Breast magnetic resonance imaging

November 2006

MSAC application 1098

Assessment report

ISBN (Print) 1-74186-203-5

ISBN (Online) 1-74186-204-3

ISSN (Print) 1443-7120
ISSN (Online) 1443-7139

First printed April 2007
Paper-based publications

Electronic copies of the report can be obtained from the Medical Service Advisory Committee's Internet site at http://www.msac.gov.au/
Printed copies of the report can be obtained from:
The Secretary
Medical Services Advisory Committee
Department of Health and Ageing
Mail Drop 106
GPO Box 9848
Canberra ACT 2601

Enquiries about the content of the report should be directed to the above address.

The Medical Services Advisory Committee (MSAC) is an independent committee which has been established to provide advice to the Minister for Health and Ageing on the strength of evidence available on new and existing medical technologies and procedures in terms of their safety, effectiveness and cost-effectiveness. This advice will help to inform government decisions about which medical services should attract funding under Medicare.

MSAC recommendations do not necessarily reflect the views of all individuals who participated in the MSAC evaluation.

This report was prepared by the Medical Services Advisory Committee with the assistance of Sarah Lord, Wei Lei, Alison Griffiths, Silke Walleser, Sharon Parker and Suchaya Thongyoo from the NHMRC Clinical Trials Centre, The University of Sydney and Simon Eckermann from the Flinders Centre for Clinical Change and Health Care Research, Flinders University. The report was edited by Bruce Howarth, PhD. This recommendation was endorsed by the Minister for Health and Ageing on 5 February 2007.
Publication approval number: P3-1289

Contents

Tables

Figures

Executive summary

The procedure

Magnetic resonance imaging (MRI) uses a strong external magnetic field to produce images of biological tissues. It is particularly well suited to imaging blood vessels, other fluid filled structures and soft tissues and has been used in the diagnosis of breast cancer.

Breast MRI is performed using an MRI machine fitted with a dedicated breast coil. A radiographer with specialised training in breast MRI is required for set-up and scanning. The Supervising Radiologist should be part of a multi-disciplinary team with expertise in breast imaging and MRI. The patient is placed within the cylindrical chamber of the MRI machine and an intravenous contrast agent (gadolinium) is used. The examination takes approximately 30 minutes.

MRI is free of ionising radiation and its associated risks. It is contraindicated in patients with ferromagnetic implants (iron and nickel materials that can be easily magnetised) or electronic implants that are incompatible with exposure to magnetic fields.

Medical Services Advisory Committee – role and approach

The Medical Services Advisory Committee (MSAC) was established by the Australian Government to strengthen the role of evidence in health financing decisions in Australia. MSAC advises the Minister for Health and Ageing on the evidence relating to the safety, effectiveness and cost-effectiveness of new and existing medical technologies and procedures and under what circumstances public funding should be supported.

A rigorous assessment of evidence is thus the basis of decision making when funding is sought under Medicare. A team from the NHMRC Clinical Trials Centre was engaged to conduct a systematic review of literature on breast MRI. An advisory panel with expertise in this area then evaluated the evidence and provided advice to MSAC.

MSAC's assessment of breast MRI

This review addresses the following questions:

Primary review question

What is the safety, effectiveness and cost-effectiveness of annual breast MRI in addition to annual mammography with or without breast ultrasound for screening asymptomatic high-risk women under the age of 50 years?

Secondary review questions

1. What is the safety, effectiveness and cost-effectiveness of annual breast MRI in addition to annual mammography with or without breast ultrasound for screening asymptomatic high-risk women aged 50 years or older?

2. What is the safety, effectiveness and cost-effectiveness of annual breast MRI versus mammography with or without breast ultrasound for screening asymptomatic high-risk women under the age of 50 years?

A systematic review of the medical literature published up to March 2006 identified nine systematic reviews and 12 primary studies reporting evidence relating to these questions. Three studies published after the preparation of the present report, including updated guidelines from the National Institute for Health and Clinical Excellence (NICE) in the United Kingdom for the classification and care of women at risk of familial breast cancer (National Collaborating Centre for Primary Care (NCCPC) 2006), were included as an appendix to this report for MSAC's consideration.

Clinical need

About 9% of Australian women will develop breast cancer during their lifetime. Men can also develop breast cancer, although the risk is very low. Overall, 11,886 cases were diagnosed in Australia in 2001 (Australian Institute of Health and Welfare (AIHW) 2006). Of those diagnosed, 84% survive at least five years (Cancer in Australia 2001) with higher survival rates reported for cases detected at an early stage.

Incidence increases with age and mammographic screening programs targeting women from the age of 40-74 years have demonstrated reductions in mortality of around 24% (Duffy et al 2002). However there is a small group (less than 1% of Australian women) with a known gene mutation or strong family history of breast cancer who account for 5-10% of breast cancer cases (NHMRC 1999). These women carry a cumulative lifetime risk of developing breast cancer of 40-80% and are more likely to develop breast cancer before the age of 50 years than the general population. Current guidelines recommend that women with a familial or genetic predisposition commence an annual mammography screening program from the age of 35-40 years, or at least 5 years earlier than the earliest age at diagnosis of the youngest affected family member. However, existing mammography screening programs may be inadequate for detecting breast cancer in these women because of its low sensitivity in women under the age of 50 years (Kerlikowske et al 1996). Breast MRI is proposed as an addition to mammography to overcome this problem.

If breast MRI is introduced for surveillance of high-risk women between the ages of 35 and 50 years (or earlier if the youngest affected family member was diagnosed before 40 years of age), utilisation is estimated at around 5,000 women per year. This estimate is based on the assumption that possibly only 20% of high-risk women will be identified by surveillance programs and thus actual utilisation may vary substantially from this figure should higher uptake be achieved.

Safety

Breast MRI is a safe procedure in patients without contraindications to exposure to magnetic fields. Ten accuracy studies comparing the performance of breast MRI with mammography in 4,040 patients (8,691 scans) found no adverse events due to either test, although three of these studies reported that 1 to 5% of women suffered claustrophobia or refused MRI for other reasons (Kriege et al 2004; Leach 2005; Tilanus-Linthorst et al 2000). Allergy to the intravenous contrast agent gadolinium is very rare. One large series observed moderate to severe gadolinium reactions in approximately 1 in 10,000 patients (De Ridder et al 2001).

MRI does not involve ionising radiation and its associated risks; however, the risks of repeated exposure to ionising radiation from annual mammography in young high-risk women may not be insignificant. There is theoretical evidence to suggest that this risk increases with repeated exposures and younger age at exposure, and is greater for women with a family history of breast cancer. However, two recent epidemiological studies have reported conflicting results, with one retrospective cohort study finding an increased risk of breast cancer with early and repeated exposure to chest x-rays, whereas a large case-control study found no association between mammography exposure and breast cancer risk in high-risk women.

Effectiveness

Evidence about the relative effectiveness of breast MRI as a screening test for breast cancer is limited to evidence from studies of test accuracy. These studies are designed to assess the test performance of breast MRI, not its impact on patient outcomes. Three level III-1 and III-2 studies of test accuracy investigating the relative accuracy of MRI plus mammography versus mammography alone in young high-risk women were identified (Leach et al 2005, Kuhl et al 2005, Warner et al 2004). Together, these studies report data from 3792 screening MRI scans performed on 1,414 high-risk women with an average age of less than 50 years (range mean/median age 40-47 years), reflecting the proposed target screening population in Australia. Despite limitations in the quality and/or reporting of these studies; the size, statistical precision and consistency of the results provides strong evidence that the addition of breast MRI increases the early detection of breast cancer compared to mammography alone. Meta-analysis showed a statistically significant 2.6-fold increase in test sensitivity (MRI sensitivity 94% [95% CI 86-98%]; mammography sensitivity 36% [95% CI 25-48%; incremental sensitivity for breast MRI of 58% [95% CI 46-70%]).

Variation in the specificity of breast MRI reported by these studies does not allow a reliable estimate of the relative specificity of the combination of breast MRI and mammography versus mammography alone. Some of this variation can be attributed to differences in the threshold used to classify false positive test results across studies. One study reported detailed data about patient recall rates (Leach et al 2005). This study indicated that the addition of breast MRI leads to a three-fold increase in the patient recall rate for further investigation and biopsy due to false positive findings (Leach et al 2005).

Two studies also provided evidence that breast MRI will also detect more cancers when added to screening programs that include the routine use of ultrasound, although the incremental sensitivity of adding MRI to a combination of mammography and ultrasound was less than adding MRI to mammogram alone (Kuhl et al 2005, Warner et al 2004).

No clinical trials or observational studies of breast screening programs in high-risk women were identified to assess the magnitude of any clinical benefits associated with early versus later detection of breast cancer in this population. A recently published analysis used a mathematical model of the natural history of breast cancer to estimate that breast MRI will reduce the risk of breast cancer mortality by 23% in high-risk women. However, considerable uncertainty surrounds this estimate. Although it is plausible that the extra cases detected by breast MRI represent an earlier stage of disease which will benefit from earlier treatment, the evidence identified in this review does not

demonstrate that MRI screen-detected invasive cancers are statistically significantly smaller or more likely to be node negative than mammogram screen-detected cases.

Secondary review questions

1. This review identified no evidence for additional conclusions about the relative effectiveness or accuracy of adding breast MRI for surveillance of high-risk women aged 50 years or older. However, existing evidence that mammography has a higher sensitivity in older women suggests the incremental accuracy of breast MRI is likely to be lower in this age group.

2. This review identified six level III-1 and III-2 studies of test accuracy that compared the accuracy of breast MRI as a replacement for mammography for screening young women at high risk of breast cancer (Leach et al 2005, Lehman et al 2005, Kriege et al 2004, Kuhl et al 2005. Podo et al 2002. Warner et al 2004). These studies provide consistent evidence that breast MRI is a more sensitive and less specific test than mammography for detecting breast cancer, although reports of MRI specificity varied widely (meta-analysis: MRI sensitivity 81% [95% CI 74-87%], mammography sensitivity 36% [95% CI 28-44%]; MRI specificity range: 81-99%; mammography specificity range 93-100%).

Economic considerations

This review identified one published evaluation of the cost-effectiveness of MRI for screening women at high risk of breast cancer in the United States (US) (Plevritis et al 2006a). This evaluation applied evidence of the accuracy of MRI and mammography to a mathematical model of the natural history of breast cancer to estimate incremental effects, incremental cost and an incremental cost-effectiveness ratio for the addition of MRI relative to mammography alone in different age groups and different risk groups in the US. Costs were evaluated from a societal perspective and included indirect costs associated with lost production time due to investigations and treatment and a 3% discount factor.

The results of this evaluation suggests that breast MRI is potentially cost-effective for screening very high-risk women in the US such as BRCA1 mutation carriers aged 35-54 years, but is unlikely to be cost-effective for screening BRCA2 carriers or a wider risk or age population such as that under consideration in Australia.

To translate this economic evaluation to an Australian setting and comply with Australian Government decision-making processes, a secondary economic analysis was performed based on the Plevritis model with the same assumptions about the diagnostic performance and effects of screening, but excluding indirect costs and applying Australian relative prices. Results of this secondary analysis suggest the potential for breast MRI to be cost-effective in a select subgroup (BRCA1 mutation carriers aged 35-54 years) of the proposed screening population in Australia.

However, while indirect costs could be excluded and Australian prices applied, this secondary analysis was not able to allow for:

1. the expected higher sensitivity of standard screening practice in Australia where ultrasound may be combined with mammography versus mammography alone as

modelled by Plevritis et al (2006a) (estimated test sensitivity 49% versus 35% respectively);

2. inclusion of women for screening in Australia with a lower baseline risk of breast cancer (≥ 30% versus 45-65% lifetime risk); and

3. a higher discount rate (5% in Australia versus 3% in the Plevritis model).

Accounting for each of these factors would increase the modelled incremental cost effectiveness ratio. Aside from these translation issues, uncertainty surrounding variables, clinical assumptions and the structure of the model used by Plevritis et al (2006a) in extrapolating evidence over time also need to be considered. It is therefore suggested that the secondary exploratory analysis, despite improving on purchase power parity, does not provide enough certainty of evidence to determine the baseline breast cancer risk at which the addition of breast MRI might be cost-effective in Australia.

Using the modelled Australian analysis, the total cost to government of implementing MRI as an additional test for screening young high-risk women for breast cancer and associated investigations up to the point of diagnosis has been estimated as AU$3,759,002 per year, an additional $3,233,600 above the estimated equivalent cost for mammography alone. This figure would vary significantly with different assumptions about the cost of breast MRI, uptake of the procedure, the sensitivity of standard mammography screening protocols that include the option of performing a screening ultrasound, patient baseline risk and other criteria used to determine eligibility for MRI.

Other considerations

Consideration of the specific expertise needed to perform breast MRI and the unique needs of young women at high risk of breast cancer are relevant to discussions about the proposed implementation of breast MRI. It was the expert opinion of the Advisory Panel that if breast MRI is introduced, assessment of patient eligibility and testing should be performed by specialists working within a multi-disciplinary team with expertise in breast imaging and MRI. This would help ensure appropriate referral and co-ordination of screening and diagnostic services including the counselling of young women identified as being at high risk about the potential benefits and harms of testing and would support the ability of MRI facilities to gain and maintain adequate skills in the technique.

International guidelines

The current American Cancer Society guidelines (2003) conclude that women at increased risk of breast cancer may benefit from earlier initiation of screening, shorter screening intervals or the addition of ultrasound or MRI to mammography and clinical breast examination. However, the guidelines state the evidence available is insufficient to justify recommendations for any of these screening approaches. An update of these guidelines was expected in late 2006.

The National Institute for Health and Clinical Excellence (NICE) in the United Kingdom released a partial update of guidelines for the classification and care of women at risk of familial breast cancer after the preparation of this report (National Collaborating Centre for Primary Care (NCCPC) 2006). This information was reviewed with two other recent publications for MSAC's consideration as an appendix to the present report.

The economic evaluation produced similar conclusions to the study by Plevritis et al (2006a) about the cost-effectiveness of MRI screening in subgroups of women at high risk of familial breast cancer selected by risk level and age. The updated NICE guidelines for breast cancer surveillance recommend that annual screening with breast MRI and mammography should be offered to known BRCA1, BRCA2 and p53 mutation carriers in the age range of 30-49 years (earlier MRI for women at exceptionally high risk) and to other high-risk groups depending on their level of risk and age.

Overall conclusions

Overall, this review provides strong evidence that breast MRI is a safe test that offers a 2.6-fold increase in the detection of breast cancer in young high-risk women compared to mammography alone. It also produces a 3-fold increase in the rate of investigations for false positive findings. However, owing to the lack of clinical evidence to determine the health benefits gained by earlier detection of breast cancer in this population and uncertainty about the applicability of estimates of cost-effectiveness derived from a US economic model, this review does not provide adequate evidence to determine the potential effectiveness and cost-effectiveness of adopting breast MRI in Australia.

Other factors that may influence a decision to support funding for breast MRI for the surveillance of young high-risk women include: judgement about the validity of linking evidence of mortality reduction from the early detection of breast cancer from screening trials in average-risk women over the age of 50 years to the early detection of breast cancer in younger high-risk women; the ethics and feasibility of conducting a trial to quantify these health benefits; the estimated total costs of introducing breast MRI; and the availability of facilities with appropriate expertise and equitable access to services.

Recommendation

Breast MRI, when combined with mammography, is safe and effective in the diagnosis of breast cancer in asymptomatic women at high risk, when used as part of an organised surveillance program.

Evidence suggests that breast MRI in combination with mammography may be cost-effective when compared with mammography alone in high risk women aged less than 50 years.

MSAC recommends interim public funding for breast MRI in the diagnosis of breast cancer in asymptomatic women with a high risk of developing breast cancer when used as part of an organised surveillance program.

Evidence should be reviewed in not less than 3 years.

- The Minister for Health and Ageing endorsed this recommendation on 5 February 2007 -

Introduction

The Medical Services Advisory Committee (MSAC) has reviewed the use of breast magnetic resonance imaging (MRI), which is a screening and diagnostic test for the detection of breast cancer. MSAC evaluates new and existing health technologies and procedures for which funding is sought under the Medicare Benefits Scheme in terms of their safety, effectiveness and cost-effectiveness, while taking into account other issues such as access and equity. MSAC adopts an evidence-based approach to its assessments, based on reviews of the scientific literature and other information sources, including clinical expertise.

MSAC's terms of reference and membership are at Appendix A. MSAC is a multidisciplinary expert body, comprising members drawn from such disciplines as diagnostic imaging, pathology, surgery, internal medicine and general practice, clinical epidemiology, health economics, consumer health and health administration.

This report summarises the assessment of current evidence for breast MRI as a screening test for the surveillance of asymptomatic women under the age of 50 years at high risk of breast cancer.

Background

Breast MRI

Magnetic resonance imaging (MRI) uses a strong external magnetic field to produce images of biological tissues. This magnetic field acts on hydrogen protons (elementary particles) in body tissues and a radiofrequency pulse is used to produce signals that vary according to their local chemical, structural and magnetic environment. MRI is particularly well suited to distinguishing between blood vessels, other fluid filled structures and surrounding soft tissues.

Breast MRI is used in the diagnosis of breast cancer. It provides multiple images of the breast, cross-sectional in three dimensions (side-to-side, top-to-bottom, front-to-back), which are then read by a radiologist. Breast MRI is performed in a dedicated MRI room using an MRI machine with a magnet strength of 1.0 Tesla or greater fitted with a breast coil. A radiographer with specialised training in breast MRI is required for setup and scanning. The Supervising Radiologist should be part of a multi-disciplinary team with expertise in breast imaging and MRI.

Breast MRI is typically performed with the patient lying prone with the breast dropping into the dedicated breast coil. As breast tissue generally has similar signal intensity to tumour tissue on routine MRI, the intravenous administration of a contrast agent containing gadolinium chelate is used to enhance breast lesions. The pattern of enhancement (the change in signal intensity over time) is used to differentiate between malignant and benign lesions. The patterns for most malignant cancers show significantly faster exchange rates compared with benign breast lesions (Knopp et al 1999). An example of the techniques currently used to perform breast MRI in Australia provided by the applicant for this assessment is presented in Appendix C. Performing breast MRI within the first half of the menstrual cycle has been recommended to minimise the influence of physiological changes that are associated with greater contrast medium uptake (Rieber et al 1999).

MRI is free of ionising radiation. It is contraindicated in patients with ferromagnetic implants (iron and nickel materials that can be easily magnetised) or electronic implants that are incompatible to exposure to magnetic fields. Use in such patients may potentially cause device malfunction, dislodgement and death. Some patients are unable to tolerate MRI due to claustrophobia.

Intended purpose

Breast MRI can be used as a screening and diagnostic test for breast cancer. Potential indications include:

- Breast cancer screening in women at high risk of breast cancer due to family history or genetic predisposition
- Breast cancer screening in women with radiologically dense breasts ('mammographic breast density', see page 4).

- Investigation of patients with equivocal results on mammogram and/or breast ultrasound, or clinical symptoms but a negative mammogram
- Local staging of recently diagnosed breast cancer
- Detection of occult breast cancer in patients with lymph node metastases
- Monitoring response to neoadjuvant therapy
- Evaluation of silicone implants

This report focuses on an assessment of breast MRI as an additional test to mammography for screening a select high-risk target population of women under the age of 50 years including:

- Individuals with a known mutation of a breast cancer predisposing gene such as Breast Cancer Gene 1 (BRCA1), Breast Cancer Gene 2 (BRCA2), or tumour antigen p53 (also known as TP53).
- First-degree relatives of a known mutation carrier.
- First-degree relatives of an affected person in a family with a high genetic predisposition to breast cancer, assessed as > 30% personal lifetime risk of breast cancer by, or in consultation with, a familial cancer clinic.

It is proposed that breast MRI be offered as an addition to mammography to women aged 35-50 years who meet at least one of these criteria. If the youngest affected family member was diagnosed with breast cancer before the age of 40 years, screening should be offered 5 years prior to the age at diagnosis of this index case.

Clinical need/burden of disease

Incidence and mortality rates for breast cancer

Breast cancer is the most common registrable cancer among Australian women. One in 11 Australian women will develop breast cancer before the age of 75 years, which represents the fifth-highest incidence rate in the world (World Health Organisation (WHO) 2003). In 2001, 11,791 new cases were identified (29% of all new cancer cases) with an age standardised incidence of 117 cases per 100,000 women (Australian Institute of Health and Welfare (AIHW) 2006).

Incidence increases with age, but in 2001, 24% of all cases were diagnosed in women younger than 50 years. These comprised 55 (0.5% of total) cases in the 20-29 years age group, 619 (5.25%) cases in the 30-39 years age group, and 2,140 (18.1%) cases in the 40-49 years age group, out of a total of 11,791 cases (AIHW 2006). The proportion is higher for breast cancers diagnosed in women at high risk. The IBIS-1 trial of tamoxifen in women at increased risk of breast cancer aged between 35 and 70 years reported that 39/101 (39%) of breast cancers diagnosed in 3,566 women assigned to receive the placebo were detected in women younger than 50 years old (Cuzick et al 2002).

The five-year relative survival rate for Australian women diagnosed with breast cancer during 1992-97 was 84% (Cancer in Australia 2001). In 2001, 2,592 female deaths were attributed to breast cancer (AIHW 2006).

Risk factors for breast cancer

Other than female gender, the most important risk factors for breast cancer are increasing age and genetic predisposition. The age-specific incidence rate of breast cancer increases from 12 cases per 100,000 women aged 30-34 years to 215 cases per 100,000 women aged over 85 years (AIHW 2006). Genetic predisposition is estimated to account for 5-10% of all breast cancer cases (NHMRC 1999; UK MRI Breast screen study advisory group 2000). Genes associated with breast cancer that can be inherited include mutations of BRCA1, BRCA2 and tumour antigen p53. However, these gene mutations do not explain all cases of breast cancer in women with a strong family history of the disease.

Hormonal factors such as a woman's menstrual and reproductive history (early age at commencement of menstruation, late age at first full-term pregnancy, fewer children and late age at menopause) and long-term use of hormone therapy are also associated with a small but significant increased risk of breast cancer (Collaborative group on hormonal factors in breast cancer 1997). The appearance of radiologically dense breast tissue, referred to as 'mammographic breast density', is another independent risk factor for breast cancer (Boyd et al 2002). Mammographic breast density reflects the composition of breast tissue; epithelium and stroma appear white on the mammogram film (radiologically dense) whereas fat appears dark (radiologically lucent).

Women who have received radiotherapy of the chest area between the ages of 10 and 30 years for treatment of Hodgkin's Lymphoma or other childhood cancers also have an increased risk of breast cancer (Kenney et al 2004; Travis et al 2005). Women with a previous diagnosis of breast cancer have a three-fold increased risk of developing a second primary cancer in the contralateral breast compared to women without a history of breast cancer. This risk increases for women with a first diagnosis of breast cancer before the age of 45 years (Harvey et al 1985).

Risk of breast cancer in women with a genetic predisposition

The lifetime risk of breast cancer among women with a mutation in BRCA1 or BRCA2 has been estimated to be in the range of 40-80% (NHMRC 1999). Female carriers of BRCA mutations, in particular BRCA1 mutations, also have an increased risk of developing a second primary breast or ovarian cancer (NHMRC 1999). Inherited mutations of p53 have been associated with a relative risk of breast cancer of 18 times the general population by age 45 years (Garber et al 1991). These mutations predominantly occur within the context of the Li-Fraumeni and Li-Fraumeni-like family syndromes which are characterised by soft tissue sarcoma, early onset breast cancer and other tumours (Dickson et al 2001). Somatic (non-inherited) mutations of p53 are also found in approximately 20-40 % of breast cancers depending on tumour size and stage of the disease (Borresen-Dale 2003).

Table 1 shows the frequency of these gene mutations and the cumulative risk of breast cancer in female carriers.

Table 1 Genes associated with inherited predisposition to breast cancer, frequency and cumulative risk

Gene	Mutation frequency	Major sites at risk	Breast cancer risk to age 75 years in mutation carriers[1]
BRCA1	1/1,000	Breast Ovary	40-80%
BRCA2	1/1,000	Breast Ovary	40-80%
p53	1/10,000	Breast Bone or soft tissue	50%

Source: NBCC 2000

1. Average population cumulative risk of breast cancer to age 75 years is 9% (NHMRC 1999)

Correct identification of high-risk individuals is complex owing to technical limitations in genetic testing for BRCA1 and BRCA2 mutations and the lack of comprehensive prediction models which take into account both genetic and environmental factors to estimate future breast cancer risk. Screening protocols such as those used in studies evaluating breast MRI have tended to use documented mutations or risk predicted by family history to determine eligibility which may not adequately capture all high-risk women (Hartman 2005).

Although individuals with early onset breast cancer or a strong family history are more likely to carry inherited mutations, a substantial number of carriers are without family history (Antoniou et al 2003). In an Australian population-based case control study, Southey et al (1999) estimated that 3.8% of women developing breast cancer before the age of 40 were carriers of an inherited mutation in BRCA1.

Prevalence of women in Australia with a genetic predisposition for breast cancer

The National Health and Medical Research Council (NHMRC) estimates that women at high risk due to family history or genetic predisposition account for less than 1% of the Australian female population (although not all these women will go on to develop breast cancer) (NHMRC 1999). Using this information and data from the 2001 Australian census, up to 22,584 women aged between 35 and 50 years may be at high risk of breast cancer and eligible for annual surveillance (Australian Bureau of Statistics 2001). Based on estimates that approximately 20% of high-risk women are diagnosed before the age of 40 years (Dagan et al 2002), then an additional 2,761 women at high risk may be eligible for screening between the ages of 25 and 34 years.

Classification of breast cancer risk

There are three categories of breast cancer risk: i) at or slightly above average risk; ii) moderately increased risk; and iii) potentially high risk. Women are classified as potentially high risk if they meet one of the following criteria:

- breast or ovarian cancer diagnosed in three or more first- or second-degree relatives on the same side of the family, or

- breast or ovarian cancer diagnosed in two or more first- or second-degree relatives on one side of the family diagnosed, plus one or more of the following features (on the same side of the family):
    - bilateral breast cancer;
    - onset of breast cancer before the age of 40 years;
    - onset of ovarian cancer before the age of 50 years;
    - breast and ovarian cancer in one individual;
    - Ashkenazi Jewish ancestry;
    - breast cancer in a male relative;
    - one first or second-degree relative diagnosed with breast cancer at age 45 years or younger, plus another first or second-degree relative on the same side of the family with bone or soft tissue sarcoma at age 45 years or younger; or
    - a demonstrated mutation in a high-risk breast cancer-associated gene such as BRCA1, BRCA2 or p53 by genetic testing.

(National Breast Cancer Centre 2006a; NHMRC 1999).

In Australia, an increasing number of individuals are offered genetic screening when they present as newly diagnosed cases. This offer is made based on age at diagnosis, family history of cancer and tumour pathology. In women with a family history of breast cancer, the risk of carrying a genetic predisposition for breast cancer is inversely associated with the age at diagnosis of the family member (ie risk increases with younger age at diagnosis). Although genetic testing is readily available, an Australian population-based study that invited 94 women diagnosed with breast cancer before the age of 40 years to receive testing demonstrated some reluctance and indecision among women with regard to knowing their mutation status (Keogh et al 2004). The reasons for declining testing were not explored in this study. Similarly, a survey of 142 female members of high-risk families who did not have a diagnosis of breast cancer reported that 51% declined to receive their mutation result (Phillips et al 2006).

Expert opinion suggests that women may not accept the offer of genetic testing if there is no perceived advantage to knowing mutation status and/or a perceived disadvantage if there is a concern that this information may be used to discriminate against known mutation-carriers, for example by insurance providers or employers (Advisory Panel, 21 August 2006).

Current prevention and surveillance strategies for women at high risk of breast cancer

Strategies for primary prevention and increased surveillance offer some options for risk management for high-risk women. Primary prevention strategies include prophylactic mastectomy, oophorectomy and chemoprevention.

Prophylactic bilateral mastectomy involves the total surgical removal of both breasts. Observational evidence has shown that the procedure is associated with a 90% reduction in breast cancer incidence at a median follow-up of 14 years (Hartmann et al 1999). However, the associated physical and psychological morbidity is significant and unacceptable for many women (Robson et al 2004). Preventive oophorectomy involves the removal of both ovaries. It is undertaken to prevent ovarian cancer and reduces the risk of breast cancer by bringing on premature menopause. Observational evidence has shown that it is associated with a reduction in the relative risk of breast cancer incidence of around 50% (Rebbeck et al 2002). Unfortunately, neither of these procedures provides complete protection against breast cancer.

Chemoprevention involves the use of medication to prevent or inhibit the development of cancer and is predominantly used in the United States (US). The most commonly used medication is tamoxifen, an anti-oestrogenic agent. Chemoprevention is not normally given to women under the age of 40 years and, although it offers a reduction in breast cancer risk for those at high risk, the absolute risk of breast cancer remains significant (UK MRI Breast Screen Study Advisory Group 2000).

The IBIS-1 trial (7,139 women including participants from Australia and New Zealand) is one of four randomised placebo-controlled studies that have investigated the role of tamoxifen for breast cancer prevention. Women eligible for the trial had risk factors for breast cancer indicating at least a two-fold increase in relative risk for ages 45-70 years, a four-fold increase in relative risk for ages 40-44 years and approximately a ten-fold increase in relative risk for ages 35-39 years. This trial concluded that in a high-risk population, five years of tamoxifen reduces the risk of breast cancer events (including non invasive cancers) by 32%, and the risk of ER positive invasive breast cancer by 31% (Cuzick et al 2002). This effect, however, was also accompanied by a doubling in the risk of endometrial cancer (although this effect did not reach statistical significance) and a 2.5 fold increase in thromboembolic events. Overall mortality was significantly higher in women randomised to tamoxifen; however a causative role for tamoxifen in the excess of cancer deaths not due to breast cancer has not been established. The overall risk-to-benefit ratio for tamoxifen in breast cancer prevention remains unclear.

Intensified surveillance is advised for women at high risk of breast cancer who do not choose to undergo prophylactic mastectomy. Surveillance strategies include regular breast self examination (BSE), clinical breast examination (CBE) every 6 to 12 months and annual mammography with or without ultrasound (NHMRC 1999). The surveillance program needs to be individualised according to patient risk. In general, annual mammography screening is recommended to start from 25 years of age for women with a familial syndrome of breast and ovarian cancer or 5 to 10 years before the age at diagnosis of the youngest case for women with a strong family history or other genetic predisposition (National Comprehensive Cancer Network. 2004), NBCC 1999).

A recent survey of 142 female BRCA1 and BRCA2 mutation carriers enrolled in the Australian and New Zealand kConFab cohort of breast cancer families who did not have a diagnosis of breast cancer showed only moderate adherence to these screening recommendations (Phillips et al 2006, survey response rate 75%). Eight respondents (6%) reported having bilateral mastectomy. Of the 90 respondents eligible for mammography, 69% (95% CI 58-78%) reported having mammography at least annually. Twenty-three percent of respondents (95% CI 15-32%) also reported undergoing breast ultrasound regularly. Uptake of mammography was associated with electing to receive information about mutation status (89% uptake, odds ratio 7.2, 95% CI 3.1-20.5) and

residence in a major city versus regional area (percentage uptake not reported, odds ratio 3.9, 95% CI 1.3-11.8, Phillips et al 2006).

No trials have been conducted to assess the effectiveness of surveillance in this population. The evidence about the effectiveness of screening procedures in women at average risk of breast cancer is summarised below.

Existing procedures

Mammography

Mammography is the primary breast imaging modality for screening asymptomatic women aged 50-69 years; for the investigation of symptomatic women aged 35 years and over; and for the follow up of women with a previous history of breast cancer (National Breast Cancer Centre 2002). A conventional mammogram consists of two sets of low dose X-rays gained from placing the patient's breasts between two plates, compressing the breast and pulling the breast tissue away from the chest wall. The X-rays include a view from the side (medio-lateral oblique) and the top (cranio-caudal). The procedure takes approximately 15 minutes.

Full field digital mammography is a new mammographic technology that uses a digital receptor instead of the conventional screen film. An ability to manipulate the image acquisition and display may lead to improved cancer detection and breast lesion diagnosis (Pisano et al 2000). A large study comparing the accuracy of digital mammography with conventional mammography has recently reported improved accuracy in women under the age of 50 years and those with radiographically dense breasts, but no difference in women overall, suggesting a role for digital mammography in young women (Pisano et al 2005). The authors also reported that digital mammography resulted in a lower average dose of radiation than conventional mammography and, therefore, is a potentially safer test. MSAC is currently assessing the safety, effectiveness and cost-effectiveness of digital mammography (MSAC Reference 37).

Mammography is the only breast screening test to be evaluated by randomised controlled trials. Systematic reviews of these trials have reported that mammography reduces breast cancer mortality. The validity of these findings has been vigorously debated in the past when it was suggested that only two of the eight published trials were methodologically adequate, and neither indicated an effect of screening on breast cancer mortality or on total mortality (Gotzsche et al 2000). To address this question, a number of groups reviewed the data from each of the trials and concluded that mammographic screening is effective with estimates of reduced breast cancer mortality ranging from 16% to 35% (Boyle 2003).

A subsequent meta-analysis of the eight breast cancer trials showed a 24% mortality reduction associated with the invitation to screening (Duffy et al 2002), Table 2). The individual trials that reported the greatest mortality reductions showed a lower risk of node-positive breast cancer with screening, confirming an association between the detection of earlier stage disease and reduced mortality.

Table 2 Mortality reduction from published results of breast cancer screening trials

Study	Age range	Percentage mortality reduction with invitation to screening (95% CI)
Health Insurance Plan (HIP-New York)	40-64	24 (7-38)
Malmo	45-69	19 (-8-39)
Two-County Trial (Kopparberg, Ostergotland), Sweden	40-74	32 (20-41)
Edinburgh	45-64	21 (–2-40)
Stockholm	40-64	26 (–10-50)
Canada NBSS-1	40-49	–3 (–26-27)
Canada NBSS-2	50-59	–2 (–33- 22)
Gothenburg	39-59	16 (–39-49)
All trials combined	39-74	24 (18-30)

Source : Extracted from Smith et al 2003, based on data from Duffy et al 2002.

Reviews undertaken by Swedish investigators, the US Preventive Services Task Force (USPSTF) and the International Agency for Research on Cancer all found the effect of mammography to be higher in women over the age of 50 years (Boyle 2003). The American Cancer Society recommends screening from age 40 years (Smith et al 2003), although the evidence of reduced mortality is less clear for women aged 40-49 years. Pooled trial data have suggested a 16% reduction in mortality for women 40-49 years at 10-14 years after the initiation of screening (Kerlikowske 1997), although the USPSTF review highlighted the difficulty in estimating the benefits and harms of screening across all age groups using data from these trials (Humphrey et al 2002).

With the exception of the Canadian National Breast Screening Study (NBSS-1), none of the above trials were specifically designed to evaluate mammography in age-specific subgroups. Tabar et al (1999) found a more modest reduction in mortality in women aged 40-49 years in a re-analysis of the Swedish Two-County trial data but no difference in survival over 16 years of follow up between women over 50 years and those younger than 50 years when tumours were grouped by size, nodal status or histological grade. The authors observed a higher prevalence of more aggressive tumours in younger women compared to women over 50 years and suggested a shorter screening interval may be appropriate for younger women (Tabar et al 1999).

The National Health Service (NHS) screening program is currently conducting a randomised trial, 'the Age Trial', to assess the effectiveness of screening in women aged 40-49 years. The trial has randomised 160,921 women since 1991 in the ratio 1:2 to intervention (annual mammography) and control arms. The trial is due to report in late 2009 (Moss et al 2005).

None of these trials have specifically assessed the impact of screening in high-risk women. However, there is evidence that mammography performs less well with reduced sensitivity in this population. Factors associated with reduced mammography sensitivity and specificity are outlined below.

Factors influencing the sensitivity of mammography

Patient, tumour and technical factors have been shown to influence the sensitivity of mammography. The major patient-related factor is increased mammographic breast

density, which can obscure the radiological features of cancer (Buist et al 2004; Kerlikowske et al 1996). Increased mammographic density is more common in pre-menopausal women than post-menopausal women but does not appear to be associated with BRCA mutation carrier status (Warner et al 2001). Decreased mammography sensitivity has also been generally observed in younger women and those with a genetic predisposition to breast cancer compared to older average-risk women. This difference, as determined by higher rates of cancers presenting in the interval between routine annual mammograms (referred to as 'interval cancers'), may be attributed to the faster growth of breast cancers in these populations (Buist et al 2004; Halapy et al 2005). One small retrospective study observed that 6 of 13 (46%) of women with BRCA mutations undergoing annual mammography screening presented with palpable cancers in the interval between screening tests (Komenaka et al 2004).

Tumour factors other than disease aggression may also contribute to the decreased sensitivity of mammography to detect BRCA-associated cancers. It has been reported that some of these cancers give a benign appearance on mammography (Tilanus-Linthorst et al 2002). In addition, ductal carcinoma in situ (DCIS), a localized early-stage breast cancer that is often associated with micro-calcifications, shows up well on mammography but is not common in BRCA-associated cancers (Marcus et al 1996).

Technical factors, such as the quality of the mammogram film and reader interpretation, may also influence the sensitivity of mammography to detect breast cancer.

Factors influencing the specificity of mammography

The risk of experiencing a false positive result with mammography has been reported to vary according to a woman's individual risk profile (age, family history of breast cancer, mammographic density and history of previous biopsies) and factors related to the interpreting radiologist (Christiansen et al 2000; Smith-Bindman et al 2005). An observational study from the United States (209 physicians, 1,220,046 screening mammograms) found that after adjusting for patient factors, lower mammography specificity was associated with interpretation by radiologists with less experience in clinical practice; exposure to a lower volume of mammograms each year; and a lower load of screening versus diagnostic mammograms (Smith-Bindman et al 2005). Another study suggested that lower mammography specificity was associated with a low threshold for reporting an abnormal result, the lack of comparison of the index mammogram with previous films and a longer interval between films (Christiansen et al 2000).

Disadvantages of mammographic screening

In addition to the reduced sensitivity of mammography in detecting breast cancer in young women with increased mammographic density, the two other main disadvantages of mammographic screening are the potential harms of false positive results and the potential carcinogenic effect of repeated exposure to ionising radiation. Women carrying BRCA1 and BRCA2 mutations are believed to be at higher risk of radiation-induced breast cancer because these mutations result in a deoxyribose nucleic acid (DNA) repair defect.

Women who undergo screening by mammography have a cumulative risk of experiencing a false positive result. A false positive finding may cause marked anxiety until further testing excludes cancer. US data has suggested that a 60 year old women screened annually has a 47% chance of having one false positive after ten mammograms

and after 10 years of follow up testing has a 19% chance of a false positive result leading to biopsy (Elmore et al 1998). These figures represent US practice patterns where it is standard to screen women annually. In Australia, it is standard practice to screen women every two years. Generally women are aware of the possibility of a false positive result and there is some evidence to suggest this risk is regarded as an acceptable consequence of screening mammography (Schwartz et al 2000).

Mammography also confers a small risk of radiation-induced breast cancer. This risk increases with the amount of exposure and with younger age at exposure, and is potentially greater for women with a family history of breast cancer (Berrington de Gonzalez et al 2005). Therefore, mammography may be more harmful for screening high-risk women who begin screening before the age of 50 years and present for more screening tests over a lifetime than women at average risk.

Another potential disadvantage of mammography screening is the possibility of diagnosis of clinically non-significant disease (Duffy et al 2005). There is evidence from evaluations of screening programs that this may occur largely as a consequence of increased detection of ductal carcinoma in situ (DCIS), which accounts for up to 20% of cancers detected with mammographic screening. The natural history of DCIS is unclear, but once detected, an increasing number of women undergo invasive treatment for lesions that may never have become clinically apparent (Ernster et al 1996). A disease 'reservoir' of DCIS has been found in 10% of women undergoing autopsy who were not known to have breast cancer during life (Welch et al 1997).

Breast ultrasound

Breast ultrasound is commonly used to complement mammography, particularly in the investigation of breast symptoms in women younger than 35 years of age and in women who are pregnant or lactating (NBCC 2002). Ultrasound is also used to evaluate palpable lesions not seen on mammography; detect an associated underlying mass; evaluate breast implants; and to guide the biopsy or surgical excision of breast tumours (NBCC 2002).

The role of breast ultrasound in screening young women at high risk of breast cancer has not been established (NBCC 2002) and its use varies by centre in Australia. Some clinicians use it routinely to screen all young high-risk women; others use it selectively, for example in young women with increased mammographic density (Advisory Panel, March 2006). A systematic review of the accuracy of screening tests for breast cancer has identified evidence that ultrasound increases the sensitivity of mammography in detecting cancers for women with mammographically dense breasts and those assessed as at high risk of breast cancer, but also results in an increase in the rate of false positive findings (Irwig et al 2004).

Emerging procedures

Breast tomosynthesis

Breast tomosynthesis involves taking serial X-Ray slices of the breast from different angles and reconstructing them to produce high resolution 'slices' for a three dimensional image of the breast (Smith 2005). Its potential advantages over conventional mammography include: improved imaging of dense breasts; reduced discomfort; and non-invasive assessment of lymph node spread (Reddy et al 2005). A potential

disadvantage of tomography is that it exposes women to a higher radiation dose than conventional mammography. Currently tomosynthesis is still being investigated and is used for research purposes only.

Classification of imaging abnormalities

Methods for the standardised reporting and classification of mammography findings have been developed to improve communication among clinicians and other health care professionals. In Australia, this involves the use of a summary of essential information presented in a checklist format with classification of abnormalities into five groups: (1) no significant abnormality; (2) benign findings; (3) indeterminate/equivocal findings; (4) suspicious findings of malignancy; (5) malignant findings (NBCC 2002).

The American College of Radiology recommends a five-level Breast Imaging Reporting and Data System (BI-RADS): (0) requires further tests; (1) negative; (2) benign; (3) indeterminate, probably benign; (4) suspicious of malignancy; (5) highly suggestive of malignancy (American College of Radiology 2003). This is the most commonly used international system for the conduct and reporting of clinical trials. The major difference between the Australian and the American system is in the classification of level 3 'indeterminate' abnormalities. In Australia, level 3 findings are always followed up with further investigations while this is not the case in the American classification system.

Breast Cancer Screening in Australia

Current Recommendations from the National Breast Cancer Centre

The NBCC guidelines recommend mammographic screening for asymptomatic women who are:

1. at average risk and aged 50-69 years (screening recommended every two years)
2. at higher risk of breast cancer including those with a:
    - previous diagnosis of invasive cancer
    - previous diagnosis of ductal carcinoma in situ
    - previous diagnosis of lobular carcinoma in situ or atypical hyperplasia
    - strong family history of breast cancer

For women of all ages who are at increased risk of developing breast cancer it is recommended that an individualised surveillance program be developed in consultation with the woman's general practitioner and specialist.

BreastScreen Australia

BreastScreen Australia is the national population-based screening program which is targeted to asymptomatic women at average risk of breast cancer. It provides free screening mammograms at two-yearly intervals for women aged 50-69, although women aged 40-49 and 70 years and older are also eligible for screening.

The State and Territory Governments have primary responsibility for the implementation of the program at their local level. The Australian Government provides overall coordination of policy formulation, national data collection, quality control, monitoring and evaluation.

Further assessment following a positive mammogram may include:

- clinical examination
- further mammographic work-up
- breast ultrasound
- fine needle aspiration cytology
- core biopsy
- open surgical biopsy.

Familial and high-risk cancer clinics

Individuals with a strong family history of breast and/or ovarian cancer may be referred for assessment at a specialised familial cancer clinic. Genetic counselling and risk management services are tailored according to individual risk. Women assessed at high risk for breast cancer require long-term management. Specialised services to focus on the ongoing management of this patient group are being established.

Currently women assessed as high risk are offered annual screening with mammography (see page 8). Ultrasound may also be included for screening pre-menopausal women, or used selectively in those with increased mammographic density (see page 11).

Breast MRI for women at high risk of breast cancer

Potential advantages of Breast MRI

Breast MRI is proposed as a more sensitive test than mammography for detecting early breast cancer. This may lead to improved health outcomes if early detection and treatment reduces mortality from breast cancer in the screened population.

As discussed on page 9, mammography performs less well in young women, in particular in those with increased mammographic breast density. In contrast, areas of increased radiological density do not affect MRI-acquired images. Breast MRI is also proposed as a safer test than mammography because it does not use ionizing radiation.

Possible disadvantages of Breast MRI

The potential advantages of using breast MRI for screening need to be assessed against the possible disadvantages. These include:

- Lower test specificity requiring more women to undergo further unnecessary imaging tests and biopsies to assess false-positive results.

- Increased anxiety due to these false-positives.
- Anxiety related to the MRI procedure itself. Up to 30% of women experience anticipatory anxiety with MRI, similar to the anxiety of elective surgery (Melendez et al 1993), although a comparison with other screening tests has not been undertaken. Women who suffer claustrophobia may be unable to tolerate MRI.
- Detection of inconsequential disease. It is not known whether breast MRI may detect some cancers that would not have otherwise become clinically apparent.
- The necessary exclusion of some groups of women from having MRI due to contraindications such as cardiac pacemakers.
- Lack of information about the effects of repeated imaging in high-risk groups.
- The added cost of the MRI procedure.

Considerations for Breast MRI screening protocols

Other issues relevant to the consideration of breast MRI as a screening test for young women at high risk of breast cancer include:

- Method of recruitment.
- Timing of breast MRI screening considering the potential for interval cancers in high-risk women screened annually.
- Optimal sequence of screening tests: mammography, ultrasound and breast MRI.
- Optimum age to commence screening with breast MRI.
- Potential application in other groups, for example:
    - Women at increased risk of breast cancer due to non-genetic factors such as prior personal history of breast cancer; or
    - Detection of cancer recurrence after breast conserving surgery.
- Appropriate education and counselling of women considering undergoing MRI as a screening procedure.

Breast cancer diagnosis, staging, treatment and prognosis

The purpose of medical screening is to prevent disability and death and improve quality of life (Strong et al 2005). As described on page 13, implementation of breast MRI within a breast screening program for young high-risk women is proposed to achieve these outcomes by detecting cases at an earlier stage of disease that will benefit from treatment. General information about the diagnosis, staging, treatment and prognosis of early breast cancer relevant to the assessment of evidence about the effectiveness of breast MRI is presented on page 15.

Pathological diagnosis of breast cancer

The diagnosis of breast cancer is made by pathological examination of suspicious lesions detected clinically or by imaging tests. This is undertaken using fine needle aspiration cytology, core needle biopsy and/or open surgical biopsy of the lesion. These procedures may sometimes require imaging guidance by mammography, ultrasound, or MRI to locate the lesion.

Breast cancer staging and grading systems

The three main factors that determine breast cancer prognosis are tumour size, histological grade and extent of spread (Haybittle et al 1982). Molecular and biological markers such as estrogen receptor expression are also used to assess prognosis and guide treatment (Fitzgibbons et al 2000).

Tumour staging is undertaken using information about the size of the primary tumour and extent of spread.

Table 3 shows the criteria used to classify cancers into five stages using the TNM system (tumour, lymph node, metastases).

Table 3 Staging of breast cancer[1]

Stage	TMN classifications
Stage 0 Tis, N0, M0 Stage I T1, N0, M0 Stage II IIa T0-1, N1, M0 T2, N0, MO IIb T2, N1, M0 T3, N0, M0 Stage III IIIa T0-2, N2, M0 T3, N1-2, M0 IIIb T4, N0-2, M0 IIIc Any T, N3, M0 Stage IV any T, any N, M1	Tumour (T) TX: primary tumour cannot be assessed Tis: carcinoma in situ, intraductal carcinoma, lobular carcinoma in situ or Paget's disease of the nipple with no associated tumour mass T1: ≤ 2 cm T2: 2-5 cm T3: > 5 cm T4: any size growing into the chest wall or skin Lymph node spread[2] (N) NX: Regional lymph nodes cannot be assessed N0: no spread to the lymph nodes N1: spread to moveable axillary lymph nodes on the same side as the breast cancer. (pathological staging includes pN1a: 1-3 axillary lymph nodes) N2: spread to fixed or matted axillary lymph nodes; or clinically apparent in internal mammary nodes in the absence of axillary node spread nodes – all on the same side as the breast cancer (pathological staging includes pN2a: 4-9 axillary lymph nodes) N3: spread to infraclavicular lymph nodes or clinically apparent internal mammary nodes in the presence of positive axillary lymph nodes; or to infraclavicular lymph nodes with or without axillary lymph nodes – all on the same side as the breast cancer. (pathological staging includes pN3a: ≥ 10 axillary lymph nodes) Metastatic disease (M) MX: distant spread cannot be assessed M0: no distant spread M1: distant spread

1 Adapted from Singletary et al (2006) from 6th edition of American Joint Committee on Cancer Staging Manual

[2] Classification based on clinical examination and imaging, refer to Singletary et al (2006) for further information about pathological staging

Tumour grading is undertaken using information about the histological appearance of the tumour cells. This involves microscopic examination of breast cancer tissue to classify tumour cells according to how abnormal they look. Tumours are classified into three grades using the Elston-modified Scarf–Richardson–Bloom system according to the percentage of tubule formation, shape of the nucleus and the percentage of tumour cells that are dividing (Elston et al 1991). Grade 1 tumours are well-differentiated cancers with relatively normal-looking cells that do not appear to be growing rapidly; grade 2 tumours are moderately differentiated; and grade 3 tumours are poorly differentiated cancers that lack normal features. Grade 3 tumours tend to grow and spread more aggressively with a 5-year survival of 50% compared to 95% and 75% for grade 1 and 2 tumours respectively (Elston et al 1991).

Treatment and prognosis

The NHMRC has produced Australian guidelines for the treatment of early breast cancer (NHMRC 2001). Treatment varies according to the stage of disease at diagnosis and other prognostic factors such as the hormone receptor status of the cancer. Standard treatment involves surgery to remove the primary cancer and excision of the axillary lymph nodes under the arm to assess and manage local spread of the disease. Depending on the stage and receptor status of disease, surgery is followed by radiotherapy, chemotherapy and/or endocrine therapy to control further spread of the disease or prevent recurrence. Breast-conserving surgery followed by local radiotherapy is offered as an alternative to total excision of the breast (total mastectomy) for the removal of small unifocal cancers that show clear histological margins (NHMRC 2001).

Less invasive procedures such as breast-conserving surgery may potentially reduce the side effects related to surgery without compromising the effectiveness of treatment for women with early breast cancer. A specialised procedure to identify and examine the first lymph node that receives drainage from the lymph vessels of the breast (sentinel node biopsy) is also being investigated as an alternative to complete axillary dissection for identifying patients with lymph node spread. Sentinel node biopsy was reviewed by MSAC in 2005 to assess whether the procedure could identify patients for whom axillary clearance is not indicated without increasing axillary recurrence rates or decreasing long-term survival. MSAC recommended that sentinel node biopsy appears to be safe and effective in identifying sentinel lymph nodes resulting in the reduction of complications due to axillary lymph node dissection, in particular lymphoedema; however, the long-term health outcomes are uncertain (Medical Services Advisory Committee 2005).

The Australian Sentinel Lymph Node Biopsy versus Axillary Clearance (SNAC) trial aims to determine whether breast cancer outcomes following sentinel node biopsy are equivalent to those following axillary clearance with reduced morbidity. Accrual to the trial ceased in May 2005 with 1,088 women recruited. Interim results suggest that sentinel node biopsy has a sensitivity of 92%, a specificity of 100% with a false negative rate of approximately 8% (Wetzig 2005).

Treatment response varies according the stage, grade and receptor status of disease, age at diagnosis and other factors. The five-year relative survival rate for Australian women diagnosed with breast cancer during 1992-97 was 84% (Cancer in Australia 2001). Stage-specific survival rates are not available nationally. Data from the South Australian Cancer Registry and 138,495 cases recorded on the National Cancer Database in the United States are shown in Table 4. Relative 5-year survival rates (excluding deaths from other causes) are higher for localized disease. More recent data from the US National Cancer

Database show a relative survival rate of 100% at 5 years for women with ductal carcinoma in situ and Stage I cancers, reducing to 20% for Stage IV cancers (American Cancer Society 2005).

Table 4 5-year survival for women diagnosed with breast cancer by stage at diagnosis

Stage	South Australian Centre for Cancer Control Research 1997-1998[1]	US National Cancer Database 1998-2001[2]
I	92%	89%[3]
II	78%	80%
III	51%	56%
IV	19%	19%

1 Roder & Centre for Cancer Control Research 2002 (data from other states not available);

2 US National Cancer Database 2006,

3. 5-year survival for stage 0 breast cancer 94%

Comparator

This report will compare the addition of breast MRI to a surveillance program that involves annual mammography with or without annual ultrasound, clinical breast examination (CBE) every six months and regular self breast examination.

Marketing status of the technology

MRI is currently available in public and private facilities in major centres in each state and territory. One hundred MRI sites have been licensed by the Department of Health and Ageing to provide services that are eligible for funding under the Medicare Benefits Schedule (MBS). Not all of these sites currently have the facilities to perform breast MRI (DoHA 2006). Unlicensed sites that are ineligible for MBS funding may also provide breast MRI.

Current reimbursement arrangement

Breast MRI is not currently reimbursed as a screening or diagnostic test under the MBS. MRI is funded for a range of head and spine, musculoskeletal and cardiovascular conditions under the MBS (item numbers 90901-00 to 90901-08), although this funding is restricted to licensed providers.

Mammography is funded by the joint Australian and State/Territory Government BreastScreen Australia Program as a population-based screening test targeted to women aged 50-69 years. Women aged 40-49 years and aged 70 years and over are also eligible for a screening mammogram every two years under this program. Mammography is also subsidised under the MBS as a diagnostic test for individuals with symptoms, a past history of breast cancer, or those at high risk of breast cancer due to family history (MBS item numbers 59300-59318).

Methodological issues relevant to the evaluation of breast cancer screening programs

The WHO Consultation Group on the methodology of noncommunicable disease screening specified nine requirements that must be satisfied before implementing a new screening program (Table 5). These guidelines broadly define medical screening as testing, including targeted screening, undertaken with the goal "to benefit the individuals being screened, reducing disability and the risk of death" (Strong et al 2005).

Table 5 Requirements for implementing a medical screening program

1. The disorder is well defined.
2. The prevalence of the disorder is known.
3. The natural history of the disorder has a significant impact on the quality of life and survival of a large proportion of the screened population.
4. The screening test is safe and simple.
5. The performance of the screening test has been established by assessing the detection rate (or sensitivity) and the false positive rate (or specificity) of the test.
6. The costs of the screening program can be justified by the health gains achieved.
7. Facilities exist to support the screening program. The total resources and costs required for screening, diagnosis, treatment and follow-up will need to be determined to assess whether the screening program is feasible.
8. The screening test is acceptable to the target population.
9. There is equity of access to screening services. If there are inadequate resources to provide basic preventive and treatment services for the entire population, the introduction of a new screening program may increase existing inequities in the allocation of health resources away from those in greatest need.

Source: Strong et al (2005)

Breast MRI is proposed for the targeted screening of young women assessed as having a high incidence of breast cancer (≥ 30% lifetime risk), rather than as a population-based screening test. For those affected, the clinical burden of disease is well documented and substantial. Breast MRI therefore clearly satisfies the first three criteria for implementing screening programs listed above. Decisions about the implementation of breast MRI may also be influenced by other factors relevant to this specific technology and target population (see page 20).

Criteria 4 to 8 are addressed by the systematic review presented in the following sections of this report. Criterion 9 emphasizes the need for health policy decisions about screening services to be based on careful consideration of health service equity.

The WHO consultation group emphasized that a screening program should only be implemented when the value of screening can be determined (criterion 6). This involves quantifying the impact of the test on patient morbidity and mortality and weighing these benefits against the financial costs and harms related to testing such as patient anxiety, discomfort, and adverse effects of any follow-up investigations and treatments (Strong et al 2005). As such, evidence that a screening test leads to the early detection of disease is insufficient, by itself, for conclusions about its clinical value.

Randomised controlled trials comparing breast cancer morbidity and mortality for screening strategies with and without breast MRI would provide the most valid evidence about the clinical value of breast MRI for screening high-risk women. As described on page 8, randomised trials have demonstrated that mammography screening programs in women aged 39 to 77 years at average risk of breast cancer lead to reductions in breast

cancer mortality compared to no screening. However, it is not clear whether the addition of new tests for the earlier detection of breast cancer targeted to young high-risk women will produce a similar advantage compared to mammography.

In the absence of randomised trials to establish the relative effectiveness of adding breast MRI; this review presents additional relevant evidence. In theory, a new test is more effective than an old test if it is more accurate than the old test, leads to a change in treatment and this treatment is effective or at least less harmful. Given the assumption that all MRI-detected cancers will be treated according to standard breast cancer treatment protocols, the two critical issues for the evaluation of the addition of breast MRI versus mammography alone are therefore:

1. does MRI detect more cases of breast cancer than mammography through earlier detection without increasing the harms of testing? and

2. does early detection of these extra cases lead to improved treatment response and/or reduced treatment harms?

Studies of test safety and accuracy comparing the new test with mammography will suffice to address the first question. Accuracy studies can show whether the new test is more sensitive than the old test (detects extra cases of true disease) and/or whether it is more specific (detects fewer false positives). If the addition of breast MRI is more sensitive but less specific than mammography, the potential harms associated with the extra false positive diagnoses need to be considered.

The second question requires evidence from randomised trials comparing patient mortality for cases detected early by breast MRI versus cases detected later by screening mammography. Trials assessing patient mortality from the time patients are allocated to screening programs that include MRI versus those that use mammography alone (trials of the test-and-treatment strategy) will provide the most valid evidence to address this question. Trials comparing patient mortality from the point of diagnosis (as occurs in treatment trials) may be subject to lead-time bias (where the new test detects an earlier stage of disease but there is no true survival benefit if the lead-time between detection by new test and detection by mammography or clinical presentation is taken into account). Length bias, where a screening test detects less aggressive cancers than those presenting clinically, may also be relevant to the comparison of breast MRI and mammography.

Therefore, in the absence of randomised trials comparing different screening strategies, we need to be convinced that treatment response is improved for any extra cases detected by breast MRI, and/or treatment harms are reduced, to conclude that breast MRI will improve patient outcomes. Evidence showing reduced mortality for screen-detected versus clinically detected cancers from mammographic trials in other patient groups may not apply to breast MRI-detected cases in young high-risk women.

Randomised trials assessing the impact of breast MRI screening on the incidence of cancers presenting with symptoms in the intervals between screening tests (referred to as 'interval cancers') can be used to investigate the possibility of length or time bias. A reduction in the incidence of interval cancers excludes the possibility that MRI just detects less aggressive cases of disease that would never have reached clinical significance. Unfortunately, avoidance of lead-time bias is more difficult to achieve without randomised trials of the entire test-and-treatment strategy.

If accuracy studies show that breast MRI leads to detection at an earlier stage of disease, then trials or observational studies comparing treatment response for early versus late stage disease in young high-risk women may provide evidence to support the value of adopting breast MRI. Trials comparing treatment decisions for breast MRI-detected cases versus mammography-detected cases may also provide evidence to support the use of breast MRI if it can be shown that early detection leads to more conservative treatment and fewer adverse events without compromising patient survival.

In summary, this review relies on a combination of evidence about the relative accuracy and safety of adding breast MRI to mammography screening programs and the relative effectiveness of early versus later detection of breast cancer in young high-risk women to determine the relative effectiveness of breast MRI. Expert opinion is used to assess the generalisability of evidence about breast MRI accuracy to its proposed use in Australia and the validity of linking evidence about test accuracy and treatment response for conclusions about potential health gains using breast MRI.

Other considerations

In addition to the findings of the systematic review of the evidence, other issues that may be relevant to decisions about breast MRI funding are discussed below.

The size of the target population

It is currently proposed that breast MRI should be offered to high-risk women aged 35-50 years or commencing at 5 years prior to the index case for those with a family member diagnosed before the age of 40 years. Australian census data from 2001 estimate there are 2,258,381 women aged 35-50 yrs (ABS 2001). Genetic epidemiological studies indicate that less than 1% of this population will be at high risk of breast cancer due to genetic disposition (NHMRC 1999).

In the absence of evidence of its effectiveness, population screening for breast cancer susceptibility genes is not recommended (Nelson et al 2005). Therefore, not all women with a genetic predisposition that would make them eligible for breast MRI screening will be identified. Expert opinion suggests that familial and high-risk cancer clinics may be expected to be able to identify approximately 20% of these women (Australian Breast MRI Working Group 2005). Based on this estimate, and the assumption that an additional 20% of high-risk women aged 25-34 years would be eligible for earlier screening because a family member was diagnosed younger than 40 years (0.2% of 1,380,683 women aged between 25 and 34 years from ABS 2001), approximately 5,000 women will be identified as eligible for a breast MRI screening program. This estimate should be considered with caution and is dependent on factors including: the capacity of familial and high-risk cancer clinics to identify and recruit close relatives of women with a known genetic predisposition; the specific criteria used to predict breast cancer risk and define eligibility for breast MRI screening; the level of awareness among health professionals and the community about the appropriate referral of women with a family history of breast and ovarian cancer; and patient acceptance of and access to breast MRI.

Access to the technology

The number of MRI machines funded to provide services under the MBS has been continually increasing since 1998. There are currently 100 funded MRI machines in operation in Australia, and a further 12 sites planned (The Diagnostic Imaging Section,

Diagnostics and Technology Branch, DoHA, 31 July 2006). Information about the number of unfunded MRI machines is not collected.

Many of the existing MRI facilities may not currently have the technology required to perform breast MRI (dedicated breast coil and software) or personnel experienced in the technique. A proportion of individuals screened by breast MRI will show a lesion that cannot be localised by mammography or ultrasound and therefore will require an MRI-guided biopsy. Additional MRI hardware and MRI-compatible needles are required to undertake this procedure. MRI-guided biopsy is not currently specified under the MBS, although it would be eligible for reimbursement under item 31506 (excision biopsy detected by mammography or ultrasound where guidewire or other localisation procedure is performed, DoHA 11 July 2006).

Provision of breast MRI within multidisciplinary cancer screening and detection services

Mammography screening is currently mostly undertaken by experienced health professionals working within dedicated breast cancer screening services (see page 12). This service model has been developed to support appropriate and efficient investigation, counselling and access to treatment. If introduced, breast MRI screening is likely to benefit from the same model of service provision. This would require the development of explicit methods for the recruitment and assessment of women potentially at high risk for breast cancer and criteria to determine their eligibility for screening breast MRI.

Identification of high-risk women is currently based on the referral of women with a diagnosis of breast cancer and a family history that suggests a genetic predisposition to specialised familial and high-risk cancer clinics for genetic assessment and counselling, and the identification or self-referral of close relatives of women assessed as having a genetic predisposition. Women classified as high risk are offered intensive surveillance. It is proposed that breast MRI screening could be implemented as part of these existing surveillance programs. This would support the appropriate referral and counselling of women eligible for testing, as well as ongoing development of expertise in the technique.

The feasibility of a classical randomised controlled trial to assess reduction in mortality due to the addition of breast MRI

A prospective randomised controlled trial reporting on breast cancer mortality would provide the most valid assessment of the relative effectiveness of different screening programs for high-risk women. However, there are two important practical barriers to conducting such a trial to assess reduction in breast cancer mortality due to the addition of breast MRI: (i) sample size and follow-up period requirements; and (ii) patient acceptability/ethics of randomisation.

To illustrate the first point: based on an estimated breast cancer mortality rate of 6 per 1,000 among high-risk women, sample size calculations suggest that a trial would need to recruit approximately 12,800 women for screening over a 5 year period, with at least 10 years follow-up, to have an 80% power to detect a 25% reduction in patient mortality. This estimate assumes that 100% compliance can be achieved and accepts a 5% chance of detecting a reduction in mortality when a true reduction does not exist. While this type of trial design is necessary to quantify the benefits of breast MRI in terms of patient outcomes, comparison of interval cancer rates (a shorter-term outcome) in patients randomised to each screening strategy may be a more feasible method to investigate the

efficacy of breast MRI. If the addition of breast MRI reduces the rate of interval cancers compared to conventional mammography alone, this provides strong evidence that breast MRI detects clinically significant disease that will benefit from early treatment.

The second critical point is that given the evidence discussed in this report that the addition of breast MRI leads to the early detection of breast cancer, it may no longer be ethical to randomise high-risk women to conventional mammography alone to quantify the benefits in terms of reduced mortality.

If randomised controlled trials are considered unethical or unfeasible, alternative methods for evaluating the impact of breast MRI screening include large accuracy studies designed to test the hypothesis that the addition of breast MRI detects cases at an earlier stage and/or lower grade than cases detected by mammography alone. This information may provide supportive evidence for conclusions about the effectiveness of breast MRI.

If funded, services that offer breast MRI may also collect prospective data about cancer detection rates, false positive recall rates, interval cancer detection rates and patient mortality. However, historical comparisons between these data and data collected prior to the introduction of breast MRI (if available) would only provide weak evidence about the relative effectiveness of breast MRI, owing to the potential for confounding bias from factors other than breast MRI that could be attributed to any differences detected (for example, improved screening practices other than MRI and improved treatments).

Approach to assessment

Research questions

Evaluators from the National Health and Medical Research Council (NHMRC) Clinical Trials Centre worked with members of the Advisory Panel to develop research questions to assess the value of breast MRI as a screening test for breast cancer in high-risk women. These questions were formulated *a priori* from information provided by the applicant and the Advisory Panel using a flowchart depicting the clinical pathway for screening for breast cancer to illustrate the potential role of breast MRI as a screening test (Appendix D).

Primary review question

1. What are the safety, effectiveness and cost-effectiveness of annual breast MRI in addition to annual mammography with or without breast ultrasound for screening asymptomatic high-risk women under the age of 50 years?

Secondary review questions

2. What are the safety, effectiveness and cost-effectiveness of annual breast MRI in addition to annual mammography with or without breast ultrasound for screening asymptomatic high-risk women aged 50 years or older?

3. What are the safety, effectiveness and cost-effectiveness of annual breast MRI versus mammography with or without breast ultrasound for screening asymptomatic high-risk women under the age of 50 years?

Assessment strategy

The evaluation team conducted a systematic review of the medical literature to address these three research questions. As described on page 18, the effectiveness of a new screening test depends on whether it improves patient outcomes. This is determined by the accuracy of the new test compared to standard screening strategies, whether the information provided by the new test changes patient management and the relative effectiveness of early treatment versus treatment at the time of diagnosis using standard screening strategies. In the absence of randomised controlled trials assessing the effectiveness of adding MRI to standard mammography screening protocols, this review included studies that assessed one or more of the following outcomes:

- the relative safety of MRI compared to mammography alone;
- the relative accuracy of MRI with or without mammography compared to mammography alone;
- the relative effectiveness of treatment for cases detected early by breast MRI plus mammography versus cases detected later by mammography; and

- patient tolerance and preferences for the addition breast MRI to mammography screening programs.

Using this approach, conclusions about the effectiveness of breast MRI depend on judgement about the validity of linking evidence of test accuracy with the evidence available about treatment effects.

It is assumed that all MRI-detected cancers will be offered treatment according to standard breast cancer treatment protocols. It is possible that MRI may bring additional patient benefits if the extra cases detected are at an earlier stage of disease where more conservative therapies may be possible. Therefore, if available, studies evaluating the impact of adding MRI on clinical management decisions were also eligible for this review, for example, studies investigating whether early detection of breast cancer by MRI results in the reduced use of systemic chemotherapy.

Assessment of the cost-effectiveness of MRI for breast cancer screening was limited by the lack of direct trial or observational evidence to quantify the effects of early detection for young high-risk women. Therefore, this report appraises the validity of a cost-effectiveness analysis undertaken in the US that uses a mathematical model of the natural history of breast cancer to estimate the effects of early detection. An assessment of the financial implications of introducing MRI in Australia for screening young women at high risk of breast cancer is also presented.

Review of literature

A systematic review was undertaken to identify relevant studies published up to March 2006. Websites of the international health technology assessment (HTA) agencies were searched for existing HTA reports research (Table 6) and electronic databases of published research (Table 7) were searched for original research papers, including systematic reviews.

Table 6 Electronic databases and HTA websites searched in this review.

Organisation	Database/website
NHS Centre for reviews and dissemination databases	
Economic evaluation database (EED)	www.york.ac.uk/inst/crd/
Database of abstracts of reviews of effectiveness (DARE)	
Heath Technology Assessment (HTA)	
Cochrane Database of Systematic Reviews and Cochrane Controlled Trials Register	www.cochrane.org
British Columbia Office of Health Technology Assessment (Canada)	www.chspr.ubc.ca
Swedish Council on Technology Assessment in Healthcare (Sweden)	www.sbu.se
Oregon Health Resources Commission (US)	www.ohppr.state.or.us/index.html
Minnesota Department of Health (US)	www.health.state.mn.us/htac/index.htm
Canadian Coordinating Office for Health Technology Assessment (Canada)	www.ccohta.ca
Alberta Heritage Foundation for Medical Research (Canada)	www.ahfmr.ca
National Library of Medicine Health Service/Technology Assessment Text (US)	www.ncbi.nlm.nih.gov
Institute for Clinical Evaluative Science (Canada)	www.ices.on.ca
DIMDI - German Institute for Medical Documentation and Information	www.dimdi.de
National Information Centre of Health Services Research and Health Care Technology (US)	www.nlm.nih.gov/nichsr
Finnish Office for Health Technology Assessment (FinOHTA) (Finland)	www.stakes.fi/finohta/linkit/
Institute Medical Technology Assessment (Netherlands)	www.imta.nl
Agence Nationale d'Accreditation et d'Evaluation en Sante (France)	www.anaes.fr
Agence d'evaluation des technologies et des modes d'intervention en sante (AETMIS)	www.aetmis.gouv.qc.ca/en/index.php
National Institute for Health and Clinical Excellence (NICE) (UK)	www.nice.org.uk
National Coordinating Centre for HTA (NCCHTA) (UK)	www.hta.nhsweb.nhs.uk
Centre for Health Program Evaluation (Australia)	Chpe.buseco.monash.edu.au

Table 7 Electronic databases searched

Database	Period covered
EMBASE.com (includesMedline)	All years up to March 2006
Premedline	As at March 2006
Current Contents	January-March 2006

Search strategy

The search strategy was developed using the key elements of the clinical question. The search strategy shown in Table 8 was used to identify papers in the databases described in Table 7.

Table 8 Search strategy

Number	Search Strategy
1	Magnetic Resonance Imaging.mp. or exp Magnetic Resonance Imaging/
2	Magnetic Resonance Imag$.mp.
3	MRI.mp.
4	MR Imag$.mp.
5	or/1-4
6	breast cancer.mp. or exp Breast Neoplasms/
7	breast tumour.mp.
8	breast carcinoma.mp.
9	breast screening.mp.
10	or/6-9
11	5 and 10
12	limit 11 to humans
13	limit 12 to English language

Reference lists of included publications were also checked and experts in the area were contacted for relevant citations that may have been inadvertently missed in the searches of major databases.

Search results

Existing health technology assessment reports

The searches of the HTA agency databases and websites (listed in Table 6) identified six systematic reviews including three HTA reports and two clinical guidelines about the use of breast MRI for high-risk women. These six reports were all eligible for inclusion in the present review.

Published literature

The search strategy retrieved a total of 2,847 citations. The number of citations retrieved from each database is shown in Table 9.

Table 9 Number citations retrieved from each database

	Embase.com (including Medline)	Pre-Medline	Current Contents	Total
Number of citations	2,774	33	40	2,847

Eligibility criteria for studies

After excluding duplicate publications, 2,789 citations were appraised by two independent reviewers to determine eligibility using the criteria listed in Table 10. Discrepancies between reviewers about study eligibility were resolved by discussion.

Studies reporting on a head-to-head comparison of breast MRI plus mammography versus mammography (with ultrasound available at radiologist's discretion) provide direct evidence about the relative effects of these tests to address the primary question. Studies comparing breast MRI with mammography have also been included to address the secondary question about the potential use of breast MRI as a first-line replacement test for mammography. Information from studies that only report data for one of the tests of interest has not been used in the comparison of tests because indirect comparisons of results from different studies can introduce bias that may overestimate the effect of the intervention (or accuracy of the test) (Bucher et al 1997).

Studies investigating MRI in fewer than 20 patients were excluded because small studies are unlikely to achieve sufficient precision in their estimates of test outcomes for reliable interpretation. Large studies that identify few cancers would be unlikely to provide precise estimates of the sensitivity of a test but may still provide useful information about the relative specificity of tests and were included in the review.

Table 10 Study exclusion criteria

1. Not an appropriate clinical study

Reports describing animal, laboratory or scientific studies, technical reports or case reports will be excluded. Non-systematic narrative reviews, letters and conference abstracts will also be excluded in this category.

Case series where the use or reporting of the reference standard is based on the breast MRI result (positive/negative) will be excluded.

Case-control studies where patients were selected for inclusion in the study based on their known disease status will be excluded.

Retrospective case referent studies (reporting on subjects all known to have the condition of interest) will be excluded.

2. Wrong patient group

Studies must include asymptomatic women at high risk of breast cancer due to family history or genetic predisposition and no prior history of breast cancer, or women with an intact contralateral breast following mastectomy for a primary breast cancer. Studies including > 50% average risk asymptomatic patients and studies with < 20 patients undergoing breast MRI will be excluded.

3. Wrong diagnostic test

Studies are to perform breast MRI with or without mammography and breast ultrasound.

4. Wrong reference standard or comparator

Studies are to use a minimum of histology as the reference standard for positive tests and a consensus of all tests as the reference standard for negative tests.

Studies comparing two or more different techniques of breast MRI without performing a reference standard will be excluded.

Studies are to use mammography with or without breast ultrasound as a comparator.

5. Wrong outcomes

Studies must report on at least one of the following:

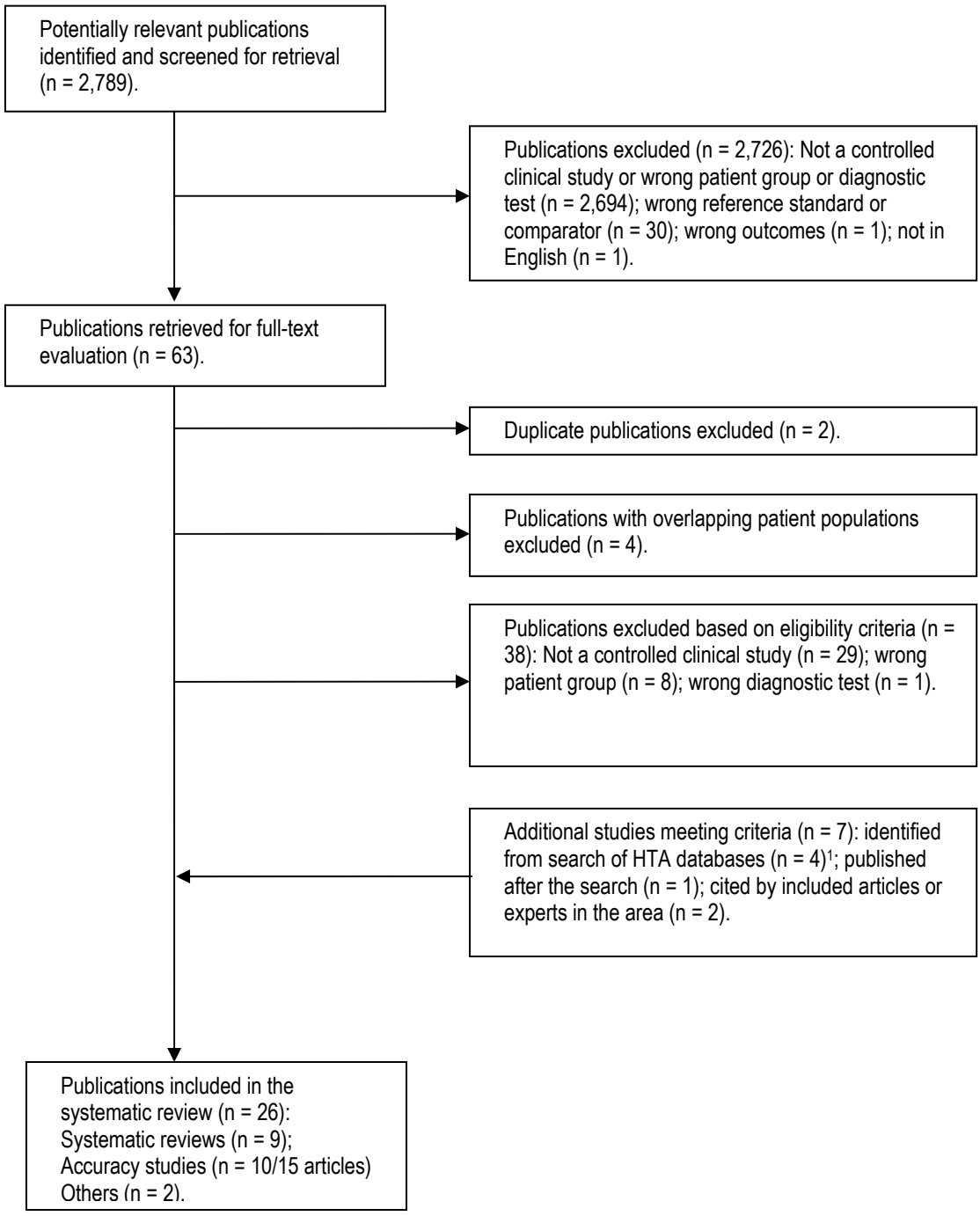
- number of true positives and false positives detected by breast MRI plus mammography versus mammography for calculation of the positive predictive value
- test recall rate and/or biopsy rate among non-diseased
- diagnostic accuracy with sufficient data to calculate sensitivity and specificity (or proportion of positive results that are false positives)
- stage, grade, size and/or nodal status of cancers detected
- interval cancer rate
- impact on clinical management
- patient outcomes (morbidity, mortality, adverse events, quality of life)

6. Not in English

Due to time constraints, only studies published in English will be eligible for inclusion.

Based on these criteria, 2,763 citations were excluded from the review. The QUOROM flowchart (Figure 1) summarises the results of the literature search and the application of the study exclusion criteria. A list of the studies that were retrieved for full-text evaluation but subsequently excluded because they did not meet the eligibility criteria for this review is available in Appendix I.

Figure 1 QUOROM flowchart summarising the results of the literature search and the application of entry criteria



[1] Two studies were identified in both literature and HTA database searches

The 26 publications meeting criteria for inclusion in the review are: HTA reports (3), other systematic reviews/meta-analysis (6), studies of diagnostic test accuracy (10 studies, 15 publications) and studies of patient preferences and/or quality of life (2 studies), one of which also reported on changes in patient attitudes toward prophylactic mastectomy.

No studies assessing of the impact of MRI plus mammography or MRI versus mammography on clinical management or treatment outcomes were identified.

Evidence appraisal

An appraisal of the evidence was conducted by appraising the strength, size and relevance of evidence from individuals studies included in the review as well as assessing the overall body of evidence available for conclusions about the safety, effectiveness and cost-effectiveness of breast MRI for screening young women at high risk of breast cancer. The evidence was also interpreted using the WHO criteria for medical screening (Strong et al 2005).

Appraisal of individual studies

The evidence presented in the selected studies was classified and appraised using the NHMRC Dimensions of Evidence (NHMRC 1999, 2005) and the MSAC Diagnostic Test Guidelines (MSAC 2005). These dimensions (Table 11) include three main domains: strength of the evidence, size of the effect and relevance of the evidence. The strength of evidence refers to the level and quality of study design and the precision of results reported. This clinical importance of the size of effect and relevance of the evidence was determined with advice from the Advisory Panel.

Table 11 Dimensions of Evidence

Type of evidence	Definition
Strength of the evidence	
Appropriate comparison	Did the study evaluate a direct comparison of the index test strategy versus the comparator test strategy?
Applicable population	Did the study evaluate the index test in a population that is representative of the subject characteristics (age and sex) and clinical setting (disease prevalence, disease severity, referral filter and sequence of tests) for the clinical indication of interest?
Study level and quality	The study design[2] and the methods used by investigators to minimise bias within the study design[3]
Statistical precision	The *p*-value or, alternatively, the precision of the estimate of the effect. It reflects the degree of certainty about the existence of a true effect
Size of effect	The distance of the study estimate from the "null" value and the inclusion of only clinically important effects in the confidence interval.
Relevance of evidence	The usefulness of the evidence in clinical practice, particularly the appropriateness of the outcome measures used.

[1] adapted from NHMRC 2000 and MSAC guidelines 2005.

[2] validity of study design ranked according to the NHMRC levels of evidence I-IV as outlined in Table 12.

[3] validity of methods used within the study design as outlined in Table 14.

Appraisal of study quality

The quality of a study refers to the extent to which it is has been designed and conducted to reduce bias in the estimation of the outcome. The potential sources of bias vary according to whether the study is designed to estimate the impact of the test on health outcomes (where the ideal is a randomised trial of alternative tests) or to estimate the diagnostic accuracy of the test (where the ideal is a cross-sectional analytic study of

consecutive patients tested using both the test of interest and a valid reference standard). A structured appraisal of each study was performed to classify studies according to the type of study design (levels of evidence) as shown in (Table 11) and to appraise the quality of the study methods using prespecified criteria (described below and listed in Tables 13 and 14).

Table 12 Designations of levels of evidence1

Level of evidence	Study design
Studies of effectiveness	
I	A systematic review of level II studies
II	A randomised controlled trial
III-1	A pseudorandomised controlled trial (alternate allocation or some other method)
III-2	A comparative study with concurrent controls: non-randomised experimental trial, cohort study, case-control study, or interrupted time series with a control group
III-3	A comparative study without concurrent controls: historical control, two or more single arm studies, or interrupted time series without a parallel control group
IV	Case series with either post-test or pre-test/post-test outcomes
Studies of test accuracy	
I	A systematic review of level II studies
II	A study of test accuracy with: an independent, blinded comparison with a valid reference standard, among consecutive patients with a defined clinical presentation
III-1	A study of test accuracy with: an independent, blinded comparison with a valid reference standard, among non-consecutive patients with a defined clinical presentation
III-2	A comparison with a reference standard that does not meet the criteria required for Level II and III-1 evidence
III-3	Diagnostic case-control study
IV	Study of diagnostic yield (no reference standard)

[1] Modified from NHMRC 1999 & 2005

The quality of the methods used in each study was assessed using a standard checklist. This checklist included different criteria for different study designs as outlined below.

Quality of systematic reviews

Criteria for appraising the quality of studies of patient preferences and quality of life outcomes were adapted from NHS Centre for Reviews and Dissemination guidelines (2001)(Table 13).

Table 13 Quality Assessment of systematic reviews1

- Was the research question specified?
- Was the search strategy documented and adequate?
- Were the inclusion and exclusion criteria specified, appropriate and applied in an unbiased way?
- Was a quality assessment of included studies undertaken?
- Were the methods of the study appraisal reproducible?
- Were the characteristics and results of the individual studies summarised?
- Were the methods for pooling the data appropriate?
- Were sources of heterogeneity explored?
- Was a summary of the main results and precision estimates reported?

1 Modified from NHS Centre for Reviews and Dissemination 2001

Quality of studies of diagnostic test accuracy

The quality of studies of diagnostic test accuracy was assessed using a checklist of 13 items adapted from the QUADAS tool developed by Whiting et al (2003) (Table 14).

Three items from the QUADAS tool were not included in this assessment because they were less relevant to distinguishing between high and low quality studies in this review (QUADAS items 5, 7, 9). For example, the reference standard was defined as histopathology for positive results and clinical follow-up for negative test results, therefore by definition it was never applied independently of the test results (item 7). Item12, which assesses whether the same clinical data was available when the test results were interpreted as would be available when the test is used in practice was also excluded. It was assumed that the radiologist reading the tests would generally have no clinical information other than the women's high-risk status, given that only asymptomatic women would receive screening and a clinical breast examination would either be negative or not performed prior to screening.

Table 14 Quality assessment of studies of breast MRI accuracy, adapted from the QUADAS tool

Item	Yes	No	Unclear
QUADAS item 1: Was the spectrum of patients representative of the patients who will receive the test in practice?			
Was the study a prospective study?	( )	( )	( )
Were women consecutively recruited?	( )	( )	( )
Did the study include women under the (average) age of 50 years at high risk of breast cancer?	( )	( )	( )
QUADAS item 2: Were selection criteria clearly described?	( )	( )	( )
QUADAS item 3: Is the reference standard likely to correctly classify the target condition?			
Was a film review performed to confirm classification of interval cancers as false negatives?	( )	( )	( )
QUADAS item 4: Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?			
Was the time period between biopsy and MRI and mammography ≤ 2 weeks	( )	( )	( )
QUADAS item 6: Did patients receive the same reference standard regardless of the index test result?			
Was the same reference standard used for MRI and mammography? (including at the same threshold eg. at BI-RADS 3)	( )	( )	( )
QUADAS item 8: Was the execution of the index test described in sufficient detail to permit replication of the test?			
Was the test threshold of MRI and mammography stated clearly?	( )	( )	( )
QUADAS item 10: Were the index test results interpreted without knowledge of the results of the reference standard?			
Were MRI results interpreted without knowledge of histopathology results/clinical follow-up? mammography? vice versa?	( )	( )	( )
QUADAS item 11: Were the reference standard results interpreted without knowledge of the results of the index test?			
Were the histological results interpreted without knowledge of the results of MRI? mammography?	( )	( )	( )
QUADAS item 13: Were uninterpretable/intermediate test results reported?	( )	( )	( )
QUADAS item 14: Were withdrawals from the study explained?	( )	( )	( )
Was sufficient information reported to construct a 2x2 table?	( )	( )	( )

Source: (Whiting et al 2003)

Studies were required to meet all 10 criteria to be assessed as high quality. Retrospective studies and studies that did not report the threshold used to define positive MRI and mammography results, or reported a test interval of greater than one month were classified as low quality.

Quality of studies assessing patient preferences and quality of life

Criteria for appraising the quality of studies of patient preferences and quality of life outcomes were adapted from NHS Centre for Reviews and Dissemination guidelines (2001) (Table 15).

Table 15 Quality Assessment of studies of health outcomes[1]

• Was the study based on a representative sample selected from a relevant population?
• Were the criteria for inclusion and exclusion explicit?
• Were the data collection methods used adequately described?
• Were outcomes assessed using valid and reliable instruments?

[1]Modified from NHS Centre for Reviews and Dissemination 2001

Representativeness of the study population was assessed by considering whether consecutive patients were enrolled. The relevance of the population was assessed by considering whether the characteristics of the study population were applicable to the population specified for this report.

The methods used to assess patient preferences were appraised as high quality if they were adequately described to allow repeatability. The methods used to assess quality of life were only appraised as high quality if validated instruments were used.

Appraisal of applicability of results

The three key criteria used to appraise of the generalisability of individual study results to young high-risk women and their applicability to Australian clinical practice were:

- The patients' characteristics and relevance to the intended test population (asymptomatic high-risk women < 50 years of age)
- The type of MRI technique performed and relevance to Australian practice (MRI with contrast, ≥ 1.5 Tesla magnet field strength, dedicated breast coil)
- The use of ultrasound at the radiologist's discretion, for example to supplement mammography for imaging women with dense breasts

Appraisal of body of evidence

In addition to the appraisal of individual studies, an appraisal of the overall body of evidence about the safety, effectiveness and cost-effectiveness of breast MRI was conducted using the same principles. This appraisal was based on the assessment of five criteria suggested by the NHMRC guidelines for the developers of guidelines (NHMRC 2005):

1. the volume of evidence – the number of studies sorted by their methodological quality and relevance to patients;
2. the consistency of the study results – whether the better quality studies had results of a similar magnitude and in the same direction ie homogenous or heterogenous findings;
3. the potential clinical impact – appraisal of the precision, size and clinical importance or relevance of the primary outcomes used to determine the safety and effectiveness of the test;
4. the generalisability of the evidence to the target population; and

5. the applicability of the evidence – integration of this evidence for conclusions about the net clinical benefit of the index test in the context of Australian clinical practice.

Data analysis

The characteristics of the study population, type of diagnostic test, reference standard, comparator, study quality and relevant endpoints were extracted for each study. Where appropriate, the results of eligible studies were statistically synthesized and pooled results presented.

Measurement of test accuracy

The accuracy of a test is determined by its ability to identify the target condition compared to a reference standard test that is used as a proxy for true disease status. Subjects who test positive using the reference standard are classified as having the disease and those who test negative are classified as disease-free.

Results of the index test and reference standard for a group of tested subjects can be summarised in a two-by-two table as shown in Figure 2.

Figure 2 Two by two table displaying the data used to determine test accuracy

		Reference standard: disease +	Reference standard: disease -
Index test	+	true positive (TP)	false positive (FP)
	-	false negative (FN)	true negative (TN)
		TP + FN	TN + FP

Total number of subjects tested = TP + TN + FP + FN

Number of subjects with disease = TP + FN

Number of subjects without disease = TN + FP

As shown, subjects who test positive for the disease of interest by both the index test and the reference standard are recorded as true positives (TP). Subjects without the target condition who test negative by both tests are recorded as true negatives (TN). When there is discordance between the results of the index test and reference standard, the index test result is recorded as a false positive (FP) if it detects the target condition and the reference standard does not. A false negative (FN) is recorded if the reference standard detects the target condition and the index test does not.

Sensitivity and specificity

The *sensitivity* of a test is the probability of a positive test in subjects with the disease of interest. The *specificity* of a test is the probability of a negative result in subjects without the disease.

Calculation:

Sensitivity = TP/(TP + FN)

Specificity = TN/(TN + FP)

The sensitivity and specificity of a test are always considered together and vary according to the threshold used to define a positive test. Sensitivity and specificity are known to vary according to the spectrum of disease (for example, variation in disease severity) in the patient group tested. If the sensitivity of a test is sufficiently high, a negative result rules out the disorder. Therefore, high sensitivity is particularly important if the penalty for missing disease is high. If the specificity of a test is sufficiently high, a positive result rules in the disorder. Therefore, high specificity is particularly important if a false positive result can harm the patient.

Data extraction

Data was extracted using a standardised instrument designed for this review. Data extraction was performed by one reviewer and checked by a second reviewer. Any discrepancies were resolved by discussion. The data extraction tables are provided in Appendices E-G. Where the publications reported percentages only, raw numbers were determined from the number of patients on which each test was performed. Where only raw numbers were reported, percentages were calculated using the number of patients reported to have had the test performed. Where possible, two by two tables were reconstructed from study data to estimate sensitivity, specificity and associated 95% confidence intervals for each test using Meta-DiSc software (Zamora et al 2004).

Assessment of heterogeneity and threshold effect

The true positive rate (sensitivity) and false positive rate (1 – specificity) from studies assessing the same test strategy were plotted in receiver operating characteristic (ROC) space for the assessment of the presence of a threshold effect for a positive test. A threshold effect occurs when studies use different thresholds to define a positive test, resulting in different estimates of test sensitivity and specificity. When heterogeneity between studies can be explained by a threshold effect, estimates of test accuracy can be summarised using a summary ROC curve and 95% confidence intervals for each test.

Meta-analysis

Data from studies reporting absolute numbers of true positive, true negative, false positive and false negative results for each test strategy were presented graphically in receiver operating characteristic (ROC) space (*Y*-axis = the true positive rate [sensitivity]; *X* axis = false positive rate [1 – specificity]).

Meta-analysis was undertaken to produce summary estimates of the sensitivity and specificity of each test where heterogeneity (non-random variation) of study results was not statistically significant (p value ≥ 0.05). Summary estimates of the percentage difference between the sensitivity and specificity of these tests were also calculated. Meta-analysis was performed using a random effects model (DerSimonian–Laird method) and heterogeneity was assessed statistically using chi-square (χ^2) tests using the Meta-Disc program (Zamora et al 2004). Only two studies fully reported data for a comparison of breast MRI plus mammography versus mammography alone and thus summary ROC analysis was not undertaken.

Ongoing studies

A search of clinical trial databases, the American College of Radiology Imaging Network database, the UK National Health Service's Health Technology Assessment Programme database and information provided by the applicant was undertaken to identify ongoing studies of breast MRI (Table 16).

Table 16 Databases searched to identify ongoing studies[1]

• www.controlled-trials.com
• clinicaltrials.gov
• www.actr.org
• www.acrin.org
• www.cancer.gov/search/clinical_trials/
• www.ncchta.org/ProjectData/1_project_select.asp

Two relevant studies were identified as summarised below:

- Pilot Screening Study of Breast Imaging Outcome Measures in Women at High Genetic Risk of Breast Cancer (National Cancer Institute Clinical Centre Project 01-C-0009, ClinicalTrials.gov Identifier NCT00006421).

  Aim: To compare the performance of MRI versus mammography in an annual breast cancer screening program for young high-risk women (from age 25 years or 5 years younger than index case). Positron emission tomography used to assess indeterminate results. This study will also investigate the performance of ductal lavage for the early detection of breast cancer.

  Participants are either known carriers of BRCA1 or BRCA2 mutation or have a first or second degree relative known to carry a BRCA mutation, or with a diagnosis of Breast-Ovarian Cancer Syndrome in a family with a known BRCA mutation.

  Principal Investigator: Jennifer Loud

  Expected enrolment 200 (www.cancer.gov/search/clinical_trials/ last update April 2006).

  Recruitment site: Warren Grant Magnuson Research Center, National Institutes of Health Clinical Center, Bethesda, Maryland, USA

- A Study of Magnetic Resonance Imaging (MRI) and ultrasound as methods of screening for breast cancer in women at high risk

  Aim: To compare the sensitivity and specificity of contrast-enhanced breast MRI and ultrasound to mammography and clinical breast examination in high-risk women below the age of 50 years.

  Principal Investigator: Christobel Saunders

  Enrolment completed 2004: 72 participants undertaking annual screening

  Recruitment sites: Western Australia

Expert advice

An advisory panel with clinical expertise in breast cancer screening, diagnosis, treatment and consumer issues was established to evaluate the evidence and provide advice to the MSAC. In selecting members for advisory panels, the MSAC's practice is to approach the appropriate medical colleges, specialist societies and associations and consumer bodies for nominees. Membership of the panel is listed in Appendix B.

Results of assessment

Is it safe?

Included studies

Health technology assessment reports and other systematic reviews

This review identified four HTA reports assessing the use of MRI for breast cancer screening in high-risk women (Australia and New Zealand Horizon Scanning Network 2005; Blue Cross Blue shield 2003; Institute for Clinical Systems Improvement 2003; National Breast Cancer Centre 2006b). Overall, these reports concluded that breast MRI is a safe procedure when used in women without contra-indications. None of the reports provided an assessment of the safety of breast MRI as an addition or replacement for annual mammography.

Safety issues about MRI raised in these reports included:

- adverse effects of false positive findings (unnecessary investigation)
- use in patients with contra-indications to exposure to magnetic fields
- allergy to gadolinium contrast agent
- claustrophobia which may preclude use in some patients
- patient discomfort due to the noise of the machine
- avoidance advised in pregnant women due to limited evidence about the safety of MRI on the developing foetus

The rates and consequences of false positive findings at breast MRI with or without mammography are discussed in the effectiveness section of this report and compared to mammography alone (page 41).

Existing HTAs do not report any cases where breast MRI was mistakenly performed when contra-indicated, nor rates of MRI cancellation due to patient claustrophobia. An earlier Australian HTA report cited a review of 30,000 MRI procedures at one institution that showed moderate to severe gadolinium allergy occurred in approximately 1 in 10,000 patients {De Ridder F., 2001 73 /id;Australia and New Zealand Horizon Scanning Network, 2005 613 /id}.

The two clinical guidelines and four other published systematic reviews identified did not report any additional evidence about the safety of breast MRI.

Primary studies

This review identified 10 accuracy studies (15 articles) reporting on the use of breast MRI in 4,040 patients (8,691 scans). Most studies only presented results for the subset of eligible patients who successfully received both tests. No adverse events due to breast

MRI or mammography were reported. Three studies reported the proportion of patients unable to have a breast MRI due to claustrophobia or other reasons. Tilanus-Linthorst et al (2000) reported that MRI was interrupted in one patient due to claustrophobia (0.9% of 109 participants) and was reported as the main reason for MRI refusal in "a few others". Leach et al (2005) reported that 12 of 649 study participants (1.8%) withdrew due to claustrophobia or discomfort caused by MRI. Kriege et al (2004) reported that 89 of 1,909 participants (4.7%) refused breast MRI due to claustrophobia or other unspecified reasons. No adverse events due to contrast allergy were reported.

Additional evidence

Additional evidence about the most serious safety issues for each procedure is outlined below.

MRI contra-indications

Other than the relative accuracy of breast MRI and the potential harms of a false diagnosis, which are discussed in later sections of this document, the most important safety issue for breast MRI is ensuring that patients with contraindications to MRI are excluded from screening. These contraindications include patients with ferromagnetic vascular clips, ferromagnetic ocular foreign bodies, cardiac pacemakers, automatic cardiac defibrillators, implanted electromagnetic devices, breast tissue expanders and cochlear implants. Although extremely rare, performing MRI in patients with these contraindications is potentially fatal (Schenck 2000).

Patient information sheets and detailed test referral forms that outline each of the contra-indications and require signed clearance by the attending doctor are used in MRI facilities in Australia to help avoid this problem. Each patient must be assessed on an individual basis, as the ferromagnetic properties and the anatomical position of the implant will determine whether the patient is a suitable candidate for MRI. It is also recommended that patients wear ear protection during the procedure, in order to prevent auditory discomfort and temporary hearing loss.

In patients without contraindications, exposure to the type of magnetic fields used in breast MRI procedures in Australia is not associated with any irreversible or hazardous biological effects (Schenck 2000).

Radiation exposure

Unlike MRI, mammography exposes women to ionising radiation. Although the risk of radiation-induced breast cancer is low, it increases with repeated exposures and younger age at exposure, and is potentially greater for women with a genetic predisposition to breast cancer (Berrington de Gonzalez et al 2005). Therefore, the potential risks of ionising radiation will become more important as siblings and daughters of women with breast cancer are identified as being at high risk earlier and therefore commence screening at a younger age due to increased community awareness and genetic testing of index cases.

Using observational data about the effects of radiation exposure on breast cancer risk, Berrington de Gonzalez et al (2005) estimated that the risk of radiation-induced breast cancer in women with breast cancer in two first degree relatives undergoing a decade of annual mammographic screening from age 40 years was 1.24 cases per 1,000 women,

compared to 0.11 cases per 1,000 women at average risk commencing a decade of screening at age 50 years.

To date, no epidemiological studies have demonstrated that participation in mammographic screening programs increases the risk of breast cancer for high-risk women. A case-control study that included 1,600 breast cancers cases and 1,600 age-matched controls, all women with a known mutation of BRCA1 or 2, found no increased breast cancer risk for those women who had participated in mammography screening (odds ratio 1·03 [95% CI 0·85-1·25]) after adjusting for parity, oral-contraceptive use, ethnic origin, and bilateral oophorectomy (Level III-2 evidence, (Narod et al 2006). A subgroup analysis suggested a modest increased risk in women who initiated screening before age 40 years (unadjusted odds ratio 1.55 [95%CI 1.08-2.23], $p = 0.02$), however this finding may be due to chance alone, given the large number of statistical comparisons performed on this dataset.

In contrast, a retrospective cohort study conducted in 1,601 women known to carry a BRCA1 or BRCA2 mutation demonstrated a strong association between past history of chest x-ray exposure and breast cancer risk in this population. This study reported a higher risk for women aged up to 40 years (hazard ratio 1.97, $p < 0.001$), women who reported more than five chest x-rays (hazard ratio 2.69), and those exposed to chest x-rays before the age of 20 years (hazard ratio 5.21) (Andrieu et al 2006). This observational study design is vulnerable to recall bias due to possibility of differential reporting of x-ray exposure by women according to their breast cancer status and potential confounding by (unknown) factors associated with both x-ray exposure and breast cancer risk. Additional evidence from prospective studies reporting on the association between mammography exposure and breast cancer risk are required to validate these findings and estimate the magnitude of risk for women with BRCA mutations participating in mammography screening.

Is it effective?

This review did not identify any studies that compared health outcomes for women at high risk for breast cancer participating in annual mammography screening programs with and without breast MRI. As a result, this section summarises the evidence about the relative accuracy of adding breast MRI versus mammography alone in high-risk women under the age of 50 years and evidence about the relative effectiveness of treatment for cases detected early by MRI versus cases detected later by mammography or clinical examination to address the primary review question.

Evidence about the use of breast MRI as an additional test in high-risk women over the age of 50 years and as a replacement test for mammography is also presented to address the secondary review questions.

The evidence is presented under the following headings:

- Test accuracy
    - What is the relative sensitivity and specificity of breast MRI plus mammography versus mammography alone for the detection of breast cancer in high-risk women younger than 50 years; and for those aged 50 years or older?

- What is the relative sensitivity and specificity of breast MRI versus mammography alone for the detection of breast cancer in high-risk women younger than 50 years; and for those aged 50 years or older?

- Stage shift
  - Does the addition of breast MRI lead to a stage-shift in cancer detection?
- Interval cancer:
  - Does the addition of breast MRI lead to a reduction in interval cancers?
- Patient preferences and attitudes
  - Do patients prefer the addition of breast MRI versus mammography alone?
- Treatment effect
  - What is the relative effectiveness of treatment for cases detected early by breast MRI versus cases detected later by mammography in this patient group?

This review did not identify any studies reporting on the impact of breast MRI on decisions about patient management.

Included studies

The literature search identified 9 systematic reviews and 12 primary studies (24 publications) investigating MRI as an additional or replacement screening test to annual mammography that were eligible for the current review (Table 17). Of the 12 primary studies identified, 8 were included in one or more of the existing systematic reviews. The remaining 4 were recent publications of test accuracy (Kuhl et al 2005) or patient tolerance and attitudes (Kurian et al 2005); or did not meet the eligibility criteria for earlier reviews (Hartman et al 2004; Lehman et al 2005b). A table showing the objectives, methods and patient characteristics of included studies is presented in Appendix E (systematic reviews) and Appendix F (primary studies).

Table 17 Summary of included studies1

Type of evidence	*N*	Test comparisons	Test populations
Systematic review[2]	9	• MRI + mammography vs mammography (3) • MRI vs mammography (8)	High-risk women (8 reviews) • Age not specified (7 reviews) • Subgroup analysis by age (1 review) Average-risk population (1 review)
Accuracy study[3]	10	• MRI + mammography vs mammography (2) • MRI + mammography + ultrasound vs mammography+/- ultrasound(1) • MRI +/- mammography vs mammography + ultrasound (1) • MRI vs mammography (10) • MRI vs mammography+US (1)	High-risk women • Average age < 50 years (8 studies)
Stage shift	6	• MRI vs mammography (6)	High-risk women • Average age < 50 years (5 studies)
Patient preferences	2	• MRI vs mammography (2)	High-risk women • Average age < 50 years (2 studies)

1. No studies reporting on therapeutic impact, interval cancer rate or health outcomes were identified.

2. 3 HTA, 2 clinical guidelines, 3 peer-reviewed systematic reviews.

3. Studies may report data for more than one comparison.

Systematic reviews

Study characteristics

Nine systematic reviews evaluated the accuracy of breast MRI as a screening test for high-risk populations. One systematic review evaluated breast MRI for general population screening; however, all four studies assessed were conducted in high-risk women (Irwig et al 2004). Five of the nine systematic reviews assessed breast MRI as a screening test as part of a broader review that included an assessment of breast MRI for: breast cancer diagnosis and staging (Institute for Clinical Systems Improvement 2003); risk reduction strategies for women at high risk of breast cancer (Calderon-Margalit et al 2004; McIntosh et al 2004); and clinical guidelines for breast cancer screening (Elmore et al 1998; Smith et al 2003). None of these systematic reviews were specifically designed to assess the addition of breast MRI to annual screening mammography in young high-risk women.

Studies included in each review varied due to differences in publication dates and inclusion criteria for the study comparator, reference standard and study quality between reviews. A total of 10 primary studies published between 2000 and 2005 were assessed in these reviews; each review included at least 4 of these studies. The Australian and New Zealand Horizon Scanning Network (ASNZSN) report (ASNZSN 2004) included all ten studies to provide the broadest summary of the evidence published up to 2004. The NBCC review was the most up-to-date, including studies published up to July 2005 but excluding studies that identified fewer than 10 cases of breast cancers, including a recent study by Lehman et al (2005) that reported on a screening population of 367 women (NBCC 2006).

Quality and applicability

The nine systematic reviews met each of the pre-specified criteria for high-quality systematic reviews (Appendix E). Most reviews reported that the quality of included studies was suboptimal, which may limit the validity of the results. Overall, variation in the methods, quality and results of included studies together with small study sizes limited data synthesis and provided weak evidence for review conclusions.

Most of these reviews investigated the accuracy of MRI versus mammography and were conducted before the publication of two large studies that compared the addition of breast MRI to mammography in high-risk women (Leach et al 2005, Kuhl et al 2005). Prior to 2004, evidence about the impact of MRI as an additional test was limited to a study reporting on a highly select population of 109 high-risk women with greater than 50% mammographic density and no abnormalities detected at prior mammography (Tilanus-Linthorst et al 2000). The NBCC review (2006) provides the most applicable evidence to address the primary research question for the current review due to the inclusion of the recent study by Leach et al 2005. However, the authors did not identify any direct evidence for conclusions about the effectiveness of the addition of breast MRI on patient outcomes.

The present review has been conducted to update and broaden the scope of these earlier reviews.

Primary studies of test accuracy

Study characteristics

Ten eligible studies of diagnostic test accuracy were identified (reported in 15 articles). Study size ranged from 20 to 1,909 women (reporting on a total of 8,691 MRI scans). Eight of these studies were designed to compare the accuracy of breast MRI versus mammography. Only three studies included a comparison of breast MRI as an additional test versus mammography with or without breast ultrasound (Table 18). The remaining two studies compared the cancer detection rate for MRI and mammography but did not provide sufficient data to calculate the sensitivity and specificity or area under the ROC curve (AUC) of mammography (Tilanus-Linthorst et al 2000, Trecate et al 2003).

Table 18 Characteristics of included studies of test accuracy

Author, year, setting	*N*[1] *(MRI)*[2]	Test comparison	Population	Outcomes
Hartman 2004 USA Single centre 2001-2003	41 *(56)*	MRI vs mammography	> 10% risk of breast cancer at 10 years or BRCA1/2 mutation Median age 43 years (range 27-72 years) Prior history of breast cancer: 29%	MRI sensitivity & specificity[3] Biopsy rate
Kriege 2004, 2006 Netherlands 6 sites 1999-2003	1,909 *(4,169)*	MRI vs mammography	≥ 15% cumulative lifetime risk Mean age 40 years (range 19-72 years) Prior history of breast cancer: 0%	Sensitivity & specificity, AUC Subgroup analyses: 1st versus subsequent screening rounds Biopsy rate Cancer size, grade & lymph node status
Kuhl 2005 German Single centre 1996-2001	529 *(1,452)*	MRI + mammography vs mammography MRI + mammography vs mammography + US MRI vs mammography	≥ 20% cumulative lifetime risk Median age 40 years (range 27-59 years) Prior history of breast cancer: 26%	Sensitivity & specificity Cancer stage, size, lymph node status & grade in women with no prior history of breast cancer Subgroup analyses: by risk group, mutation status, prior history of breast cancer
Leach 2005 (MARIBS) UK 22 sites 1997-2004	649 *(1,881)*	MRI + mammography vs mammography MRI vs mammography	Annual risk of breast cancer ≥ 0.9% or 1st degree relative with 60% chance BRCA1/2 mutation carrier Median age 40 years (range 31-55 years) Prior history of breast cancer: 0%	Sensitivity & specificity, AUC Subgroup analyses:1st versus subsequent screening rounds Recall rate, biopsy rate Cancer size, grade & lymph node status
Lehman 2005 (IBMC) USA, Canada, Germany 13 sites	367 *(367)*	MRI vs mammography	> 25% cumulative lifetime risk; or patients with prior history breast cancer ≤ 5 years ago or > 5 years if at high risk Mean age 45 years ± std deviation 9.7 Prior history of breast cancer: 10%	Sensitivity & specificity Recall rate Biopsy rate Cancer size & lymph node status
Podo et al 2002 Italian Multicentre 2000-2002	105 *(105)*	MRI vs mammography MRI vs mammography + US	Known or suspected BRCA1/2 mutation carrier, prior history of unilateral breast cancer Median age 51 years (range 25-77 years) Prior history of breast cancer: 38%	Sensitivity & specificity[3] Cancer size, grade & lymph node status
Stoutjesdijk 2001 Netherlands Single centre 1994-2001	75 *(75)*	MRI vs mammography	≥ 15% cumulative lifetime risk age range 21-71, 87% ≤ 50 years Prior history of breast cancer: 0%	AUC
Tilanus-Linthorst 2000 Single centre 1994-1998	109 *(193)*	MRI vs mammography	≥ 25% cumulative lifetime risk and > 50% breast density Mean age: 41.5 years (range 22-68 years) Prior history of breast cancer: 0%	MRI sensitivity & specificity[3] Cancer size & lymph node status
Trecate 2003 Italian Single centre Year NR	20 *(20)*	MRI vs mammography	High risk based on known or suspected BRCA1/2 mutation carrier status Age range 30-61 years Prior history of breast cancer: 25%	MRI sensitivity & specificity[3]
Warner 2004 Canada single centre 1997-2003	236 *(457)*	MRI, mammography, US, CBE vs combin-ations without MRI MRI vs mammography	BRCA1/2 mutation carriers Median age: 47 years (range 25-65 years) Prior history of breast cancer: 30%	Sensitivity & specificity Subgroup analyses for 1st and Subsequent screening rounds Cancer size & lymph node status

AUC: area under the ROC curve.

1. Number of asymptomatic patients undergoing MRI and mammography.

2. Total number of screening MRI scans undertaken at first and subsequent screening rounds.

3. Breast MRI sensitivity and specificity calculated from data reported in text.

Five studies reported data from more than one screening round (Kriege et al 2004, Kuhl et al 2005, Leach et al 2005, Tilanus-Linthorst et al 2000, Warner et al 2004). Three of these studies also presented a subgroup analysis to investigate whether the relative sensitivity and specificity of breast MRI versus mammography varied between the initial and subsequent screening rounds (Kriege et al 2004, Leach et al 2005, Warner et al 2004). Kriege et al (2004) also investigated the relative accuracy of breast MRI between participants with and without a history of prior mammography at the initial screening round.

Most studies only presented the results for the subset of eligible patients who successfully received both tests. Based on information provided by four studies, 9-21% of eligible participants invited for screening withdrew from the screening program and/or were lost to follow-up (median 11%, Kriege et al 2004, Kuhl et al 2005, Leach et al 2005, Warner et al 2004). Reasons for withdrawal included: subsequent gene test or carrier status negative (2-5%), prophylactic mastectomy (3-7%); developed primary or metastatic breast cancer (2-5%) lost to follow-up (1-8%); claustrophobia or MRI refusal (1-5%). Other reasons for withdrawal included pregnancy, too large to fit into MRI machine, MRI contra-indicated, serious illness and death (Leach et al 2005, Warner et al 2004).

Quality

Eight studies used a prospective design, one was retrospective (Stoutjesdijk et al 2001) and another did not report sufficient information to determine whether it was conducted prospectively or retrospectively (Trecate et al 2003). None of the studies reported that a consecutive sample of patients was tested. Most studies only presented the results for the subset of eligible patients who successfully received both tests.

None of the studies of test accuracy were classified as high quality. Three studies were classified as fair quality (Kriege et al 2004, Leach et al 2005, Warner et al 2004). These three studies all reported that breast MRI and mammography were prospectively conducted within 2 weeks of each other and interpreted blinded to the results of the reference standard and comparator test, using a pre-specified threshold for classifying a positive test result. Data for each test were fully reported including the classification of indeterminate results. These studies were not classified as high quality because they used a suboptimal definition and/or application of the reference standard. Two studies only used histopathology to verify a positive test result after review of all imaging tests (Leach et al 2005, Warner et al 2004). Two studies classified interval cancers as false negatives but did not report whether this assessment was based on a review of films (Kriege et al 2004, Leach et al 2005). Warner et al (2004) relied on a consensus between tests to exclude false negative results. Two prospective studies that met most other quality criteria were classified as low quality because they included results from tests performed more than one month apart in their assessment of test accuracy (Kuhl et al 2005, Lehman et al 2005). Five other studies classified as low quality did not meet several quality criteria, including three studies that did not report sufficient data to calculate and compare the sensitivity and specificity of breast MRI and mammography (Stoutjesdijk et al 2001, Tilanus-Linthorst et al 2000, Trecate et al 2003).

A summary of the quality of included accuracy studies is presented in Table 19 and in more detail in Appendix F.

Table 19 Quality and applicability of included studies

Author, year,	N	Patient selection	Selection and application of reference standard	Test execution & interpretation	Data analysis	Appraisal
Primary review question: MRI + mammography versus mammography						
Kuhl 2005 German Single centre 1996-2001	529	Prospective Consecutive recruitment NR Median age 40 years	• Histopathology for all test positives, 6-month follow-up for indeterminate findings • No film review for interval cancers	• 1.5-Tesla magnet • Test interval within 8 wks • Blinded • Test threshold stated	Exclusions reported Indeterminate results reported Test failures reported Data for 2 x 2 table reported	Level III-I, low quality Applicable intervention and population
Leach 2005 (MARIBS) UK 22 sites 1997-2004	649	Prospective Consecutive recruitment NR Median age 40 years 99% ≤ 50 years	• Histopathology not performed for all test positives, based on combination of all tests • No film review for interval cancers	• 1.0-1.5-Tesla magnet • Test interval not specified but preferably 'same day' • Blinded, double reporting • Test threshold stated	Exclusions reported Indeterminate results reported Test failures reported Data for 2 x 2 table reported	Level III-2, fair quality Applicability limited • 1-1.5 T MRI • Double reporting
Warner 2004 Canada 1 site 1997-2003	236	Prospective Consecutive recruitment NR Mean age: 47 years	• Histopathology not performed for all test positives – MRI repeated in 1 month if positive and discordant with other tests. • No film review for interval cancers	• 1.5-Tesla MRI magnet • Test interval within 2 wks • Blinded • Test threshold stated	Exclusions partly reported Indeterminate results reported Test failures partly reported Data for 2 x 2 table reported	Level III-2, fair quality Applicable intervention and population
Secondary review question: MRI versus mammography						
Hartman 2004 USA Single centre 2001-2003	41	Prospective Consecutive recruitment NR Median age: 43 years	• Histopathology for all test positives, follow-up imaging for indeterminates • Film review for interval cancers – NR, 15 of 16 patients with negative results followed up at one year to date	• 1.5-Tesla magnet • Test interval within 8 wks • Blinding to comparator results NR • Test threshold varied for individual patients	Exclusions NR Indeterminate results reported Test failures NR Data for 2 x 2 table reported for MRI	Level III-I, low quality Applicable intervention and population
Kriege 2004 Netherlands 6 sites 1999-2003	1909	Prospective Consecutive recruitment NR Mean age: 40 years	• Histopathology for all test positives • Film review for interval cancers – NR	• MRI magnet strength NR • Test interval within 2 wks • Blinded • Test threshold stated	Exclusions reported Indeterminate results reported Test failures reported Data for 2 x 2 table reported	Level III-I, fair quality Applicable intervention and population

Author, year,	N	Patient selection	Selection and application of reference standard	Test execution & interpretation	Data analysis	Appraisal
Lehman 2005 (IBMC) USA, Canada, Germany 13 sites	367	Prospective Consecutive recruitment NR Mean age: 45 years	• Histopathology not performed for all test positives 9 women with discordant MRI/mammography results did not have FNAC or biopsy • Film review for interval cancers – NR	• MRI magnet strength NR • Test interval within 90 days • Blinded • Test threshold-stated	Exclusions reported Indeterminate results reported Test failures NR Data for 2 x 2 table reported	Level III-I, low quality Applicable intervention and population
Podo et al 2002 Italian Multi-center 2000-2002	105	Prospective Consecutive recruitment NR Median age: 51 years	• Histopathology not performed for all test positives • No film review for interval cancers	• ≥ 1.0-Tesla magnet • Test interval NR • Blinded NR • Test threshold unclear	Exclusions NR Indeterminate result NR Test failures NR Data for 2 x 2 table reported	Level III-2, low quality Applicability limited • 1T MRI • median age 51yr
Stoutjesdijk et al 2001 Netherlands Single centre 1994-2001	75	Retrospective Consecutive recruitment: No Mean age NR 87% ≤ 50 years	• Histopathology for all test positives • Film review for interval cancers	• 1.5-Tesla magnet • Test interval within 4 months • Blinded • Test threshold stated	Exclusions reported Indeterminate results reported Test failures NR Data for 2 x 2 table NR	Level III-2, low quality Applicable intervention and population
Tilanus-Linthorst et al 2000 Single centre 1994-1998	109	Prospective Consecutive recruitment – NR Mean age: 42 years > 50% breast density	• Histopathology not performed for all test positives • No film review for interval cancers	• 1.5-Tesla magnet • Test interval NR • Blinded NR • Test threshold not specified	Exclusions NR Indeterminate results NR Test failures reported Data for 2 x 2 table NR	Level III-2, low quality Applicability limited • > 50% breast density
Trecate et al 2003 Italian Single centre Year NR	20	Retrospective/prospective – NR Consecutive recruitment – NR (Data from 3 patients with clinical signs of cancer not extracted) Mean/median age NR Age range 30-61 years	• Histopathology not performed for all test positives – performed based on combination of MRI + mammography, US and clinical findings • No film review for interval cancers – follow-up of negative results undertaken but results NR	• 1.5-Tesla magnet • Test interval NR • Blinded NR • Test threshold NR	Exclusions NR Indeterminate results reported Test failures NR Data for 2 x 2 table NR	Level III-2, low quality Applicability limited • Age distribution NR

NR: not reported

Verification bias: where the reference standard is applied differently for breast MRI and mammogram.

Applicability

Most included studies enrolled women that generally represent the proposed population for breast MRI screening in Australia and used breast MRI techniques similar to those proposed in Australia.

All studies enrolled asymptomatic women with a known family history or genetic predisposition for breast cancer (or reported data for a subset of such women). The risk criteria used to select patients differed between studies, ranging from women assessed as having at least a 15% cumulative lifetime risk (Kriege et al 2004, Stoutjesdijk et al 2001) to women with a known BRCA1 or BRCA2 mutation (Warner et al 2004, see Table 18). However, the majority of participants in the two studies accepting women with ≥ 15% cumulative lifetime risk were assessed as having at least a 30% cumulative risk of breast cancer, which is the risk level proposed for the selection of women to breast MRI screening programs in Australia. The proportion of known BRCA1 or BRCA2 mutation-carriers included in each study ranged from 8% (Kuhl et al 2005) to 100% (Warner et al 2004).

No studies restricted enrolment to women less than 50 years of age; however, the average age of women recruited was between 40 and 50 years in 8 of the 10 studies. Podo et al (2002) reported a median patient age of 51 years; whereas Trecate et al (2002) did not report an age distribution. Six studies included women with a prior history of breast cancer requiring screening of the contralateral breast. The proportion of women included for screening the contralateral breast ranged from 10% (Lehman et al 2005) to 38% (Podo et al 2002).

All studies were performed using MRI machines fitted with a dedicated breast coil and intravenous contrast. Most studies were performed using MRI machines with ≥ 1.5-Tesla magnets. Two studies included tests performed on MRI machines with 1.0-1.5-Tesla magnets, which may compromise test performance (Leach et al 2005, Podo et al 2002, Table 17). The magnet strength for MRI procedures conducted in the studies by Kriege et al (2004) and Lehman et al (2005) was not reported.

The major problem with the applicability of the study results to the primary review question is the lack of studies investigating the addition of breast MRI to a mammography screening program where ultrasound is available at the discretion of the radiologist, as currently occurs in Australia. Two studies compared the combination of breast MRI plus mammography versus mammography alone (Kuhl et al 2005, Leach et al 2005). A third study investigated the accuracy of adding breast MRI to a combination of mammography, ultrasound and clinical breast examination, although data for calculation of the 95% confidence intervals were not reported (Warner et al 2004). This latter study may provide the most applicable evidence for an assessment of the proposed addition of breast MRI in Australia.

Other included studies

Two studies assessed the tolerance, preferences and changes in health-related quality of life of high-risk women participating in breast cancer screening programs that include breast MRI (Kurian et al 2005; Rijnsburger et al 2004). Both studies reported findings for breast MRI versus mammography, rather than breast MRI in addition to mammography. One of these studies also assessed the impact of intensified screening with

mammography, breast MRI and ductal lavage on women's decisions about prophylactic mastectomy (Kurian et al 2005).

Both studies were conducted as sub-studies of larger studies investigating breast MRI accuracy. Kurian et al (2005) surveyed 34 women (representing 84% of women invited) using an 8-item questionnaire. This study was appraised as fair quality due to the use of an unvalidated questionnaire. The small patient numbers did not allow precise estimates of patient discomfort and preferences. Rijnsburger et al (2004) administered a validated questionnaire to 334 women at three different time points before and after screening to assess changes in health-related quality of life scores. The authors reported response rates of 67%, 64% and 59% for the first, second and third questionnaire respectively, with no significant difference in age or breast cancer risk classification between participants and non-participants. This study was appraised as high quality (Appendix F).

Test accuracy

The results of all eligible primary studies that compare the sensitivity and specificity of MRI as an additional (3 studies) and/or replacement (6 studies) test to mammography for screening high-risk women are summarised in Table 20 and discussed below.

Table 20 Accuracy studies comparing breast MRI plus mammography versus mammography alone for detecting breast cancer in asymptomatic high-risk women[1]

Study author, year	*N*	Prevalence breast cancer (1st screening round)	Breast MRI		Mammography		Test of comparison	Study level & quality
			Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)		
Breast MRI as an additional test								
Kuhl et al 2005	529	14/529 (2.6%)	93% (81-99%)	96% (95-97%)	33% (19-49%) *M+US* 49% (33-65%)	97% (96-98%) *M+US* 89% (87-91%)	NR	NHMRC Level III-1 Low
Leach et al 2005	649	20/649 (3%)	94% (81-99%)	77% (75-79%)	40% (24-58%)	93% (92-95%)	NR	NHMRC Level III-2 Fair
Warner et al 2004[1]	236	13/236 (6%)	NR *MRI+M+US +CBE*[1] 95%	NR *MRI+M+US + CBE* NR	36% (17-59%) *M+US+CBE*[1] 64%	99.8% (99-100%) *M+ US+CBE*[1] NR	NR	NHMRC Level III-2 Fair
Breast MRI as a replacement test for mammography								
Kriege et al 2004, 2006	1909	26/1,723 (2%)	71% (56-84%)	90% (89-91%)	40% (26-56%)	95% (94-96%)	Difference in AUC 0.14 (95% CI 0.02-0.26) $p < 0.05$	NHMRC Level III-1 Fair
Kuhl et al 2005	529	14/529 (3%)	91% (78-97%)	97% (96-98%)	33% (19-49%)	97% (96-98%)	Sensitivity: p < 0.001 Specificity: p > 0.5	NHMRC Level III-1 Low
Leach et al 2005	649	20/649 (3%)	77% (60-90%)	81% (80-83%)	40% (24-58%)	93% (92-95%)	Sensitivity: p = 0.01 Specificity: p < 0.0001	NHMRC Level III-2 Fair
Lehman et al 2005[2]	367	4/367 (1%)	100% (40-100%)	93% (89-95%)	25% (0.6-81%)	98% (96-99%)	NR	NHMRC Level III-1 Low
Podo et al 2002[2]	105	7/105 (7%)	100% (63-100%)	99% (94-100%)	13% (0.3-53%)	100% (96-100%)	NR	NHMRC Level III-2 Low
Warner et al 2004	236	13/236 (6%)	77% (55-92%)	95% (93-97%)	36% (17-59%) *M+US+CBE* 64%[1]	99.8% (99-100%) *M+ US+CBE* NR	Sensitivity MRI vs M $p = 0.02$	NHMRC Level III-2 Fair

AUC=area under the curve, CBE=clinical breast examination, M=mammography, MRI-magnetic resonance imaging, NR: not reported, US=ultrasound

1. data insufficient for reconstruction of 2 x 2 table

2. data reported in article text used to construct 2 x 2 table

Breast MRI plus mammography versus mammography alone

Two studies comparing the accuracy of breast MRI plus mammography versus mammography alone demonstrated that the addition of breast MRI significantly increases the sensitivity of mammography for the detection of breast cancer (Kuhl et al 2005, Leach et al 2005, Table 20). Meta-analysis showed a sensitivity of 94% (95% CI 86-98%, test for heterogeneity $p = 0.82$) for breast MRI plus mammography compared to a sensitivity of 36% (95% CI 25-48%, test for heterogeneity $p = 0.50$) for mammography

and an incremental sensitivity for breast MRI of 58% (95% CI 46-70%, $p < 0.0001$, test for heterogeneity $p = 0.61$). Kuhl et al (2005) also reported that the addition of ultrasound increased the sensitivity of mammography to 49% (95% CI 33-65%) but did not report on the sensitivity of the combination of MRI, mammography and ultrasound.

A third study found that including MRI in a screening program of mammography, ultrasound and clinical breast examination increased the sensitivity of screening from 64% to 95% (Warner et al 2004, data for reconstruction of 2 × 2 table and calculation of 95% confidence interval not reported).

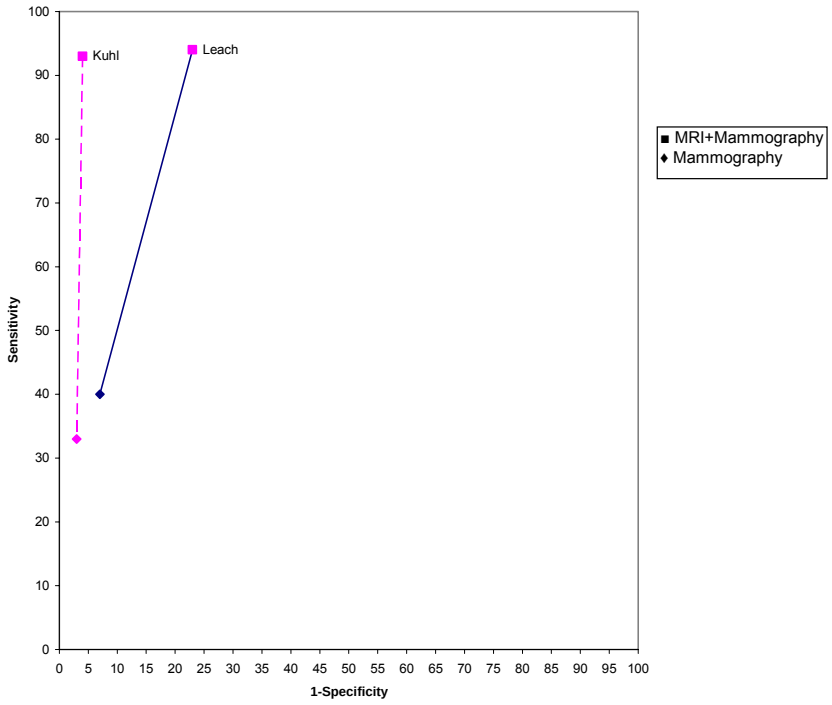
Findings about the specificity of breast MRI plus mammography were less consistent. Leach et al (2005) reported that the addition of MRI resulted in a significant decrease in test specificity (breast MRI plus mammography 77% [95% CI 75-79%] versus mammography 93% [95% CI 92-95%]), while Kuhl et al (2005) observed a smaller difference (MRI plus mammography 96% [95% CI 95-97%] versus mammography 97%, [95% CI 96-98%], Table 20). Warner et al (2004) did not report on the specificity of the combination of breast MRI plus routine screening tests. The variation in test specificity between the studies reported by Leach et al (2005) and Kuhl et al (2005) prevent synthesis of these data for a summary estimate of the 'true' specificity of breast MRI plus mammography versus mammography alone.

The use of different thresholds for positive test results may account for some of the difference in MRI specificity observed between studies. In Leach et al (2005), all patients that were 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 categories 3-5). Furthermore, all tests were double read and the most conservative score was selected for calculation of test accuracy. In contrast, Kuhl et al (2005) only classified patients as having a positive test if they had a 'suspicious' or 'malignant' finding. Patients recalled for additional review at 6 months due to an indeterminate test or probable benign finding were not classified as having a positive test unless further testing gave a 'suspicious' or 'malignant' result. Kuhl et al (2005) did not report the recall rates for short-term follow-up when a combination of MRI and mammography was used, although the authors noted no significant difference in recall rates for short-term follow-up or biopsy between MRI (11.5%) and mammography (9.5%). Differences in recall rates for MRI and mammography are discussed further on page 55.

The results of the two studies that reported sensitivity and specificity data are plotted in the ROC space to illustrate improvement in test performance achieved with the addition of MRI (Kuhl et al 2005, Leach et al 2005, Figure 3). These data are insufficient to plot a summary ROC curve to calculate a summary estimate of the relative accuracy of MRI plus mammography versus mammography alone.

Both studies collected data from women screened annually over a period of at least two years and up to 7 years. A secondary analysis conducted by Leach et al (2005) suggested that the performance of both tests was maintained over subsequent screening rounds (see page 53 for a full summary of results from studies comparing the accuracy of MRI and mammography at first versus subsequent screening rounds).

Figure 3 Accuracy of Breast MRI plus mammography vs mammography alone for the detection of breast cancer in ROC space



Breast MRI plus mammography versus breast MRI alone

Data reported by Leach et al (2005) can also be used to compare the accuracy of using MRI as an additional test (MRI plus mammography) versus as a replacement test (breast MRI alone) for the detection of breast cancer. These findings show that the combination of MRI plus mammography increases the sensitivity of MRI with a small but significant decrease in test specificity (MRI plus mammography sensitivity 94% [95% CI 81-99%], specificity 77% [95% CI 75-79%]; MRI sensitivity 77% [95% CI 60-90%], specificity 81% [95% CI 80-83%]) (Table 20).

Breast MRI versus mammography

Six studies compared the accuracy of breast MRI as a replacement test for mammography alone. All reported that breast MRI is a more sensitive but less specific test than mammography for the detection of breast cancer, although there was a wide variation in the results between studies (range: breast MRI sensitivity 71-100%, specificity 81-99%; mammography sensitivity 13-40%, specificity 93-100%).

Five of these studies reported that MRI has a statistically significantly higher sensitivity for detecting breast cancer than mammography (Kriege et al 2004; Kuhl et al 2005; Leach 2005; Podo et al 2002; Warner et al 2004). Four studies also showed that MRI has a statistically significantly lower specificity than mammography (Leach et al 2005, Lehman et al 2005, Kriege et al 2004, Warner et al 2004, Table 20). These data are plotted in ROC space to illustrate the improvement in test performance using MRI versus mammography (Figure 4). Meta-analysis using the random effects model showed: MRI sensitivity 81% (95% CI 74-87%, test for heterogeneity $p = 0.05$) compared to mammography sensitivity 36% (95% CI 28-44%, test for heterogeneity $p = 0.66$). Differences in the sensitivity of MRI versus mammography between these studies were

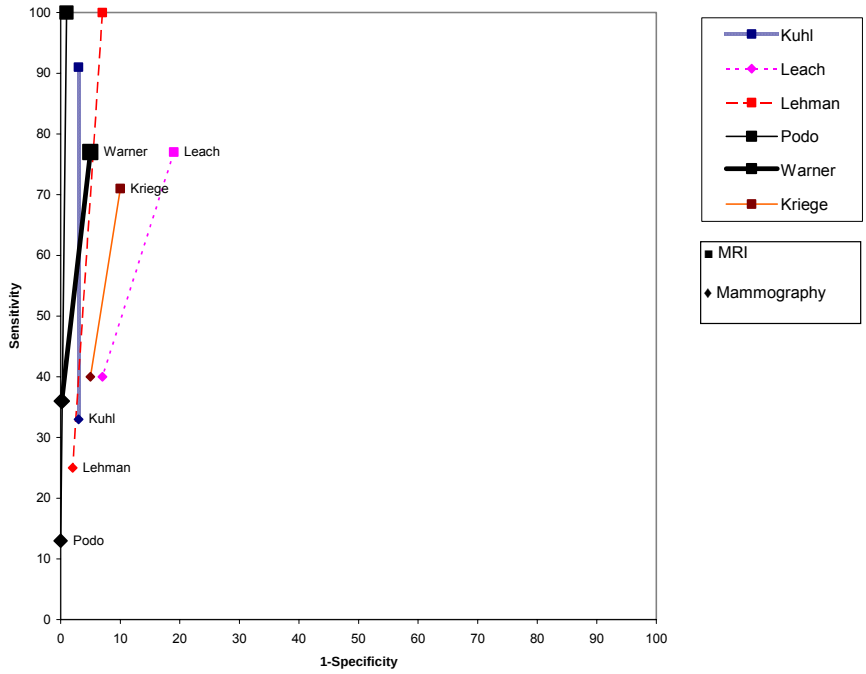
heterogenous (percentage difference in sensitivity = 52% [95% CI 35-69%], $p < 0.0001$, test for heterogeneity $p = 0.01$).

Heterogeneity of test specificity between studies was also statistically significant (MRI specificity 90% [95% CI 89-90%], test for heterogeneity $p < 0.0001$; mammography specificity 95% [95% CI 95-96%], test for heterogeneity $p < 0.0001$). Owing to the different classifications of a 'false positive' result between studies, a comparison of the recall rates initiated by each test for patients in whom further testing excluded cancer provided the most clinically relevant information for comparisons of the false positive rates between tests and is discussed further below.

The largest accuracy study collected data from 1,909 women screened for a median of 2.9 years Kriege et al (2004). A total of 39 evaluable invasive breast cancers were detected with a higher rate of cancer detection at the first screening round (22 'prevalent' cancers, 1,723 tests) than at all subsequent screening rounds (17 'incident' cancers, 2,431 tests). The authors reported that MRI detected more cancers than mammography at the initial and subsequent screening rounds, maintaining a statistically significantly higher sensitivity but lower specificity than mammography at each screening round (Appendix G). Subgroup analysis showed the magnitude of the difference between the sensitivity of MRI and mammography was highest at the initial screening round (initial screening round: MRI sensitivity 93%, mammography sensitivity 20%, $p = 0.001$; subsequent screening rounds: MRI sensitivity 77%, mammography sensitivity 29%, $p = 0.02$). Warner et al (2004) reported similar findings; however, a third study did not observe a difference in the sensitivity of MRI or mammography between the initial and subsequent screening rounds (Leach et al 2005).

Kriege et al (2004) suggested that the prior use of mammography screening in the study population may cause the relative sensitivity of MRI versus mammography to be highest at the initial screening round. They proposed that if screening mammography was already routinely performed, breast MRI would detect a pool of mammographically occult cancers at the initial screen, which may exaggerate its relative accuracy to detect incident cancers at subsequent screening rounds (Kriege et al 2006). Subgroup analyses appeared to support this hypothesis. The authors reported that the relative sensitivity of breast MRI was highest in the first screening round for women who had prior mammography but showed a similar sensitivity to mammography in the group of women who had no prior screening tests (Kriege et al 2006). However, the latter findings are surrounded by a large range of uncertainty owing to the small number of women with no prior history of mammography ($N = 303$, 7 breast cancer cases).

Figure 4 Accuracy of Breast MRI vs mammography for the detection of breast cancer in ROC space



Recall rates

Two studies reported patient recall rates following screening (Lehman et al 2005, Leach et al 2005). The proportion of screening participants recalled for any further investigation ranged from 10% (Lehman et al 2005) to 15% (Leach et al 2005). The proportion of screening participants recalled for fine needle aspiration cytology or biopsy ranged from 6% (Leach et al 2005) to 7% (Lehman et al 2005). Leach et al (2005) reported data for a comparison of recall rates for MRI plus mammography versus mammography alone; whereas Lehman et al (2005) only reported recall rates for MRI versus mammography. The results are presented in Table 21.

Table 21 False positive recall rate and biopsy rate reported by accuracy studies[1]

Study author, year	Number screened and recall rates	False positive recall rate	False positive biopsy rate
Leach et al 2005 MARIBS	649 (1,881 scans) recall rate 279/1,881 (15%) resulting in 33 screen-detected cancers • MRI+M 12.7% per woman yr • M 3.9% per woman yr • MRI 10.7% per woman yr biopsy rate 107/1,881 = 6%	246 false positive recalls (13%) 8.5 recalls per cancer detected False positive rate • MRI ± M: 428/1,881=23% • M: 121/1,881 = 6% • MRI: 344/1,881 = 17% No. of further tests performed for benign or normal findings Classified by test generating initial recall (patients recalled received 1-5 additional tests): • *MRI ± M:* 376/1,881 = 20% • *M:* 119/1,881 = 6% • *MRI:* 325/1,881 = 17% Comparison *MRI+M vs M*: $p < 0.001$[2]	96 procedures for normal/benign findings (5%) (57 FNAB[4], 21 core biopsy, 11 MR-guided biopsy, 7 surgical biopsy) 0.21 benign surgical biopsies per cancer detected False positive biopsy rate • MRI ± M 88/1,881 = 5% • M 29/1,881 = 1.5% • MRI 80/1,881 = 4% Comparison *MRI+M vs M:* $p < 0.001$[2]
Lehman et al 2005	367 recall rate 38/367 (10%) • MRI 31/367 (8%) • M 8/367 (2%) biopsy rate 27/367 (7%) • MRI 24/367 (6.5%) • M 4/367 (1%)	• MRI 27/367 (7%) • M 7/367 (2%) Comparison *MRI vs M*: $p = 0.001$[3]	• MRI 20/367 (5%) • M 3/367 (1%) Comparison *MRI vs M;* $p < 0.001$[3]

M = mammogram, US = ultrasound

1. Breast Screen Australia reports recall rates of 8.7% for average risk women over the age of 40 years attending their first round of mammographic screening and 4.2% for women attending subsequent rounds (Australian Institute of Health and Welfare 2005).
2. χ^2 test of proportions.
3. 2-sided Fisher's exact test of proportions.
4. 57 FNAB reported in published table, 47 FNAB reported in published text.

These results show that a statistically significantly greater proportion of women without breast cancer are recalled for further investigation and biopsy when breast MRI is introduced as an additional or replacement test to mammography. Based on the data reported by Leach et al (2005), the relative risk of being recalled for further investigation where cancer was excluded was 3.5 (95% CI 2.9-4.3, $p < 0.0001$) after MRI plus mammography versus mammography alone and similarly, the relative risk of being recalled for a biopsy where histopathology excluded malignancy was 3.0 (95% CI 2.0-4.6, $p < 0.0001$).

Stage shift

Six studies reported on the size, lymph node status and/or grade of invasive cancers detected by each test (Table 22). Another study reported this information without distinguishing between the detection modality (MRI, mammogram or interval cancer, Kriege et al 2004).

Table 22 Cancer stage, grade and nodal status for cases detected at screening of high-risk women[1]

Author	Number of invasive breast cancer cases (number of in-situ cancers)	Invasive cancer cases < 10 mm (% of all invasive cancers detected by modality)	Negative lymph node/micrometastases (% of all invasive cancers detected by modality)	Histological grade
Breast MRI plus mammography versus mammography				
Leach et al 2005	29 (6) *MRI + M:* 28 *MRI*: 25 *M:* 9	11/29 (38%) cases < 10 mm *MRI + M:* 10/28 (36%) *MRI*: 7/25 (28%) *M*: 4/9 (44%) *p* = 0.71[1]	21/26 (81%) cases node negative *MRI + M:* 19/24 (79%) *MRI:* 18/21 (86%) *M:* 6/9 (67%) *p* = 0.65[1]	*MRI + M:* G1 3/28 (11%), G3 18/28 (64%) *MRI:* G1 3/25 (12%), G3 17/25 (68%) *M:* G1 0/9 (0%), G3 7/9 (78%)
Warner et al 2004	16 (6) *MRI + M:* 13 *MRI:*13 *M:* 5	5/16 (31%) cases < 10 mm *MRI + M*: 5/13 (38%) *MRI*: 5/13 (38%) *M:* 1/5 (20%) *p* = 0.62[1]	13/15 (87%) cases node negative *MRI + M:* 11/12 (92%) *MRI:* 11/12 (92%) *M*: 4/4 (100%) *p* = 1.00[1]	NR
Breast MRI versus mammography				
Kuhl et al 2005	34 (9) *MRI*: 31 (8) *M*: 10 (4) *M + US:* 16 (5)	25/43 (58%) cases DCIS or invasive cancer < 10 mm with no lymph nodes *MRI*: 23/39 (59%) *M:* 5/14 (36%) *M + US*: 6/21 (29%) *p* = 0.21[2]	% of 34 cases node negative NR Women without a history of breast cancer: *MRI:* 26/31 (84%) *M*: 6/10 (60%) *M + US:* 11/16 (69%) *p* = 0.19[2]	NR by imaging modality
Lehman et al 2005	3 (1) *MRI:* 3 *M*: 1	1/3 (33%) cases < 10 mm *MRI:* 1/3 (33%) *M:* 1/1 (100%)	3/3 (100%) cases negative nodes *MRI:* 3/3 (100%) *M:* 1/1 (100%)	NR by imaging modality
Podo et al 2002	5 (3) *MRI:* 5 *M:* 1	3/3 (100%) cases < 10 mm *MRI:* 3/5 (60%) *M:* 1/1 (100%)	5/5 (100%) cases node negative *MRI:* 5/5 (100%) *M:* 1/1 (100%)	*MRI:* G1: 2/5 (40%), G3 1/6 (17%) *M:* G1: 1/2 (50%), G3 0/1 (0%)
Tilanus-Linthorst et al 2000	3 (0) *MRI:* 3 *M:* 0	0/3 (0%) cases < 10 mm	3/3 (100%) cases node negative *MRI:* 3/3 (100%) *M:* 0/0	NR by imaging modality
No test comparison				
Kriege et al 2004	44 (6) *MRI:* 32 *M:* 18	19/44 (43%) of cases < 10 mm *MRI:* NR *M:* NR	28/42 (67%) of cases node negative *MRI:* NR *M:* NR	G1: 19/43 (44%) of cases G3: 16/43 (37%) of cases

G = histological grade of tumour, M = mammography, US = ultrasound

1. 2-sided Fisher's exact test MRI plus mammography versus mammography.

2. 2-sided Fisher's exact test MRI versus mammography.

Interpretation of these findings is limited by the small number of cases identified in each study (3-44 invasive cancers per study, total 134 invasive cancers) and the variability in tumour characteristics between studies. Two studies reported on the characteristics of cancers detected by a combination of MRI and mammography versus mammography alone (Leach et al 2005, Warner et al 2004). Leach et al (2005) observed that the addition of breast MRI detected a similar proportion of invasive cancers < 10 mm as mammography (breast MRI 36% of cases detected, mammography 44%) but a higher proportion of node negative invasive cancers (breast MRI 79%, mammography 67%). In contrast, Warner et al (2004) observed that breast MRI detected a higher proportion of cancers < 10 mm than mammography (breast MRI 38% of cases detected, mammography 20%) but a similar proportion of node negative disease (breast MRI 92%, mammography 100%). None of these differences were statistically significant. Studies comparing tumour characteristics for cancers detected by MRI versus mammography also reported inconsistent results, although the two larger studies (Kuhl et al 2005 N = 34 cases, Leach et al 2005 N = 29 invasive cancers) both suggested that MRI may detect a higher proportion of node negative cases than mammography alone.

Rather than comparing cases detected by MRI versus mammography, Kriege et al (2004, 2006) investigated whether a breast MRI screening program detected earlier disease than standard clinical care. The authors compared the size, nodal status and grade of 44 invasive breast cancers detected by MRI and mammography in high-risk women participating in a screening study with age-matched cases from a study of women with breast cancer gene-mutations (high-risk breast cancer cases) and a national registry of all breast cancer cases (Table 23). The screening history of cases in these two 'control' groups was not known.

Table 23 Comparison of cancer stage, grade and nodal status for 44 invasive cases detected at screening of high-risk women versus cases detected in a general high-risk population and the general population1

Cancer stage	High-risk women screening cases: Breast MRI + mammogram (%)	Non-screening high-risk cases (%) *p* comparison with screening cases	Non-screening population cases (%) *p* comparison with screening cases
Invasive cancer < 10 mm	19/44 (43%)	5/39 (13%) $p = 0.04$	193/1,380 (14%) $p < 0.001$
Negative lymph node/micrometastases	28/42 (67%)	17/39 (44%) $p = 0.001$	657/1,380 (48%) $p < 0.001$
Histological Grade 1	19/43 (44%) Exception: BRCA1, BRCA2 grade 3: 12/19 (63%)]	4/37 (11%) $p = 0.01$	99/900 (11%) $p < 0.001$

1 Source: Kriege et al 2004

This study demonstrated a statistically significant stage shift toward earlier disease (smaller cancers, fewer cases with lymph node spread, more cases with Grade 1 histological classification) for invasive cancers detected in women participating in a breast cancer screening program with breast MRI and mammography (including the four interval cancers) compared to cases detected in two non-screening populations (although some of these cases may have been detected by screening). One exception was that a higher proportion of grade 3 cases were observed among the subgroup of BRCA1 and BRCA2 mutation carriers versus the two control groups (63% versus 51% respectively, Kriege et al 2004).

These findings indicate that screening high-risk women results in the early detection of breast cancer. However, these data do not allow a comparison of the cancer stage of cases detected by the addition of MRI versus cases detected by mammography alone in young high-risk women.

Interval cancer rates

No studies compared interval cancer rates for high-risk women participating in breast cancer screening programs that included breast MRI with rates for high-risk women participating in screening programs that did not include MRI.

Patient tolerance, health-related quality of life and attitudes

Two studies compared patient tolerance of breast MRI versus mammography. These studies reported mixed results. Overall, breast MRI appeared to be tolerated at least as well as mammography by most patients. In a survey of 288 women, 30% reported mammography as quite to very painful versus 1% who found breast MRI painful (Rijnsburger et al 2004). However participants also reported that they were less anxious about mammography (5% reported anxiety) than breast MRI (10% reported anxiety). The other study surveyed 36 women and thus may not be large enough to detect a true difference between tests (Kurian et al 2005). This study reported similar ratings of patient discomfort for both tests (moderate to maximal discomfort 42% for mammography versus 48% for breast MRI). In this survey, 61% of participants reported the experience of breast MRI was the same (19%) or better (42%) than mammography, whereas 39% reported that it was worse (Kurian et al 2005).

Rijnsburger et al (2004) also found a small significant reduction in self-rated health on visual analogue score over time for women participating in an MRI breast cancer screening program ($p \leq 0.01$). However, no other generic quality of life score changed over time.

Kurian et al (2005) also reported on the impact of an intensified breast cancer screening program that included breast MRI and ductal lavage on changing patient attitudes toward prophylactic mastectomy (Kurian et al 2005). Twenty one women (62%, 95% CI 44-78%) reported that participation did not change their attitude to prophylactic mastectomy. Of those who did change their attitude, more women reported that participation made them more opposed to prophylactic mastectomy rather than less opposed (24% versus 15% respectively, $p = 0.02$). No studies reported on the impact of breast MRI screening on prophylactic surgery rates.

Treatment effect

This review did not identify any studies comparing treatment outcomes for screen-detected breast cancer cases detected early by MRI plus mammography versus cases detected later by mammography or clinical examination alone in young high-risk women.

This section therefore discusses relevant evidence from studies reporting on:

(i) outcomes for mammogram screen-detected versus clinically detected cases in average-risk patients; and

(ii) outcomes for cases detected in high-risk women by stage (tumour size, axillary node status) and histopathological grade.

Outcomes for mammogram-screen detected cases versus clinically detected cases in the general population.

As outlined on page 9, a meta-analysis of randomised controlled trials has established that the early detection of breast cancer by mammography results in a 24% (95% CI 18-30%) reduction in breast cancer mortality in patients of average risk aged between 39 years and 74 years compared to cases presenting clinically (Duffy et al 2002). Screening trials including women aged 40-49 years have suggested the benefits of mammographic screening are more modest in this age group (Kerlikowske 1997). This may be partly attributed to the lower sensitivity of mammography in women under the age of 50 years (Kerlikowske et al 1996) and the higher risk of more aggressive tumours in this age group (Remvikos et al 1995; Tabar et al 1999).

Thus evidence of reduced mortality for mammogram screen-detected cases versus clinically-detected cases in the general population may not apply to the extra MRI screen-detected cases in young high-risk women. Factors limiting the applicability of evidence from these mammography trials include:

- Breast cancers in women with a genetic predisposition show different characteristics, including a higher proportion of aggressive tumours, than sporadic breast cancers and early detection may not bring the same survival benefits as are observed for sporadic cancers.

- The extra cases of breast cancer detected by the addition of breast MRI may represent a different spectrum of disease to the extra cases detected by mammography versus clinical presentation. Differences in size, grade and stage of disease for MRI screen-detected versus mammogram screen-detected cases are not clearly defined.

The benefits of mammography screening programs have been attributed to the detection of cases at an earlier stage of disease compared with cases detected clinically. Randomised controlled trials that show the highest proportion of node negative disease in mammography-detected cases have reported the greatest reduction in mortality to support this hypothesis (Smith et al 2003).

The same stage shift has not been consistently observed for mammogram screen-detected cased in women under the age of 50 years. One observational study of mammography screening for women less than 50 years with a family history of breast cancer reported a similar proportion of node negative disease among 23 mammogram-detected invasive cancers and 52 non-screen detected cancers (65% versus 60% respectively, $p = 0.8$, Kollias et al 1998). However, this study did find a statistically significant difference in the proportion of in-situ cancers detected by mammography in asymptomatic women (6 cases, 21%) versus clinically detected cases in symptomatic women (2 cases, 4%, $p = 0.01$). The authors proposed that mammography screening of young high-risk women was justified based on the potential benefits of early detection of in-situ cancers (Kollias et al 1998). Another observational study reported a statistically significant difference in the proportion of early stage disease in 26 mammogram screen-detected versus 24 clinically-detected cases, in high-risk women (81% versus 46% respectively, $p = 0.018$. These figures include cases classified as in-situ and stage I disease with similar findings for women aged 50 years of younger (Tilanus Linthorst et al 2000b).

Outcomes for cases detected in young high-risk women by tumour stage

This review did not identify any studies comparing treatment outcomes for high-risk women by stage of disease at diagnosis. Epidemiological data from population-based cancer registries in Australia and the United States have confirmed a survival benefit for breast cancers detected at stage I (T1N0) versus stage II (T1N1, T2N0) in the general population (Roder et al 2002, US National Cancer Database 2006). One large observational study has indicated that tumour size and nodal status are also important prognostic factors for BRCA1-associated breast cancers (Brekelmans et al 2006). This evidence suggests that, similar to screening programs in the general population, screening programs in high-risk women that detect a higher proportion of early stage disease will bring improved survival benefits. Although it is plausible that the extra cases detected by breast MRI represent an earlier stage of disease and therefore will benefit from earlier treatment, this review does not provide convincing evidence that MRI screen-detected invasive cancers are significantly smaller or more likely to be node negative than mammogram screen-detected cases (see page 56). This question still needs to be addressed.

What are the economic considerations?

Background

Economic evaluation of new health care technologies is particularly important where the new technology offers health benefits at additional cost. It is clear there will always be a limit to the additional cost which would be paid for a given health gain. Economic evaluation is generally aimed at determining whether such incremental costs represent value for money.

The usual process for an economic evaluation is first to consider the additional benefits accrued with the new technology relative to the comparator (the incremental effectiveness), and then determine cost differences between the new technology and the comparator (the incremental costs). Effectiveness is measured in clinically appropriate natural units or a multidimensional measure such as quality adjusted life years (QALYs). When both costs and effects are known, then an incremental cost-effectiveness ratio (ICER) can be determined. The calculation of an incremental cost-effectiveness ratio is shown below:

$$\text{ICER} = \frac{\text{Cost (new test)} - \text{Cost (comparator)}}{\text{Effectiveness (new test)} - \text{Effectiveness (comparator)}}$$

In cases where a new technology offers inferior or equal health benefits at a higher cost it does not provide value for money. This technology is "dominated" by the comparison technology. In cases where the new technology offers superior health benefits at a lower cost than the comparator, the new technology is said to be "dominant".

Existing Literature

A broad literature search was conducted to identify published economic evaluations of MRI for breast cancer screening. Medline, EMBASE, Pre-medline and Current Contents databases were searched using the clinical search terms listed in Table 8 combined with the economic search terms (cost$ or econ$).mp.

This search identified two studies: one study reported on a costing of conducting MRI and follow-up investigations in 109 high-risk women to calculate a cost per case detected (Tilanus-Linthorst et al 2000); the second study was an economic evaluation that estimated the effect of breast MRI using a mathematical model of the natural history of breast cancer and calculated an estimate of the incremental cost-effectiveness for performing breast MRI in different age groups and different risk groups (Plevritis et al 2006a). The main results of these studies are outlined in Table 24 and described below with a detailed appraisal of applicability to the present review.

Table 24 Existing published economic evaluations of breast MRI

Author / year	Type of analysis	Main outcomes
Plevritis et al 2006a	Modelled analysis using estimates of screening effects derived from a mathematical model of the natural history of breast cancer. Comparison of MRI plus mammography versus mammography alone for screening BRCA1/2 mutation carriers with sensitivity analyses including variation by age group and risk classification.	Cost per QALY gained by adding MRI from age 35-54 years: • US$ 55,420 for BRCA1 mutation carriers • US$ 130,695 for BRCA2 mutation carriers • US $ 98,454 for BRCA2 mutation carriers who have mammographically dense breasts. Annual MRI screening is not cost-effective among younger women (aged 25-34 years) because of their lower breast cancer incidence, or among older women (aged ≥ 55 years) because of declining quality of life and competing risk of death from other causes.
Tilanus-Linthorst et al 2000	Simple costing and calculation of cost per case detected following the addition of MRI in women with over 25% risk of breast cancer, negative mammography and > 50% dense breast tissue at mammography.	The total cost of MRI in the 109 women was €32,842 versus mammography €10,557 Additional cost to detect 3 breast cancers in 109 women • 193 MRIs performed at €170 • 51 ultrasound examinations at €61 • 29 FNAC at €127 • 2 benign excision biopsies at €1,026. The additional detection cost = €13,930 per detected patient.

In the study by Tilanus-Linthorst et al (2000), breast MRI was undertaken on 109 women with a cumulative risk of breast cancer of 25% or greater, negative mammography and > 50% breast density at mammography. MRI detected an additional three cases of breast cancer in this population. A simple costing was presented but no cost-effectiveness analysis was undertaken. The average cost per extra patient detected with breast cancer by MRI was estimated using resource use data for the 109 women and included the costs for the MRI scan and all associated follow up investigations such as ultrasound, fine needle aspiration cytology and biopsy. These costs were then compared to average costs incurred for mammography. The analysis indicated that MRI would incur an additional €13,930 euros per patient detected with breast cancer. The figures are assumed to be in 2000 values although this is not explicitly stated. It is difficult to assess the validity of this cost estimation and the applicability of unit estimates to Australian practice as the analysis could not be reproduced.

Published analysis of the incremental cost-effectiveness of breast MRI plus mammography versus mammography alone

The recently published economic analysis by Plevritis et al (2006a) used a model-based analysis to establish the cost effectiveness of adding breast MRI to annual mammography for screening BRCA1/2 mutation carriers from ages 25 to 69 years, compared to annual

mammography screening alone. In the base-case analysis, costs were evaluated from a societal perspective and included indirect costs associated with lost production time. Cost-effectiveness was reported both in terms of incremental cost per life year and incremental cost per QALY gained for MRI plus mammography versus mammography alone.

In addition to the published results, the author has provided additional cost-effectiveness data specific to the 35-54 year age group that excludes indirect costs for lost production time associated with investigations and treatment for the purposes of the present report (S Plevritis, personal communication 2 August 2006). These data are more applicable to the proposed use of breast MRI in Australia (women under the age of 50 years) from a government funding perspective. The published results and the secondary exploratory analysis using these additional data are presented below.

Methods

Plevritis et al (2006a) developed a cost-effectiveness model based on a continuous-time Monte Carlo simulation in C++. Outcomes were modelled using patient survival following cancer detection, estimated using a previously published mathematical model of the natural history (stage progression, detection and survival) of invasive breast cancer (Plevritis et al 2006b). This model used data from the Surveillance, Epidemiology and End Results (SEER) program, which were collected before population-based mammography screening (1975-1981). This natural history model was modified to adjust for the characteristics of BRCA1/2 mutation carriers.

Utility estimates were based on evidence from published literature. Resource use estimates included direct costs as well as indirect costs (Table 27). In particular, costs associated with time lost from work following screening were included in the cost-effectiveness ratios. Screening and related procedures were based on Medicare re-imbursement costs. Costs of cancer therapy were estimated from the literature. Costs were based on 2005 US dollars and a 3% discount rate for both costs and effects was applied.

Resource use associated with screening, diagnosis and breast cancer therapy was modelled based on information from the published literature. Assumptions regarding the percentage of diagnostic tests prompted by screening are presented in Table 25.

Table 25 Diagnostic tests prompted by screening (from Plevritis et al 2006a)

Test	Type of follow up prompted	Follow up as % screening population	Type of biopsy	Biopsy type as % of biopsies undertaken
MRI	MRI short interval 6 months	15% (30% initial screening)	–	–
	Biopsy	5%	Surgical, MRI-guided	0.25
			Surgical, mammography-guided	0.20
			Core needle biopsy, MRI-guided	0.25
			Core needle biopsy, ultrasound-guided	0.20
			Fine needle aspiration, ultrasound-guided	0.10

Table 25 cont.

Test	Type of follow up prompted	Follow up as % screening population	Type of biopsy	Biopsy type as % of biopsies undertaken
Mammography	Additional mammography views[1]	6%	–	–
	Ultrasound	2%		
	Biopsy	2%	Surgical, mammography guided	0.30
			Core needle biopsy, ultrasound-guided	0.17
			Stereotactic biopsy	0.49
			Fine needle aspiration, ultrasound-guided	0.04

1. 3% additional mammography views plus 3% 6 month follow up mammography

Additional assumptions regarding the cancer treatment for patients after diagnosis are:

- All cancer patients undergo mastectomy with breast reconstruction, 60% of women under the age of 50 years are assumed to have unilateral mastectomy and 40% bilateral mastectomy. Among women aged 50 years and older, the proportions are 50% each for unilateral and bilateral mastectomy.
- All breast cancer patients receive adjuvant chemotherapy with an anthracycline–taxane-based combination.
- Patients with oestrogen receptor-positive disease receive adjuvant tamoxifen.
- All patients receive follow up of three half-day physician visits per year for 4 years and two visits per year for the next 5 years.

Results

The model showed a higher expected survival rate for both BRCA1 and BRCA2 mutation carriers using MRI plus mammography than for mammography alone. The incremental cost per QALY for women screened between ages 34 and 55 years was AU$ 75,609 and AU$ 173,789 for BRCA1 and BRCA2 mutation carriers respectively (excludes indirect costs, Table 26).

Table 26 Incremental cost per QALY saved, breast MRI plus mammography versus mammography, as modelled by Plevritis et al (2006a)

	Costs USD	Cost AUD	QALYs saved	Incremental cost per QALY saved USD	Incremental cost per QALY saved AUD
Base case (25-69)[1]					
BRCA1	$27,321	$37,703	0.308	$88,651	$122,338
BRCA2	$29,605	$40,855	0.158	$188,034	$259,487
Subgroup analysis (34-55)[2]					
BRCA1	$12,541	$17,307	0.229	$54,789	$75,609
BRCA2	$13,194	$18,208	0.105	$125,934	$173,789

USD = 1.38 AUD (purchase parity rate 2005 OECD http://www.oecd.org/dataoecd/61/54/18598754.pdf).

1. Includes direct and indirect costs.

2. Direct costs only. Indirect costs excluded from the analysis.

Sensitivity analysis

The incremental cost-effectiveness of MRI screening was shown to vary by patient age group. Analyses on varying age groups showed the lowest cost per QALY for a screening population between 40 and 49 years. The results were sensitive to variations in:

- breast cancer incidence
- discount rate
- MRI costs
- sensitivity of mammography for the detection of breast cancer

Higher breast cancer incidence increased the incremental cost-effectiveness of MRI. Conversely higher MRI costs, sensitivity of mammography and the discount rate all reduced the incremental cost-effectiveness. Results were not shown to be sensitive to the detection threshold of MRI (varied from 3-8 mm). This can be explained by the fact that modelling of survival is based on SEER survival curves that do not differentiate survival in tumours < 1 cm. This assumption was applied because reliable survival curves for tumours this size are not available: these tumours were not frequently observed in the pre-mammography screening era the SEER data is based on (S. Plevritis, personal communication, 6 June 2006).

Discussion of paper

This economic evaluation has been based on a simulation model which applied evidence from the SEER database, published literature and assumptions based on expert opinion to US resource cost estimates for breast cancer screening investigations and treatment. It is clear that there are limitations to any modelled approach, particularly when modelling the highly uncertain variables required to estimate disease progression and potential survival benefit. Differences are likely to exist between actual patient outcomes and those indicated by a modelled cohort and the modelled survival benefits may or may not reflect real health gains. However, the general limitations of this analysis reflect the lack of current trial data linking evidence about the accuracy of MRI to improved patient survival and this modelled approach currently offers the best available evidence about the potential cost-effectiveness of MRI plus mammography versus mammography alone.

Interpretation of results for an Australian setting

Whilst the model provided by Plevritis et al (2006a) offers the best available evidence of the cost-effectiveness of adopting breast MRI for practice, population and relative prices in the US, the results need to be translated to allow interpretation in an Australian setting. A direct exchange rate conversion of the incremental cost-effectiveness from US to Australian dollars is an inappropriate measure of potential cost-effectiveness of adding MRI to mammography screening protocols in Australia. Variations in resource costs, clinical practice and the screening population between those assumed in Plevritis et al (2006a) and those applicable in Australia could influence results substantially. A discussion of these factors has been presented below along with a modelled analysis to help improve the interpretation of results.

Resource costs

The costs applied in the model by Plevritis et al (2006a) were evaluated on the basis of the resource use associated with breast MRI for the American health care system and society. A review of these resource inputs revealed significant differences in unit costs between the US and Australia for the screening, diagnostic follow up and treatment of breast cancer.

Table 27 shows the comparison of resource costs for screening and diagnostic followup, while breast cancer therapy costs are presented in Appendix H. All US figures have been converted to Australian dollars. Australian resource costs have been based on current MBS reimbursement fees and are consequently taken from a government perspective. It should be noted that the fees for breast MRI and MRI biopsy have been estimated because neither of these procedures are currently reimbursed through the MBS. The estimated fee for breast MRI has been based on an existing MBS code provided by the applicant for a similarly resource intensive procedure (631012, scan of head and neck). The cost for an MRI biopsy has been modelled by applying estimates for biopsy capital equipment, needles and MRI time (Table 27).

The comparison of cost estimates shows that the direct cost of mammography is similar for the US and Australia ($118 versus $90 respectively). However, the direct cost of MRI could be expected to be around half that charged in US ($538 vs $1,086). Diagnostic follow up procedures and breast cancer therapies are also estimated to be less costly. These differences will influence incremental costs and hence the incremental cost-effectiveness ratio.

Table 27 Resource cost estimates (Australian dollars)

		Australian cost data AU$			
	Plevritis et al (2006a) costs in AU$	MBS items included	Unit Costs	No	Total costs
Screening tests					
MRI	$1,432.44				$537.60
MRI[1]		*631012*	*$492.80*	1	
Gadolinium contrast		*63491*	*$44.80*	1	
Short interval follow-up MRI	$1,086.06				$537.60
Mammogram, both breasts[2]	$118.68	59300	$89.50	1	$89.50
Short interval follow-up mammogram	$107.64	59300	$89.50	1	$89.50
Costs of follow-up prompted by screening test					
Breast ultrasound (both breasts)	$96.60	55076	$109.10	1	$109.10
Surgical biopsy, mammogram-guided	$1,738.80	31506	$337.60	1	$337.60
Fine needle aspiration, ultrasound-guided	$638.94				$217.60
Fine needle aspiration biopsy		*31513*	*$119.35*	1	
Ultrasound guidance		*55070*	*$98.25*	1	
Core needle biopsy, ultrasound-guided	$721.74				$228.45
Core needle biopsy		*31548*	*$119.35*	1	
Ultrasound guidance		*55054*	*$109.10*	1	
Stereotactic biopsy	$1,073.64	31545	$515.50	1	$515.50
Surgical biopsy, MRI-guided	$2,315.64			1	$1,601.60
Surgical biopsy		*31506*	*$337.60*	1	
MRI-guidance[3]		*Estimated fee*	*$1,264*	1	
Core needle biopsy, MRI-guided	$1,324.80				$1,383.35
Core-needle biopsy		*31548*	*$119.35*	1	
MRI-guidance		*Estimated fee*	$1,264	1	

MRI and MRI guided biopsy are not currently reimbursed through the MBS.

1. The cost for MRI was provided by the applicant based on MBS item for a similarly resource intensive procedure.

2. Screening mammography is funded through BreastScreen Australia and not through the MBS. This cost refers to the MBS item for bilateral diagnostic mammography.

3. The cost of MRI-guided biopsy is based on estimated MBS cost of breast MRI ($492.80 x 1.5), plus estimated per MRI needle cost of $225 (expected range $150-$300 based on expert opinion) and capital cost of $300 per patient for biopsy equipment (expected cost $75,000 for biopsy equipment (range $50,000 - $100,000) assumed 5 year life expectancy treating 50 patients per annum, ANZHSN 2004) plus fee for biopsy.

Modelled analysis

To help interpret results from Plevritis et al (2006a) in an Australian setting, and in particular to assess the impact of the differences in resource costs indicated in Table 27, a more limited decision-analytic model was developed. This model maps the expected incremental costs of mammography versus MRI plus mammography for one round of screening and all investigations up to the point of diagnosis and applies estimates from Plevritis et al (2006a) for screening test sensitivity and specificity, cancer incidence and

resource utilisation prompted by screening (diagnostic follow up procedures). Using this framework, separate incremental cost estimates for one round of screening were modelled using US direct costs and Australian costs. The US incremental costs for one round of screening estimated using this model were compared to US lifetime incremental costs reported by Plevritis et al (2006a). The resulting ratio was then applied to one round of treatment costs for Australia to proportion up these costs to lifetime incremental costs from an Australian perspective (see Table 29).

Results of modelled analysis

The modelled analysis for one round of screening up to the point of diagnosis results in an incremental cost per cancer detected of AU$75,800 for BRCA1 and AU$175,013 for BRCA2 (Table 28). The results of the modelled analysis indicate that applying Australian resource costs in the Plevritis et al (2006a) model could reduce the reported incremental cost per QALY saved for women aged 35-54 from AU$75,609 to AU$28,974 for BRCA1 patients and AU$173,789 to AU$66,598 for BRCA2 patients. Costs are reported in 2005 Australian dollars.

Table 28 Results of modelled analysis, women aged 35-54, incremental cost of one round of screening, MRI plus mammography versus mammography (Australian dollars)[1]

	Total incremental cost 5,000 women	Incremental cost per patient	Extra cancers detected 5,000 women[2]	Incremental cost per cancer detected
BRCA1	$3,233,600	$647	43	$75,800
BRCA2	$3,233,600	$647	18	$175,013

1. Model is based on 5,000 women to illustrate the difference in extra cases detected for a surveillance population of 5,000 depending on whether the cumulative breast cancer incidence rate is estimated at 65% (BRCA1) or 45% (BRCA2) by age 70. The actual surveillance population risk is not known.

2. Calculated by converting cumulative breast cancer incidence rates of 65% (BRCA1) and 45% (BRCA2) mutation-carriers by age 70 years into annual probabilities and using the following estimates for test sensitivity and specificity from Plevritis et al (2006a)

BRCA1: Mammography sensitivity = 35%, specificity = 96%; MRI + mammography sensitivity 72%, specificity 87% (ages 35-54 years)

BRCA2: Mammography sensitivity = 42%, specificity = 95%; MRI + mammography sensitivity = 70%, specificity = 86% (ages 35-54 years)

Table 29 Modelled incremental cost-effectiveness analysis, MRI plus mammography versus mammography (Australian dollars)

	Average annual screen modelled incremental cost per woman	Lifetime reported Incremental costs per woman	Annual to lifetime cost ratio	Estimated Lifetime Incremental costs per woman	Estimated lifetime incremental QALYs per woman	Estimated incremental cost per QALY per woman
BRCA1						
AUD	$647	–	0.10	$6,632	0.229	$28,974
US direct costs	$1,688	$17,307	0.10	$17,307	0.229	$75,609
BRCA2						
AUD				$6,977	0.105	$66,598
US direct costs				$18,208	0.105	$173,789

These estimates suggest that, allowing for differences in relative and absolute prices in isolation, there is the potential for a lower incremental cost-effectiveness ratio than that suggested by a simple purchase power parity conversion. However, to translate evidence

on the incremental cost-effectiveness from the US-based Plevritis model, the influence of factors reflecting differences in clinical practice (in particular the sensitivity of standard screening protocols, baseline risk of the patient population and discount rate) in Australia from that modelled by Plevritis et al (2006a) needs to be considered.

Test accuracy

The incremental effect of MRI could change if different assumptions regarding test accuracy are applied. If an Australian screening program routinely combined ultrasound with mammography as currently performed in some clinics, the sensitivity of standard testing (mammography plus ultrasound) could be expected to increase from that applied in the Plevritis et al (2006a) paper (35%) to potentially as high as 49% (Kuhl et al 2005). This increased sensitivity would reduce the potential gain in effects from adding MRI. Conversely, the sensitivity of mammography plus MRI applied in Plevritis was conservative compared to some other studies (MRI + mammography sensitivity 70% and 72% assumed for BRCA1 and BRCA2 patients). Leach et al 2006, reported a sensitivity of 94% when both mammography and MRI were combined. A higher sensitivity would increase the incremental gain in the effects of mammography plus MRI. It is not possible to estimate exactly how either of these variations in test accuracy may affect the incremental cost per life year saved without recalibrating the Plevritis model.

Screening population

The Plevritis model specifically investigated BRCA1/2 mutation carriers and assumed a relatively high baseline breast cancer risk of 65% for BRCA1 and 45% for BRCA2. The screening population under study in our assessment is broader and is generally likely to be a lower-risk group. Breast cancer incidence was the most sensitive variable in the Plevritis model. Assuming a lower baseline cancer risk will lower absolute incremental effects and cost offsets and hence increase the incremental cost per life year for MRI.

Discount factor

The 3% discount factor applied in Plevritis is lower than the 5% rate currently recommended in Australia for economic evaluations by the Pharmaceutical Benefits Advisory Committee (PBAC). Increasing the discount factor applied to both costs and effects will drive a higher incremental cost per life year saved because screening costs are assumed to occur upfront while survival benefits are delayed and therefore incur heavier discounting. Sensitivity analyses reported in Plevritis et al (2006a) (including indirect costs of lost production time) indicated that the incremental cost per QALY of MRI would increase by over 50% if a discount factor of 5% rather than 3% was applied.

Discussion

The cost effectiveness analysis reported by Plevritis et al (2006a) needed to be translated to an Australian setting to reflect Australian clinical practice, proposed target screening population and Australian prices, applying an appropriate discount rate and excluding indirect costs associated with lost production time. The remodelled Australian analysis accounts for differences in costs between Australia and the US, but important differences in screening population, clinical practice and discount factor have not been accounted for. These differences would each increase the incremental cost per life year in Australia above that of the remodelled analysis. In particular, the incremental cost-effectiveness ratio will increase if:

1. MRI were compared with mammography plus ultrasound with a test sensitivity of 49% rather than the 35% modelled;

2. a lower breast cancer incidence is applied to reflect the lower-risk target population proposed in Australia (> 30% cumulative risk) than that modelled by Plevritis et al (2006a); and

3. a discount rate of 5% rather than 3% were applied to reflect PBAC guidelines, which according to sensitivity analysis undertaken by Plevritis et al (2006a) would increase the incremental cost effectiveness ratio by more than 50%.

While these additional considerations would clearly significantly increase the incremental cost effectiveness ratios above that of the remodelled analysis, the inability to adjust these variables in the Plevritis model does not allow the net effect of these three variables to be robustly estimated.

Total financial costs to government

The total financial costs to government of subsidising breast MRI screening have been estimated taking into account the costs of providing MRI and the costs of diagnostic follow-up (additional imaging and biopsy) prompted by MRI screening. Costs of cancer therapy post diagnosis have not been included. It has been assumed that MRI would be used together with mammography to replace the current practice of performing a mammography alone. In addition the following assumptions regarding the potential eligibility and uptake have been applied.

1. 1% of women 25-50 years of age have a genetic predisposition for breast cancer.

2. Only women 25-34 years of age who have an index case with breast cancer below the age of 35 and a genetic predisposition for breast cancer will be eligible for screening. This has been estimated as 20% of the at-risk women in the 24-35 age group.

3. All women 35-50 years of age with a genetic predisposition for breast cancer will be eligible for screening.

4. Only 20% of women 25-50 years of age who are eligible for screening will be identified.

Applying these assumptions, it has been estimated that approximately 5,000 women per year will present for additional MRI screening if government funding for MRI were to be provided. The estimated cost of MRI and the associated diagnostic follow-up per year per woman screened would be AU$647 based on the modelled calculations described in the previous section (see Table 28). Applying this figure, adding MRI to mammography in a screening population of 5,000 women gives an estimated total cost to government of AU $3,759,000 per year, a further AU$3,233,600 above the modelled equivalent cost for mammography alone (see base case Table 30).

Table 30 Financial implications: total cost to government per annum (Australian dollars)

	Expected number of women screened per annum	Total cost per annum addition of MRI	Incremental cost per annum addition of MRI
Base case 1% prevalence of genetic predisposition 20% of eligible women identified for screening	5,000	$3,759,002	$3,233,600
0.75% prevalence of genetic predisposition for breast cancer	3,800	$2,856,841	$2,457,536
40% of eligible women identified for screening	10,000	$7,518,004	$6,467,201

However, total government-perspective costs depend on criteria used to determine patient eligibility and uptake of the procedure. Variations in assumptions on the prevalence of having a genetic predisposition for breast cancer as well as the proportion of women subsequently identified for screening would clearly impact the figures presented. The number of women screened per annum may reduce if, for example, the prevalence of genetic factors resulting in a predisposition for breast cancer or uptake of screening is lower than expected or, as this review indicated, if around 1-5% of women refuse MRI due to claustrophobia. Equally, however, screening numbers may increase with improved identification of high-risk women and/or population growth. The financial impact for plausible higher and lower estimates for the number of women screened per annum based on variation in these assumptions is presented in Table 30.

It should be emphasised that in addition to uncertainty in screening numbers there is also uncertainty surrounding the cost of screening and follow-up investigations per patient, as outlined previously. Total cost to government will clearly vary according to the modelled assumptions considered in the cost effectiveness analysis. Variations in estimates of breast cancer incidence among those identified for screening, the sensitivity of standard screening protocols, discount factor and limitations in modelling of effects may all have an important impact on the estimated total cost per annum. Furthermore, it should be noted that these estimates are based on the assumptions used by Plevritis et al (2006a) on the diagnostic follow-up prompted by MRI. Total costs may therefore also vary according to the type and frequency of diagnostic follow-up prompted by MRI in Australia.

Discussion

This review presents strong evidence that MRI with gadolinium contrast is a safe test when contraindications to its use are duly observed. However, as discussed below, there is a limited body of evidence for conclusions about its effectiveness and cost-effectiveness for screening young women at high risk of breast cancer.

Evidence about the relative effectiveness of adding MRI to standard mammography protocols is limited to studies reporting on test accuracy. These studies are designed to demonstrate differences in the performance of different screening strategies and do not provide evidence about the impact of these strategies on patient outcomes. This review identified three level III-1/2 studies of test accuracy investigating the relative accuracy of screening protocols with and without breast MRI in young high-risk women. Together,

these studies provide a moderate volume of evidence (based on the performance of 3,792 scans in 1,414 women).

All three studies provided evidence applicable to the proposed use of breast MRI plus mammography versus mammography alone in Australia. All studies investigated women classified as having a lifetime risk of breast cancer of 20% or greater. Although only one study was designed to assess breast MRI in women under the age of 50 years (Leach et al 2005) (median age 40 years), both other studies reported the average age of participants was less than 50 years. The median age of participants in Kuhl et al (2005) was 40 years and the mean age of participants in Warner et al (2004) was 47 years.

None of the studies were assessed as high quality, which needs to be taken into consideration when interpreting the results, as it may compromise their ability to estimate the true accuracy of the tests. Two studies were classified as fair quality, level III-2 evidence because histopathology was only used to verify a positive test finding after a review of all imaging tests (Leach et al 2005, Warner et al 2004). This method may introduce partial verification bias, although the direction of this potential bias on estimates of the relative accuracy of adding breast MRI is unknown. The third study met the criteria for level III-1 evidence but was classified as low quality due to the investigators' acceptance of an interval of up to 8 weeks between mammography and breast MRI (Kuhl et al 2005). If disease progression is rapid, as may be anticipated in young women with a genetic predisposition to breast cancer, a long interval between tests may introduce the possibility of bias favouring the test performed last.

Despite these limitations in study quality, the consistency and precision of estimates of test sensitivity across these studies provide strong evidence that the combination of breast MRI and mammography is a highly sensitive test for the detection of breast cancer (meta-analysis: MRI plus mammography sensitivity 94% [95% CI 86-98%]) and offers approximately a 2.6-fold increase in the early detection of breast cancer cases compared to the use of mammography alone (meta-analysis: mammogram sensitivity 36% [95% CI 25-48%] in screening programs for high-risk women.

Unfortunately, less evidence was identified for an assessment of the relative accuracy of adding breast MRI to a mammography screening protocol that includes the use of ultrasound, either routinely or in selected women with increased mammographic breast density as is currently performed in Australia. Two of the above-mentioned studies reported that the sensitivity of mammography increased when combined with ultrasound (49% [95% CI 33-65%] Kuhl et al 2005) or ultrasound and clinical breast examination (64%, data not available to calculate confidence interval, Warner et al 2004). These data indicate that the impact of adding breast MRI on the overall sensitivity of breast cancer screening programs will be lower if these programs include the routine use of ultrasound and clinical breast examination.

Evidence about the specificity of screening protocols that include breast MRI was less consistent. The two studies that reported on the specificity of the combination of breast MRI and mammography found conflicting results which may be attributed, at least in part, to the different thresholds used to define false positives (Kuhl et al 2005, Leach et al 2005). Leach et al's (2005) classification of a false positive as a test finding that initiated further testing to exclude malignancy provides the most relevant data for estimates of the clinical and financial impact of false positives initiated by breast MRI. The authors found the addition of breast MRI reduced the specificity of screening to 77% compared to 93% for mammography alone (Leach et al 2005). These findings corresponded to a false

positive rate of 23% for the combination of MRI and mammography versus 6% for mammography alone (relative risk for being recalled for a false positive result 3.5 (95% CI 2.9-4.3, $p < 0.0001$, based on data reported by Leach et al 2005). The biopsy rate for false positive imaging findings was 5% for MRI and mammography versus 1.5% for mammography alone. Leach et al (2005) also showed that the combination of breast MRI plus mammography provided optimal sensitivity, with a small but statistically significant reduction in specificity versus MRI alone. As a result, any clinical benefits associated with earlier detection will need to be weighed against the potential distress and costs of additional investigations in women recalled for false positive MRI findings.

This review did not identify any studies that assessed the accuracy of including MRI versus mammography alone for screening high-risk women over the age of 50 years, or that compared the performance of these tests between age groups. Thus, there is insufficient evidence for conclusions about the accuracy of adding breast MRI to mammography screening programs in high-risk women aged 50 years or older. However, existing evidence that mammography is less sensitive in younger women (Kerlikowske et al 1996) suggests that the relative accuracy of breast MRI plus mammography versus mammography alone is likely to be greater in women younger than 50 years than older women.

This review also assessed the relative accuracy of breast MRI as a replacement test for mammography in order to provide a broad review of the evidence available about breast MRI. A larger volume of evidence was available to address this question than the question of breast MRI as an additional test. Six studies reporting level III-1 to III-2 evidence of test accuracy reported on data collected from a total of 3,795 patients (8,487 scans). These studies were appraised as fair to low quality. Meta-analysis estimated MRI sensitivity at 81% (95% CI 74-87%) compared to mammography sensitivity at 36% (95% CI 28-44%). Four of the six studies reported that the sensitivity of MRI was statistically significantly higher than mammography. One of these studies also showed that this difference was maintained at subsequent screening rounds (Kriege et al 2004). The other two studies detected a total of less than 10 cancers and thus had less statistical power to detect differences in test sensitivity (Lehman et al 2005; Podo et al 2002). Estimates of MRI specificity varied significantly across studies (MRI specificity range 81-99%, mammography specificity range 93-100%), which may in part reflect differences in the definition of a false positive used and precludes the statistical synthesis of results. Four studies showed that breast MRI was statistically significantly less specific than mammography. Of the two studies that did not find a significant difference between tests: Podo et al (2002) was a relatively small study (105 participants, 105 scans) and therefore lacked statistical power to detect a difference; Kuhl et al (2005) used a higher threshold for classifying a positive result than other studies. Overall, these studies provide a strong body of evidence that MRI is a more sensitive but less specific test than mammography in detecting breast cancer in young high-risk women.

No studies have directly assessed the impact of including breast MRI in mammogram screening programs on patient survival, interval cancer rates or other clinical outcomes. Two fair to high quality studies assessing patient tolerance of different breast screening procedures indicate that breast MRI is tolerated at least as well as mammography by most patients (Kurian et al 2005, Rijnsburger et al 2004).

Epidemiological evidence of improved survival for breast cancer cases, including cases associated with a BRCA mutation diagnosed at an early stage of disease (small primary tumour with no spread to axillary nodes), versus cases diagnosed at a later stage (larger

primary tumour and/or spread to axillary lymph nodes) suggest the potential benefits of screening tests that can detect cancers at an earlier stage of disease. In addition, mammography screening trials reporting a higher proportion of node negative disease in mammogram screen-detected cases have reported greater reductions in morbidity. These findings provide supportive evidence for the potential benefits of early detection of breast cancer, although the applicability of this evidence to MRI-screen detected cases versus mammography screen-detected cases in young high-risk women is unknown. Two large accuracy studies comparing the stages of cases detected by MRI with or without mammography versus mammogram screen-detected cases provide the most applicable evidence for conclusions about the potential clinical impact of adding breast MRI to annual mammography screening programs (Leach et al 2005, Kuhl et al 2005). Both studies reported a favourable but non-significant reduction in the proportion of cases detected with spread to axillary node using breast MRI. These data are insufficient to determine whether breast MRI produces a stage shift in the detection of breast cancer.

Overall, this review presents a strong body of evidence that the addition of breast MRI is safe and will produce a statistically significant and substantial 2.6-fold increase in the early detection of breast cancer. Evidence about the relative specificity of breast MRI is less consistent; however, one study that reports the most clinically relevant data indicates breast MRI produces a three-fold increase in recalls for further investigation, including biopsy of false positive findings (Leach et al 2005). Expert opinion suggests that improved early detection will reduce mortality from breast cancer based on evidence from mammography trials conducted in older average populations, although the magnitude of the clinical benefits associated with early detection by MRI versus later detection by standard mammography screening programs in young high-risk women cannot be estimated and thus an economic evaluation to estimate cost-effectiveness could not be performed.

Investigators in the US have recently published a modelled economic analysis that relies on estimates of the effectiveness of screening derived from a mathematical model of the natural history of breast cancer (Plevritis et al 2006a). This analysis suggests that early detection of breast cancer by MRI leads to a 23% reduction in breast cancer mortality in high-risk women aged between 25 and 69 years (corresponding to a survival benefit of 1.4 to 2.1 years for BRCA1 and BRCA2 mutation carriers respectively). The validity of the mathematical model used to derive these estimates is difficult to assess without observational evidence about the natural history of breast cancer and appropriate comparative data from clinical studies. Thus, considerable uncertainty surrounds these estimates of effectiveness. Furthermore, the estimates of resource use and costs reported in this study are based on clinical practice in the US and thus the estimates of cost-effectiveness may not apply to Australia. A secondary exploratory analysis of data from this study for women aged 35-54 years, that only takes direct costs into consideration and applies Australian costs to the point of diagnosis, suggests the addition of breast MRI results will produce an incremental cost per QALY saved of AU$75,609 to AU$28,974 for BRCA1 patients and AU$173,789 to AU$66,598 for BRCA2 patients. However, differences between Australia and the US in the use of follow-up tests, the sensitivity of existing screening protocols that include ultrasound, the incidence of breast cancer in the proposed screening population, and discount factors are likely to overestimate the cost-effectiveness of breast MRI markedly. Despite these important limitations, these data indicate that the addition of breast MRI may potentially be cost-effective for screening a highly select population at high risk of breast cancer. The Plevritis model also provides a useful indication of the key factors that may influence the cost-effectiveness of MRI

(MRI costs, incidence of breast cancer in target population, sensitivity of standard screening tests and discount factor).

Finally, owing to the uncertainty about the health benefits gained and cost-effectiveness of adopting breast MRI in screening programs for young women at high risk of breast cancer, the existing evidence does not satisfy the WHO Consultation Group's criteria for implementing a medical screening program (Strong et al 2006). Theoretical evidence that early detection by breast MRI will lead to an, as yet unmeasured, reduction in breast cancer mortality and information estimating the financial implications of implementing a screening program provide sub-optimal evidence for a rational decision about funding.

As raised in the background section of this report, other factors that may influence a decision to implement breast MRI include: judgement about the validity of linking evidence of mortality reduction from the early detection of breast cancer from screening trials in average-risk women over the age of 50 years to the early detection of breast cancer in younger high-risk women; the ethics and feasibility of conducting a trial to quantify these health benefits; the total costs of introducing breast MRI based on the potential size of the target population and anticipated uptake of the procedure; and the availability of MRI facilities with appropriate equipment and experience to perform the service.

Other issues raised by the Advisory Panel relevant to discussions about the proposed implementation of breast MRI were based on consideration of the specific expertise needed to provide the service and the unique needs of the proposed target population. It was the expert opinion of the Advisory Panel that, if breast MRI is introduced, assessment of patient eligibility and testing should be performed by specialists working within multi-disciplinary teams with expertise in breast imaging and MRI. This would help ensure appropriate referral and co-ordination of screening and diagnostic services including the counselling of young women identified as being at high risk about the potential benefits and harms of testing and support the ability of MRI facilities to gain and maintain adequate skills in the technique.

Conclusions

Safety

The most important safety consideration for breast MRI is ensuring that patients with ferromagnetic or electromagnetic devices or other contraindications to exposure to magnetic fields are identified and excluded from testing.

In patients without contraindications, breast MRI is a safe procedure with no irreversible or hazardous biological effects other than the potential risk of allergy to the contrast medium gadolinium, estimated to occur in approximately 1 in 10,000 patients. Some women cannot tolerate MRI due to claustrophobia. Three studies reviewed reported that 1 to 5% of women refused MRI due to claustrophobia or other reasons (Tilanus-Linthorst et al 2000, Kriege et al 2004, Leach et al 2005).

The proposed screening protocol of breast MRI plus mammography carries a risk of radiation-induced breast cancer due to the ionizing radiation used for mammography. This risk is low but there is theoretical evidence to suggest that it increases with repeated exposures, younger age at exposure, and is potentially greater for women with a family history of breast cancer. However, there is ongoing uncertainty about the true level of risk in this patient population. Evidence from one large level III-2 retrospective cohort study suggests that exposure to x-rays before the age of 40 years produces a two-fold increase in the relative risk of breast cancer for women with a BRCA1 or BRCA2 gene mutation, and a higher risk if exposure occurs at younger ages, or on multiple occasions (Andrieu et al 2006). However, another level III-2 case-control study that included a larger number of breast cancer cases did not identify an overall association between mammography exposure and breast cancer risk in this group of women (Narod et al 2006).

Effectiveness

Primary review question: breast MRI as an additional screening test for the detection of breast cancer in high-risk women under the age of 50 years

Evidence assessing the effectiveness of adding MRI to standard mammography screening protocols is limited to three studies reporting level III-1 and III-2 evidence of test accuracy. Despite limitations in the quality of study design and reporting of these studies, the size, statistical precision and consistency of estimates of improved test sensitivity provide strong evidence that the addition of breast MRI increases the early detection of breast cancer over mammography alone in young high-risk women. Meta-analysis shows a statistically significant 2.6-fold increase in the early detection of breast cancer (MRI sensitivity 94% [95% CI 86-98%] versus mammography sensitivity 36% [95% CI 25-48%]).

Variation in the reported specificity of breast MRI between studies and differences in the thresholds used to report false positives do not allow a reliable estimate of the relative specificity of the combination of breast MRI and mammography versus mammography alone; however, one study provided evidence that this strategy leads to a threefold

increase in the patient recall rate for further investigation and biopsy due to false positive findings (Leach et al 2005).

Two studies also provided evidence that breast MRI will also detect more cancers when added to screening programs that include the routine use of ultrasound, although the relative sensitivity of adding MRI will be lower than for screening programs based on mammography alone.

No evidence was identified from clinical studies of breast screening programs in high-risk women or epidemiological studies to estimate the magnitude of the clinical benefits associated with early detection versus later detection in young high-risk women. A recently published analysis uses a mathematical model of the natural history of the disease to estimate that breast MRI will reduce the risk of breast cancer mortality by 23% in high-risk women (thus providing similar clinical benefits to those demonstrated by breast cancer screening trials in older average-risk women) (Plevritis et al (2006a). However, considerable uncertainty surrounds this estimate. Although it is plausible that the extra cases detected by breast MRI represent an earlier stage of disease and therefore will benefit from earlier treatment, this review does not demonstrate that MRI screen-detected invasive cancers are statistically significantly smaller or more likely to be node negative than mammogram screen-detected cases.

Secondary review question: High-risk women aged 50 years and older

This review identified no evidence for additional conclusions about the relative effectiveness or accuracy of breast MRI in high-risk women aged 50 years or older, because the majority of women enrolled in MRI accuracy studies were under the age of 50 years. Mammography has a higher sensitivity in older women and therefore the relative accuracy of breast MRI is likely to be lower in this age group.

Secondary review question: Breast MRI as a replacement test for mammography

This review provides strong evidence from six level III-1 and III-2 studies of test accuracy that breast MRI is a more sensitive but less specific test than mammography for detecting breast cancer in young high-risk women, although reports of MRI specificity varied widely (meta-analysis: MRI sensitivity 81% [95% CI 74-87%], mammography sensitivity 36% [95% CI 28-44%]; MRI specificity range: 81-99%; mammography specificity range 93-100%). As before, the magnitude of the clinical benefits associated with early detection of breast cancer in this population are uncertain.

Economic considerations

The economic evaluation published by Plevritis et al (2006a) suggests that breast MRI is potentially cost-effective for screening very high-risk women in the US such as BRCA1 mutation carriers between the ages of 35 and 54 years, but is unlikely to be cost-effective for screening BRCA2 carriers or a wider risk or age population such as that under consideration in Australia.

A secondary exploratory economic analysis for Australia, based on the Plevritis model with the same assumptions about diagnostic performance and the effects of screening, but excluding indirect costs and applying Australian relative prices, suggests the potential

for breast MRI to be cost-effective in a select high-risk subgroup of the proposed screening population in Australia eg. BRCA1 mutation carriers aged 35-54 years.

However, while indirect costs could be excluded to comply with Australian Government decision-making processes and Australian prices applied, this secondary analysis has not been able to allow for important factors that would increase the modelled incremental cost effectiveness ratios:

- Inclusion of women for screening in Australia with a lower baseline risk of breast cancer (≥ 30% versus 45-65% lifetime risk).

- A higher discount rate (5% versus 3%).

Furthermore, differences between the sensitivity of standard screening practice in Australia (up to 49% for a combination of mammography and ultrasound) and that assumed in the Plevritis model (35% for mammography alone) could also increase the modelled cost-effectiveness ratio. However, this effect may be offset if the sensitivity of MRI plus mammography reaches 94% as estimated in the present systematic review compared to the conservative estimates applied in the Plevritis model (MRI plus mammography sensitivity 70% for BRCA1 and 72% for BRCA2 for women aged 35-54 years).

The inability to adjust these variables in the Plevritis model does not allow the net effect of these three additional considerations to be robustly estimated. Therefore, even if the methods and assumptions used to model the incremental effects of MRI are accepted, the secondary exploratory analysis does not provide sufficiently valid data to determine an explicit threshold range of baseline breast cancer risk at which the addition of breast MRI may and may not be cost-effective in Australia.

Recommendation

Breast MRI, when combined with mammography, is safe and effective in the diagnosis of breast cancer in asymptomatic women at high risk, when used as part of an organised surveillance program.

Evidence suggests that breast MRI in combination with mammography may be cost-effective when compared with mammography alone in high risk women aged less than 50 years.

MSAC recommends interim public funding for breast MRI in the diagnosis of breast cancer in asymptomatic women with a high risk of developing breast cancer when used as part of an organised surveillance program.

Evidence should be reviewed in not less than 3 years.

- The Minister for Health and Ageing endorsed this recommendation on 5 February 2007 -

Appendix A. MSAC terms of reference and membership

MSAC's terms of reference are to:

- advise the Minister for Health and Ageing on the strength of evidence pertaining to new and emerging medical technologies and procedures in relation to their safety, effectiveness and cost-effectiveness and under what circumstances public funding should be supported;
- advise the Minister for Health and Ageing on which new medical technologies and procedures should be funded on an interim basis to allow data to be assembled to determine their safety, effectiveness and cost-effectiveness;
- advise the Minister for Health and Ageing on references related either to new and/or existing medical technologies and procedures; and
- undertake health technology assessment work referred by the Australian Health Ministers' Advisory Council (AHMAC) and report its findings to AHMAC.

The membership of MSAC comprises a mix of clinical expertise covering pathology, nuclear medicine, surgery, specialist medicine and general practice, plus clinical epidemiology and clinical trials, health economics, consumers, and health administration and planning:

Member	Expertise or Affiliation
Dr Stephen Blamey (Chair)	general surgery
Associate Professor John Atherton	cardiology
Professor Syd Bell	pathology
Dr Michael Cleary	emergency medicine
Dr Paul Craft	clinical epidemiology and oncology
Dr Kwun Fong	thoracic medicine
Dr David Gillespie	gastroenterology
Dr Debra Graves	medical administrator
Professor Jane Hall	health economics
Professor John Horvath	Chief Medical Officer, Department of Health and Ageing
Dr Terri Jackson	health economics
Professor Brendon Kearney	health administration and planning
Dr Ray Kirk	health research
Associate Professor Frederick Khafagi	nuclear medicine
Professor Alan Lopez	medical statistics and population health

Associate Professor Donald Perry-Keene	endocrinology
Dr Ewa Piejko	general practice
Ms Sheila Rimmer	consumer health issues
Ms Samantha Robertson	Department of Health and Ageing representative
Professor Jeffrey Robinson	obstetrics and gynaecology
Professor Ken Thomson	radiology
Dr Douglas Travis	urology
Dr Mary Turner	Australian Health Ministers' Advisory Council representative
Dr David Wood	orthopaedics

Appendix B. Advisory Panel, Evaluator and Project Manager

Advisory panel Application 1089 Breast MRI

Dr Paul Craft (Chair) MBBS, MPH, FRACP Director Medical Oncology Unit The Canberra Hospital	MSAC member
Associate Professor Richard Bell MBBS, FRACP, FRCPA, MRACMA Director of Cancer Services Director of Medical Oncology The Geelong Hospital Barwon Health, Victoria	Medical Oncology Group of Australia nominee
Dr John Buckingham MBBS, MS (Minn), FRACS, FACS Department of Surgery Calvary Hospital	Royal Australasian College of Surgeons nominee
Associate Professor Jennifer Cawson MBBS, FRANZCR, MPH, MD Director, St Vincent's BreastScreen Consultant Radiologist, St Vincent's Hospital	Royal Australian and New Zealand College of Radiologists nominee
Dr Debra Graves MBBS MHA FRACMA CEO Royal College of Pathologists of Australasia	MSAC member
Dr Bronwen Harvey MBBS BA MPH(Hons) GCHE Medical Adviser Targeted Prevention Programs Branch Population Health Division Australian Government Department of Health and Ageing	Medical Adviser Department of Health and Ageing
Mr Ian Morris MSc, GradDipHlthSc(Admin), FIR, CD Chief Radiographer Womens and Childrens Health Service, Western Australia	Australian Institute of Radiography nominee
Ms Margaret Tassell Consumer Representative Breast Cancer Network Australia	Consumers' Health Forum nominee

Evaluators

Ms Alison Griffiths
BA (Hons) Economics

NHMRC Clinical Trials Centre
University of Sydney

Ms Wei Lei
BClinMed (China), MMed

Dr Sarah Lord
MBBS, MS (Epi)

Ms Sharon Parker
BHSci, MPH

Ms Suchaya Thongyoo

Ms Silke Walleser
BSc (Hons) MPH

Members from the Department of Health and Ageing

Ms Alex Lloyd
Project Manager

Health Technology Section
Medicare Benefits Branch

Ms Marlene Williamson
Project Manager

Appendix C. Example of a protocol for breast MRI for high-risk women

This example of a breast MRI protocol has been extracted from information provided by Professor Christobel Saunders from the University of Western Australia on behalf of the Australian Breast MRI Working Group to illustrate the technical requirements of scanning.

Rationale

- Breast MRI is sensitive but not specific. Specificity relies on observing the rate of early enhancement and the morphology of the enhanced lesion.
- The examination must provide serial T1-weighted images to display and plot the enhancement rates, particularly in the first 8 minutes after injection.

This protocol collects a pre-contrast axial scan and a series of post-contrast axial scans, followed by an optional high resolution fat-suppressed sequence.

Summary

All of these sequences can be performed using 1.5T or 3T magnets with phased-array compression breast coils, with administration of gadolinium contrast. The patient is scanned prone with breasts in the bilateral phased array breast coil. If there is the possibility of parallel imaging (on most 3T units with multi-element coils and on several 1.5T units) the sequences are shorter. Even with nonparallel imaging it is possible to do sagittal imaging of both breasts, sequentially, toggling back and forward from one breast to the other. It is possible to do the imaging axially but the spatial resolution is not as good as with sagittal scanning.

Contrast

- 0.16-0.20 mmol/kg gadolinium
- Volume = Weight (kg) divided by three
- The correct dose is critical to allow standard levels of lesion detection/optimal lesion conspicuity.
- Contrast must be injected rapidly into the antecubital vein preferably using a power injector (> 1 ml/sec) and followed with a saline flush of 20 mls.

Sequences

1. Axial T2 FSE both breasts, partly as a localizer but at sufficient resolution to allow identification (in retrospect, if necessary) of high signal foci which are more likely fibroadenomas than cancers.

2. Sagittal (or axial) T1 non-fat-suppressed both breasts.

3. Sagittal (or axial) STIR both breasts.

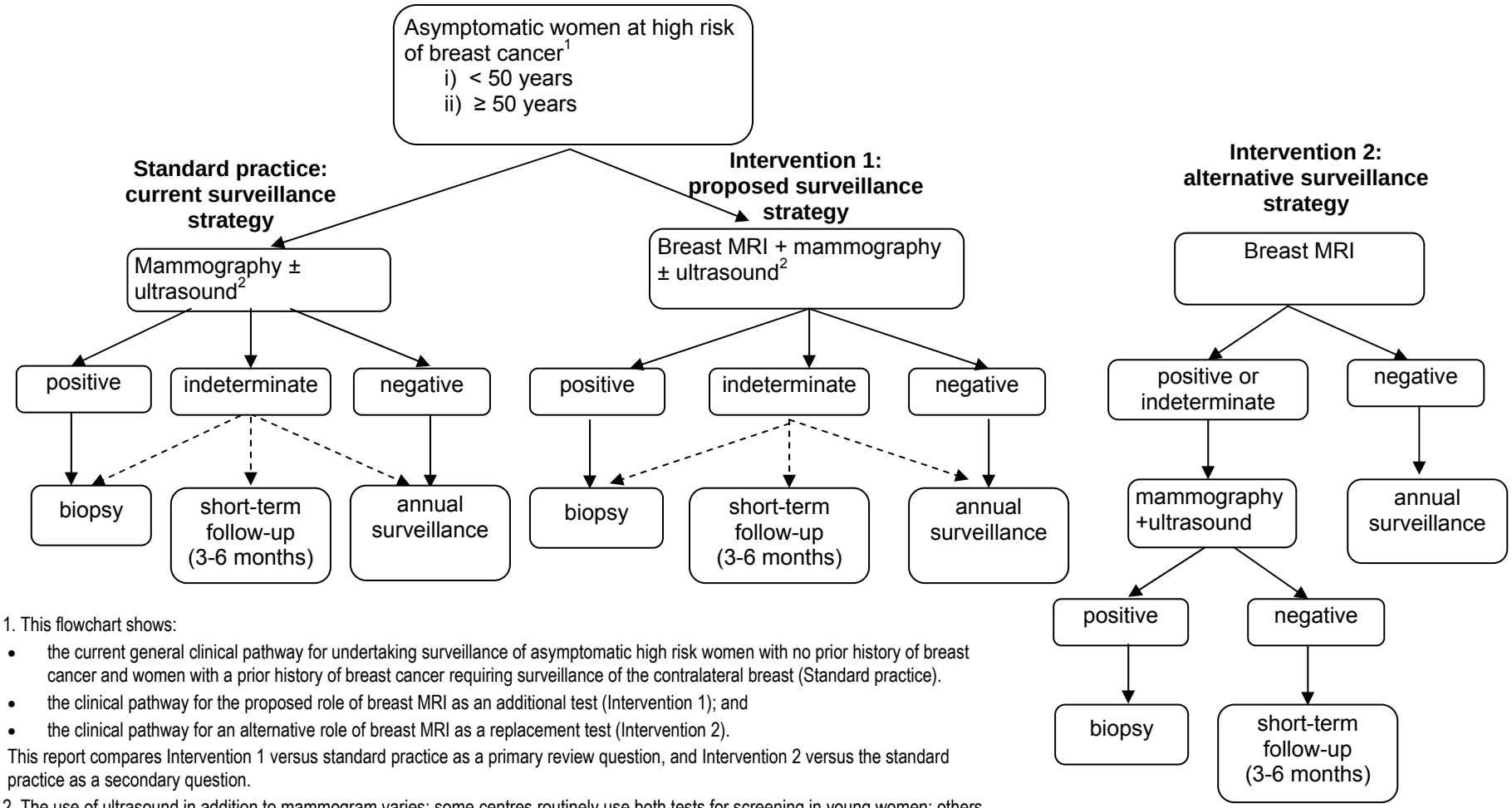
4. Dynamic Series: pre and 3 or 4 post 0.16 mmol per kg Gadolinium as Magnevist (or the equivalent iodinated Gadolinium contrast agent). These can be performed axially or as split sagittal. If they are fat-suppressed the acquisition time is about 90 seconds with either technique. A reasonable alternative is to do the dynamic series non-fat-suppressed and subtract the pre-contrast series from each of the dynamic post-contrast series.

   The post contrast series is done immediately post injection and then immediately again, for a total of three post-contrast series (a fourth post-contrast set about 6 minutes post gadolinium may be done).

5. Sagittal 3D volume acquisition fat suppressed with ≤ 1 mm thickness. This series can be used to generate a Maximum Intensity Projection (MIP) which is useful for further localizing any suspicious lesion (the major demonstration in many dynamic series being in the sagittal plane it is sometimes necessary to know exactly where the lesion is in the other planes).

(The protocol for finding lesions and filming and storage of digital images is not shown here).

Appendix D. Clinical flowchart



1. This flowchart shows:

- the current general clinical pathway for undertaking surveillance of asymptomatic high risk women with no prior history of breast cancer and women with a prior history of breast cancer requiring surveillance of the contralateral breast (Standard practice).
- the clinical pathway for the proposed role of breast MRI as an additional test (Intervention 1); and
- the clinical pathway for an alternative role of breast MRI as a replacement test (Intervention 2).

This report compares Intervention 1 versus standard practice as a primary review question, and Intervention 2 versus the standard practice as a secondary question.

2. The use of ultrasound in addition to mammogram varies: some centres routinely use both tests for screening in young women; others use it selectively eg for women with dense breasts or follow-up of positive mammography.

Appendix E. Characteristics, appraisal and results of included systematic reviews

Title/Author/Year	Objective of report	Number & publication dates of included studies	Population considered in included studies Test comparison	Conclusions/recommendation	Quality Assessment
HTA and horizon scanning reports					
Australia and New Zealand Horizon Scanning Network Health Policy Advisory Committee on Technology Australia, 2004	Horizon scan report To assess the current state of development of MRI screening for breast cancer, present use, potential future application and likely impact on the Australian health care system in genetically high-risk women	10 studies published 2000-2004 (Reported in 12 articles and 3 abstracts)	Population Asymptomatic women: • with no prior history of breast cancer • with prior history of breast cancer • at high risk of breast cancer Test comparison: MRI vs mammography	MRI appears to be a more sensitive screening test than mammography with comparable false positive rates (based on 6 test accuracy studies). Breast MRI may have an application in the screening of pre-menopausal women at high risk of breast cancer including women with a family history of breast or ovarian cancer germ-line mutation associated with breast cancer Introduction of an MRI screening program may require the purchase of additional MRI scanners.	High quality
Breast MRI for detection or diagnosis of primary or recurrent breast cancer Blue Cross Blue Shield Association Assessment Program Volume 18, Number 15 USA Dec 2003	To assess the effectiveness of breast MRI screening for asymptomatic women at high risk of breast cancer due to genetic predisposition	5 studies Published 2000-2002	Patients with genetic high risk for breast cancer MRI vs mammography	MRI is recommended for screening high-risk women due to evidence that it is more sensitive with equivalent or slightly inferior specificity than mammography (based on 2 test accuracy studies). Other conclusions: • There is no evidence linking evidence of improved test accuracy to improved patient outcomes; however, based on inferences of treatment effectiveness, improved sensitivity could either provide equivalent or improved health outcomes in a high-risk genetic group. Increased false positives are a possible risk but may be of lesser consequence in this population. • There is no evidence about the effectiveness of screening in women under	High quality

Title/Author/Year	Objective of report	Number & publication dates of included studies	Population considered in included studies Test comparison	Conclusions/recommendation	Quality Assessment
				40 years of age, thus current guidelines for mammography screening young high-risk women are based on inference from existing evidence in other populations.	
MRI for the detection of breast abnormalities Institute for Clinical Systems Improvement USA Dec 2003	Included a systematic review to assess breast MRI for screening for high-risk women Other potential indications assessed: • local staging of recurrent breast cancer • monitoring response to neoadjuvant therapy • problem solving, eg: a. questionable lesion on mammogram/US b. focal area of clinical concern with negative mammogram and US c. questions of recurrence d. A palpable axillary lymph node metastasis from presumed breast primary tumour with negative mammogram and clinical breast exam • evaluation of silicone implants	5 studies Published 2000-2003	High-risk patients, criteria not specified MRI vs mammography MRI vs US	Investigation of MRI screening of high-risk patients is ongoing in several multicentre studies. It is not recommended for the average-risk population. Other relevant conclusions: • In the absence of contraindications, MRI is a safe procedure. • MRI sensitivity for detecting invasive breast cancer range 93%-100% MRI specificity range 37%-96% (Conclusion Grade II). • MRI sensitivity for detecting DCIS range 45%-100%. • Patients should be aware of possible false-positive findings (4-63%) and false-negative findings (0%-55%), depending on the indication. • Facilities offering MRI for breast cancer evaluation must also have a breast centre where further assessment with diagnostic equipment (including biopsy) can be done.	High quality
Other conclusions: • Breast MRI accurately estimates residual disease in 97%-100% of patients after induction of chemotherapy prior to surgery (2 studies). • Breast MRI sensitivity for evaluation of silicone breast implant rupture: range 78%-87%; specificity range 76%-91%. • Breast MRI is helpful in ruling out invasive cancer in many problem-solving situations (negative predictive values 96%-100%). However, should not be used to avoid further evaluation of a suspicious lesion detected on mammogram or US.					

Title/Author/Year	Objective of report	Number & publication dates of included studies	Population considered in included studies Test comparison	Conclusions/recommendation	Quality Assessment
Clinical guidelines					
NICE Clinical Guidelines and Evidence Review for the classification and care of women at risk of familial breast cancer McIntosh et al UK. 2004 (updated October 2006, see Appendix J for supplementary information)	Clinical guidelines for the classification and care of women at high risk of familial breast cancer Included a systematic review to assess screening for high-risk women	4 studies comparing MRI versus mammography for screening high-risk women (1 study investigating addition of MRI in women with > 50% breast density and no abnormalities detected at mammography) published 2000-2001	Women at high risk due to strong family history, or known BRCA1/2 mutation MRI + mammography vs mammography MRI vs mammography MRI vs US MRI vs CBE	Women at moderate-high risk should be offered annual mammographic surveillance from age 40 years. (level C recommendation based on level III evidence or extrapolated from level I-II evidence) MRI and ultrasound should not be used in routine screening of high-risk individuals but may have a role in problem-solving mammographically detected abnormalities. (level D recommendation: based on Level IV evidence or extrapolated from level I-II evidence) This recommendation should be reviewed when ongoing studies report in the next two years. Relevant findings: • There is some evidence that MRI is more sensitive for the detection of breast cancer than mammography or combined mammography, CBE, US screening in women with a strong family history and/or BRCA1/2 mutations. • Mammography surveillance is less sensitive in younger women, women with a family history of breast cancer, and in BRCA1/2 mutation carriers.	High quality

Title/Author/Year	Objective of report	Number & publication dates of included studies	Population considered in included studies Test comparison	Conclusions/recommendation	Quality Assessment
American Cancer Society Guidelines for Breast Cancer Screening: Update 2003 Smith et al US 2003 (Further update expected late 2006).	To produce updated evidence-based guidelines for breast cancer screening Included a review of the evidence for screening women at high risk of breast cancer	4 studies Published 2000-2001	Women at high risk due to strong family history, or known BRCA1/2 mutation MRI vs mammography	Breast MRI is a more sensitive screening test in younger high-risk women than mammography. Breast MRI specificity varied according to follow-up management but was generally lower than that of screen-film mammography. Women at increased risk of breast cancer might benefit from: • earlier initiation of screening • shorter screening intervals, or • the addition of US or MRI to mammography and CBE. However, the evidence currently available is insufficient to justify recommendations for any of these screening approaches. Potential disadvantages for adopting MRI include: costs, lack of standardised techniques, inability of MRI to detect microcalcifications, variability of equipment, higher rate of false positive findings, access to MRI guided biopsies.	High quality

Author & Year	Objective of report	Number & publication dates of included studies	Population considered in included studies Test comparison	Conclusions/recommendation	Quality Assessment
Systematic reviews					
Calderon-Margalit & Paltiel 2004	To evaluate the efficacy of screening, bilateral prophylactic mastectomy, prophylactic oophorectomy, chemoprevention for prevention and improving survival in BRCA1 and BRCA2 carriers Included a review of evidence about MRI screening	4 studies investigating breast MRI for screening high-risk women Published 2000-2001	Asymptomatic high-risk women MRI versus mammography	No screening studies have assessed impact of screening on breast cancer mortality, breast cancer stage or grade, or quality of life of breast cancer patients. Accuracy studies differ in study populations and reference standard, and did not distinguish between tests performed for screening or diagnosis. Further research is needed to assess efficacy of screening MRI and to determine criteria for most efficient application to reduce unnecessary procedures and reduce costs.	High quality
Elmore et al 2005	To evaluate breast cancer screening methods Included a review of evidence about screening methods for women at high risk of breast cancer	6 studies investigating breast MRI for screening high-risk women Published 2000-2004	High-risk women MRI versus mammography +/- US	MRI has no radiation and is more sensitive and slightly less specific than mammography in high-risk women Disadvantages are: requires contrast, time-consuming, contra-indicated in some women due to ferromagnetic devices or claustrophobia. Recommendations for high-risk women not stated.	High quality
Irwig et al 2004	To assess the accuracy of breast cancer screening technologies	4 studies investigating breast MRI for screening (1 study investigating addition of MRI in women with > 50% breast density and no abnormalities detected at mammography). Published 2000-2001	Asymptomatic women MRI if mammography negative vs mammography alone MRI vs mammography MRI vs US MRI vs CBE	MRI appears to be more sensitive but less specific than mammography in selected high-risk women. There is insufficient evidence to support the use of breast MRI in population screening. Further evaluation required to assess its potential role in high-risk women.	High quality

Author & Year	Objective of report	Number & publication dates of included studies	Population considered in included studies Test comparison	Conclusions/recommendation	Quality Assessment
National Breast Cancer Collaboration 2006	To evaluate breast MRI as an additional or replacement test to other modalities in breast cancer screening programs for high-risk women Included a review of breast MRI in subgroups defined by: • age • risk factors	6 systematic reviews identified 4 primary studies investigating breast MRI for screening high-risk women included for analysis Primary studies published 2004- 2005	High-risk women MRI ± mammography versus mammography	MRI has a higher sensitivity and slightly lower specificity compared to mammography. No studies have evaluated whether MRI results in improved survival. MRI may potentially have a particular role in younger women due to the poor performance of mammography in women < 50 years. However, evidence to support targeting younger women versus all high-risk women was not identified.	High quality
Results of meta-analysis (4 studies): MRI versus mammography: MRI sensitivity 77%, specificity 87% versus mammography sensitivity 40%; specificity 94%. Similar results for all women and subgroup of women with BRCA1 or BRCA2 mutations MRI + mammogrpahy versus mammography: • if either test positive: MRI + mammography sensitivity 93%, specificity 76% versus mammography sensitivity 40%, specificity 94% • if both tests positive: MRI + mammography sensitivity 23%, specificity 98% (sensitivity 56%, specificity 90% for subgroup of women with BRCA1 or BRCA2 mutations)					

Appendix F. Characteristics and quality appraisal of included studies

Author & Year	Setting	*N*	Study objective Study design Index test, comparator and reference standard	Study Population	Outcomes	Study quality and applicability
Hartman et al 2004	USA Single centre 2001-2003	41	Study design Prospective cohort study Index test Annual MRI with contrast, performed in second week of the menstrual cycle for premenopausal women 1.5-Tesla, Breast coil Comparator test • Annual mammography (2 views) • Annual ductal lavage Biannual CBE also performed Reference standard Positive test: Histopathology for all test positives 6-month follow-up for indeterminate findings Negative test: NR Test threshold for positive result Not prespecified, varied for individual patients according to imaging findings and patient history. In general: Positive test: dominant lesion ≥ 5 mm with suspicious features were biopsied Indeterminate test: focally enhancing lesions ≥ 5 mm were reviewed at 6 month follow-up Average follow-up: NR	Inclusion criteria • Asymptomatic women • high risk • documented BRCA1/BRCA2 mutation Included women with prior diagnosis of breast cancer Criteria for high risk: > 10% risk of breast cancer at 10 years using Claus model Patient characteristics Median age: 42.5 years (range 27-72 years) BRCA1/BRCA2 mutation: 24/41 (58.5%) High risk: 17/41 (41.5%) BRCA1: 19/41 (46.3%) BRCA2: 5/41 (12.2%) Prior diagnosis of breast cancer: 12/41 (29.3%) Prior diagnosis of ovarian cancer: 3/41 (7.3%)	Biopsy rate Rate of cancer detection (Breast MRI sensitivity and specificity calculated from data reported in text, mammography data not fully reported.)	NHMRC Level III-1 Quality: LOW Prospective study Consecutive recruitment NR Reference standard: suboptimal: Histology for all test positives: yes Film review for interval cancers: NR Test interval: suboptimal (within a 2-month period (initially) or within a 2-week period). Test interpretation: blinding to results of comparator test NR Test threshold: suboptimal Varied for individual patient's history and imaging findings Exclusions: NR Test failures: NR Applicability: Fair Applicable population – yes: high-risk women, median age: 42.5 years Applicable intervention – yes: breast MRI with contrast, 1.5-Tesla, breast coil Applicable comparison – suboptimal: relative accuracy of MRI vs mammography reported MRI + mammography vs mammography NR Comparison of mammography/MRI + ultrasound NR

Kriege et al 2004, 2006 Magnetic Resonance Imaging Screening Study Group (MRISC)	Netherlands 6 sites 1999-2003	1,909	Study design • Prospective test accuracy study • Substudy of cases detected: cohort study with historical control to assess stage shift in detection of breast cancer • Subgroup analysis: comparison of test accuracy and stage of cancers detected between 1st and subsequent screening rounds (Kriege et al 2006) Index test Annual MRI with contrast Magnet strength, breast coil NR Comparator test • Annual mammogram (2 views) • Biannual CBE Reference standard Positive test Histopathology for all test positives Indeterminate finding for MRI or mammography verified by US ± biopsy or repeated test Negative test Based on combination of all tests. Negative finding for MRI/mammography, positive for CBE verified by additional investigation. Interval cancers classified as false negatives Film review of interval cancers to verify false negative classification NR Test threshold for positive result BI-RADS category 0, 3-5 Average follow-up: median 2.9 years (range 0.1-3.9 years)	Inclusion criteria • Asymptomatic women • High risk • Age 25-70 years • No prior diagnosis of breast cancer Criteria for high risk • ≥ 15% cumulative lifetime risk of breast cancer using Claus model Patient characteristics Mean age: 40 years (range 19-72 years) Risk classification: BRCA1/2 or other mutations (50-85% risk) 358/1,909 (19%) High risk (30-49% risk) 1,052/1,909 (55%) Moderate-risk (15-29% risk): 499/1,909 (26%) Pre-menopausal: 1,365/1,909 (75%) Prior diagnosis of breast cancer: 0/1,909 (0%) Age-matched unscreened controls Group 1: Average-risk nonscreening: all women who had breast cancers diagnosed in 1998 in the Netherlands Group 2: High-risk nonscreening: unselected women with ≥ 15% cumulative lifetime risk of breast cancer, primary diagnosis of breast cancer, from 2 cities, 1996-2002	Test accuracy Sensitivity Specificity AUC Biopsy rate Rate of cancer detection Mortality Cancer characteristics Subgroup analysis with age-matched unscreened controls • size: ≤ 10 mm • lymph node spread/micrometastases • histological grade 1-3 Subgroup analyses Test sensitivity and specificity at: • 1st versus subsequent screening round • 1st screening round with vs without prior history of mammography	NHMRC Level III-1 Quality: FAIR Prospective accuracy study Consecutive recruitment NR Reference standard suboptimal: Histopathology for all test positives: yes Film review for internal cancers: NR Test interval: adequate (same day or same time period, days 5-15 of menstrual cycle) Test interpretation: blinded Tests read without results of CBE Test threshold: stated and consistent between tests Exclusions reported: 43/1,952 • 35/1,952 excluded as not high-risk • 8/1,952 withdrawals prior to screening Other withdrawals: 177/1,909 • 89/1,909 (5%) refused MRI • 65/1,909 (3%) prophylactic mastectomy • 23/1,909 (1%) lost to surveillance/follow-up 5/50 cancers excluded from the analysis due to MRI/mammography not performed (3), cancer detected at CBE (1) or off-site mammogram (1) Test failures reported Applicability: Fair Applicable population – yes: high-risk women, mean age 40 years Applicability of intervention: suboptimal: MRI with contrast, but breast coil, magnet strength NR Applicable comparison – suboptimal: relative accuracy of MRI vs mammography reported MRI + mammography vs mammography NR Comparison of mammography/MRI + ultrasound NR

Kuhl et al 2005	Germany Single centre 1996-2001	529	Study design Prospective test accuracy study Index test 1. Annual MRI with contrast 2. Combination of annual MRI plus mammography 1.5-Tesla magnet, breast coil Comparator test • Annual mammography (2 views) • Biannual ultrasound Biannual CBE also performed Mammography accuracy reported separately and in combination with ultrasound Reference standard Positive test: Histopathology for all test positives Indeterminate findings followed-up by 6-month follow-up Negative test 6-month follow-up with CBE and US Interval cancers classified as false negatives (identified by histo-pathology for 6 women undergoing prophylactic mastectomy) No film review for interval cancers Test threshold for positive test: BI-RADS 4-5 Same threshold for MRI and mammography Average follow-up time Mean follow up 5.3 years (range 2-7 years)	Inclusion criteria: • Asymptomatic • High-risk • Starting at age 30, or 5 years prior to the youngest family member affected by breast cancer Excluded if prior history of breast cancer with bilateral mastectomy, chemotherapy in the last 12 months or metastatic disease. Criteria for high risk: ≥ 20% cumulative lifetime risk of breast cancer using Claus model Patient characteristics: Median age: 40 years (range 27-59 years) Sex: 1 male with prior history of breast cancer (0.2%) Risk classification: Lifetime risk of 20%: 110/529 (21%) Lifetime risk of 21%-40%: 241/529 (46%) BRCA1/BRCA2 mutation: 43/529 (8%) Prior diagnosis of breast cancer: 139/529 (26%) Premenopausal: NR	Test accuracy Sensitivity Specificity Cancer detection rate Cancer characteristics • Stage • Lymph node status • Histological grade (In women without prior history of breast cancer) Subgroup analyses • All women • Lifetime risk of 20% • Lifetime risk of 21%-40% • BRCA mutation carriers • No history of breast cancer • Prior diagnosis of breast cancer	NHMRC Level III-1 Quality: LOW Prospective accuracy study Consecutive recruitment NR Reference standard: suboptimal Histopathology for all test positives: yes No film review for internal cancers Test Interval: Suboptimal (within 8 week period) Test interpretation: blinded to comparator test results Tests read with patient history and CBE Test threshold stated and same for MRI and mammography Exclusions reported : 61/590 (12 clinical abnormality on first visit, 49 lost to follow up after first round surveillance) Missing data: 249 surveillance rounds collected in 86/529 participants during the first 2 years were incomplete (no mammogram). These data sets were not included in the analysis Test failures NR Applicability: high Applicable population – yes: high-risk women, median 40 years Applicable intervention – yes: Annual MRI with contrast, 1.5 Tesla. Breast coil Applicable comparison – high: relative accuracy of MRI vs mammography and MRI + mammography vs mammography reported Comparison of mammography + ultrasound reported

Leach MO et al 2005 Warren et al 2002 Magnetic Resonance Imaging Breast Screening (MARIBS)	UK 22 sites 1997-2004	649	Study design Prospective test accuracy study Index test Annual MRI with contrast, double reporting 1.0-1.5-Tesla magnet Breast coil Comparator test: Annual mammogram (1 or 2 views), double reporting Reference standard Positive test: Histopathology undertaken based on combination of all tests. Negative test 1 year follow-up Interval cancers classified as false negatives. No film review for interval cancers Test threshold for positive test: BI-RADS 0, 3-5 Same threshold for MRI and mammography Average follow-up ≥ 2 years (range: 2-7 years)	Inclusion criteria • Asymptomatic women • High-risk • Age 35-49 years Excluded if prior history of breast cancer or if tested negative to genetic test during study Criteria for high risk • Known carriers of BRCA1/2, or TP53 mutation • First degree relative with a BRCA1/2 or TP53 mutation • Strong family history of breast and/or ovarian cancer • Family history consistent with classic Li-Fraumeni syndrome Patient characteristics Median age 40 years (range 31-55 years) Risk classification: BRCA1 mutation: 82/649 (13%) BRCA2 mutation: 38/649 (6%) Family history BRCA1/2: 105/649 (16%) Li-Fraumeni 5/649 (0.8%) Other family history 419/649 (65%), 30 subsequently excluded due to negative genetic tests Prior history of breast cancer: 0/649 (0%)	Test accuracy Sensitivity Specificity AUC Recall rate Biopsy rate Subgroup analyses • Test accuracy for BRCA1 versus BRCA2 carriers • Test accuracy at 1st versus subsequent screening round Cancer characteristics • Lymph node status • Grade • Size Psychological impact study and cost-effectiveness analysis underway	NHMRC Level III-2 Quality: FAIR Prospective accuracy study Consecutive recruitment NR Reference standard: suboptimal, biopsy histopathology based on interpretation of all results, interval cancers classified as false negatives No film review for internal cancers Test Interval: NR but 'preferably same day' Test interpretation: blinded CBE not performed. Test threshold stated and same for MRI and mammography. Exclusions reported 189/838 • Not tested due to logistic problems 106/838 (13%) • Did not received both tests 83/838 (10%) 134 other withdrawals reported including: • 30/649 (5%)screened but excluded from analysis when non-carrier status confirmed • 35/649 (5%) breast cancer • 28/649 (4%) prophylactic mastectomy • 19/649 (3%) personal reasons/stress • 12/649 (2%) refused MRI due to claustrophobia or discomfort • lost to follow-up: 12/649 (2%) • MRI contraindicated 3/649 (0.5%) • MRI size restriction 2/649 (0.3%) Test failures reported Applicability: yes, suboptimal Applicable population – yes: high-risk women, 99% ≤ 50 years Applicable intervention – suboptimal: MRI with contrast, 1.0-1.5 Tesla, dedicated breast coil Applicable comparison – yes: relative accuracy of MRI vs mammography and MRI +

						mammography vs mammography reported Comparison of mammography/MRI + ultrasound NR
Lehman CD et al 2005a, 2005b International Breast MRI Consortium (IBMC)	USA, Canada 13 sites 1999-2002 Substudy: USA, Germany 10 sites, 2002	367 Substudy: 103	Study design Prospective test accuracy study Index test MRI with contrast Magnet strength, breast coil NR Comparator test Mammogram (number of views NR) CBE also performed Reference standard: Positive test Histopathology not undertaken for all test positives. 9 women with discordant MRI/mammography results did not have FNAC or biopsy Negative test NR Film review for interval cancers: NR Film review for discordant results Test threshold for positive result BI-RADS 4,5 Average follow-up NR	Inclusion criteria • Asymptomatic • High-risk • Age ≥ 25 years *or:* • Prior history of breast cancer Criteria for high risk • > 25% based on family history or genetic testing (Claus, Gail, Couch or Berry model) • Prior history of breast cancer if - ≤ 5 years ago; or - > 5 years with > 50% risk of breast cancer; or - BRCA1, BRCA2 mutation carrier Substudy (overlap with larger study NR): • Age ≥ 25 years • Prior history of breast cancer within 6 months Excluded if prior biopsy in the study breast within 6 months or receiving chemotherapy or hormonal therapy Patient characteristics Mean age 45 years, standard deviation ±9.7 Substudy: Mean age 52 years, standard deviation ±9.9 Risk classification: BRCA1/BRCA2 mutation: NR Pre-menopausal: 180/367 (49%) Proportion high risk: 329/367 (90%) Prior history of breast cancer: 38/367 (10%)	Test accuracy Sensitivity Specificity Recall rates Diagnostic yield Cancer characteristics (4 cases) • Histology • Size • stage	NHMRC Level III-1 Quality: LOW Prospective accuracy study Consecutive recruitment NR Reference standard: suboptimal Histopathology for all test positives: No Film review for internal cancers : NR Test Interval: suboptimal (within 90 days) Test interpretation: blinded Test threshold stated and same for MRI & mammography: BI-RAD ≥ 4 Exclusions reported 23/390 (5.9%) • 2 withdrawals prior to completion of screening • 14 excluded as mammogram not performed or not available • 7 excluded as breast MRI not performed or not available Applicability: yes, suboptimal Applicable population – yes: High-risk screening population 90%, mean age: 45 years Applicable intervention – suboptimal: MRI with contrast, magnet strength and use of breast coil NR. Applicable comparison – yes: relative accuracy of MRI vs mammography and MRI + mammography vs mammography NR mammography/MRI + ultrasound NR Substudy of patients with a recent diagnosis of breast cancer, test performance assessed in the contralateral breast

				(Substudy participants all had prior history of breast cancer)		
Podo et al 2002	Italian Multicentre 2000-2002	105	Study design Prospective test accuracy study Index test Annual MRI with contrast Performed within the second week of the menstrual cycle ≥ 1.0 Tesla, breast coil Comparator test Annual mammography (2 views) Annual ultrasound Reference standard Positive test FNAC/biopsy and histopathology for all mammography positives Repeat MRI at 1-2 months for MRI positives and discordant with mammography. Negative test Clinical examination and follow-up. Classification & film review of interval cancers NR Test threshold for positive result Mammography: BI-RADS NR MRI: score ≥ 3 (scale scored 0-8) but discordant findings repeated in 1-2 months Average follow-up: 2 year follow-up planned 14/105 (13%) attended follow-up at 1 year at time of publication	Inclusion criteria • high risk • age: women ≥ 25 years; men ≥ 50 years Criteria for high risk • Known or suspected carrier of BRCA1/2 mutation • Women with personal history of breast cancer with contralateral breast intact Excluded if pregnancy, breast-feeding, current chemotherapy, terminal illness or specific contraindications to MR examinations. Patient characteristics Median age: 51 years (range 25-77 years) Risk classification: BRCA1/BRCA2 mutation: NR Prior history of breast cancer: 40/105 (38%) Prevalence of cancer at first screening round: 7/105 (7%)	Cancer detection rate TPR FPR	NHMRC Level III-2 Quality: LOW Prospective accuracy study Consecutive recruitment NR Reference standard: Invalid Histopathology not performed for all test positives No film review for internal cancers Test interval NR Test interpretation: NR Test threshold: unclear Exclusions: NR Test failures: NR Applicability: LIMITED Applicable population – suboptimal: high-risk women, median age: 51 years Applicable intervention – suboptimal: contrast-enhanced MRI, ≥ 1.0 Tesla, breast coil Applicable comparison – suboptimal: relative accuracy of MRI vs mammography reported Comparison of MRI vs mammography + ultrasound reported MRI + mammography vs mammography not reported

Stoutjesdijk et al 2001	Netherlands Single centre 1994-2001	75	Study design Retrospective cohort study Index test Annual MRI with contrast, Performed in the second week of the menstrual cycle for premenopausal women 1.5-Tesla magnet Breast coil Comparator test Annual mammogram (oblique and craniocaudal views) Reference standard Positive test Histopathology for all test positives Indeterminate finding verified by repeated test 3-6 months later Negative test Verified by next 2 yrs follow-up Film review for interval cancers Test threshold for positive result BI-RADS 4 -5 further investigation if BI-RADS-3 Average follow-up time: 2 years	Inclusion criteria • Asymptomatic women • High-risk • No prior diagnosis of breast cancer Criteria for high risk • ≥ 15% cumulative lifetime risk of breast cancer owing to familial/ genetic predisposition Patient characteristics Mean/Median: NR (range 21-71years) 87% ≤ 50 years. Risk classification: 1 (50-85% risk): 15/75 (20%) 2 (30-50% risk): 26/75 (35%) 2 or 3 (15-50% risk): 15/75 (20%) 3 (15-30% risk): 19/75 (25%) BRCA1/BRCA2 mutation: NR Prior history of breast cancer: 0/75 (0%)	Sensitivity Specificity AUC Rate of cancer detection Sub-group analyses: Received mammography and MRI within a 4-month period	NHMRC Level III-2 Quality: LOW Retrospective study Distorted selection: second test performed as diagnostic test for 10/75 women. Nonconsecutive recruitment of women opting to participate in surveillance Reference standard: adequate Histopathology for all test positives: yes Film review for internal cancers: yes Test interval: suboptimal (within a 4-month period) Test interpretation: blinded Test threshold: stated and same for MRI and mammography. Exclusions reported: • 66/245 no histology or adequate follow-up information • 104/245 only had MRI or mammogram, or both tests > 4 months apart. Test failures: NR Applicability: yes Applicable population – yes: high-risk women, 87% ≤ 50 years Applicable intervention – yes: MRI with contrast, 1.5 T system, standard bilateral dedicated breast coil Applicable comparison – yes: relative accuracy of MRI vs mammography and MR I+ mammography vs mammography reported Ultrasound was performed in some women Comparison of mammography/MRI + ultrasound NR

Tilanus-Linthorst et al 2000 (Substudy of Tilanus-Linthorst 2000b)	Netherlands Single centre 1995-1998	109	Study design Prospective accuracy study Index test MRI with contrast 1.5-Tesla magnet Breast coil Comparator test Annual mammogram (2 views) Performed with biannual CBE Reference standard Positive test MRI guided ultrasound ± FNAC. Histopathology if positive FNAC. Negative test MRI or FNAC followed with routine surveillance Interval cancers classified as false negatives No film review for interval cancers Test threshold for positive result Not specified other than 'Focal enhancement'	Inclusion criteria • Asymptomatic • High-risk • >50% breast density at mammography • No abnormalities 6 months prior at clinical examination and mammography Criteria for high risk: • ≥ 25% cumulative risk of breast cancer owing to familial predisposition (using Claus model) Patient characteristics: Mean age: 41.5 years (range 22-68 years) Risk classification: BRCA1/BRCA2 mutation: 12/109 (11%) Prior history of breast cancer: 0/109 (0%) (assumed from study eligibility criteria)	TPR FPR Recall rate FNAC rate Biopsy rate Costs Cancer characteristics • Stage	NHMRC Level III-2 Quality: LOW Consecutive recruitment NR Reference standard: invalid – histopathology not performed for all test positives No film review for internal cancers Test Interval: NR Test interpretation: Blinding NR Test threshold: NR Exclusions: NR Test failure: reported Applicability: limited Applicable population – no: high-risk women, mean age 41.5 years with breast density > 50% and normal mammogram 6 months ago Applicable intervention – yes: MRI with contrast, 1.5 Tesla, double breast surface coil Applicable comparison – no: MRI vs mammography NR MRI + mammography vs mammography NR mammography/MRI + ultrasound NR
Trecate et al 2003	Italy Single centre Recruitment over 7 month period, year NR	20 (23 women enrolled, 3 patients with clinical signs excluded from this review)	Study design Retrospective/prospective: NR Index test MRI with contrast 1.5 Tesla, breast coil Comparator test Mammogram (2 views) Ultrasound Reference standard Positive test Histopathology not performed for all test positives: performed based on combination of all tests and clinical findings Negative test	Inclusion criteria • BRCA1/2 genes mutational carrier • high-risk *or:* • prior history of breast/ovarian cancer carrying mutational BRCA1/2 genes ‡ 3/23 patients clinically symptomatic Criteria for high risk • > 50% of carrying a susceptibility gene for breast cancer based on family history Patient characteristics Mean/Median age: NR (range 30-61 years) Risk classification:	Rate of cancer detection TPR FPR FNR	NHMRC Level III-2 Quality: LOW Retrospective/prospective NR Consecutive recruitment NR Reference standard: invalid – histopathology not performed for all test positives Test interval NR Test interpretation: blinding NR Test threshold: NR Exclusion reported: NR Test failure: NR Applicability: Limited Poorly reported Applicable population – suboptimal: high-risk women, ?% ≤ 50 years

			Clinical follow-up, but results NR Film review for interval cancers NR Test threshold for positive result NR	BRCA1/BRCA2 mutation: NR Prior history of breast cancer: 5/20 (25%)		Applicable intervention – yes: MRI with contrast 1.5 Tesla and dedicated breast coil Applicable comparison – suboptimal: relative accuracy of MRI vs mammography reported MRI + mammography vs mammography NR Comparison of mammography + ultrasound NR
Warner et al 2001, 2004	Canada 1 site 1997-2003	236	Study design prospective test accuracy study Index test Annual MRI with contrast 1.5 Tesla, breast coil Performed during the 2nd week of the menstrual cycle (for premeno-pausal women), included additional diagnostic scan for potentially suspicious areas Comparator test Annual mammogram (4 view) Annual ultrasound Biannual CBE CBE 6 monthly Reference standard Positive test Histopathology not performed for all positive tests. If MRI finding positive and discordant with other tests, MRI repeated in 1 month. Negative test Based on combination of all tests Indeterminate test: BI-RADS category 3 (probably benign) followed-up with repeat examinations in 6 months, 1 year and 2 years. Film review for interval cancers	Inclusion criteria • Asymptomatic? • High-risk • Age 25-65 years *or:* • Prior history breast cancer Excluded if history of bilateral breast cancer and currently undergoing chemotherapy, or having metastatic disease, or weighed > 91 kg Criteria for high risk • BRCA1/2 mutation carriers Patient characteristics Mean age: 47 years (range 26-65 years) Risk classification: BRCA1 mutation 137/236 (58%) BRCA2 mutation 99/236 (42%) Pre-menopausal: 106/236 (45%) Prior history of breast cancer: 70/236 (30%)	Accuracy Sensitivity Specificity AUC Cancer characteristics • Lymph node status • Size	NHMRC Level III-2 Quality: FAIR Prospective accuracy study Consecutive recruitment: NR Reference standard: invalid – histopathology not performed for all test positives Test Interval: adequate (same day) Test interpretation: blinded Tests read without results of CBE Test threshold: BI-RAD ≥ 4 but confirmation differed for MRI and mammography if discordant results Exclusions: 100% participation in first screening round Study withdrawals before round 3: 31/236 (13%), reasons reported Test failures: partly reported Applicability: High Applicable population – yes: High-risk screening population, mean age 47 years (includes women with history of breast cancer) Applicable intervention – yes: MRI with contrast, 1.5 Tesla, dedicated breast coil Applicable comparison – high: relative accuracy of MRI vs mammography reported Combination breast MRI, mammogram, US, CBE vs MRI or mammography or US reported

			Test threshold for positive result BI-RADS 4 or 5 Average follow-up Median 1year (range 1-3 years)			
Patient preference, tolerance and therapeutic impact studies						
Kurian et al 2005 (Subset of Hartman 2004)	USA Single centre 2001-2003	36 (43 invited)	Study objective To evaluate tolerance of intensive breast screening versus prophylactic mammography in high-risk women. Study design cross-sectional survey Intensive screening Annual MRI with contrast, 1.5-Tesla, 4-coil phased-array breast coil Annual mammogram Annual Ductal lavage(DL) 6 monthly CBE Questionnaire: 8 items to evaluate test discomfort, comparison of MRI versus mammogram and versus ductal lavage, use of sedatives before testing, change in attitude toward prophylactic mastectomy. Self-administered survey at median: 13 months on trial (range 1-29 months)	Inclusion criteria • Asymptomatic women • High-risk Women with prior diagnosis of breast or ovarian cancer included if completed adjuvant therapy > 1 year ago. Criteria for high risk • ≥ 10% cumulative risk of breast cancer at 10 years • BRCA1 or BRCA2 mutation Patient characteristics Median age 41 years 36/46 (78%) < 50 years. 20/46 (44%): premenopausal Prior diagnosis of breast cancer: 12/46 (26%) Prior diagnosis of ovarian cancer: 3/46 (7%) Response rate: 36/43 (84%)	Procedure discomfort rating Preference MRI versus mammogram Change in attitudes towards PM	Quality: FAIR Based on a representative sample: women at high inherited risk of breast cancer Explicit inclusion criteria Survey, no pretest/post-test High response rate of questionnaire 36/43 (84%) Validation of questionnaire not reported Exclusions reported 3/46 (7%) lost to follow-up Small sample size limits ability to assess statistically significant differences in responses Applicability: yes Applicable population – yes: high-risk women, 78% ≤ 50 years Applicable intervention – yes: Annual MRI with contrast, 1.5-Tesla, 4-coil phased-array breast coil Applicable comparison – yes: MRI vs mammography reported

Rijnsburger et al 2004 (substudy of Kriege et al 2004)	Netherlands 6 centres 2000-2002	334 (519 invited)	Objective To evaluate the short-term effects of breast cancer screening in high-risk women on generic health-related of life and distress Study design Pretest-posttest questionnaires Screening Annual MRI with contrast, Magnet strength, breast coil NR Annual mammogram 6-monthly CBE Questionnaire Completed at 2 time points T0 2 months prior to screening T1 immediately prior to screening T2 1-4 weeks after screening when all test results available Items included: • Medical Outcomes Study 36-item short form (SF-36) to measure health status • EQ-5D preference-based measure of health-related quality of life • Somatic subscale (SOM) of Symptom Checklist-90 • Other screen-specific items Women who did not return the questionnaire within 4 wks were sent a reminder	Inclusion criteria • Asymptomatic women • High risk • Age 25-70 years • No prior diagnosis of breast cancer Patient characteristics Mean age: 41 years Risk classification: 1 BRCA1, 2 mutation carriers (50-85% risk): 11% 2 high risk (30-50%): 56% 3 moderate risk (15-30%): 33% Prior diagnosis of other cancer: 4%	Health-related quality of life Distress Pain Discomfort Anxiety Measured at 3 different time points Subgroups: By screening modality • CBE alone • CBE, mammography, MRI • Additional investigations required or not	Quality: HIGH Based on a representative sample: women at high inherited risk of breast cancer Explicit inclusion criteria Response rate adequate: 334/492 (68%) of those contacted consented, 329/492 (67%) completed first questionnaire, 288/492 (59%) completed 3 questionnaires. No significant difference in age or breast cancer risk classification between participants and non-participants. Validated standard tool used for assessing quality of life. Applicability: yes, suboptimal Applicable population – yes: high-risk women, Mean age: 41 years Applicable intervention – suboptimal: Annual MRI with contrast, magnet strength and use of breast coil NR Applicable comparison – yes: MRI vs mammography reported

AUC=area under the ROC curve, CBE=clinical breast examination, US=ultrasound

Appendix G. Results of included primary studies

Author & Year	N (*no. of MRI performed*)	Index test comparator	Cancer detection, prevalence, incidence	MRI outcomes	Comparator outcomes	Other
Results of studies of diagnostic accuracy						
Hartman et al 2004	41 (56) Tests performed on 41 women (56 MRI). 15/16 women with initial negative results returned for a second annual screen at time of reporting)	MRI versus mammography	Cancers detected: 1 DCIS Prevalence of breast cancer at initial screening round: 1/41 (2.4%)	MRI TP 1 FP 10 FN 0 TN 30 Sensitivity 100% (95% CI 25-100%) Specificity 40% (95% CI 25-57%) If MRI required follow-up imaging and subsequently shown to be benign, classified as false positive. Specificity 75% (59-87%) if MRI requiring biopsy and subsequently shown to be benign, classified as false positive. Complete data for 2nd screening round not available Recall rate: 25/41 (61% [95% CI 45-76%]) Biopsy rate: 11/41 (27% [95% CI: 13-40%]) Rate of cancer detection: 1/41 (2.4% [95% CI 0.06-12.9%])	Mammography Rate of cancer detection: 0/41 (0%)	Cancer characteristics NR Test failures MRI: Reasons for non-completion of MRI NR Mammography: NR Exclusions NR
Kriege et al 2006; Kriege et al 2004	1909 *(4,169)*	MRI versus mammography	Cancers detected: 44 invasive cancers (39 evaluable) 6 DCIS 4 interval cancers Overall cancer detection rate per 1,000 women years: All cancers 9.5 (95% CI: 7.1-12.3%) Invasive cancers 8.4	MRI TP 32 FP 420 FN 13 TN 3,704 Sensitivity 71% (95% CI 56-84%) Specificity 90% (95% CI 89-91%) AUC 0.83 Difference AUC MRI vs mammography 0.14 (95% 0.02-0.26, $p < 0.05$) Subgroup analysis: • 1st screen overall - sensitivity 18/22 = 82%	Mammography TP 18 FP 207 FN 27 TN 3,917 Sensitivity 40% (95% CI 26-56%) Specificity 95% (95% CI 94-96%) AUC 0.69 Subgroup analysis: • 1st screen overall - sensitivity 8/22 = % • 1st screen without prior mammography:	Cancer characteristics Invasive cancer < 10 mm High-risk screening: 19/44 (43%) Control group 1 average risk non-screening: 193/1,380 (14%), $p < 0.001$ Control group 2 high-risk non-screening: 5/39 (13%), $p = 0.04$ Prevalent cancer: 9/22 (41%) Incident cancer: 9/17 (53%) Prevalent vs incident $p = 0.52$ Lymph node negative

			(95% CI: 6.1-11.3%) Subgroup analysis for invasive cancers: 1st round: prevalent cancers: 22/1,723 (12.8 per 1,000) ≥ 2nd round: incident cancers: 17/2,431 (7.0 per 1,000) Prevalence of breast cancer at initial screening round: 26/1,723 (1.5%)	• 1st screen without prior mammography: - sensitivity 4/7 = 57% - specificity 93% - false positive rate 7.1% • 1st screen with prior mammography: - sensitivity 14/15 = 93% MRI vs M sensitivity: $p = 0.003$ specificity 86% false positive rate 14% MRI vs M false positive rate: $p < 0.001$ • Subsequent rounds: - sensitivity 13/17 = 76% - MRI vs M sensitivity: $p = 0.02$ - specificity 92% - false positive rate 8.2% - MRI vs M false positive rate: $p < 0.001$ Significant trend toward a decrease in MRI sensitivity between 1st and subsequent rounds, $p = .009$ Detection rate for invasive cancer: 1st round in women without prior mammography: 13.2 per 1,000 tests 1st round in women with prior mammography: 9.9 per 1,000 tests Subsequent round: 5.3 per 1,000 tests $p = 0.05$ Recall rate: NR by test False positives proceeding to biopsy: 24/4,169 tests (0.6%)	- sensitivity 5/7 = 71% - specificity 94 % - false positive rate 6% • 1st screen with prior mammography: - sensitivity 3/15 = 20% - specificity 95% - false positive rate 5.5% MRI vs M specificity: $p < 0.001$ • Subsequent rounds: - sensitivity 5/17 =29 % - specificity 95% MRI vs M specificity: $p < 0.001$ false positive rate 4.6% Borderline significant trend toward a decrease in mammography sensitivity between 1st and subsequent rounds, $p = 0.05$ Detection rate for invasive cancer: 1st round in women without prior mammography = 16.5 per 1,000 tests 1st round in women with prior mammography = 2.1 per 1,000 tests subsequent round: 2.1 per 1,000 tests, $p = 0.003$ Recall rate: NR by test False positives proceeding to biopsy =7/4,169 tests (0.2%)	High-risk screening: 28/42 (67%) Control group 1: 657/1,380 (48%) $p < 0.001$ Control group 2:17/39 (44%) $p = 0.001$ Prevalent cancer: 19/22 (86%) Incident cancer: 11/16 (69%) Prevalent vs incident $p = 0.24$ Histological grade I High-risk screening: 19/43 (44%) Control group 1: 99/900 (11%) $p < 0.001$ Control group 2: 4/37 (11%) $p = 0.01$ [note BRCA1/2 screening grade 3: 12/19 (63%)] Prevalent cancer :5/22 (23%) Incident cancer: 8/17 (47%) Prevalent vs incident p=0.10 Mortality High-risk screening: no breast cancer deaths with median 1.5 year follow-up Test failures MRI: 89/1,909 (4.7%) refused MRI due to claustrophobia or other Mammography: NR Exclusions Initial exclusions: 43/1,952 (2%) Withdrawal/Incomplete follow-up: Prophylactic mastectomy: 65/1,909 (3.4%) Lost to surveillance: 23/1909 (1.2%)

Kuhl et al 2005	529 *(1,452)*	MRI + mammography versus mammography MRI + mammography versus mammography + US MRI versus mammography	Cancers detected: 44 cancers 43 screen-detected cancers (in 41 patients) including 9 DCIS 1 Interval cancer (But 2 screen-detected cases found at 6 month US) Prevalence of breast cancer at initial screening round: 14/529 (2.6%) 26.5/1,000 women Mean incidence of breast cancer at years 2-6: 29/1,172	MRI + mammography TP 40 FP 55 FN 3 TN 1,354 Sensitivity 93% (95% CI 81-99%) Lifetime risk of 21%-40%: 100% Mutation carriers: 100% Specificity 96% (95% CI 95-97%) Lifetime risk of 21%-40%: 96% Mutation carriers: 94% MRI TP 39 FP 39 FN 4 TN 1,370 Sensitivity 91% (78-97%) Lifetime risk of 21%-40%: 100% Mutation carriers: 100% Specificity 97% (96-98%) Lifetime risk of 21%-40%: 98% Mutation carriers: 98% MRI vs M sensitivity: $p = 0.001$ MRI vs M specificity: $p > 0.05$	Mammography TP 14 FP 45 FN 29 TN 1,364 Sensitivity 33% (95% CI 19-49%) Lifetime risk of 21%-40%: 25% Mutation carriers: 25% Specificity 97% (95% CI 96-98%) Lifetime risk of 21%-40%: 97% Mutation carriers: 97% Mammography + US TP 21 FP 155 FN 22 TN 1,254 Sensitivity 49% (95% CI 33-65%) Lifetime risk of 21%-40%: 45% Mutation carriers: 38% Specificity 89%(95% CI 87-91%) Lifetime risk of 21%-40%: 90% Mutation carriers: 89%	Cancer characteristics: Invasive cancer <10mm Overall: 25/43 (58%) MRI: 23/39 (59%) Mammography: 5/14 (36%) Mammography + US: 6/21 (29%) $p = 0.21$ Lymph node negative MRI: 26/31 (84%) Mammography: 6/10 (60%) Mammography + US: 11/16(69%) $p = 0.19$ Histological grade NR by imaging modality Test failures NR Exclusions 49/578 (8%) lost to follow-up
Leach MO et al 2005	649 *(1,881)*	MRI + mammography versus mammography MRI versus mammography	Cancers detected: 35 cancers: 33 screen-detected cancers including 6 DCIS 2 Interval cancers Prevalence of breast cancer at initial screening round: 20 /649 (3%) Incidence of breast cancer at subsequent screening rounds: 15/629 Cancer detection rate per 1,000 women years At initial screen: 27	MRI + mammography TP 33 FP 428 FN 2 TN 1,418 Sensitivity 94% (81-99%) Specificity 77% (75-79%) MRI TP 27 FP 344 FN 8 TN 1,502 Sensitivity 77% (60-90%) Specificity 81% (80-83%) AUC 0.85 (0.84-0.87) Test for difference MRI vs mammography accuracy Sensitivity: $p = 0.01$ Specificity: $p < 0.0001$ AUC: $p = 0.035$ Recall rate *MRI ± M* 12.7% per women year	Mammography TP 14 FP 121 FN 21 TN 1,725 Sensitivity 40% (24-58%) Specificity 93% (92-95%) AUC 0.70 (0.68-0.72) Recall rate 3.9% per woman year Number of further tests for benign or normal findings by test generating initial recall (1-5 tests ordered per patient recalled) 119/1,881 (6%) Biopsy rate 29/1,881 (1.5%)	Subgroup analyses BRCA1 or 1st degree relative with BRCA1 MRI sensitivity 92% (64-100%), Mammography sensitivity 23% (5-54%) MRI versus mammography, $p = 0.004$ MRI specificity 79% (75-83%) Mammography specificity 92% (88-94%), MRI versus mammography, $p < 0.0001$ BRCA2 or 1st degree relative with BRCA2 MRI sensitivity 58% (28-84%), Mammography sensitivity 50% (21-79%) MRI versus mammography, $p = 1.0$ MRI specificity 82% (77-87%) Mammography specificity 94% (91-97%) MRI versus mammography, $p=0.0001$ Cancer characteristics

			At subsequent screening: 13 Recall rate: 279/1,881 (14.8%) False positive recall rate: 246/1,881 (13%) 8.5 recalls per cancer detected Biopsy rate: 107/1,881 (6%) False positive biopsy rate: 96/1,881 (5%) 0.21 benign surgical biopsies per cancer detected	Number of further tests for benign or normal findings by test generating initial recall (1-5 tests ordered per patient recalled) 376/1,881 (20%) *MRI* Number of further tests for benign or normal findings by test generating initial recall (1-5 tests ordered per patient recalled) 10.7% per woman year 325/1,881 (17%) Comparison MRI + M vs M : $p < 0.001$ Biopsy rate *MRI ± M* 88/1,881 (5%) *MRI* 80/1,881 (4%) Comparison MRI+M vs M:, $p < 0.001$		Invasive cancer < 10 mm 11/29 (38%) MRI + mammography: 10/28 (36%) MRI: 7/25 (28%) Mammography: 4/9 (44%), $p = 0.71$ lymph node negative invasive cancer 21/26 (81%) MRI + mammography: 19/24 (79%) MRI: 18/21 (86%) Mammography: 6/9 (67%), $p = 0.65$ histological grade I MRI+mammography: 3/28 (11%) MRI: 3/25 (12%) Mammography: 0/9 (0%) histological grade 3 MRI+mammography: 18/28 (64%) MRI: 17/25 (68%) Mammography: 7/9 (78%) Test failures *MRI:* • 12/649 (2%) refused MRI due to claustrophobia or discomfort • 3/649 (0.5%) MRI contraindicated • 2/649 (0.3%) MRI size restriction Mammogram failures: NR Exclusions 189/838 (22.6%) Other withdrawals:134/649 (21%) including: negative genetic test: 30/649 (5%) prophylactic mastectomy: 28/649 (4%) breast cancer: 35/649 (5%) lost to follow-up: 12/649 (2%) Other reasons for incomplete follow-up: stress or personal reasons; serious illness; death.

(Lehman et al 2005b; Lehman et al 2005a)	367 *(367)* substudy: 103 women with past history of breast cancer	MRI versus mammography	Cancers detected: 4 cancers: 4 screen-detected cancers including 1 DCIS 0 Interval cancers Prevalence of breast cancer at initial screening round: 4/367 (1%) Recall rate: 38/367 (10%) Biopsy rate: 27/367 (7%) Substudy Cancer prevalence: 4/103 (3.9%)	MRI TP 4 FP 27 FN 0 TN 336 Sensitivity 100% (40-100%) Specificity 93% (89-95%) Concordance of MRI/mammography 330/367 (90%) Additional diagnostic yield 0.8% (–0.3-2.0%) Recall rate MRI: 31/367 (8.4%) Biopsy rate 24/367 (6.5%) Substudy (women with past history of breast cancer) TP 4 FP 8 FN 0 TN 91 Sensitivity: 100% (40-100%) Specificity: 92% (85-96%) Biopsy rate 23/367 (6%)	Mammography TP 1 FP 7 FN 0 TN 336 Sensitivity 25% (0.6-81%) Specificity 98% (96-99%) Recall rate Mammography: 8/367 (2.2%) Biopsy rate 4/367(1.1%) Substudy (women with past history of breast cancer) TP 0 FP 3 FN 4 TN 96 Sensitivity: 0% (0-60%) Specificity: 97% (91-99%)	Invasive cancer < 10 mm 1/33 (33%) MRI: 1/3 (33%) Mammography: 1/1 (100%) Lymph node negative 3/3 (100%) MRI: 3/3 (100%) Mammography: 1/1 (100%) Histological grade: NR Test failures: MRI: 7/390 (1.8%) not completed or results not available. Reasons for non-completion not reported Mammogram: NR Exclusions 23/390 (5.9%)
Podo et al 2002	105 *(105)*	MRI versus mammography MRI versus mammography + US	Cancers detected: 8 cancers: 4 screen-detected cancers including 3 DCIS 0 Interval cancers Prevalence of breast cancer at initial screening round: 7/105 (7%) Incidence of breast cancer in 2nd round 1 (14 women presenting for 2nd round)	MRI TP 8 FP 1 TN 96 FN 0 Sensitivity 100% (63-100%) Specificity 99% (94-100%)	Mammography TP 1 FP 0 TN 97 FN 7 Sensitivity 13% (0.3-53%) Specificity 100% (96-100%) Mammography + US TP 1 FP 0 TN 97 FN 7 Sensitivity 13% (0.3-53%) Specificity 100% (96-100%)	Cancer characteristics Invasive cancer <10mm 3/5 (60%) MRI: 3/5 (60%) Mammography: 0/1 (0%) Lymph node negative 5/5 (100%) MRI: 5/5 (100%) M: 1/1 (100%) Histological grade I MRI: 2/5 (40%) Mammography: 1/2 (50%) Histological grade 3 MRI: 1/6 (17%) Mammography: 0/3 (0%) Test failures: NR Exclusions: NR

Stoutjesdijk et al 2001	75 *(75)*	MRI versus mammography	Cancers detected: 12 cancers: 12 screen-detected cancers including 2 DCIS 1 lymphoma 0 Interval cancers Prevalence of breast cancer at initial screening round: 12/75 (16%)	MRI AUC 0.98 (95%CI 0.95-1.0) Difference AUC MRI vs mammography 0.28 (95% CI: 0.17-0.39, *p* = 0.02) Adjusted for age and risk category: AUC 0.99 (95% CI: 0.96-1.0) Difference AUC MRI vs mammography 0.19 (95% CI: 0.09-0.29, *p* = 0.05)	Mammography + MRI within a 4-month period: AUC 0.70 (95%CI 0.60-0.80) Adjusted for age and risk category: AUC 0.80 (95% CI: 0.70-0.90)	Cancer characteristics: NR Test failures: NR Reasons for only receiving MRI or mammography NR Exclusions 66/245 (27%): no histology or adequate follow-up information 104/245 (42%): MRI or mammography not performed or undertaken > 4 months apart
Tilanus-Linthorst et al 2000, 2000b	109 *(193)*	MRI vs mammography	Cancers detected: 3 cancers: 3 screen-detected cancers including 0 DCIS 0 Interval cancers Prevalence of breast cancer at initial screening round: NR Overall detection rate (over 1-3 annual screening rounds) 3/193 (1.6%)	MRI TP 3 FP 6 FN 0 TN 100 Sensitivity 100% (29-100%) Specificity 94% (88-98%) False positive MRI rate reported as 6/193 (3%) of tests Biopsy rate with benign finding 2/193 (1%) However: • Additional US undertaken following 48/193 (25%) of screening tests in patients who did not have cancer. • FNAC undertaken following 26/193 (13%) of screening tests in patients who did not have cancer.	Mammography Not fully reported MRI detected 3 cancers not seen at mammography	Cancer characteristics Invasive cancer < 10 mm 0/3 (0%) Lymph node negative 3/3 (100%) MRI: 3/3 (100%) Mammography: 0/0 (0%) Test failures MRI: interrupted by claustrophobia in 1/109 (0.9%) of patients In addition “a few women” refused MRI “mainly because of claustrophobia” Mammography failures: NR Exclusions: NR Cost Extra cost of breast MRI per detected cancer €13 930 (cost of MRI, US, FNAC, biopsy per detected case)
Trecate et al 2003	20	MRI versus mammography	Cancers detected: 1 cancer: 1 screen-detected cancer 0 DCIS Interval cancers NR	MRI Cancer detection rate: 1/20 (5%) TP 1 FP: 0 FN: 0	Mammography Cancer detection rate: 0/20 (0%)	Cancer characteristics: NR Test failures: NR Exclusions: NR

Warner et al 2001, 2004	236 *(457)*	MRI + mammography + US + CBE versus mammography MRI versus mammography	Cancers detected: 22 cancers in 21 women: 21 screen-detected cancers including: 6 DCIS 1 Interval cancer Prevalence of breast cancer at initial screening round: Prevalence screen 13 cancers/236 (6%) Cancers detected in rounds 2-3: 9/221	Combination breast MRI, mammogram, US, CBE Sensitivity 95% (data for reconstruction of 2 x 2 table NR) Specificity NR MRI TP 17 FP 20 FN 5 TN 416 Sensitivity 77% (95% CI 55-92%) Specificity 95% (95% CI 93-97%) False positives proceeding to biopsy based on MRI alone: 32/236 (14%)	Mammography TP 8 FP 1 FN 14 TN 434 Sensitivity 36% (95% CI 17-59%) Specificity 99.8% (95% CI 99-100%)	Cancer characteristics Invasive cancer < 10 mm 5/16 (31%) MRI+mammography: 5/13 (38%) MRI: 5/13 (38%) Mammography: 1/5 (20%) Lymph node negative 13/15 (87%) MRI + mammography: 11/12 (92%) MRI: 11/12 (92%) Mammography: 4/4 (100%) Test failures MRI:3/236 (1.3%) too large to fit into MRI Mammogram: NR Exclusions Study withdrawals before round 3: 31/236 (13%) Reasons for withdrawal included: Bilateral mastectomy 16/236 (6.8%) Pregnancy 3/236 (1.3%) Developed metastatic cancer 4/236 (1.7%) Lost to follow-up 4/236 (1.7%) No longer wish to participate 1/236 (0.4%)

Patient preference, tolerance and therapeutic impact studies						
Kurian et al 2005	36 (43 invited)	Mammography vs MRI	1/46 (2%) breast cancer (DCIS) identified over study period 1/46 (2%) ovarian cancer 6/46 (13%) PM	MRI Maximal discomfort: 2/36 [6% (0-19%)] Moderate discomfort: 15/36 [42% (26-59%)] Minimal discomfort: 19/36 [53% (36-70%)] Comparison of MRI to mammography: Overall same or better: 61% Much worse: 4/36 [11% (3-26%)] Somewhat worse 10/36 [28% (14-45%)] Same 7/36 [19% (8-36%)] Somewhat better 6/36 [17% (6-33%) Much better 9/36 (25% (12-42%)	Mammography Maximal discomfort: 1/36 [3% (0-15%)] Moderate discomfort: 13/36 [36% (21-54%)] Minimal discomfort: 22/36 [61% (44-77%)]	Subgroup analysis of discomfort rating Little difference in discomfort rating of MRI and mammogram reported by • BRCA1 or BRCA2 mutation status • Prior history of breast or ovarian cancer Changes in attitudes towards prophylactic mastectomy Less opposed: 15% (5-31%) No change in opinion: 62% (44-78%) More opposed: 24% (11-41%) Comparison of proportion of patients who were/were not less opposed to prophylactic mammography, p=0.02
Rijnsburger et al 2004	334	Mammography vs MRI		MRI Pain Quite: 0.9%; Very: 0% Discomfort Quite: 4.6%; Very: 4.6% Anxiety Quite: 7.4%; Very: 2.8%	Mammography Pain Quite: 21.1%; Very: 9.0% Discomfort Quite: 15.8%; Very: 6.0% Anxiety Quite: 4.5%; Very: 0.7%	Health-related quality of life Small significant reduction in self-rated health score on visual analogue score over time ($p \leq 0.01$). No other generic quality of life scores changed over time. Study group showed higher SF-32 scores than age and sex-adjusted scores from Dutch and American reference population.

Appendix H. Breast cancer therapy costs

Comparison of US costs for breast cancer therapy used by Plevritis et al (2006a), including indirect costs, and Australian cost estimates.

Resources	Plevritis et al (2006a) costs in AU$	Australian Cost data AU$				
		Source	MBS/DRG/PBS item number [3]	Unit costs[4]	No	Total costs
Unilateral mastectomy with breast reconstruction						
Unilateral mastectomy			J06A	$6,125	1	$6,125
Breast reconstruction			J14Z	$17,923	1	$17,923
Total costs of unilateral mastectomy/breast reconstruction	*$41,827*					*$24,047*
Bilateral mastectomy with breast reconstruction						
Bilateral mastectomy [1]			J06A	$6,125	1.5	$9,187
Breast reconstruction			J14Z	$17,923	1.5	$26,884
Total costs of bilateral mastectomy/breast reconstruction	*$62,741*					*$36,070*
Hormone positive adjuvant therapy [2]						
Tamoxifen (20mg/day)			2110C		365*5	$2,400
Adjuvant chemotherapy		Seymour et al 1998, Eckerman 2003		$1,647	1	$1,647
			DRGR63Z	$757		
Adjuvant radiotherapy		Seymour et al 1998, Eckerman 2003		$6,825	1	$6,825
			DRGR64Z	$2,093		
Total costs of hormone positive adjuvant therapy	$32,664					$10,873
Hormone negative adjuvant therapy [2]						
Adjuvant chemotherapy		Seymour et al 1998, Eckerman 2003		$1,647	1	$1,647
			DRGR63Z			
Adjuvant radiotherapy		Seymour et al 1998, Eckerman 2003		$6,825	1	$6,825
			DRG R64Z			
Total costs of hormone positive adjuvant therapy	$19,416					$8,472

Table continued

Resources	Plevritis et al (2006a) costs in AU$	Australian Cost data AU$				
		Source	MBS/DRG/PBS item number [3]	Unit costs[4]	No	Total costs
Metastatic breast cancer treatment	$55,916					
		Seymour et al 1998, Eckermann et al 2003 [5]		*$40,578*	1	*$40,578*
		Wai et al (2001) (Canada)		*$46,263*	1	*$46,263*
Ovarian cancer treatment	*$144,608*					
		AIHW 2006 (based on data from 2000--01)[3]		*$22,607*	1	*$22,607*
Surveillance costs (annual)						
3 physician visits (attendance with specialist physician)			MBS 104	*$74.05*	3	$222
1 mammogram of both breasts			MBS 59300		1	$90
total annual surveillance costs	*$725*[6]					*$474*

1. Bilateral mastectomy assumed to incur an additional 50% over unilateral mastectomy.

2. Adjuvant therapy costs are assumed to be incurred over period of 1 year, based on Seymour et al (1998) and Eckermann et al (2003), tamoxifen for five years.

3. MBS and PBS costs are from 2005 versions, DRG costs are from round 8 cost round (2003-2004) indexed to 2005.

4 All costs were adjusted to 2003/2004 prices using health price index and to 2005 using consumer price index.

5. Assuming 5 years of treatment: 4 years cancer treatment and 1 year palliative care.

6. These are median annual costs, based on 3 half-day physician visits each year in first 4 years followed by 2 per year in the next 5 years.

Appendix I. Excluded studies

The following studies were retrieved by the literature search for full-text evaluation but excluded from the review because they did not meet the eligibility criteria listed in Table 10.

Beattie, M., Kerlikowske, K., 2005. 'The addition of MRI greatly improved the sensitivity of screening for early breast cancer in mutation carriers', *Evidence-based Obstetrics and Gynecology*, 7 (2), 100-102.

Bland, K. I., 2005. 'Added cancer yield of MRI in screening the contralateral breast of women recently diagnosed with breast cancer: Results from the International Breast Magnetic Resonance Consortium (IBMC) trial. Commentary', *Journal of Surgical Oncology*, 92 (1), 15-16.

Boetes, C., 2002. 'The evaluation of women with familial risk of breast cancer', *Journal of Experimental and Clinical Cancer Research*, 21 (3 SUPPL.), 97-101.

Bradbury, J., 2002. 'Breast-imaging pilot studies show promise', *Lancet Oncology*, 3 (8), 454.

Choi, S., 2005. 'Should women at high risk of breast cancer be screened with MRI?', *Canadian Medical Association Journal*, 173 (1), 32.

Feig, S. A., 2003. 'Screening strategy for breast cancer', *Seminars in Breast Disease*, 6 (4), 161-172.

Furman, B., Gardner, M. S. et al, 2003. 'Effect of 0.5 Tesla magnetic resonance imaging on the surgical management of breast cancer patients', *American Journal of Surgery*, 186 (4), 344-347.

Gilbert, F. J., 2005. 'Screening for breast cancer in women at moderate and high risk', *Clinical Oncology*, 17 (4), 240-243.

Goffin, J., Chappuis, P. O. et al, 2001. 'Re: Magnetic resonance imaging and mammography in women with a hereditary risk of breast cancer [2]', *Journal of the National Cancer Institute*, 93 (22), 1754-1755.

Gur, D., 2004. 'MRI in breast cancer', *The New England journal of medicine*, 351 (21), 2235-2236.

Hachiya, J., Seki, T. et al, 1991. 'MR imaging of the breast with Gd-DTPA enhancement: Comparison with mammography and ultrasonography', *Radiation Medicine - Medical Imaging and Radiation Oncology*, 9 (6), 232-240.

Hartman, A. R., Daniel, B. L. et al, 2004. 'Breast magnetic resonance image screening and ductal lavage in women at high genetic risk of breast carcinoma', *Women's Oncology Review*, 4 (2), 153-154.

Heywang, S. H., Fenzl, G., Hahn, D., 1986. 'MR imaging of the breast: Comparison with mammography and ultrasound', *Journal of Computer Assisted Tomography*, 10 (4), 615-620.

Hylton, N., 2002. 'Contrast-enhanced MRI of the breast', *Applied Radiology*, 31 (9 SUPPL.), 30-35.

Ingle, D. J., 2001. 'Magnetic resonance imaging of the breast: Current and potential applications for breast cancer detection, diagnosis, and staging', *Applied Radiology*, 30 (4 SUPPL.), 12-16.

Kriege, M., Brekelmans, C. T. M. et al, 2005. 'Efficacy of magnetic resonance imaging and mammography for breast cancer screening in women with a familial or genetic predisposition', *Obstetrical and Gynecological Survey*, 60 (2), 107-109.

Kuhl, C. K., 2004. 'Screening of women with hereditary risk of breast cancer', *Clinical Breast Cancer*, 5 (4), 269-271.

Kuzmiak, C. M., Pisano, E. D., 2002. 'Breast MRI: Not ready for routine practice', *Applied Radiology*, 31 (9), 16-19.

Liberman, L., 2004. 'Breast cancer screening with MRI – What are the data for patients at high risk?', *New England Journal of Medicine*, 351 (5), 497-500.

Love, S., Fischer, U., Chersevani, R., 2002. 'Early diagnosis and imaging', *Breast*, 11 (2), 196-197.

Mokbel, K., Elkak, A. E., 2001. 'Magnetic resonance imaging for screening of women at high risk for hereditary breast cancer [3]', *Journal of Clinical Oncology*, 19 (21), 4184-

Morris, E. A., Liberman, L. et al, 2003. 'MRI of occult breast carcinoma in a high-risk population', *American Journal of Roentgenology*, 181 (3), 619-626.

Morris, E., 2005. 'Breast MRI for cancer screening in high-risk patients', *Applied Radiology*, 34 (5 SUPPL.), 4-9.

Nelson, R., 2004. 'MRI better than mammography for detection of breast cancer?', *Lancet Oncology*, 5 (9), 520.

Orel, S. G., 1998. 'High-resolution MR imaging for the detection, diagnosis, and staging of breast cancer', *Radiographics : a Review Publication of the Radiological Society of North America, Inc*, 18 (4), 903-912.

Pathak, A. P., Artemov, D. et al, 2005. 'Characterizing extravascular fluid transport of macromolecules in the tumor interstitium by magnetic resonance imaging', *Cancer Research*, 65 (4), 1425-1432.

Pediconi, F., Venditti, F. et al, 2005. 'CE-Magnetic Resonance mammography for the evaluation of the contralateral breast in patients with diagnosed breast cancer', *Radiologia Medica*, 110 (1-2), 61-68.

Robson, M. E., Offit, K., 2004. 'Breast MRI for women with hereditary cancer risk', *Journal of the American Medical Association*, 292 (11), 1368-1370.

Sakorafas, G. H., Tsiotou, A. G., 2002. 'Contrast-enhanced magnetic resonance imaging has a high sensitivity for detection of breast lesions', *The Breast Journal*, 8 (2), 121.

Sim, L. S. J., Hendriks, J. H. C. L., Fook-Chong, S. M. C., 2004. 'Breast ultrasound in women with familial risk of breast cancer', *Annals of the Academy of Medicine Singapore*, 33 (5), 600-606.

Smith, R. A., Saslow, D. et al, 2003. 'ACS guidelines for breast cancer screening: update 2003', *CA A Cancer Journal for Clinicians*, 53 (3), 141-169.

Teifke, A., Hlawatsch, A. et al, 2002. 'Undetected malignancies of the breast: dynamic contrast-enhanced MR imaging at 1.0 T', *Radiology*, 224 (3), 881-888.

Tilanus-Linthorst, M. M., Obdeijn, I. M., Bartels, K. C., 2005. 'MARIBS study', *Lancet*, 366 (9482), 291-292.

Tilanus-Linthorst, M., Verhoog, L. et al, 2002. 'A BRCA1/2 mutation, high breast density and prominent pushing margins of a tumor independently contribute to a frequent false-negative mammography', *International Journal of Cancer*, 102 (1), 91-95.

Wang, L., 2003. 'Mammography and beyond: Building better breast cancer screening tests', *Journal of the National Cancer Institute*, 95 (5), 344-346.

Warner, E., Causer, P. A., 2005. 'MRI surveillance for hereditary breast-cancer risk', *Lancet*, 365 (9473), 1747-1749.

Warren, R. M. L., Pointon, L. et al, 2005. 'Reading protocol for dynamic contrast-enhanced MR images of the breast: Sensitivity and specificity analysis', *Radiology*, 236 (3), 779-788.

Zielinski, S. L., 2004. 'MRI more sensitive than mammography for women at high risk of breast cancer', *Journal of the National Cancer Institute*, 96 (17), 1275.

Appendix J. Eligible studies published after the preparation of this report

The present report is based on evidence identified from a systematic review of the literature conducted up to March 2006 (see literature review methods page 24). Two of the included studies published additional evidence between the time of preparing the final draft of this report in August 2006 and its consideration by MSAC in November 2006 (Griebsch et al 2006; Kriege et al 2006b). The National Institute for Health and Clinical Excellence (NICE) in the United Kingdom also produced a partial update of guidelines for the classification and care of women at risk of familial breast cancer in October 2006 that was informed by a new systematic review and economic evaluation of breast MRI for breast cancer screening (National Collaborating Centre for Primary Care, 2006).

This appendix presents a summary of this additional evidence. It was prepared for MSAC in October 2006 as a supplement to the final draft report. Please note that a further systematic review of the literature was not undertaken and so this appendix may not provide a summary of all relevant evidence published after March 2006.

The aim, design and results of the three studies are summarised in the table on pages 122-125 and discussed below.

1. Systematic review and cost-utility analysis (National Collaborating Centre for Primary Care (NCCPC) 2006)

This systematic review investigated the effectiveness of MRI plus mammography versus mammography alone and versus MRI alone for breast cancer screening in women at increased risk compared to the general population. The authors identified five test accuracy studies that were also included in the present MSAC review (Leach et al 2005, Kriege et al 2004, Warner et al 2004, Kuhl et al 2005, Lehman et al 2005). The NCCPC review satisfies some but not all criteria for a high quality systematic review and was therefore classified as fair quality. The authors presented a descriptive summary of the results of included studies but did not report an appraisal of the validity of each study. Heterogeneity between study results was not explored and the authors did not undertake a meta-analysis of these data.

The cost-utility analysis used a Markov model to estimate the incremental cost per QALY saved using MRI plus mammography versus mammography alone for screening high-risk women aged 30-49 years.

The costs of screening used in this model were based on estimates published by Griebsch et al (2006, results described below) from the MARIBS study (Leach et al 2005). The effects of screening were based on data about the relative sensitivity and specificity of MRI and mammography from the MARIBS study and included the following assumptions about the impact of screening on health outcomes:

- All cases would be identified in primary care following two false negative screening cycles.

- An individual would incur an increased mortality risk between a false negative screening test and eventual detection. This increased risk was estimated at 0.5% during the screening cycle following the false result.

- The prognosis of an individual at diagnosis was varied according to the number of false negatives experienced. For example, the 5-year survival for high-risk women who are not BRCA1 mutation carriers was assumed to be 85% if detected at the first screening test, 75% if detected at the second possible opportunity and 65% at the third possible opportunity.

These assumptions appear to have been based on expert opinion. No additional information was provided to support the estimates about the effects of early breast cancer detection used in the model. The Plevritis-modelled estimates of the impact of screening on the natural history of breast cancer cited in the present report were based on a stochastic model of the natural history of cancer that was fitted using data from a population-based cancer registry (SEER registry) in the United States. These modelled estimates may therefore provide a more robust estimate of the incremental effects of screening than the NCCPC (2006) assumptions, although there are still clear limitations to this approach.

The results of the NCCPC cost-utility analysis were sensitive to the differential estimates of 5-year survival rates following a delayed diagnosis and to the cost of MRI screening. Despite the different methods used to model the effects of screening, the results of the NCCPC base case model are similar to those presented in the draft assessment report based on the Plevritis model.

The incremental cost-effectiveness ratio in the NCCPC base model ranged from £7,781 (AU $19, 204) per QALY saved for BRCA1 mutation carriers aged 40-49 years to £29,622 (AU $73,122) per QALY saved for non-BRCA1 high-risk women aged 40-49 years (classified as ≥ 30% lifetime risk of breast cancer, includes BRCA2 carriers who were classified as having a 12% 10-year risk of breast cancer at age 40 years) and £38,919 (AU $96,124) per QALY saved for high-risk women aged 30-39 years. The draft assessment report estimates a cost per QALY saved of $28,974 for BRCA1 mutation carriers and $66,598 for BRCA2 mutation carriers aged 35 to 55 years using the Plevritis model.

The authors of the NCCPC cost-utility model drew similar conclusions to Plevritis et al (2006) about the cost-effectiveness of MRI screening in subgroups of women at high risk of familial breast cancer selected by risk level and age. The updated NICE guidelines for breast cancer surveillance recommend that known BRCA1, BRCA2 and p53 mutation carriers should be offered annual screening with breast MRI and mammography between the ages of 30 and 49 years (earlier MRI for women at exceptionally high risk) and to other high-risk groups depending on their level of risk and age (see table).

2. Economic analysis from the United Kingdom based on data collected in the MARIBS study (Griebsch et al 2006)

This study reported on the cost per additional breast cancer detected using MRI. It did not estimate the potential effects of earlier detection and thus does not allow an assessment of the impact of MRI screening on patient-relevant outcomes such as mortality and health-related quality of life.

Estimates of screening costs, including the costs of follow-up for false positive findings, were based on resource utilisation data collected from the MARIBS study. These costs included the costs of equipment, equipment maintenance, consumables and staff time. The costs associated with additional investigations were calculated based on assumptions about a clinician's threshold for recalling patients if they were given the results of mammography alone, MRI alone or both tests.

Mammography and MRI costs in this UK study were similar to the Australian estimates in the present MSAC report when converted to Australian dollars (Mammography cost $83 versus $90; and MRI costs $614 versus $538 respectively). The estimated incremental cost per screening episode was higher in the Griebsch et al study ($742) than the Australian estimate presented in the present MSAC report ($614).

Griebsch et al estimated the cancers detected by each screening method using data from the MARIBS study (33 screen detected cancers in 649 women, 1,881 scans).

The incremental cost per cancer detected reported by Griebsch et al (2006) based on the MARIBS study (and converted to Australian dollars) was $69,798 for the overall study population, $28,954 for BRCA1 mutation carriers and $37,768 for BRCA2 mutation carriers. The authors reported that their results were sensitive to estimates of the level of patient risk and unit costs of MRI. The authors concluded that the combination of MRI plus mammography is potentially cost effective in women at high risk of familial breast cancer, particularly BRCA1 and BRCA2 mutation carriers, although further research is required to determine whether early detection of breast cancer increases quality-adjusted life expectancy.

The cost per cancer detected was substantially lower than the cost per cancer detected in the present MSAC report which applied data on cancer incidence and test performance from the Plevritis model ($75,800 for BRCA1 and $175,013 for BRCA2 mutation carriers). These differences may largely be attributed to variations in assumptions regarding the sensitivity and specificity of the tests and the cancer incidence in the screening population.

The Griebsch model estimated mammography sensitivity as 23% in BRCA1 mutation carriers and 50% in BRCA2 mutation carriers from the MARIBS study (compared to estimates of 35% and 42% respectively used in the current MSAC Plevritis-based model); and MRI plus mammography sensitivity as 92% in BRCA1 and BRCA2 mutation carriers (whereas the current MSAC Plevritis-based model used estimates of 72% and 70% for women aged 35-54 years). Applying a higher sensitivity for the combination of MRI plus mammography as used in the Griebsch model would magnify the gain in effects for the addition of MRI over mammography alone and reduce the reported incremental cost per cancer detected.

The Griebsch model estimated cancer incidence based on the cancer detection rates reported in the MARIBS study (13/139 women with BRCA1 mutations, 12/86 women with BRCA2 mutations developed cancer during course of the study). These estimates were higher than the estimates used in the MSAC Plevritis-based model (cumulative incidence BRCA1 65%, BRCA2 45%), particularly for BRCA2 mutation-carriers, where Griebsch et al (2006) assumed an incidence rate 3.9 times higher than the Plevritis model. Assuming a higher incidence of cancer in a comparison of MRI plus mammography versus mammography alone would magnify the additional number of cancers detected by

the combination of MRI plus mammography and consequently lower the associated incremental cost per cancer detected.

It should be noted that the incidence of breast cancer among participants with a known BRCA2 mutation was particularly high in the MARIBS study, with 14% of BRCA2 participants developing cancer during the course of the study compared to 9% of BRCA1 participants. These results are inconsistent with the expected incidence of breast cancer in these two groups from existing evidence and may be unreliable due to relatively small numbers of participants with known BRCA1 and BRCA2 mutations in the MARIBS study. The very high incidence of breast cancer applied in the cost-effectiveness analysis by Greibsch et al may therefore limit the applicability of these results to a broader high-risk screening population.

3. Substudy of the accuracy study reported by Kriege et al (2004) (Kriege et al 2006b)

This study compared the characteristics of MRI-only detected cancers with other screen-detected cancers (Kriege et al 2006c). The authors classified the 41 evaluable screen-detected breast cancer cases as MRI-only-detected (20 cancers) or other screen-detected (21 cancers). Tumour characteristics were not available for all evaluable cancers.

Comparison of tumour size, lymph node status, histological differentiation grade, mitotic activity index, oestrogen receptor status and histology type by detection method showed that the extra cases detected by MRI alone were more often lymph node negative than the cases detected by other imaging or palpation (94% versus 59%, $p = 0.02$). MRI-only-detected cancers also appeared to be smaller than cancers detected by other methods but this difference was not statistically significant (cancers < 1 cm, MRI-only 58%, other methods 31%, $p = 0.11$).

At the time of preparing the present MSAC review, two studies provided evidence that the addition of breast MRI resulted in a favourable but nonsignificant increase in the proportion of screen-detected cancers that were lymph node negative (Leach et al 2005 29 cancers, Kuhl et al 2005 34 screen-detected cases, see Table 22). Both studies detected fewer cancers than the study reported by Kriege et al (2006b), and thus had less power to detect a true difference in tumour characteristics between MRI-screen-detected and mammography-screen-detected cancers.

The data reported by Kriege et al (2006b) therefore provides the strongest evidence to date that including breast MRI in screening programs for high-risk women will lead to a favourable stage shift in the detection of breast cancer. However, as discussed in this report, the magnitude of the clinical benefit associated with the detection of earlier stage disease has not been measured.

Summary

Overall, the results of these three studies are generally consistent with the evidence included in the body of this report. As discussed above, the recent study by Kriege et al (2006b) provides stronger evidence to support conclusions about the potential clinical benefit of adding breast MRI than the evidence available at the time of preparing this report. However, these additional studies provided no further evidence on which to base conclusions about the magnitude of clinical benefits associated with early detection of breast cancer versus later detection in young high-risk women.

The two additional economic analyses both indicated that, other than assumptions about the clinical impact of the early detection and treatment of breast cancer in young women at high risk, estimates of cost-effectiveness were most sensitive to variations in the cost of breast MRI and the underlying incidence of breast cancer in the screened population. The large differences in the estimated cost per cancer detected with the addition of breast MRI between Griebsch et al (2006) and this report highlight the sensitivity of estimates of cost-effectiveness to the expected incidence of cancer in the screened population and estimates of the incremental sensitivity and specificity of adding breast MRI to standard mammography screening programs.

Table of additional studies: study aim and design

Author, date, setting	*N*	Aim	Study design	Results	Quality appraisal
National Collaborating Centre for Primary Care Produced for NICE, October 2006 United Kingdom	5 primary studies	To assess the effectiveness and cost-effectiveness of: • MRI + mammography; • MRI; and • mammography in surveillance for breast cancer in women at increased risk compared to the general population. Undertaken as part of a broader review aiming to update NICE guidelines for the classification and care of women at risk of familial breast cancer.	Clinical guidelines: Systematic review and economic evaluation Population: asymptomatic women at increased risk of breast cancer Intervention: MRI or MRI + mammography (digital or film) Comparator: mammography (digital or film) Outcomes: • Sensitivity, specificity • No. cases identified • Positive and negative predictive values • Mortality • Cost effectiveness Studies excluded if: • Insufficient data to reconstruct 2 x 2 table • Did not report on women at increased risk of breast cancer Markov model constructed to estimate incremental cost per QALY saved using MRI + mammography versus mammography alone for screening moderate & high-risk women aged 30-49 years. • Test sensitivity & specificity based on MARIBS data • Breast cancer incidence by age based on published series in high-risk women • Costs based on MARIBS data and assumptions published by Griebsch et al (2006)	Included studies: 5 accuracy studies[1] Descriptive summary of individual studies presented Meta-analysis not undertaken Results of systematic review: • MRI + mammography is a more sensitive test than mammography alone in surveillance for breast cancer in women at high risk of familial breast cancer, particularly BRCA1 & BRCA2 carriers (Level 1b evidence). • 4 of 5 studies showed that MRI is a less specific test than mammography in high-risk women • Mammography is a useful addition to MRI in high-risk women, particularly for BRCA2 carriers because of their high incidence of DCIS. There is some evidence that mammography has a higher sensitivity than MRI in detecting DCIS in BRACA2 carriers (Level Ib). • 2 studies showed a greater differential in sensitivity in favour of MRI over mammography in BRCA1 carriers. (Level Ib) Results of cost utility model for annual MRI + mammography versus mammography: Incremental cost-effectiveness ratio (ICER) • BRCA1 mutation - 30-39 years: £13 486 per QALY saved - 40-49 years: £7 781 per QALY saved - *Conclusion*: cost-effective intervention in all women aged 30-49 years with a BRCA1 mutation. • high risk, non-BRCA1 mutation - 30-39 years: £38,919 per QALY saved - 40-49 years: £29,622 per QALY saved - *Conclusion:* equivocal cost-effectiveness in a high risk (non-	Systematic review Fair quality ✓ Explicit review question ✓ Explicit & appropriate eligibility criteria ✓ Explicit & comprehensive search strategy ✗ Validity of included studies appraised ✗ Methods of study appraisal reproducible ✗ Heterogeneity between studies assessed ✗ Summary of main results and precision estimate Economic evaluation Model limited by uncertainty about effects ✓ Explicit economic question ✓ Comprehensive description of alternative strategies *Costs* ✓ All important and relevant costs for each alternative strategy identified ✓ Costs measured accurately ✓ Costs valued credibly Perspective: NHS & personal social services costs

			Univariate sensitivity analyses and probabilistic sensitivity analyses conducted to test: • Variation in costs +/- 20%, and variation in MRI and mammography costs from MARIBS study • Variation in assumptions about effects of screening • Variation in utilities +/- 0.1 • Variation in radiation-induced cancer incidence	BRCA1) population aged 30-49, especially in those aged 30-39. Cost-effective intervention in non-BRCA1 women aged 40-49 with a 20% or greater 10-year risk (ICER ~ £ 20,000). Sensitivity analyses showed conclusions robust in BRCA1 and other high-risk populations In addition to patient risk level, ICER for MRI + mammography vs mammography most sensitive to assumptions about relative 5-year survival rates after delayed diagnosis and MRI unit cost.	*Outcomes* ? All important and relevant outcomes for each alternative strategy identified ? Clinical effectiveness established ✗ Outcomes measured accurately ? Outcomes valued credibly *Analysis* ✓ Discounting applied to costs and effects = 3.5% per annum ✓ Incremental analysis of costs and consequences? ✓ Sensitivity analyses conducted to investigate uncertainty in estimates of cost or consequences? ? Results generalisable to the setting of interest in the review

Recommendations for care of women at high risk extracted from the NICE updated guidelines (2006):

"● Access to psychological support and assessment is a key part of the package of care needed for many women covered by this guideline.

● All women aged 40-49 years satisfying referral criteria to secondary or specialist care (at raised risk or greater) should be offered annual mammographic surveillance.

● Surveillance should only be undertaken after provision of information about its potential advantages and disadvantages for the early detection of breast cancer, and where offered, this should be of high quality (equivalent to NHS Breast Screening Programme standard) and audited.

●● Women who are known to have a genetic mutation should be offered annual MRI surveillance if they are:

– *BRCA1* and *BRCA2* mutation carriers aged 30-49 years

– *TP53* mutation carriers aged 20 years or older.

●● MRI surveillance should be offered annually when indicated:

From 30-39 years: to women at a 10-year risk of greater than 8%

From 40-49 years: to women at a 10-year risk of greater than 20%, or to women at a 10-year risk of greater than 12% where mammography has shown a dense breast pattern.

● Genetic testing is appropriate only for a small proportion of women who are from high-risk families.

● Risk-reducing surgery (mastectomy and/or oophorectomy) is appropriate only for a small proportion of women who are from high-risk families and should be managed by a multidisciplinary team."

Author, date, setting	*N*	Aim	Study design	Results	Quality appraisal
Griebsch et al 2006 (using data from clinical study published by Leach et al 2005) United Kingdom	649 (1,881 screens) from Leach et al (2005)	To assess the cost-effectiveness of MRI + mammography versus mammography or MRI alone	Modelled estimate of incremental cost per case detected Sensitivity analysis using Markov Chain Monte Carlo method conducted to test: • Variations in assessment costs, including MRI and mammography unit costs • Variation in underlying incidence of cancer, including prevalent screens, BRCA1 subgroup, BRCA2 subgroup	Incremental cost per cancer detected for MRI + mammography versus mammography alone All women: £28,284 BRCA2: £15,302 BRCA1: no extra cases detected by addition of mammography to MRI. Incremental cost per case detected by MRI versus mammography = £11,731 Results most sensitive to MRI unit cost.	Economic evaluation Model does not assess impact of screening on health outcomes ✓ Explicit economic question ✓ Comprehensive description of alternative strategies *Costs* ✓ All important and relevant costs for each alternative strategy identified ✓ Costs measured accurately ✓ Costs valued credibly Perspective: NHS costs in £2003 *Outcomes* ✗ All important and relevant outcomes for each alternative strategy identified ✗ Clinical effectiveness established ✓ Outcomes measured accurately (based on one study) ? Outcomes valued credibly *Analysis* ✗ Discounting applied to equipment costs = 3.5% per annum over 10 years, no discounting of costs and effects of screening ✓ Incremental analysis of costs and consequences ✓ Sensitivity analyses conducted to investigate uncertainty in estimates of cost or consequences – but limited reporting of impact of variations in effects ? Results generalisable to the setting of interest in the review

Author, date, setting	N	Aim	Study design	Results	Quality appraisal
Kriege et al 2006b (using data from clinical study published by Kriege et al 2004) The Netherlands	1,909	To assess the tumour characteristics by detection screening method: MRI versus mammography	Prospective cross-sectional study Outcomes included: proportions of MRI-only-detected versus other-screen-detected tumours with characteristics: • Size < 1 cm • Negative lymph node • Type of histology • Oestrogen receptor status	45 evaluable breast cancers 41 screen-detected cancers MRI-only-detected cancers = 20 (49%) Other screen-detected cancers = 21 (51%) including: • Detected by mammogram only = 6 • Detected at MRI and mammography = 10 • Detected by CBE only = 1 • Detected by CBE ± MRI or mammography= 4 Tumour characteristics for MRI-only versus other screen-detected cases • Tumour size < 1 cm 11/19 (58%) versus 5/16 (31%) *p* = 0.11 • Lymph node negative 16/17 (94%) versus 5/16 (31%) *p* = 0.02 • Grade III 5/19 (26%) versus 5/16 (31%) *p* = 0.52 (comparison of grades I-III)	Fair quality ✓ Explicit description of patients ✓ Explicit eligibility criteria ✗ All patients included in analysis ✓ Outcomes assessed objectively ✓ Explicit description of techniques

1. Leach et al (2005), Kriege et al (2004), Warner et al (2004), Kuhl et al (2005), Lehman et al (2005) – all included in present review.

Abbreviations

AIHW	Australian Institute of Health and Welfare
AUC	Area under curve
BI-RADS	Breast Imaging Reporting and Data System
BRCA1	Breast Cancer Gene 1
BRCA2	Breast Cancer Gene 2
BSE	Breast self examination
CBE	Clinical breast examination
CI	Confidence interval
DCIS	Ductal carcinoma in situ
DoHA	Department of Health and Ageing
DNA	Deoxyribonucleic acid
FN	False negative
FNA	Fine needle aspiration
FP	False positive
HIC	Health Insurance Commission
HTA	Health technology assessment
IARC	International Agency for Cancer Research
IBMC	International Breast MRI Consortium
MBS	Medicare Benefits Schedule
MARIBS	Magnetic Resonance Imaging Breast Screening
MRI	Magnetic resonance imaging
MRISC	Magnetic Resonance Imaging Screening Study Group
MSAC	Medical Services Advisory Committee
NBSS	National Breast Screening Study
NHMRC	National Health and Medical Research Council
p53	Tumour antigen p53

ROC	Receiver operating characteristic
RR	Relative risk
SROC	Summary receiver operating characteristic
TN	True negative
TNM	Tumour, lymph node, metastases
TP	True positive
TP53	Tumour antigen p53
US	United States
USPSTF	United States Preventive Services Task Force
WHO	World Health Organisations

References

American Cancer Society 2005, *Breast cancer survival by stage*. Available from: http://www.cancer.org [Accessed 30 May 2006].

American College of Radiology 2003, *Breast Imaging Reporting and Data System Atlas*, American College of Radiology, Reston, Virginia.

Andrieu, N., Easton, D. F. et al, 2006. 'Effect of chest X-rays on the risk of breast cancer among BRCA1/2 mutation carriers in the international BRCA1/2 carrier cohort study: a report from the EMBRACE, GENEPSO, GEO-HEBON, and IBCCS Collaborators' Group', *Journal of Clinical Oncology* 24(21):3361-6.

Antoniou, A., Pharoah, P. D. et al, 2003. '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.', *American Journal of Human Genetics* 72(5):1117-30.

Australia and New Zealand Horizon Scanning Network 2005, *MRI screening for breast cancer in genetically high-risk women*, Commonwealth of Australia, Canberra.

Australian Bureau of Statistics 2001, *2001 Census Data : Australia*. Australian Bureau of Statistics, Available from: http://www.abs.gov.au [Accessed 3 March 2006].

Australian Institute of Health and Welfare 2005, *BreastScreen Australia monitoring report 2001-2002*, Australian Institute of Health and Welfare, Canberra.

Australian Institute of Health and Welfare (AIHW) 2006, *Interactive Cancer Data*. AIHW, Available from: http://www.aihw.gov.au/cognos/cgi-bin/ppdscgi.exe?DC=Q&E=/Cancer/canceragerates v7 [Accessed 17 February 2006].

Berrington de Gonzalez, A., Reeves, G., 2005. 'Mammographic screening before age 50 years in the UK: comparison of the radiation risks with the mortality benefits', *British Journal of Cancer* 93(5):590-6.

Blue Cross Blue Shield 2003, 'Magnetic Reasonance Imaging of the breast in screening women considered to be a high genetic risk of breast cancer', *Technology Evaluation Assessment Program*.

Borresen-Dale, A. L., 2003. 'TP53 and breast cancer', *Human Mutation* 21(3):292-300.

Boyd, N. F., Dite, G. S. et al, 2002. 'Heritability of mammographic density, a risk factor for breast cancer.', *New England Journal of Medicine* 347(12):886-94.

Boyle, P., 2003. 'Global summit on mammographic screening', *Annals of Oncology*, 14 (8), 1159-1160.

Brekelmans, C. T. M., Seynaeve, C. et al, 2006. 'Survival and prognostic factors in BRCA1-associated breast cancer', *Annals of Oncology*, 17 (3), 391-400.

Bucher, H. C., Guyatt, G. H. et al, 1997. 'The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials', *Journal of Clinical Epidemiology*, 50 (6), 683-691.

Buist, D. S., Porter, P. L. et al, 2004. 'Factors contributing to mammography failure in women aged 40-49 years', *Journal of the National Cancer Institute* 96(19):1432-40.

Calderon-Margalit, R.,Paltiel, O., 2004. 'Prevention of breast cancer in women who carry BRCA1 or BRCA2 mutations: A critical review of the literature', *International Journal of Cancer*, 112 (3), 357-364.

Christiansen, C. L., Wang, F. et al, 2000. 'Predicting the cumulative risk of false-positive mammograms', *Journal of the National Cancer Institute* 92(20):1657-66.

Collaborative group on hormonal factors in breast cancer, 1997. 'Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer', *Lancet*, v350 (n9084), 1047-1059.

Cuzick, J., Forbes, J. et al, 2002. 'First results from the International Breast Cancer Intervention Study (IBIS-I): a randomised prevention trial.', *Lancet* 360(9336):817-24.

Dagan, E.,Gershoni-Baruch, R., 2002. 'Anticipation in hereditary breast cancer', *Clinical Genetics* 62(2):147-50.

De Ridder F., De Maeseneer, M. et al, 2001. 'Severe adverse reactions with contrast agents for magnetic resonance: clinical experience in 30,000 MR examinations', *Jbr-Btr: Organe de la Societe Royale Belge de Radiologie* 84(4):150-2.

Dickson, R. B. 2001, 'Molecular basis of breast cancer.', In: Mendelsohn, J., Howley, P.M., Israel, M. A., Liotta, L.A. (eds), *The Molecular basis of cancer.* Philadelphia: W.B. Saunders Company.

Duffy, S. W., Tabar, L., Smith, R. A., 2002. 'The mammographic screening trials: commentary on the recent work by Olsen and Gotzsche', *CA: a Cancer Journal for Clinicians* 52(2):68-71.

Eckermann, S. D., Martin, A. J. et al, 2003. 'The benefits and costs of tamoxifen for breast cancer prevention', *Australian & New Zealand Journal of Public Health* 27(1):34-40.

Elmore, J. G., Barton, M. B. et al, 1998. 'Ten-year risk of false positive screening mammograms and clinical breast examinations.', *New England Journal of Medicine* 338(16):1089-9.

Elston, C. W.,Ellis, I. O., 1991. 'Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up', *Histopathology*, 19 (5), 403-410.

Ernster, V. L., Barclay, J. et al, 1996. 'Incidence of and treatment for ductal carcinoma in situ of the breast', *The Journal of the American Medical Association*, 275 (12), 913-918.

Fitzgibbons, P. L., Page, D. L. et al, 2000. 'Prognostic factors in breast cancer. College of American Pathologists Consensus Statement 1999', *Archives of Pathology & Laboratory Medicine* 124(7):966-78.

Garber, J. E., Goldstein, A. M. et al, 1991. 'Follow-up study of twenty-four families with Li-Fraumeni syndrome', *Cancer Research*, 51 (22), 6094-6097.

Gotzsche, P. C.,Olsen, O., 2000. 'Is screening for breast cancer with mammography justifiable?', *Lancet* 355(9198):129-34.

Griebsch, I., Brown, J. et al, 2006. 'Cost-effectiveness of screening with contrast enhanced magnetic resonance imaging vs X-ray mammography of women at a high familial risk of breast cancer', *British Journal of Cancer* 95(7):801-10.

Halapy, E., Chiarelli, A. M. et al, 2005. 'Accuracy of breast screening among women with and without a family history of breast and/or ovarian cancer', *Breast Cancer Research and Treatment*, 90 (3), 299-305.

Hartman, A. R., 2005. 'The problems with risk selection; scientific and psychosocial aspects', *Recent Results in Cancer Research* 166:125-44.

Hartman, A. R., Daniel, B. L. et al, 2004. 'Breast Magnetic Resonance Image Screening and Ductal Lavage in Women at High Genetic Risk for Breast Carcinoma', *Cancer*, 100 (3), 479-489.

Hartmann, L. C., Schaid, D. J. et al, 1999. 'Efficacy of bilateral prophylactic mastectomy in women with a family history of breast cancer.', *New England Journal of Medicine* 340(2):77-84.

Harvey, E. B.,Brinton, L. A., 1985. 'Second cancer following cancer of the breast in Connecticut, 1935-82', *National Cancer Institute Monographs* 68:99-112.

Haybittle, J. L., Blamey, R. W. et al, 1982. 'A prognostic index in primary breast cancer', *British Journal of Cancer* 45(3):361-6.

Humphrey, L. L., Helfand, M. et al, 2002. 'Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force.', *Annals of Internal Medicine* 137(5 Part 1):347-60.

Institute for Clinical Systems Improvement 2003, *MRI for the detection of breast abnormalities*, Available at: http://www.icsi.org [Accessed on 30 May 2006].

Irwig, L., Houssami, N., van Vliet C., 2004. 'New technologies in screening for breast cancer: a systematic review of their accuracy', *British Journal of Cancer* 90(11):2118-22.

Kenney, L. B., Yasui, Y. et al, 2004. 'Breast cancer after childhood cancer: a report from the Childhood Cancer Survivor Study.', *Annals of Internal Medicine* 141(8):590-7.

Keogh, L. A., Southey, M. C. et al, 2004. 'Uptake of Offer to Receive Genetic Information about BRCA1 and BRCA2 Mutations in an Australian Population-Based Study', *Cancer Epidemiology Biomarkers Prevention*, 13 (12), 2258-2263.

Kerlikowske, K., 1997. 'Efficacy of screening mammography among women aged 40 to 49 years and 50 to 69 years: comparison of relative and absolute benefit', *Journal of the National Cancer Institute Monographs*, 22 79-86.

Kerlikowske, K. M., Grady, D. M. et al, 1996. 'Effect of Age, Breast Density, and Family History on the Sensitivity of First Screening Mammography', *JAMA: The Journal of the American Medical Association*, 276 (1), 33-38.

Knopp, M. V., Weiss, E. et al, 1999. 'Pathophysiologic basis of contrast enhancement in breast tumors', *Journal of Magnetic Resonance Imaging* 10(3):260-6.

Komenaka, I. K., Ditkoff, B. A. et al, 2004. 'The development of interval breast malignancies in patients with BRCA mutations', *Cancer 100(10):2079-83,*

Kriege, M., Brekelmans, C. T. et al, 2006a. 'Differences between first and subsequent rounds of the MRISC breast cancer screening program for women with a familial or genetic predisposition', *Cancer 106(11):2318-26,*

Kriege, M., Brekelmans, C. T. et al, 2006. 'Tumor characteristics and detection method in the MRISC screening program for the early detection of hereditary breast cancer', *Breast Cancer Research & Treatment*, E-pub ahead of print

Kriege, M., Brekelmans, C. T. M. et al, 2004. 'Efficacy of MRI and mammography for breast-cancer screening in women with a familial or genetic predisposition', *New England Journal of Medicine*, 351 (5), 427-437+519.

Kuhl, C. K., Schrading, S. et al, 2005. 'Mammography, breast ultrasound, and magnetic resonance imaging for surveillance of women at high familial risk for breast cancer', *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 23 (33), 8469-8476.

Kurian, A. W., Hartman, A. R. et al, 2005. 'Opinions of women with high inherited breast cancer risk about prophylactic mastectomy: an initial evaluation from a screening trial including magnetic resonance imaging and ductal lavage', *Health Expectations*, 8 (3), 221-233.

Leach, M. O., 2005. '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*, 365 (9473), 1769-1778.

Lehman, C. D., Blume, J. D. et al, 2005a. 'Added cancer yield of MRI in screening the contralateral breast of women recently diagnosed with breast cancer: Results from the International Breast Magnetic Resonance Consortium (IBMC) trial', *Journal of Surgical Oncology*, 92 (1), 9-15.

Lehman, C. D., Blume, J. D. et al, 2005b. 'Screening women at high risk for breast cancer with mammography and magnetic resonance imaging', *Cancer*, 103 (9), 1898-1905.

Marcus, J. N., Watson, P. et al, 1996. 'Hereditary breast cancer: pathobiology, prognosis, and BRCA1 and BRCA2 gene linkage', *Cancer* 77(4):697-709.

McIntosh, A., Shaw, C., et al. 2004, *Clinical Guidelines and Evidence Review for The Classification and Care of Women at Risk of Familial Breast Cancer*, National Collaborating Centre for Primary Care/University of Sheffield, London.

Medical Services Advisory Committee 2005, *Sentinel node biopsy*, Commonwealth of Australia, Canberra.

Melendez, J. C.,McCrank, E., 1993. 'Anxiety-Related Reactions Associated With Magnetic Resonance Imaging Examinations', *The Journal of the American Medical Association*, 270 (6), 745-747.

Moss, S., Waller, M. et al, 2005. 'Randomised controlled trial of mammographic screening in women from age 40: predicted mortality based on surrogate outcome measures', *British Journal of Cancer* 92(5):955-60.

Narod, S. A., Lubinski, J. et al, 2006. 'Screening mammography and risk of breast cancer in BRCA1 and BRCA2 mutation carriers: a case-control study', *Lancet Oncology*, 7 (5), 402-406.

National Breast Cancer Centre 2006b, *Magnetic resonance imaging for the early detection of breast cancer in women at high risk: A systematic review of the evidence. Camperdown, NSW: NBCC (in press).*, NBCC, Camperdown, NSW.

National Breast Cancer Centre 2006a, *Advice about familial aspects of breast cancer and epithelial ovarian cancer: A guide for health professionals.* NBCC, Sydney.

National Breast Cancer Centre 2002, *Breast Imaging: a guide for practice. Camperdown, NSW: NBCC.*, Sydney.

National Collaborating Centre for Primary Care (NCCPC) 2006, *Familial breast cancer: The classification and care of women at risk of familial breast cancer in primary, secondary and tertiary care.* NICE, Available from: http://www.nice.org.uk/page.aspx?o=317667 [Accessed 12 October 2006c].

National Comprehensive Cancer Network. 2004, *Genetic/familial high-risk assessment: breast and ovarian*. National Comprehensive Cancer Network, Available from: http://www.nccn.org/professionals/physician_gls/PDF/genetics_screening.pdf. [Accessed 23 June 2006].

Nelson, H. D., Huffman, L. H. et al, 2005. 'Genetic risk assessment and BRCA mutation testing for breast and ovarian cancer susceptibility: Systematic evidence review for the US Preventive Services Task Force', *Annals of Internal Medicine*, 143 (5), 362-379.

NHMRC 1999, *Clinical Practice Guidelines. Familial aspects of cancer: A guide to clinical practice.*, Commonwealth of Australia, Canberra.

NHMRC 2001, *Clinical practice guidelines for the management of early breast cancer: Second edition.* Canberra.

NHMRC 2005, *Additional levels of evidence and grades for recommendations for developers of guidelines. Pilot Program 2005*, National Health and Medical Research Council, Australian Government, Canberra.

Phillips, K. A., Jenkins, M. A. et al, 2006. 'Risk-reducing surgery, screening and chemoprevention practices of BRCA1 and BRCA2 mutation carriers: a prospective cohort study', *Clinical Genetics* 70(3):198-206.

Pisano, E. D., Gatsonis, C. et al, 2005. 'Diagnostic performance of digital versus film mammography for breast-cancer screening.', *New England Journal of Medicine* 353(17):1773-83.

Pisano, E. D.,Parham, C. A., 2000. 'Digital mammography, sestamibi breast scintigraphy, and positron emission tomography breast imaging', *Radiologic Clinics of North America* 38(4):861-9.

Plevritis, S. K., Kurian, A. W. et al, 2006a. 'Cost-effectiveness of Screening BRCA1/2 Mutation Carriers With Breast Magnetic Resonance Imaging', *JAMA: The Journal of the American Medical Association*, 295 (20), 2374-2384.

Plevritis, S. K., Salzman, P. et al, 2006b. 'A natural history model of stage progression applied to breast cancer', *Statistics in Medicine*, published online 5 April 2006.

Podo, F., Sardanelli, F. et al, 2002. 'The Italian multi-centre project on evaluation of MRI and other imaging modalities in early detection of breast cancer in subjects at high genetic risk', *Journal of Experimental and Clinical Cancer Research*, 21 (3 SUPPL.), 115-124.

Reddy, D. H.,Mendelson, E. B., 2005. 'Incorporating new imaging models in breast cancer management', *Current Treatment Options in Oncology* 6(2):135-45.

Remvikos, Y., Magdelenat, H., Dutrillaux, B., 1995. 'Genetic evolution of breast cancers. III: Age-dependent variations in the correlations between biological indicators of prognosis', *Breast Cancer Research & Treatment* 34(1):25-33.

Rieber, A., Nussle, K. et al, 1999. 'MR mammography: influence of menstrual cycle on the dynamic contrast enhancement of fibrocystic disease', *European Radiology* 9(6):1107-12.

Rijnsburger, A. J., Essink-Bot, M. L. et al, 2004. 'Impact of screening for breast cancer in high-risk women on health-related quality of life', *British Journal of Cancer*, 91 (1), 69-76.

Robson, M. E.,Offit, K., 2004. 'Breast MRI for Women With Hereditary Cancer Risk', *JAMA: The Journal of the American Medical Association*, 292 (11), 1368-1370.

Schenck, J. F., 2000. 'Safety of strong, static magnetic fields', *Journal of Magnetic Resonance Imaging*, 12 (1), 2-19.

Schwartz, L. M., Woloshin, S. et al, 2000. 'US women's attitudes to false-positive mammography results and detection of ductal carcinoma in situ: cross-sectional survey', *Western Journal of Medicine* 173(5):307-12.

Smith, A., 2005. 'Full-field breast tomosynthesis', *Radiology Management* 27(5):25-31.

Smith, R. A., Saslow, D. et al, 2003. 'American Cancer Society guidelines for breast cancer screening: update 2003', *CA: a Cancer Journal for Clinicians* 53(3):141-69.

Smith-Bindman, R., Chu, P. et al, 2005. 'Physician predictors of mammographic accuracy', *Journal of the National Cancer Institute* 97(5):358-67.

Southey, M. C., Tesoriero, A. A. et al, 1999. 'BRCA1 mutations and other sequence variants in a population-based sample of Australian women with breast cancer', *British Journal of Cancer* 79(1):34-9.

Strong, K., Wald, N. et al, 2005. 'Current concepts in screening for noncommunicable disease: World Health Organization Consultation Group Report on methodology of noncommunicable disease screening', *Journal of Medical Screening* 12(1):12-9.

Tabar, L., Vitak, B. et al, 1999. 'Update of the Swedish Two-County Trial of breast cancer screening: histologic grade-specific and age-specific results', *Swiss Surgery* 5(5):199-204.

Tilanus-Linthorst, M., Verhoog, L. et al, 2002. 'A BRCA1/2 mutation, high breast density and prominent pushing margins of a tumor independently contribute to a frequent false-negative mammography', *International Journal of Cancer* 102(1):91-5.

Tilanus-Linthorst, M. M. A., Obdeijn, I. M. M. et al, 2000. 'First experiences screening women at high risk for breast cancer with MR imaging', *Breast Cancer Research and Treatment*, 63 (1), 53-60.

Travis, L. B., Hill, D. et al, 2005. 'Cumulative absolute breast cancer risk for young women treated for Hodgkin lymphoma.', *Journal of the National Cancer Institute* 97(19):1428-37.

Trecate, G., Vergnaghi, D. et al, 2003. 'Breast MRI screening in patients with increased familial and/or genetic risk for breast cancer: A preliminary experience', *Tumori*, 89 (2), 125-131.

UK MRI Breast screen study advisory group, 2000. 'Rationale for a national multi-centre study of magnetic resonance imaging screening in women at genetic risk of breast cancer', *The Breast*, 9 (2), 72-77.

Warner, E., Plewes, D. B. et al, 2004. 'Surveillance of BRCA1 and BRCA2 mutation carriers with magnetic resonance imaging, ultrasound, mammography, and clinical breast examination', *Journal of the American Medical Association*, 292 (11), 1317-1325.

Warner, E., Plewes, D. B., Shumak, R. S., 2001. 'Comparison of breast magnetic resonance imaging, mammography, and ultrasound for surveillance of women at high risk for hereditary breast cancer', *Journal of Clinical Oncology*, (15), 3524-3531.

Welch, H. G. M.,Black, W. C. M., 1997. 'Using Autopsy Series To Estimate the Disease "Reservoir" for Ductal Carcinoma in Situ of the Breast: How Much More Breast Cancer Can We Find?', *Annals of Internal Medicine*, 127 (11), 1023-1028.

Wetzig, N. 2005, *Sentinal lymph node biopsy versus axillary clearance*. Available from: http://www.ctc.usyd.edu.au/cochrane/index.html [Accessed 1 June 2006].

Whiting, P., Rutjes, A. et al, 2003. 'The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews', *BMC Medical Research Methodology*, 3 (1), 25-

World Health Organisation (WHO) 2003, *The World Health Report, 2003. Shaping the future*, WHO, Geneva.

Zamora, J., Muriel, A., Abraira, V. 2004, *Meta_DiSc for Windows: a software package for the meta-analysis of diagnostic tests*. Available from: www.hrc.es/investigacion/metadisc.html [Accessed 13 June 2006].