res Treat* 100:247–254
16. National Institute for Health and Clinical Excellence (2009) NICE CG80 Early and locally advanced breast cancer: full guideline. National Institute for Health and Clinical Excellence, London. Available via <http://guidance.nice.org.uk/CG80/Guidance/pdf/English>. Accessed December 2010.
17. National Cancer Institute (2009) Breast cancer screening: Q&A . National Cancer Institute, Bethesda, MA. Available via <http://www.cancer.gov/newscenter/qa/2009/breastscreenqa>. Accessed December 2010
18. Zakhireh J, Fowble B, Esserman LJ (2008) Application of screening principles to the reconstructed breast. *J Clin Oncol* 28:173–180
19. Barnsley GP, Grunfeld E, Coyle D et al (2007) Surveillance mammography following the treatment of primary breast cancer with breast reconstruction: a systematic review. *Plast Reconstr Surg* 120:1125–1132