EUROPEAN JOURNAL OF CANCER 43 (2007) 1905–1917

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

Review

A systematic review of the effectiveness of magnetic resonance imaging (MRI) as an addition to mammography and ultrasound in screening young women at high risk of breast cancer

S.J. Lord[a,*], W. Lei[a], P. Craft[b,c], J.N. Cawson[d], I. Morris[e], S. Walleser[a], A. Griffiths[a], S. Parker[a], N. Houssami[f]

[a]National Health and Medical Research Council Clinical Trials Centre, The University of Sydney, Level 5, Building F, 88 Mallett Street, Locked Bag 77, Camperdown, New South Wales 2050, Australia
[b]Medical Oncology Unit, The Canberra Hospital, Yamba Drive, Woden, Australian Capital Territory 2605, Australia
[c]Academic Unit of Internal Medicine, Australian National University Medical School, Frank Fenner Building 42, Canberra Australian Capital Territory 0200, Australia
[d]St Vincents Hospital, University of Melbourne, 41 Victoria Parade, Fitzroy, Melbourne, Victoria 3065, Australia
[e]Department of Diagnostic Imaging, Princess Margaret Hospital for Children and King Edward Memorial Hospital for Women, Subiaco, Western Australia 6008, Australia
[f]Screening and Test Evaluation Program (STEP), School of Public Health, The University of Sydney, A27, Sydney, New South Wales 2006, Australia

ARTICLE INFO

Article history:
Received 1 May 2007
Received in revised form 9 June 2007
Accepted 20 June 2007
Available online 2 August 2007

Keywords:
Magnetic resonance imaging
Breast neoplasms
Mass screening
Sensitivity and specificity
BRCA1 protein
BRCA2 protein

ABSTRACT

Breast magnetic resonance imaging (MRI) has been proposed as an additional screening test for young women at high risk of breast cancer in whom mammography alone has poor sensitivity. We conducted a systematic review to assess the effectiveness of adding MRI to mammography with or without breast ultrasound and clinical breast examination (CBE) in screening this population. We found consistent evidence in 5 studies that adding MRI provides a highly sensitive screening strategy (sensitivity range: 93–100%) compared to mammography alone (25–59%) or mammography plus ultrasound +/– CBE (49–67%). Meta-analysis of the three studies that compared MRI plus mammography versus mammography alone showed the sensitivity of MRI plus mammography as 94% (95%CI 86–98%) and the incremental sensitivity of MRI as 58% (95%CI 47–70%). Incremental sensitivity of MRI was lower when added to mammography plus ultrasound (44%, 95%CI 27–61%) or to the combination of mammography, ultrasound plus CBE (31–33%). Estimates of screening specificity with MRI were less consistent but suggested a 3–5-fold higher risk of patient recall for investigation of false positive results. No studies assessed as to whether adding MRI reduces patient mortality, interval or advanced breast cancer rates, and we did not find strong evidence that MRI leads to the detection of earlier stage disease. Conclusions about the effectiveness of MRI therefore depend on assumptions about the benefits of early detection from trials of mammographic screening in older average risk populations. The extent

* *Corresponding author:* Tel.: +61 2 9563 5322; fax: +61 2 9565 1863.
E-mail address: slord@ctc.usyd.edu.au (S.J. Lord).
0959-8049/$ - see front matter
doi:10.1016/j.ejca.2007.06.007

to which high risk younger women receive the same benefits from early detection and treatment of MRI-detected cancers has not yet been established.

1. Introduction

Women with a strong family history of breast cancer or a predisposing gene mutation such as BRCA1 or BRCA2 have a cumulative lifetime risk of developing breast cancer of 21–65%[1,2] with a substantial proportion of these cancers diagnosed before the age of 50 years. Prophylactic mastectomy, oophorectomy and chemoprevention with a selective oestrogen receptor modulator such as tamoxifen can reduce this risk but are associated with adverse events and may be unacceptable options for some women.[3–5] Surveillance of high risk women with breast imaging is recommended as a secondary preventive measure on the assumption that early diagnosis and treatment will confer similar benefits of reduced breast cancer mortality in this population as reported from randomised controlled trials of population screening programs using mammography in average risk women.[6–8]

Although mammography is a reasonably sensitive test for screening postmenopausal women, it is less sensitive in younger women and those with a genetic predisposition to breast cancer.[9] This has been attributed to increased mammographic density in premenopausal women which can obscure the radiological features of early breast cancer and the faster growth of breast cancers in these populations. Further, it has been suggested that cancers associated with BRCA mutations, in particular BRCA1, are more likely to have a benign appearance on mammography.[10,11]

Emerging evidence that magnetic resonance imaging (MRI) is more sensitive than mammography for the detection of breast cancer, supported by expert opinion about the potential associated patient benefits have led to recent guidelines recommending its use for the surveillance of some high risk women in the United Kingdom (UK) and United States (US),[12,13] while some countries are contemplating whether there is sufficient evidence to implement such recommendations. Given this ongoing uncertainty about the trade-offs between potential benefits and harms, a review of the evidence on this issue is timely to guide screening decisions. Few accuracy studies and no existing systematic reviews have specifically investigated the incremental value of MRI as an addition to conventional imaging with mammography and ultrasound in young high risk women. We present a systematic review on the incremental accuracy of breast MRI as an addition to annual mammography with or without breast ultrasound and clinical breast examination (CBE) in screening women under 50 years, provide pooled estimates of screening accuracy where appropriate, and evaluate evidence relating to early detection and outcomes in this population. This work was undertaken as part of a broader review conducted for the Australian Medical Services Advisory Committee to assist with decisions about public funding of screening MRI for this group of women.

2. Materials and methods

2.1. Search strategy

Electronic databases (MEDLINE, Pre-Medline, EMBASE, the Cochrane Library) and websites of health technology assessment agencies were searched to identify relevant studies published in English between 1966 and March 2007 using Medical Subject Heading terms and text words for MRI and breast cancer. A total of 3367 citations were identified. Two independent reviewers scanned citation abstracts to identify studies that compared the addition of MRI to mammography with or without ultrasound and CBE in asymptomatic high-risk women, and reported sufficient data for calculation of the incremental sensitivity (probability of a positive test in women with breast cancer) and specificity (probability of a negative test in women without breast cancer) of MRI using histology as the reference standard for positive tests and a consensus of all tests as the minimum reference standard for negative tests using data tables (see data analysis). Comparative studies reporting on prognostic tumour characteristics (tumour size and/or axillary lymph node involvement) of invasive cancers detected with the addition of MRI versus conventional imaging alone, interval cancer rates, or relevant patient outcomes were also eligible for inclusion.

Five accuracy studies were eligible for review (see Appendix for flowchart of search strategy).[14–18] Four systematic reviews investigating MRI for breast cancer screening were identified: three were conducted prior to the publication of one or more eligible accuracy studies[12,19,20]; the fourth included all eligible studies but did not specifically investigate the incremental value of MRI versus conventional imaging.[13] The present systematic review appraises evidence published up to March 2007 to assess the incremental accuracy of MRI, and compare prognostic characteristics of cancers detected by the addition of MRI versus those detected by conventional testing alone.

2.2. Quality appraisal

Two independent reviewers assessed the quality of included studies using the QUADAS tool.[21] Studies were classified as high quality if: they were conducted prospectively using well-defined selection criteria and recruited consecutive eligible subjects; reported on the execution of study tests and test threshold for a positive test in sufficient detail to allow test replication; applied the same reference standard to validate the results of study tests; interpreted test results without the knowledge of the reference standard or comparator tests; conducted study tests within two weeks; reported indeterminate test results; and explained study withdrawals. Studies not conducted prospectively or not meeting the criteria for an adequate reference standard or test interval were classified as low quality, other studies were classified as fair quality.

2.3. Data analysis

The characteristics of the study population (subject age and breast cancer risk), MRI magnet strength, comparator tests and reference standard at the first screening round were extracted from each study to assess applicability and variation between studies. Cancer detection rates at the first and subsequent screening rounds, number of additional cancers detected (yield), false positive patient recall and benign biopsy rates were also extracted where available. Data extraction was performed by one reviewer and checked by a second reviewer.

Two by two tables were reconstructed and per-patient sensitivity, specificity and associated 95% confidence intervals (CI) for test strategies with and without MRI were calculated using Meta-DiSc software.[22] Differences in sensitivities and specificities between test strategies were calculated. Where appropriate, summary estimates and 95%CI of sensitivity and specificity for each test strategy, and the differences between strategies, were obtained using the random-effects method of DerSimonian-Laird. Heterogeneity between studies was assessed separately for estimates of sensitivity and specificity, and differences between tests in sensitivity and specificity, using Chi-squared tests.[23] Relative risk and absolute risk differences of false positive recall and benign biopsy rates (and 95%CI) were calculated for the addition of MRI versus conventional testing alone.

3. Results

3.1. Study characteristics

There were no randomised controlled trials (RCTs) of MRI in breast screening. Five studies compared the incremental sensitivity and specificity of MRI as an additional test to the combination of mammography, ultrasound and CBE[17,18] and/or mammography alone.[14–16] Four studies reported on tumour size and/or axillary lymph node spread of invasive cancers detected by each test strategy.[14,15,17,18] We did not identify any studies that compared patient outcomes, reported mortality as an outcome or interval cancer rates in high risk women screened with and without MRI.

Subjects participated in an average of one[16,17] to five[14] screening rounds with some subjects followed up with annual screening to 7 years[14,15] providing data from a total of 4,534 MRI scans in 2059 women. Criteria for eligibility varied across studies (Table 1). Mean/median age ranged from 40[14,15] to 47 years,[18] although only one study restricted enrolment to women $\leqslant$50 years.[15] Risk criteria used to select subjects ranged from estimation of $\geqslant$20% cumulative lifetime risk[14] to known BRCA1 or BRCA2 mutation carriers (estimated 65% cumulative lifetime risk).[18] One study excluded participants with a prior history of breast cancer.[15] The rate of breast cancers detected in the first screening round (prevalence) for each study ranged from 1% to 6%.

Studies were performed using MRI equipment with a field strength of 1.0–1.5[15,17] or 1.5 Tesla,[14,18] a dedicated breast coil and intravenous gadolinium contrast. Each study only reported data for the subset of eligible subjects who successfully received MRI and comparator tests. Four studies reported between 6% and 21% of enrolled subjects were subsequently excluded or lost to follow up.[14–16,18] Reasons for subject exclusions included: test/s not performed according to study protocol; subject ineligible due to negative gene test; bilateral mastectomy performed; diagnosis of primary or metastatic breast cancer; subject no longer wished to participate; or lost to follow up. Two studies also reported MRI scanning failed or was contraindicated in 1–3% of enrolled subjects due to: subject too large for MRI machine; claustrophobia; discomfort; pregnancy; or other contra-indication.[15,18]

3.2. Study appraisal

All studies were conducted prospectively, but did not report as to whether a consecutive or selected sample of eligible subjects was recruited. Positive test results were validated by histology, but studies reported different thresholds for classifying positive tests and different protocols for the assessment of MRI-indeterminate findings and validation of discordant MRI-positive, mammography-negative findings (Table 1). Three studies validated negative tests with follow-up at 6 months or longer and classified interval cancers as false negatives.[14,15,18] One study reported that tests were performed on the same day,[18] two studies reported tests were performed 'preferably' on the same day (one study achieved this in 92% of cases),[15,17] whereas two studies reported testing was performed within an interval of 8 weeks [14] or 90 days.[16]

3.3. Test accuracy

Studies estimated the sensitivity of MRI plus conventional tests at between 86% and 100% compared to 25–59% for mammography alone and 49–67% for combinations of mammography and ultrasound with or without CBE (Table 2). Meta-analysis of data from the three studies that compared the sensitivity and specificity of MRI plus mammography versus mammography alone showed the sensitivity of MRI plus mammography as 94% (95%CI 86–98%) and the incremental sensitivity of MRI as 58% (95%CI 47–70%, test for heterogeneity $P = 0.84$).[14–16] The incremental sensitivity of MRI was lower when added to mammography plus ultrasound (44%, 95%CI 27–61%) or to the combination of mammography, ultrasound plus CBE (31–33%, Table 2).[14,17,18]

Estimates of the specificity of MRI plus conventional testing varied widely across studies (range 77–96%) and precluded meta-analysis to estimate the 'true' relative specificity of screening strategies versus without MRI (Table 2). Reporting on the MARIBS study, Leach and colleagues found adding MRI resulted in a statistically significant 17% absolute reduction in the specificity of screening .[15] Three other studies reported smaller 1–7% absolute reductions in screening specificity.[14,16,17] Warner and colleagues did not report on specificity for the combination of MRI plus conventional tests.[18]

Two studies reported data from women screened annually over a period of at least two years and up to 7 years.[14,15] Both studies observed a higher detection rate of breast cancer at the initial screening round (prevalent cancers) than subsequent screening rounds (incident cancers). Only one study reported the proportion of study subjects who had participated in prior screening.[18] In this study of BRCA1 and 2 mutation carriers,

Table 1 – Characteristics and appraisal of eligible studies investigating the accuracy of breast MRI as an addition to screening mammography with(out) ultrasound and CBE for the surveillance of women at high risk of breast cancer

Author, year, setting, recruitment period	N[a] (tests)[b]	Index test and comparator(s)	Study population	Reference standard and classification of a positive MRI	Test interval and interpretation
Kuhl et al., 2005 Germany, Single centre 1996–2001	529 (1452)	*Index test:* 1.5 Tesla MRI *Test comparisons: conventional tests +/– MRI* • M + US • M	Median age 40 years (range 27–59 years) Prior history of breast cancer: 26% Female: 98.8% (1 male participant)	*Positive test*: Histopathology *Negative test*: 6-monthly follow-up with CBE and US. Interval cancers classified as false negatives *Indeterminate test:* (BI-RADS 3), 6-month follow-up	Test interval <8 weeks Tests interpreted with patient history and CBE, blinded to comparator test results
		Follow-up time mean 5.3 years, range 2–7 years	*Risk classification* • 20% risk: 21% • 21–40% risk: 46% • BRCA1/BRCA2 mutation carrier: 8%	*Classification of a positive MRI:* BI-RADS 4-5	
Leach et al., 2005 (magnetic resonance imaging breast screening (MARIBS)) United Kingdom 22 sites 1997–2004	649 (1881)	*Index test:* 1.5 Tesla MRI *Test comparisons: conventional test +/– MRI* • M	Median age 40 years (range 31–55 years) Prior history of breast cancer: 0% Female: 100%	*Positive test*: Histopathology performed based on combination of all tests *Negative test*: 12-month follow-up. Interval cancers classified as false negatives	Test interval not specified but preferably 'same day' Tests interpreted blinded to comparator test results, CBE not performed, double reporting of all tests
		Follow-up time range 1–7 years	*Risk classification* • Annual risk $\geqslant$0.9% or 1st degree relative with 60% chance BRCA1/2 mutation carrier: 82% • BRCA1/BRCA2 mutation carrier: 18%	*Classification of a positive MRI:* BI-RADS 0, 3-5	
Lehman et al., 2005 (International Breast MRI Consortium Working Group (IBMC)) USA, Canada 13 sites 1999–2002	367 (367)	*Index test:* MRI, Tesla NR *Test comparisons: conventional test +/– MRI*[c] • M	Mean age 45 years, standard deviation 9.7 years (range NR) Prior history of breast cancer: 10% Female: 100%	*Positive test*: Histopathology, except in 9 cases where MRI/M discordant results *Negative test*: Based on combination of tests (film review for discordant results). No interval cancers reported	Test interval $\leqslant$90 days Tests interpreted blinded to comparator test results
		Follow-up time no long-term follow-up reported	*Risk classification* • Cumulative risk >25%: 90% • Prior history of breast cancer: 10%	*Classification of a positive MRI:* BI-RADS 4-5	

Sardanelli et al., 2006 Italy 17 sites 2000–2003 (interim analysis)	278 (377)	*Index test:* 1.0–1.5 Tesla MRI *Test comparisons: conventional tests +/– MRI*[c] • M + US + CBE • M + CBE • M *Follow-up time* NR (inferred range 1–2 years)	Mean age 46 years (range 25–79 years) Prior history of breast cancer: 39% Female: 100% *Risk classification* • BRCA1/BRCA2 mutation carrier 60% • M • 1st degree relative BRCA1/ BRCA2 3% • strong family history of breast and/or ovarian cancer 37%	*Positive test*: Histopathology *Negative test*: 12-month follow-up. No interval cancers reported *Indeterminate test:* (BI -RADS 3), 4-month follow-up *Classification of a positive MRI:* BIRADS 4-5	Test interval ⩽2 weeks (96%), 2–8 weeks 4% Tests interpreted blinded to comparator test results
Warner et al., 2004 Canada Single centre 1997–2003	236 (457)	*Index test:* 1.0–1.5 Tesla MRI *Test comparisons: conventional tests +/– MRI*[c] • M + US + CBE • M + CBE • M *Follow-up time* median 2 years, range 1–3 years	Median age: 47 years (range 25–65 years) Prior history of breast cancer: 30% Female: 100% *Risk classification* • BRCA1/2 mutation carriers 100%	*Positive test*: Histopathology performed for positive tests, except if MRI BI-RADS 3-5 discordant with other tests, then MRI repeated in 1 month. Tests resulting in benign biopsy classified as false positives, but not tests resulting in additional tests without biopsy *Negative test*: based on combination of tests. Interval cancer classified as false negative after film review *Indeterminate test*: (BI-RADS 3), 6, 12 and 24 month follow-up *Classification of a positive MRI:* BIRADS 4-5	Test interval <2 weeks Tests interpreted blinded to comparator test results and CBE 10/32 biopsies (31%) could not be performed under US guidance

CBE = clinical breast examination; M = mammography; MRI = magnetic resonance imaging; US = ultrasound; BI-RADS= The American College of Radiology Breast Imaging Reporting and Data System: BI-RADS 0 = needs further work-up; 1 = negative; BI-RADS 2 = benign finding; BI-RADS 3 = probably benign finding; BI-RADS 4 = suspicious abnormality; BI-RADS 5 = highly suggestive of malignancy; NR = not reported.

a Number of asymptomatic patients undergoing MRI and mammography.

b Total number of eligible screening test rounds.

c Data for comparison inferred from data presented in published article.

Table 2 – Incremental accuracy, cancer yield and risk of false positive from the addition of MRI to conventional breast cancer screening strategies for high risk women

Study author, year	*N* (*tests*)	Total breast cancers[a] Detection rate year 1	Screening strategy sensitivity and specificity (95%CI)		Incremental cancer yield and test sensitivity (95%CI) using MRI	Relative risk and absolute risk difference (95%CI) of false positive patient recall and benign biopsies with and without MRI
			With MRI	Without MRI		
Kuhl et al., 2005	529 (1452)	43 (41 patients)	*Conventional testing = mammography + ultrasound*			
			Sensitivity 93% (81–99%)	Sensitivity 49% (33–65%)	Incremental yield 19/1452	False positive patient recall rate NR
		14/529 (3%)	Specificity NR	Specificity 89% (87–91%)	13.1 additional cancers per 1000 screening rounds	Benign biopsy rate NR
					Incremental sensitivity 44% (27–61%)	
			Conventional testing = mammography			
			Sensitivity 93% (81–99%)	Sensitivity 33% (19–49%)	Incremental yield 26/1452	False positive patient recall rate NR[b]
			Specificity 96% (95–97%)	Specificity 97% (96–98%)	17.9 additional cancers per 1000 screening rounds	Benign percutaneous biopsy (use of surgical biopsy NR) RR: 1.22 (0.83–1.80)
					Incremental sensitivity 60% (45–76%)	RD: 7 (6–20) additional benign percutaneous biopsies per 1000 screening rounds
Leach et al., 2005	649 (1881)	35	*Conventional testing = mammography*			
		20/649 (3%)[c]	Sensitivity 94% (81–99%)	Sensitivity 40% (24–58%)	Incremental yield 19/1881	False positive patient recall
			Specificity 77% (75–79%)	Specificity 93% (92–95%)	10.1 additional cancers per 1000 screening rounds	RR: 3.43 (2.59–4.54)
						RD: 78 (61–94) additional false positive recalls per 1000 screening rounds
					Incremental sensitivity 54% (36–72%)	Benign percutaneous biopsy
						RR: 3.15 (2.04–4.88) RD: 30 (19–40) additional benign biopsies per 1000 screening rounds Benign surgical biopsy RR: 2.0 (0.5–8.0) RD: 2 (1.5–5) additional benign surgical biopsies per 1000 screening rounds
Lehman et al., 2005	367 (367)	4	*Conventional testing = mammography*			
		4/367 (1%)	Sensitivity 100% (40–100%)	Sensitivity 25% (0.6–81%)	Incremental yield 3/367	False positive patient recall
			Specificity 91% (87–94%)	Specificity 98% (96–99%)	8.2 additional cancers per 1000 screening rounds	RR: 4.86 (2.18–10.82)

					Incremental sensitivity 75% (33–100%)	RD: 74 (41–106) additional recalls per 1000 screening rounds Benign percutaneous biopsy RR: 9.50 (2.23–40.49) RD: 46 (22–70) additional biopsies per 1000 screening rounds
Sardanelli et al., 2006	278 (377)	18 11/278 (4%)	*Conventional testing = mammography + ultrasound + CBE*			
			Sensitivity 100% (82–100%)	Sensitivity 67% (41–87%)	Incremental yield 6/377	False positive patient recall rate NR
			Specificity 96% (94–98%)	Specificity NR	15.9 additional cancers per 1000 screening rounds Incremental sensitivity 33% (11–55%)	Benign biopsy rate NR
			Conventional testing = mammography + CBE			
			Sensitivity 100% (82–100%)	Sensitivity 67% (41–87%)	Incremental yield 5/377	False positive patient recall rate NR
			Specificity NR	Specificity NR	13.3 additional cancers per 1000 screening rounds Incremental sensitivity 28% (4–52%)	Benign biopsy rate NR
			Conventional testing = mammography			
			Sensitivity 100% (82–100%)	Sensitivity 59% (33–82%)	Incremental yield 7/377	False positive patient recall rate NR
			Specificity NR	Specificity 99% (98–100%)	18.6 additional cancers per 1000 screening rounds Incremental sensitivity 36% (10–61%)	Benign biopsy rate NR
Warner et al., 2004	236 (457)	22 (21 patients) 13/236 (6%)	*Conventional testing = mammography + ultrasound + CBE*			
			Sensitivity 95% (77–100%)	Sensitivity 64% (41–83%)	Incremental yield 7/457	False positive patient recall rate NR Benign biopsy rate NR
			Specificity NR	Specificity NR	15.3 additional cancers per 1000 screening rounds Incremental sensitivity 31% (10–54%)	
			Conventional testing = mammography + CBE			
			Sensitivity 86% (65–97%)	Sensitivity 45% (24–68%)	Incremental yield 9/457	False positive patient recall rate NR
			Specificity NR	Specificity NR	19.7 additional cancers per 1000 screening rounds	Benign biopsy rate NR

(*continued on next page*)

Table 2 – continued

Study author, year	*N* (tests)	Total breast cancers[a] Detection rate year 1	Screening strategy sensitivity and specificity (95%CI)		Incremental cancer yield and test sensitivity (95%CI) using MRI	Relative risk and absolute risk difference (95%CI) of false positive patient recall and benign biopsies with and without MRI
			With MRI	Without MRI		
			Conventional testing=mammography			
			Sensitivity 86% (65–97%)	Sensitivity 36% (17–59%)	Incremental sensitivity 41% (16–66%)	
			Specificity NR	Specificity 99.8% (99–100%)	Incremental yield 11/457	False positive patient recall rate NR
					24.1 additional cancers per 1000 screening rounds	Benign biopsy rate NR
					Incremental sensitivity 50% (25–75%)	

CBE = clinical breast examination; M = mammography; MRI = magnetic resonance imaging; NR = not reported; RD = absolute risk difference; RR = relative risk; US = ultrasound.

a Total cancers detected including invasive and ductal cancer *in situ*.

b False positive recall rate for MRI + M versus M not reported, but no statistically significant difference in patient recall rates for short-term follow-up between mammography (9.5%) versus MRI alone (11.5%).

c 20 cancers reported in text, trial profile figure reports 17 cancers at year 1 and 1 interval cancer.

Warner and colleagues reported 87% of subjects had prior mammography screening during the last 15 months.

MRI and mammography performance was maintained at first and subsequent screening rounds in the MARIBS study.[15] Warner and colleagues reported a decrease in MRI-initiated benign biopsy rate in each subsequent screening round and a decrease in MRI sensitivity over each screening round, however participation in screening dropped to 58% and 36% at the second and third screening rounds respectively, and these differences in test sensitivity and specificity were not statistically significant.[18]

3.4. *Additional cancer yield*

The additional cancer yield of MRI in women with negative findings on conventional testing ranged from 10 to 24 additional cancers detected per 1000 screening rounds. Studies reporting data at each end of this range showed that adding MRI to mammography provided a similar incremental sensitivity (50–54%), but were conducted in subjects with different risk levels.[15,18] The highest cancer yield was observed in the study population with the highest risk and prevalence of breast cancer.[18]

Four studies observed that mammography detected cancers that were occult on MRI.[14,15,17,18] One of these studies also found that ultrasound detected cancers that were not identified by either mammography or MRI.[18] The fifth study only reported four cancers, all of which were detected by MRI.[16] The two studies that reported on CBE as a component of conventional testing observed CBE-detected cancers that were occult on mammography[17] or mammography plus ultrasound,[18] but not MRI.

3.5. *Test recall rates*

Three studies reported subject recall rates for further investigation of false positive findings[15,16] and/or benign percutaneous biopsy (core needle biopsy or fine needle aspiration biopsy) rates[14–16] for different screening strategies. One study also reported on MRI-initiated benign surgical biopsies. Based on these data, the risk of being recalled for further investigation where cancer was subsequently excluded was three to five-fold higher when MRI was added to mammography versus mammography alone (relative risk 3.43–4.86) with an estimated 71–74 additional false positive recalls per 1000 screening rounds (Table 2).

The relative risk of undergoing a benign percutaneous biopsy due to the addition of MRI varied between 1.22 and 9.50 with an estimated 7–46 additional benign biopsies per 1000 screening rounds. Data from the MARIBS study also indicated that adding MRI doubled the risk of undergoing a benign surgical biopsy (relative risk 2.0, 95%CI 0.5–8.0; 2 additional benign surgical biopsies per 1000 screening rounds).[15]

3.6. *Stage shift in cancer detection*

None of the studies reviewed compared rates of advanced cancers in successive screening rounds. There were no statistically significant differences in tumour size (<10 mm or

Table 3 – Prognostic characteristics of breast cancers detected by screening with and without MRI

Study author, year	N	Screen-detected cancers	Number of invasive cancers detected by screening strategies with and without MRI & proportion with size < 10 mm and no lymph node involvement[a]		Size and lymph node status of invasive cancers detected by MRI alone
			With MRI	Without MRI	
Kuhl et al., 2005	390[b] *(total 529)*	Invasive 31,[c] DCIS 9	*Conventional testing = mammography + ultrasound* N = 31 • size <10 mm; ⩾20 mm NR mean 12.4 mm, s.d. 6.7, median 11.0 mm • LN negative 26/31 (84%)	N = 16 • size <10 mm; ⩾20 mm NR mean 13.9 mm, s.d. 6.4, median 13.0 mm • LN negative 11/16 (69%)	N = 14 • size <10 mm; ⩾20 mm NR mean 9.0 mm, s.d. 4.3, median = 7.5 mm[d] • LN negative 14/14 (100%)[e]
			Conventional testing = mammography N = 31 • size <10 mm; ⩾20 mm NR mean 12.4 mm, s.d. 6.7, median 11.0 • LN negative 26/31 (84%)	N = 10 • size <10 mm; ⩾20 mm NR mean 13.2 mm, s.d. 7.8 median 12.0 mm • LN negative 6/10 (60%)	NR
Leach et al., 2005	649	Invasive 28, DCIS 5	*Conventional imaging = mammography* N = 28 • size <10 mm: 10/28 (36%) ⩾20 mm: 9/28 (32%) LN negative 19/24 (79%)	N = 9 • size <10 mm: 4/9 (44%) ⩾20 mm: 3/9 (33%) LN negative 6/9 (67%)	N = 19 • size <10 mm: 6/19 (32%) ⩾ 20 mm: 6/19 (32%) • LN negative13/15 (87%)
Sardanelli et al., 2007	278	Invasive 14, DCIS 4	*Conventional imaging = mammography + ultrasound + CBE* N = 14 • size <10 mm: 3/14 (21%) ⩾20 mm: 3/14 (21%) • LN negative 10/13 (77%)	N = 12 • size <10 mm: 1/12 (8%) ⩾20 mm: 3/12 (25%) • LN negative 9/11 (82%)	N = 2 • size <10 mm: 2/2 (100%) ⩾20 mm: 0/2 (0%) • LN negative 1/2 (50%)
Warner et al., 2004	236	Invasive 15, DCIS 6	*Conventional imaging = mammography + ultrasound + CBE* N = 15 • size <10 mm: 5/15 (33%) • LN negative 12/14 (86%)	N = 10 • size <10 mm: 2/10 (20%) • LN negative 8/9 (89%)	N = 5 • size <10 mm: 3/5 (60%) • LN negative 4/5 (80%)
			Conventional imaging = mammography + CBE N = 13 • size • <10 mm:5/13 (38%) ⩾20 mm: 2/13 (15%) • LN negative 11/12 (92%)	N = 6 • size • <10 mm: 1/6 (17%) ⩾20 mm: 1/6 (17%) • LN negative 5/5 (100%)	N = 7 • size • <10 mm: 4/7 (57%) ⩾20 mm: 1/7 (14%) • LN negative 6/7 (86%)

DCIS = ductal carcinoma *in situ*; LN = lymph node; NR = not reported.

a No statistically significant difference in proportion of invasive cancers with size < 10 mm or no lymph node involvement detected by screening strategies with and without MRI.

b 390 women without a prior history breast cancer from total study population of 529.

c 31 primary invasive breast cancers detected in 30 women, excludes invasive cancers detected in patients with prior history of breast cancer.

d T-test for comparison of mean size of cancers detected by conventional tests versus by MRI alone: $P = 0.02$.

e X^2-test for comparison of proportion of cancers with no lymph node involvement detected by conventional tests versus by MRI alone: $P = 0.02$.

⩾20 mm) or lymph node involvement of invasive cancers detected with and without MRI (Table 3). Data reported by Kuhl and colleagues indicated that the additional cancers detected by MRI alone were statistically significantly smaller and less likely to involve lymph nodes than cancers detected by mammography and ultrasound ($P < 0.05$). However, data from the larger MARIBS study did not show similar findings.

4. Discussion

This review provides strong evidence that adding breast MRI to conventional screening tests offers a highly sensitive strategy for the early detection of breast cancer in young high risk women. There is some evidence that adding MRI to mammography may increase patient recall rates due to increased false positive findings, although the magnitude of this trade-off is less certain.

Strategies that combined MRI plus mammography with or without ultrasound and CBE reported sensitivities of 93–100%. Screening sensitivity was 58% higher (95%CI 46–70%) when MRI was added to mammography due to the low sensitivity of using mammography alone in this population, with smaller but statistically significant net gains when MRI was added to combinations of mammography, ultrasound and CBE (incremental sensitivity 31–44%). Mammography, ultrasound and MRI each detected cancers that were not detected by the other modalities indicating that the use of all three tests in combination would provide optimum sensitivity.

Interpretation of estimates of sensitivity need to take into account the duration and completeness of subject follow-up to identify false negatives. Two studies that reported 100% sensitivity for the combination of MRI plus conventional tests either did not undertake long-term follow-up[16] or reported an interim analysis before all subjects completed 12-months of follow-up,[17] whereas other studies followed patients up for an average of two or more years and identified false negative (interval) cancers.

The increased sensitivity of screening with MRI should be considered with evidence about the potential corresponding risks of increased false positive findings. Estimates of the relative specificity of adding MRI to conventional tests were less consistently reported and varied widely across studies. This variation may be explained, at least in part, by the use of different thresholds for classifying positive test results. Screening specificity using MRI was lowest in the MARIBS study (77%, 95%CI 75–79%), where all patients referred for further tests and those with indeterminate results that were 'probably benign' were classified as having a positive test in addition to patients classified as having 'suspicious' or 'malignant' findings (BI-RADS 0.3–5).[15] Furthermore, all tests were double read and the most conservative score was selected for calculation of test sensitivity and specificity. In contrast, studies that only classified subjects as having a positive test if they had a 'suspicious' or 'malignant' finding (BI-RADS 4–5) estimated screening specificities of up to 96% using MRI, but these results do not take into account patients recalled for further investigation of indeterminate findings.

Two studies reported data indicating that adding MRI to mammography increased a woman's risk of being recalled for further investigation of a false positive result approximately three to five-fold,[15,16] with the number of additional false positive patient recalls per 1000 screening rounds varying according to the risk of the disease in the screened population. Both studies showed that adding MRI increased the risk of benign percutaneous biopsies by at least threefold, with Leach and colleagues also reporting a twofold increase in the risk of surgical biopsy.[15] In contrast, Kuhl and colleagues observed a smaller, non-significant difference in the rate of benign percutaneous biopsies when MRI was added to mammography (surgical biopsies not reported).[14] The authors attributed these favourable results to the high level of experience of study radiologists. At least in the setting of screening mammography, there is evidence that women prefer to be recalled more frequently for investigation of false positive results, provided this will lead to an increased overall chance of detecting cancer.[24]

The effectiveness of a screening test can be demonstrated by showing that it reduces patient mortality; or inferred from evidence that it detects extra cases of disease with a reduction in rates of interval cancers or the incidence of advanced cancers in subsequent screening rounds.[25] We did not identify any studies assessing the impact of MRI screening on any of these outcomes. Conclusions about the value of MRI are therefore based on evidence of improved sensitivity and assumptions about the benefits of early detection in young high risk women extrapolated from mammographic screening trials undertaken in older average risk populations.[5,9] Reductions in mortality observed in population breast cancer screening programs have been attributed to the ability of mammography screening to produce a stage shift in the detection of earlier stage breast cancer in asymptomatic women and the improved efficacy of treatment for early versus late stage disease. However, it is not clear whether early detection will produce the same benefits in high risk women because breast cancers in this population are known to have a different prognostic profile.[5]

We did not identify convincing evidence that the addition of MRI detects breast cancer at an earlier stage than conventional testing alone. However, none of the included studies were designed to test this hypothesis, and the number of cases detected may have been too small to demonstrate a true difference in tumour stage. In a large study comparing the accuracy of MRI versus mammography, Kriege and colleagues reported on tumour characteristics from 44 invasive breast cancers.[26] They found a statistically significant difference in the proportion of cases detected with favourable prognostic characteristics (size < 10 mm and absence of axillary lymph node spread and micrometastases) in high risk women participating in breast cancer screening programs that used MRI compared to a population-based sample of breast cancer cases (screening history unreported) and a sample of high risk women who did not participate in screening. Unfortunately, these findings do not allow conclusions about differences in tumour stage for the additional cases detected by adding MRI to conventional screening tests. The authors did not find a difference in tumour stage between invasive cancers de-

tected by MRI at the first versus subsequent screening rounds.[27] Therefore, although there is strong evidence that the addition of MRI results in the early detection of additional breast cancer cases compared to conventional screening tests, the benefits of early detection in terms of improved patient prognosis have not yet been quantified.

Two cost-effectiveness analyses used different methods to estimate the magnitude of incremental effects of adopting MRI in terms of quality adjusted life years saved. Plevritis and colleagues used a mathematical model of the natural history of invasive breast cancer to estimate the relative effects of screening with MRI plus mammography versus mammography alone and used US health care costs to estimate an incremental cost-effectiveness ratio.[28] The National Collaborating Centre for Primary Care applied evidence about test accuracy and UK costs from the MARIBS study to assumptions about the differential effects (in terms of 5-year survival) of early versus late detection.[12]

These different approaches both led to similar conclusions that MRI is potentially cost-effective for screening women younger than 50–54 years at very high risk of breast cancer (e.g. proven BRCA1 mutation carriers), but less cost-effective for screening populations with a wider risk or wider age distribution. Both modelled analyses also indicated that the relative effectiveness and cost-effectiveness of MRI varied substantially with variations in the incidence of breast cancer in the screened population and the estimated incremental sensitivity of MRI versus conventional screening tests. Notably, neither study investigated the incremental cost-effectiveness of MRI compared to screening programs that included the routine use of ultrasound which would be expected to reduce the incremental sensitivity and therefore cost-effectiveness of MRI.

Conclusions about the incremental accuracy of MRI compared to mammography plus ultrasound with or without CBE in this review were limited to evidence from three studies (2286 MRI studies in 1043 patients).[14,17,18] These studies reported a wide variation in the sensitivity of each modality, for example, the estimated sensitivity of mammography varied from 33% to 59%, and were conducted in populations with different levels of risk, aged from 25 to 79 years, leaving some uncertainty about the relative sensitivity of each modality in high risk women under 50 years. Furthermore, none of these studies reported sufficient data for a comparison of the specificity of screening strategies that used this combination of screening modalities. Thus, no conclusions can be drawn about the impact of MRI as an addition to mammography plus ultrasound on false positive recall rates; and any potential associated harms such as increased patient anxiety and use of healthcare resources for unnecessary biopsies.

These findings highlight the difficulty in synthesising evidence for conclusions about the value of a test from studies that have used different methods to assess test accuracy and do not report primary data for estimates of sensitivity and specificity. Prospective accuracy studies with methods and results clearly reported according to the Standards for Reporting of Diagnostic Accuracy (STARD guidelines) are needed to address outstanding questions about the performance of different screening strategies.[29]

The absence of evidence about the effects of early detection of breast cancer, in terms of reduced patient mortality, in this group of women should be judged in context of the barriers to conducting a RCT to quantify these effects (these include: large sample size and lengthy follow-up to detect a true effect; and the ethics and acceptability of randomisation to patients and clinicians). An alternate approach would be a RCT to compare screen-detected tumour stage and interval cancer rates (a shorter term outcome) for screening strategies with and without MRI as this may be more feasible.[25]

This review provides convincing evidence that including MRI in breast cancer screening programs for young high risk women detects additional breast cancers compared to conventional screening strategies. However, there is no direct evidence for estimates of the magnitude of patient benefits. If a policy of using MRI for screening young high risk women is adopted as is occurring in the UK,[12] the US[13] and under consideration in Australia,[30] challenges for clinicians include the need for comprehensive risk prediction models to accurately select high risk women who may stand to gain the most benefits. It will be critical to provide counselling and information to these women about the large uncertainty surrounding these potential benefits; the higher risk of unnecessary false positive findings; and the benefits and harms of primary prevention measures as an alternative strategy to screening.

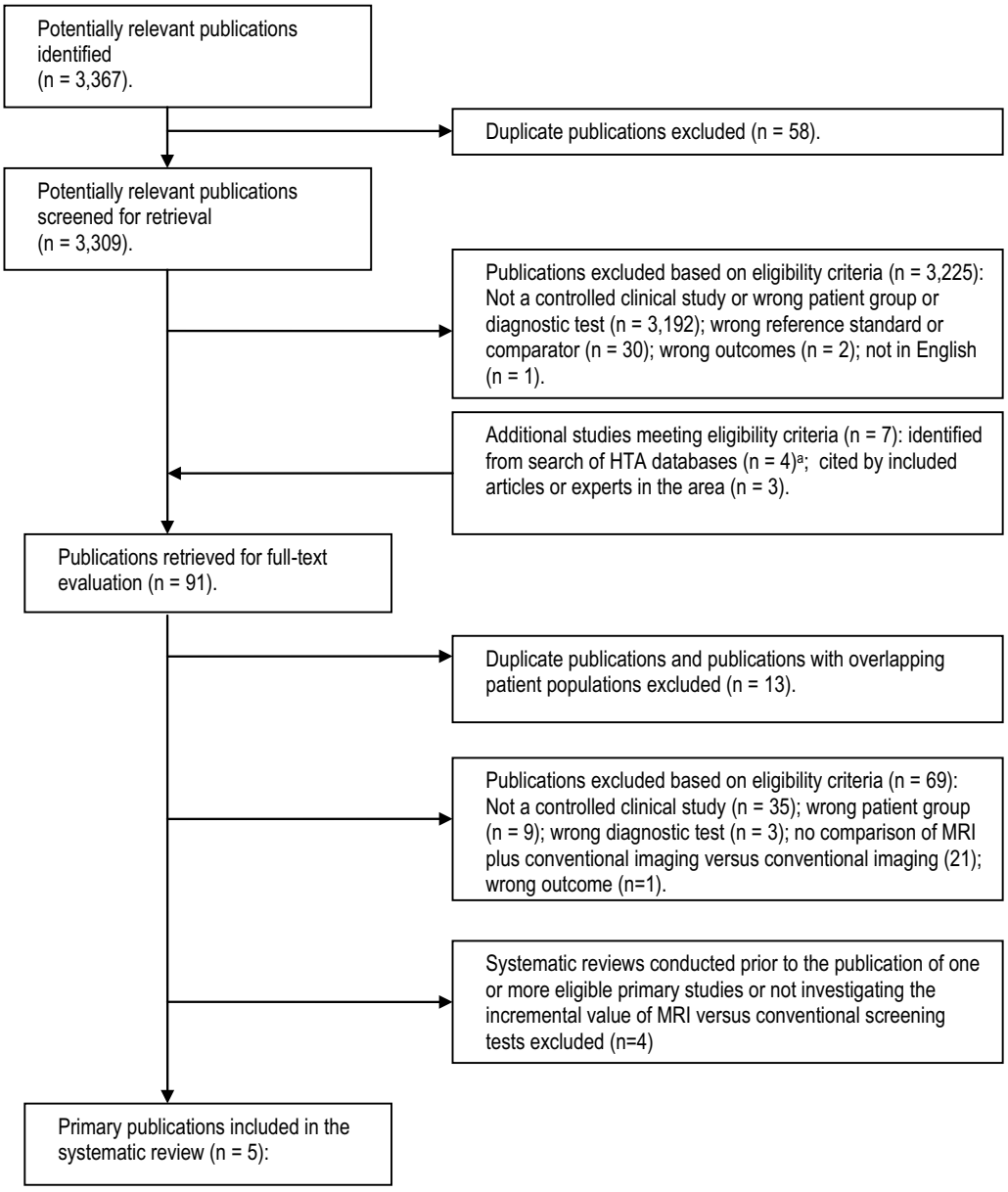
Conflict of interest statement

None declared.

Acknowledgement

We acknowledge Associate Professor Richard Bell, Dr. John Buckingham, Dr. Debra Graves, Dr. Bronwen Harvey and Ms Margaret Tassell, members of the Advisory Panel to the Australian Medical Services Advisory Committee's Breast Magnetic Resonance Imaging assessment report (application 1098), for their generous contributions to the planning and interpretation of the broader review on which this paper is based. The authors have received funding from the Department of Health and Ageing, Commonwealth of Australia for the assessment of magnetic resonance imaging of the breast.

Appendix. Quality of reporting of meta-analyses (QUOROM) flowchart summarising the results of the literature search and the application of eligibility criteria



REFERENCES

1. Antoniou A, Pharoah PD, Narod S, et al. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case series unselected for family history: a combined analysis of 22 studies. *Am J Hum Genet* 2003;**72**:1117–30.
2. Collaborative group on hormonal factors in breast cancer. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet* 2001;**358**:1389–99.
3. Cuzick J, Forbes J, Edwards R, et al. First results from the International Breast Cancer Intervention Study (IBIS-I): a randomised prevention trial. *Lancet* 2002;**360**:817–24.
4. Hartmann LC, Schaid DJ, Woods JE, et al. Efficacy of bilateral prophylactic mastectomy in women with a family history of breast cancer. *New Engl J Med* 1999;**340**:77–84.
5. Houssami N, Wilson R. Should women at high risk of breast cancer have screening magnetic resonance imaging (MRI)? *Breast* 2007;**16**:2–4.
6. Duffy SW, Tabar L, Smith RA. The mammographic screening trials: commentary on the recent work by Olsen and Gotzsche. *CA: Cancer J Clin* 2002;**52**:68–71.
7. Humphrey LL, Helfand M, Chan BK, Woolf SH. Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med* 2002;**137**:347–60.
8. International Agency for Research on Cancer (IARC). *Breast cancer screening. 7*, IARC Handbooks of Cancer Prevention, Lyon: IARC Press; 2002.

9. Liberman L. Breast cancer screening with MRI – what are the data for patients at high risk? *New Engl J Med* 2004;**351**:497–500.
10. Hamilton LJ, Evans AJ, Wilson AR, et al. Breast imaging findings in women with BRCA1- and BRCA2-associated breast carcinoma. *Clin Radiol* 2004;**59**:895–902.
11. Tilanus-Linthorst M, Verhoog L, Obdeijn IM, et al. A BRCA1/2 mutation, high breast density and prominent pushing margins of a tumor independently contribute to a frequent false-negative mammography. *Int J Cancer* 2002;**102**:91–5.
12. National Collaborating Centre for Primary Care. Familial breast cancer: The classification and care of women at risk of familial breast cancer in primary, secondary and tertiary care. Partial update of NICE clinical guideline 14. <http://www.nice.org.uk/page.aspx?o=317667 [accessed 12.10.06].
13. Saslow D, Boetes C, Burke W, et al. American cancer society guidelines for breast screening with MRI as an adjunct to mammography. *CA: Cancer J Clin* 2007;**57**:75–89.
14. Kuhl CK, Schrading S, Leutner CC, et al. Mammography, breast ultrasound, and magnetic resonance imaging for surveillance of women at high familial risk for breast cancer. *J Clin Oncol* 2005;**23**:8469–76.
15. MARIBS study group. Screening with magnetic resonance imaging and mammography of a UK population at high familial risk of breast cancer: A prospective multicentre cohort study (MARIBS). *Lancet* 2005;**365**:1769–78.
16. Lehman CD, Blume JD, Weatherall P, et al. Screening women at high risk for breast cancer with mammography and magnetic resonance imaging. *Cancer* 2005;**103**:1898–905.
17. Sardanelli F, Podo F, D'Agnolo G, et al. Multicenter comparative multimodality surveillance of women at genetic-familial high risk for breast cancer (HIBCRIT Study): interim results. *Radiology* 2007;**242**:698–715.
18. Warner E, Plewes DB, Hill KA, et al. Surveillance of BRCA1 and BRCA2 mutation carriers with magnetic resonance imaging, ultrasound, mammography, and clinical breast examination. *J Am Med Assoc* 2004;**292**:1317–25.
19. Irwig L, Houssami N, van Vliet C. New technologies in screening for breast cancer: a systematic review of their accuracy. *Br J Cancer* 2004;**90**:2118–22.
20. National Breast Cancer Centre (NBCC). *Magnetic resonance imaging for the early detection of breast cancer in women at high risk: a systematic review of the evidence*. Camperdown: NBCC; 2006.
21. Whiting P, Rutjes A, Reitsma J, Bossuyt P, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;**3**:25.
22. Zamora J, Muriel A, Abraira V. Meta_DiSc for Windows: a software package for the meta-analysis of diagnostic tests. <www.hrc.es/investigacion/metadisc.html> [accessed 03.02.07].
23. Egger M, Davey Smith G, Altman DG. *Systematic reviews in health care: meta-analysis in context*. 2nd ed. London: BMJ Publishing Group; 2001.
24. Ganott MA, Sumkin JH, King JL, et al. Screening mammography: do women prefer a higher recall rate given the possibility of earlier detection of cancer? *Radiology* 2006;**238**:793–800.
25. Irwig L, Houssami N, Armstrong B, Glasziou P. Evaluating new screening tests for breast cancer. *BMJ* 2006;**332**:678–9.
26. Kriege M, Brekelmans CTM, Boetes C, et al. Efficacy of magnetic resonance imaging and mammography for breast cancer screening in women with a familial or genetic predisposition. *Obstet Gynecol Surv* 2005;**60**:107–9.
27. Kriege M, Brekelmans CT, Boetes C, et al. Differences between first and subsequent rounds of the MRISC breast cancer screening program for women with a familial or genetic predisposition. *Cancer* 2006;**106**:2318–26.
28. Plevritis SK, Kurian AW, Sigal BM, et al. Cost-effectiveness of screening BRCA1/2 mutation carriers with breast magnetic resonance imaging. *JAMA* 2006;**295**:2374–84.
29. Bossuyt PM, Reitsma JB, Bruns DE, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Ann Intern Med* 2003;**138**:W1–W12.
30. The Medical Services Advisory Committee. *Breast magnetic resonance imaging, MSAC Application 1098*. Canberra: Commonwealth of Australia; 2007.