



NHMRC National Breast Cancer Centre

Review of the evidence
about the value of
mammographic screening
in 40-49 year old women

Professor Les Irwig
Associate Professor Paul Glasziou
Dr Alexandra Barratt
Mr Glenn Salkeld

About the authors:

Professor Les Irwig

Les Irwig is Professor of Epidemiology in the Department of Public Health and Community Medicine at the University of Sydney.

Associate Professor Paul Glasziou

Paul Glasziou is an Associate Professor in Clinical Epidemiology in the Department of Social and Preventive Medicine at the University of Queensland, and also works part-time in general practice.

Dr Alexandra Barratt

Alexandra Barratt is a Lecturer in Epidemiology in the Department of Public Health and Community Medicine at the University of Sydney.

Mr Glenn Salkeld

Glenn Salkeld is a Senior Lecturer in Health Economics in the Social and Public Health Economics Research Group within the Department of Public Health and Community Medicine at the University of Sydney.

National Library of Australia Cataloguing-in-Publication data:

Review of the evidence about the value of mammographic screening in 40-49 year old women.

Bibliography.

ISBN 0-9586732-9-2

1. Breast - Cancer - Diagnosis - Australia.
 2. Breast - Radiography - Australia.
 3. Medical screening - Australia.
 4. Middle aged women - Diseases - Diagnosis - Australia.
- I. Irwig, Les. II. National Breast Cancer Centre (Australia).

616.99449075

© 1997

NHMRC National Breast Cancer Centre
PO Box 572
KINGS CROSS NSW 1340

Ph: (02) 9334 1700

Fax: (02) 9326 9329

email: director@nbcc.org.au

website for publications: <http://www.nbcc.org.au/pages/info/resource/nbccpubs/nbccpubs.htm>

The NHMRC National Breast Cancer Centre is funded by the Commonwealth Government

Summary

Meta-analyses of the effectiveness of screening mammography in 40-49 year old women were reviewed to determine the likely effect of starting screening within this age group, and to assess the most favourable age for commencement of screening. Seven meta-analyses were identified and these show a small reduction in mortality from breast cancer after approximately eight years of screening. The size of this benefit increases with time: the meta-analysis with the longest follow-up shows a relative reduction in mortality of 15% (relative risk (RR) 0.85; 95% confidence interval (CI) 0.71-1.01) after an average follow-up period of thirteen years.

We repeated the meta-analysis using the most up-to-date data available from trials of screening among women aged 40-49 years with identical results (RR 0.85; 95% CI 0.71-1.01; risk difference -3.8 per 10,000). From this we can say that approximately 2,600 women aged between 40 and 49 years need to be screened to prevent one death from breast cancer thirteen years later. Using the effect estimates from this meta-analysis and Australian breast cancer incidence and mortality data, we modelled the number of deaths that would be prevented if women began screening at 40 rather than 50 years of age. Applying this model to a hypothetical cohort of 10,000 women who were offered screening about every two years from 40 years of age, the anticipated benefit would be a saving of about seven lives after thirteen years. To achieve this, about 2,000 women would be recalled for assessment, 230 women would require biopsies, 100 invasive cancers would be diagnosed and a further twenty-one women would be diagnosed with ductal carcinoma in situ. The marginal cost of offering second yearly screening to women from 40 years of age is estimated to be approximately \$40,000-\$65,000 per life year gained, but could be between \$13,000 and \$infinite, depending on the assumptions used in the cost-effectiveness calculations.

Contents

Summary	i
List of tables and figures	iii
Abbreviations	iv
1 Introduction	1
2 Review of existing meta-analyses	4
3 Methodological issues in relation to existing meta-analyses	7
4 Benefits and harms of screening	11
5 Weighing up benefits and harms: a balance sheet	19
6 Cost-effectiveness estimates	23
7 Conclusions	28
References	30
Appendix 1	35
Addendum: Some common questions about breast cancer screening in 40-49 year old women	38

Tables

1	Results of meta-analyses of screening women aged 40-49 years, arranged in order of increasing duration of follow-up	5
2	Characteristics of randomised trials of screening included in the meta-analyses presented in Table 1 (from Falun report 1996 ⁹)	6
3	Methodological criticisms and comments on the Canadian NBSS trial ¹⁶	10
4	Balance sheet - anticipated outcomes of ten years of screening offered to 10,000 Australian women from 40 years of age	19
5	Estimated benefits of mammographic screening among higher risk women	22
6	Cost estimates on a per screen basis for recruitment, screening and assessment	24
7	Cost per life year saved for (a) 40% and (b) 0% effect seen in trials is due to screening after 50 years of age, assuming relative risk reduction of 0.15 (95% CL 0.01, 0.29)	25

Figures

1	Benefits and harms of breast cancer screening by age	2
2	Meta-analysis of randomised trials of screening mammography among women aged 40-49. Results presented as risk differences and numbers needed to invite to screening to prevent one death	12
3	Deaths prevented by second yearly screens from 40-49 years of age assuming a) constant relative mortality effect and increasing incidence (solid circles) b) increasing relative mortality effect and increasing incidence (diamonds) c) increasing relative mortality effect (proportional to increasing sensitivity) and increasing incidence (open circles)	15

Abbreviations

AHMAC	Australian Health Ministers' Advisory Council
AHTAC	Australian Health Technology Advisory Committee
BRCA	breast cancer
CI	confidence interval
CL	confidence limits
DCIS	ductal carcinoma in situ
ERT	estrogen replacement therapy
HIP	Health Insurance Plan of Greater New York
HPI	health price inflators
MeSH	Medline search heading
MISCAN	microsimulation screening analysis
NBSS	National Breast Screening study
NNS	number needed to invite to screening to prevent one death
NPEDBC	National Program for the Early Detection of Breast Cancer
OR	odds ratio
RR	relative risk
RRR	relative risk reductions

1 Introduction

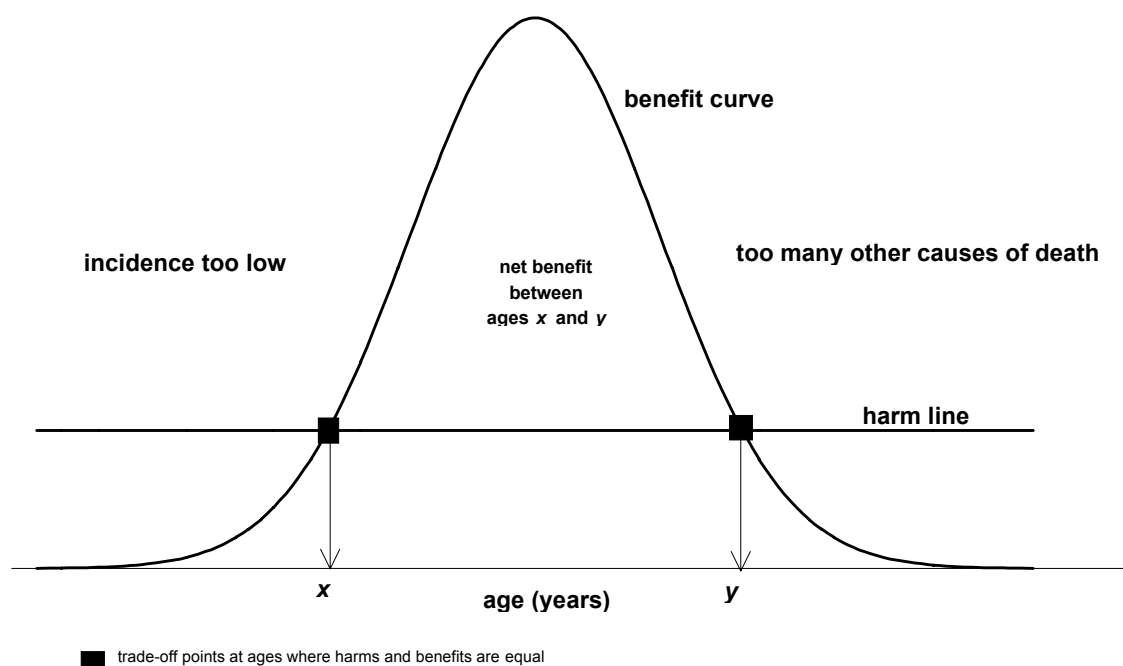
Breast cancer is the commonest cause of cancer death in Australian women with approximately 7,500 new cases annually and over 2,500 deaths annually.¹ The lifetime risk of breast cancer (up to 75 years) is now approximately one in fourteen in this country. Breast cancer incidence increased by 39% between 1973 and 1993 in NSW; over the same time period mortality rates were unchanged.² Evidence from randomised controlled trials shows that mammographic screening can substantially reduce mortality from breast cancer, at least among women aged 50-69 years.³ Although many trials included women in their forties, the trials have not generally shown a convincing reduction in mortality among these younger women and there has been considerable debate about whether screening should be offered to women aged 40-49 years.

Decisions about whether to offer mass screening are difficult and complex, and involve trading off the benefits against the harms of screening. Clearly there is no point in offering screening if early detection does not result in better outcomes; there is nothing to be gained by extending the period of time in which women know they have breast cancer unless there is some benefit of extended life or improved quality of life. If there is evidence that screening results in better outcomes, then the value of those better outcomes achieved for a relatively small number of people needs to be weighed up against the harms which may be incurred by the relatively large numbers of people who participate in a mass screening program and are never diagnosed with the disease of interest.

For most cancer screening programs, there is only a relatively brief age range during which screening is worthwhile (see Figure 1). Among very young women, breast cancer is so uncommon that the very small number of women who would benefit from screening would be vastly outweighed by the number of women who would be inconvenienced by regular screening. As the incidence of breast cancer increases with age, the benefit of screening will gradually increase with age to a point where the benefits outweigh the harms; 'trade-off point 1' (labelled x in Figure 1). For some older age groups the benefit will outweigh the harms. Eventually, however a woman has fewer and fewer years in which breast cancer can cause symptoms or death, and then a screening program would detect more cancer which would

never be clinically important. Thus we reach another critical point at which, again, the benefits of screening are outweighed by the harms: ‘trade-off point 2’ (labelled y in Figure 1).

Figure 1
Benefits and harms of breast cancer screening by age



Thus there is an age range of time during which screening for breast cancer is worthwhile. The challenge is to identify the ages at which the ‘trade-off points’ occur; to do this we need to quantify the size of the mortality benefit from screening, all the potential harms of screening (including unnecessary investigation and treatment) and the resource implications. With this information, women, policy makers and health professionals can then decide where the trade-off points occur so that mass mammographic screening can be offered to women between those ages. This document aims to summarise the best currently available information for use in that decision-making process.

Aims

The aims of this report are to:

- review the evidence on the impact of screening on breast cancer mortality among women aged 40-49 years and to review the methodological issues likely to impact on the analysis of effect;
- assess the incremental benefit of commencing screening at age 40 rather than age 50 years;
- estimate the likely relative risk reduction, absolute risk reduction and number needed to screen if screening commenced at age 40 rather than age 50 years; and,
- estimate the cost per life year gained by screening women from age 40 rather than from age 50 years.

2 Review of existing meta-analyses

As there have been several meta-analyses on breast cancer screening, we did not review all the primary studies. Rather, we located and examined the meta-analyses to determine why their results differed and what the implications were for obtaining reliable estimates of the effects of screening on breast cancer mortality.

We attempted to identify all the meta-analyses done on breast cancer screening, ie all papers in which some formal statistical pooling procedure was done to summarise the magnitude of effect of screening programs on breast cancer mortality rates in 40-49 year old women. Meta-analyses were identified in several ways. First, MedLine was searched using the MeSH heading breast neoplasm (exploded) and the MeSH heading meta-analysis (or any word beginning with meta). Second, we asked content experts for any meta-analyses of which they were aware. Third, we followed up the references in any papers we identified. We include in this report only those papers where some formal pooling procedure has been performed and exclude those which summarise the results of each study without pooling results. In the case of multiple publication of the same meta-analysis, we have used the most complete report available, usually that published first.

Each meta-analysis was examined to determine:

- which trials were included;
- the length of follow-up for each trial; and,
- the estimate of RR and its 95% CI for each trial.

Seven meta-analyses were found, with publication dates from 1992 to 1996 (see Table 1).³⁻⁹ These meta-analyses combined the results from some or all of seven randomised trials.¹⁰⁻¹⁶ No meta-analysis included the Kotka Pilot study on 40-47 year old women, which included only 4,163 women offered screening, allocated to screening because they were born in even calendar years.¹⁷ Likewise, the recently commenced Turku study, in which the effect of annual screening versus screening every three years is being compared, has not been included.⁹ For each meta-analysis, the duration of follow-up and the estimate of effect used for each primary study is shown in Tables A1-7 in Appendix 1.

Table 1
Results of meta-analyses of screening women aged 40-49 years,
arranged in order of increasing duration of follow-up

Meta-analysis (first author and year)	Studies included*	Length of follow-up in years (unweighted average)	Result (RR or OR)	95% CI
Elwood 1992 ⁴	1-3,5-7	7.0	1.08	.85-1.39
Wald 1994 ⁶	1-7	9.7	0.93	.76-1.15
Kerlikowske 1995 ³	1-7	9.8	0.92	.75-1.13
Glasziou 1995 ⁷	1-7	9.9	0.95	.76-1.18
Nystrom 1993 ⁵	1-4	10.2	0.90	.65-1.24
Smart 1995 ⁸	1-7	11.0	0.84	.69-1.02
Falun 1996 ⁹	1-7	13.1	0.85	.71-1.01

* Key to studies:

1 Malmo

2 Two county study

3 Stockholm

4 Gothenburg

5 Edinburgh

6 HIP

7 Canadian NBSS

As can be seen in Table 1, later meta-analyses have more years of follow-up. The longer the duration of follow-up, the lower the RR, in other words, the more effective screening appears to be. This mirrors the effect seen in trials for which the effects of screening have been plotted against time.^{4,5}

The most recent meta-analysis, which had the longest duration of follow-up, shows a 15% reduction in mortality (95% CI 0.71-1.01) after an average duration of follow-up of thirteen years.⁹

Characteristics of each trial are shown in Table 2.

Table 2
Characteristics of randomised trials of screening included in the meta-analyses presented in Table 1 (from Falun report 1996⁹)

Trial	Age range	N invited/control	Use of grid	Interval in months	Views	Readers
Two county	40-49	19,844/15,604	no	24	1	1
Malmo	45-49	3,795/3,769	no	21	2/1*	2
Stockholm	40-49	14,842/7,103	no	28	1	1
Gothenburg	40-49	10,879/12,800	yes	18	2/1	1
Edinburgh	45-49	11,370/10,269	yes	24	2/1	1
HIP	40-49	14,432/14,701	no	12	2	2
NBSS	40-49	25,214/25,216	Partly***	12	1/2**	1

* 2/1 = 2 views on first screen, 1 view subsequently

** 1/2 = 1 view at screening, 2 views for suspicious cases before any further assessment

*** in the latter part of the trial and not in all centres

3 Methodological issues in relation to existing meta-analyses

Incremental benefit of starting screening at age 40 rather than at age 50

A major difficulty with interpreting the evidence from these meta-analyses is that the trials on which they are based enrolled women aged between 40 and 49 years (or between 45 and 49 years in some cases) not women aged 40 years. Furthermore, they were not designed to have a comparison group offered screening from 50 years of age. Therefore we do not have a direct answer to the question ‘What is the incremental benefit of beginning screening at 40 years of age rather than 50 years of age?’. This question can only be answered directly by a randomised trial which recruits women as they turn 40. Such a trial is being established in the UK² (where the usual policy is to screen from 50 years of age) and has randomised 150,000 women thus far (S. Moss, personal communication).

Until the results of the UK trial are available, which will not be for some years, we can only indirectly estimate the effect of commencing screening at 40 rather than 50 years of age. We know from the meta-analyses that the benefit in mortality only starts after approximately eight years of follow-up, and increases with increasing duration of follow-up. As many of the women in the trials would have been over 50 years of age by the time eight or more years of follow-up was obtained, it is possible that some, or even all, of the observed benefit is due to other factors, such as screening of women after 49 years of age. There are two ways to address this issue: the first is by examining the empirical evidence and the second is by using statistical modelling techniques to estimate the proportion of the observed benefit which can be attributed to screening before 50 years of age.

Empirical approach

The interpretation of mortality data from the trials depends on whether screening was continued for the study group and whether screening was offered to the control group after the first few years of the trial. In summary, three patterns of screening can be discerned in the trials. First, one study (Malmo¹⁰) continued screening for over a decade and never screened controls; the observed effect is that of screening as from age 40-49 compared to no screening (note Malmo also recruited

women from age 45 only). Second, some studies (HIP,¹⁵ NBSS-Canada,¹⁶ Edinburgh^{14*}) screened for up to 7 years in the intervention group only; the observed effect is that of screening 40-49 year olds without any further screening in later decades. Third, some studies (Two county,¹¹ Stockholm,¹² Gothenburg¹³) introduced screening for the control groups after about 5 years of screening the intervention group. These studies most closely approximate the 'ideal' trial in which women are randomised to commence screening at 40 years of age rather than at 50 years of age. A meta-analysis of the data from these three studies is the best source of empirical data from which to estimate the effect of commencing screening before 50 years of age. We have conducted this meta-analysis and the results are reported below.

Modelling approach

Another approach to this problem is to use the available data to construct a statistical model to estimate the effect. De Koning, using the well-established microsimulation screening analysis (MISCAN) program and data on mortality as well as intermediate measures from the Swedish studies published in the meta-analysis by Nystrom,⁵ estimated that about two-thirds of the effect in the 40-49 year old age group (after seven to twelve years of follow-up) was due to screening which took place over age 50.¹⁹ While raising doubts about the fit of the model, Tabar provided extra data which indicated that 33% of the cancers detected in women aged 40-49 at the start of the Two-county program were in fact detected after age 50.²⁰

More recently, at the National Institutes of Health Consensus Conference in Washington, de Koning presented updated models with further follow-up which indicated that about 40% of the benefit of screening women in their forties is attributable to screening which occurs after 50 years of age.²¹ In a later section of this report, we have estimated the benefits and the costs that could be anticipated if screening were offered to women from 40 rather than 50 years of age using two assumptions: firstly, assuming that 40% of the benefit is attributable to screening after 50 years of age, and secondly, assuming that all of the benefit is attributable to screening before 50 years of age.

* The Edinburgh trial was of mixed design, owing to the introduction of the UK national screening program with some control women being offered screening within five to ten years of entering the trial.

Should the Canadian NBSS study be excluded?

Several authors have suggested that the Canadian study should be omitted from meta-analysis, and many meta-analyses show the effect separately with the Canadian study included and excluded. In particular, concern has been expressed over the randomisation procedure used and the possibility that subversion of the randomisation process led to an excess of women with poor prognosis breast cancer being allocated to the intervention (mammography) arm of the trial.²² However, the number of cases of breast cancer detected at baseline physical examination with no nodes, 1-3 nodes, >3 nodes and unknown number of nodes were 35, 13, 17, 0 respectively for the mammography group and 34, 16, 5 and 5 for the control group. Thus the total cancers in the mammography and control groups are similar (65 versus 60); and those without and with nodes involved in each group are similar (35, 30 versus 34, 21). These small differences are unlikely to make an important prognostic difference which would affect the results of the trial. Further an independent review of the randomisation procedure has been conducted by Dr Bailar (McGill University) and Prof MacMahon (Harvard University) with assistance from KPMG Investigation and Securities Inc. The review examined the documentation of the randomisation procedure in detail and found no credible evidence of subversion of the randomisation process. It concluded that if there were any subversions they were few in number and could only have had a trivial effect on the study results.²³ No other trial of mammographic screening has had its procedures reviewed with this level of scrutiny. Table 3 presents some other reasons that have been given for the exclusion of the Canadian study, and some comments on the appropriateness of exclusion on these grounds. In summary, we do not accept that there are any valid reasons for the exclusion of the Canadian NBSS study. Indeed, the Canadian study has published its methods in detail and this has allowed more intense scrutiny than has been accorded other trials which have generally not published so much information about their recruitment and randomisation processes. It is interesting to note that as the length of follow-up increases, the Canadian results are becoming increasingly similar to those of other trials.

Table 3
Methodological criticisms and comments on the
Canadian NBSS trial¹⁶

Reason given for exclusion	Comment
The result is an outlier, notably different from the other studies	Discrepant results do not warrant exclusion; they indicate that there is uncertainty about the findings between studies.
The trial had a chance finding of excess cancers in the few years after randomisation	It is inappropriate to exclude a study because the results are not in accord with expectation. Would one exclude a study where there was an unexpected advantage in the early years? In fact, the 'excess' may be no more than one might expect from the institution of a screening program. Moreover, comparison of pre-randomisation characteristics shows good balance between the arms of the trial. ^{16,24}
The screening procedure was of poor quality	Based on the usual criteria (sensitivity, stage distribution, interval cancers, detection rates), there are no grounds for thinking that the screening procedure was less adequate than in other studies. ⁷
The study was done on volunteers, rather than being population-based	There is no reason to expect that the effect of screening on mortality should be any less on volunteers than in the general population.

4 Benefits and harms of screening

The benefits and harms of screening are presented and discussed below and then weighed-up in a 'balance sheet'.

Benefits of screening

Benefit: incremental benefit of beginning screening at 40 rather than 50 years of age

1 Results of meta-analysis

There is evidence that the way in which research results are presented affects the reader's interpretation of the size and importance of the results.²⁵ Consequently we have conducted some meta-analyses of the available data and have explored a variety of ways of presenting the results. Below we present these results in terms of relative risks, the number of women who would need to be screened to prevent one death (NNS) and the number of deaths that would be prevented for every 10,000 women invited to screening.

2 Relative risk estimates

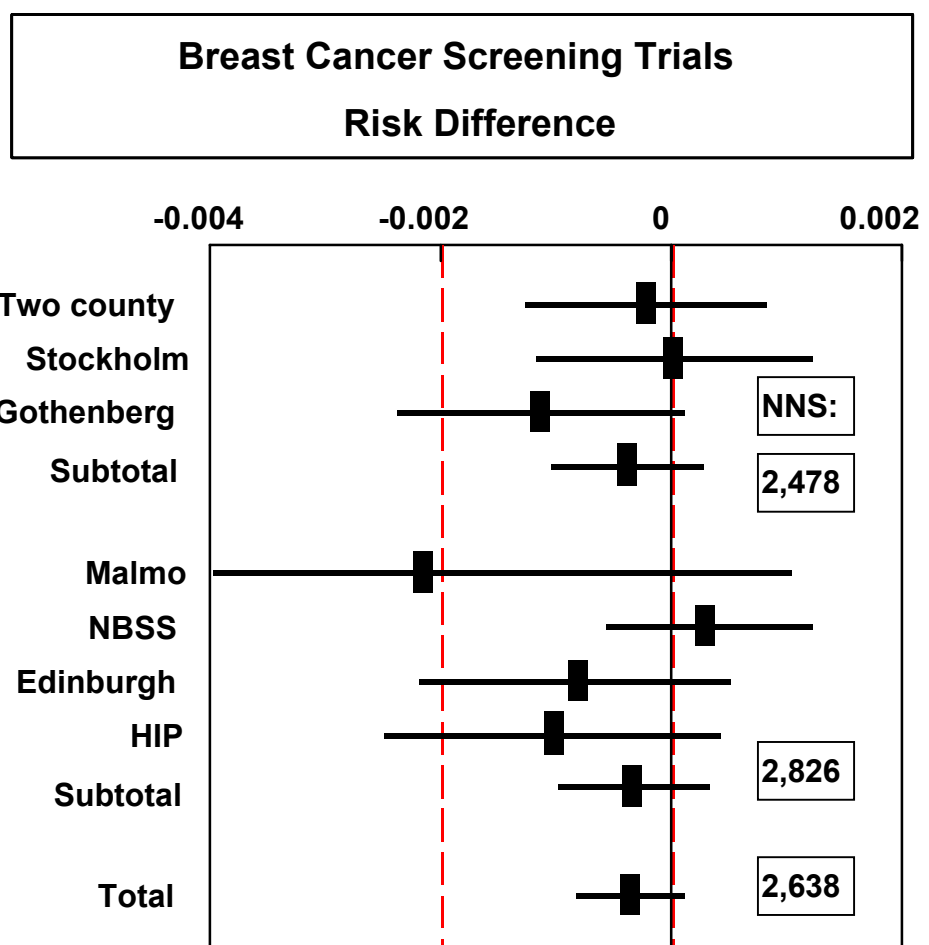
We conducted two meta-analyses based on all or a subset of the seven randomised controlled trials which have provided mortality data. Data presented at a recent conference in Falun plus updated data from the Canadian study were used.⁹ The analyses were done with MetaAnalyst (J Lau, MetaAnalyst, version 0.988, Boston 1996) using a fixed effects model. The results of these meta-analyses form the basis for the number needed to screen calculations and the cost-effectiveness analyses.

First, we conducted a meta-analysis using data from all seven trials, including updated data from the Canadian NBSS trial. After an average follow-up period of thirteen years, the result was an RR of 0.85 (95% CI 0.71-1.01). This result is identical to the result presented at Falun and reported recently in the *International Journal of Cancer*, that is an RR of 0.85 (95% CI 0.71-1.01).⁹

We then conducted a meta-analysis using only trials which introduced screening for the control groups about five years after screening was first offered to intervention group (see methodological discussion above). The point estimate was unchanged from the unrestricted meta-analysis, suggesting there is an incremental benefit of beginning screening at 40 rather than 50 years of age. It is possible that in this

estimate of effect some of the benefit is due to screening which occurs after 50 years of age.^{19,21} However, for the purposes of the present calculations we have assumed all the benefit is attributable to screening before 50 years of age.

Figure 2
Meta-analysis of randomised trials of screening mammography among women aged 40-49. Results presented as risk differences and numbers needed to invite to screening to prevent one death



NNS: Number needed to invite to screening to prevent one death

3 Risk differences and numbers needed to screen

The results of the meta-analysis can also be expressed as risk differences (and inverted to give the number of women needed to screen to prevent one death). Presenting the results as risk differences takes into account the underlying risk of breast cancer death and is a more meaningful measure of the potential benefit of screening. For the

three trials which introduced screening for the control group, there is a risk difference of 4.0 deaths per 10,000 women between the control and intervention groups over thirteen years (see Figure 2). This is equivalent to saying that 2,478 women aged 40-49 need to be invited to screening to prevent one death thirteen years later.

For the remaining four trials, in which the control group was not screened, the risk difference is 3.5 deaths prevented per 10,000 women screened (2,826 women need to be invited to screening for thirteen years to prevent one death). Overall (for all seven trials combined) the risk difference is 3.8 deaths prevented per 10,000 women screened (2,638 women need to be invited to screening for thirteen years to prevent one death)(see Figure 2).

As there is no important difference in the results of the meta-analyses using all seven studies and the subset of three trials with delayed screening, we have used the results from all trials combined for all subsequent analyses (that is a risk difference of 3.8 deaths prevented per 10,000 women invited to screening).

4 Deaths prevented per 10,000 women invited to screening from 40 years of age

The results of the meta-analysis described above and shown in Figure 2 are for the age range 40-49 years, and hence, on average, for the incremental effect of commencing screening from age 45 years. Commencing screening at 40 years would almost certainly give rise to greater absolute benefit but the extrapolation requires some assumptions and modeling. To approximate the incremental effect of commencing screening in the 40-44 year age group, we might assume the relative mortality benefit is the same, but that this is applied to the cumulative mortality from age 40 rather than age 45 years. The approximate risk difference would then be:

$$\begin{aligned} & \text{Cumulative risk difference for screening from 40-44 years} = \\ & \text{cumulative risk difference for screening from 45-49 years} \times \\ & \text{cumulative mortality (40-53)/cumulative mortality (45-58)} \end{aligned}$$

Based on NSW data, 1987-1991, the cumulative thirteen year mortality in women aged 40 is about 75% of that in women aged 45 years. Thus if 3.8 deaths are prevented per 10,000 women aged 45 at commencement of screening, then an additional $3.8 \times 0.75 = 2.9$ deaths might be prevented by commencing screening at age 40 years. This is probably somewhat inaccurate as it assumes the same relative mortality benefit in the 40-44 and 45-49 year age groups when, in fact, the benefit may be greater in the older than the younger group because of the increasing sensitivity of screening mammography with age.²⁶ However, using this estimate, a total of 6.7 deaths should be prevented over thirteen years per 10,000 women invited to begin screening at

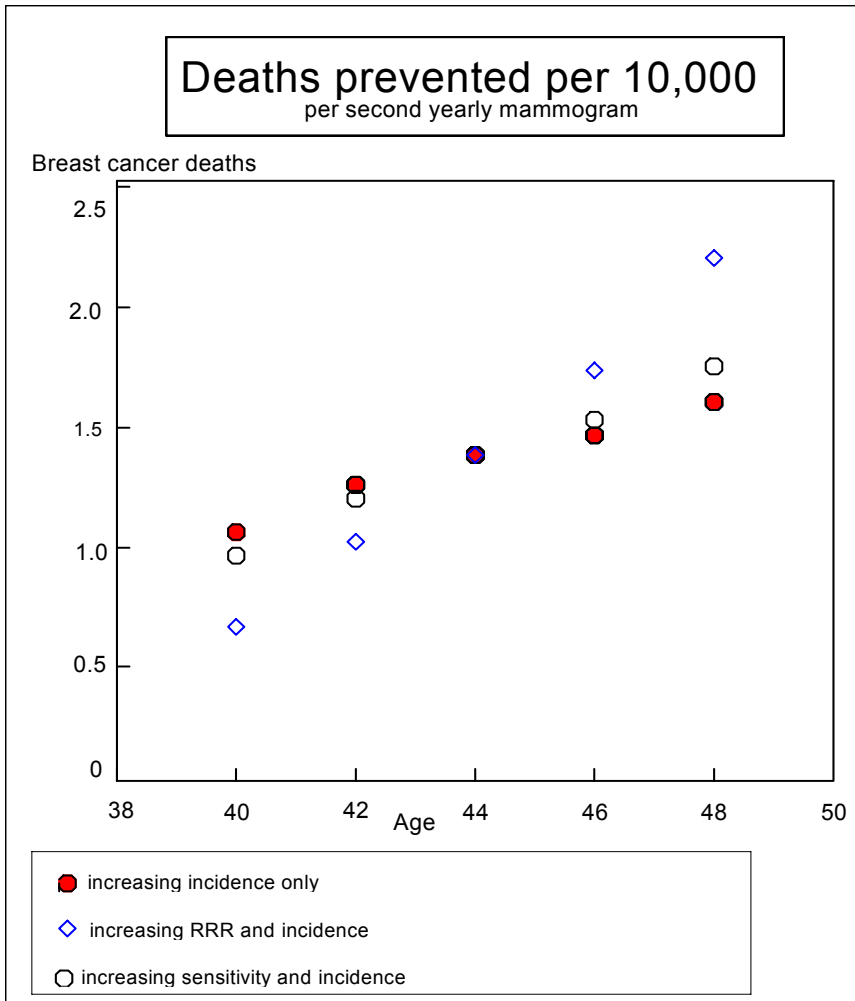
40 rather than 50 years of age. This effect occurred in trials in which attendance (compliance) was about 75-80%; thus we could anticipate that about nine deaths ($6.7 \times 100/75$) would be prevented per 10,000 women who actually attended five screens (from age 40-49) over the following thirteen years.²⁷

The benefit will vary by exact time of commencing screening. In Figure 3 the solid circles show the number of breast cancer deaths prevented for each of the second yearly screens from age 40-48 years based on the expected mortality in the thirteen years after each screen (approach a). This is probably an overestimate of the benefit of screening at the lower end of the age range and an underestimate of the benefit of screening at the upper end of the age range because the mortality benefit RRR increases with age. We can take account of the increasing relative risk reduction (RRR) by using a linear extrapolation of the RRR in 40-49 year old women of 15% and in 50-59 year old women of 28%⁵ to predict the number of deaths prevented by each screen but with the same total number of breast cancer deaths prevented as in approach a. This approach (b) is represented by the diamonds in Figure 3. An alternative approach (c) is to assume that the RRR increases linearly at the same rate that sensitivity increases based on published estimates of the sensitivity of screening mammography for invasive cancer of 58.3% at approximately 35 years of age, 75% at 45 years of age and 92.3% at 55 years of age.²⁶ This approach (c) is shown in open circles in Figure 3.

These calculations are based on the mortality from breast cancer in NSW women and the relative mortality reductions observed in trials conducted in Sweden, Scotland and Canada. There is therefore an assumption that the mortality from breast cancer in NSW is approximately similar to the mortality from breast cancer in these countries. A comparison of mortality rates shows this assumption is reasonable; NSW mortality rates in the 35-44 and 45-54 year age groups are generally slightly higher than those in Sweden, similar to Canada and lower than the mortality rates in Scotland.^{28,29}

Figure 3
Deaths prevented by second yearly screens from 40-49 years of age
assuming

- a) constant relative mortality effect and increasing incidence (solid circles)
- b) increasing relative mortality effect and increasing incidence (diamonds)
- c) increasing relative mortality effect (proportional to increasing sensitivity) and increasing incidence (open circles)



Note: deaths shown are prevented in the thirteen years after screening commences

Harms of screening

Against this benefit in mortality from breast cancer which could be achieved by introducing screening from 40 years of age, we need to balance a number of harms. These are presented individually below. Please note the harms presented below are outlined and discussed briefly; we have not undertaken a systematic review of the literature for each harm.

Harm: unnecessary investigation of false positive results

In Australia, the recall rate for abnormal screening mammograms is approximately five percent, although this varies between initial and later screens.² Women whose mammograms are reported as abnormal will be subjected to further investigation, although most of these women will not have cancer. They are therefore subjected to inconvenience, discomfort, expense and anxiety which would not have occurred had they not been screened. Local research has shown that women do experience psychological consequences of being recalled for further investigation but that these effects are short-term only,³¹ although there is some evidence suggesting greater adverse effects, for example, in one study 13% of women who had a false positive result were doing breast self-examination at least weekly.³¹

American data shows the positive predictive value (the proportion of women with abnormal screening mammography who were diagnosed as having breast cancer) increases with age; it has been estimated at 4% (95% CI 2-6%) at age 40-49 years, 9% (95% CI 6-12%) at 50-59 years and 17% (95% CI 13-21%) at 60-69 years.³³ Local data are similar, showing that cancer is diagnosed in 4%, 8% and 12% of women aged 40-49, 50-59 and 60-69 years respectively who are recalled for further assessment.³⁰ Thus a smaller proportion of women aged 40-49 years who are recalled will have cancer, compared to older women who are recalled. It is not known whether a higher rate of non-malignant lesions among women aged 40-49 years will compromise compliance among these women in later years, when the benefit of screening is greater.

Harm: overtreatment

Early detection is only beneficial if earlier treatment changes the course of the detected disease. For some women, their breast cancer may never have presented as a clinical problem. For these women, screening has not improved survival, it has merely increased their years of life with a diagnosis of cancer. This problem is best illustrated by ductal carcinoma in situ (DCIS). US data shows age-adjusted incidence of DCIS increased by over 300% between 1983

and 1992, probably largely due to the use of screening mammography.³⁴ Of women diagnosed with DCIS in the US in 1992, 44% were treated with mastectomy, 23% with lumpectomy and radiation and 30% by lumpectomy alone.³⁴ However, the natural history of this condition is not well known and optimal treatment is undetermined. In the largest follow-up series of untreated cases, Eusebi et al found that over seventeen years, nine of eighty cases developed into invasive cancers.³⁵ Thus, it is still unclear whether detection of DCIS is helpful or harmful. Data from the Australian public screening program show that DCIS accounts for 24.5% of breast cancer detected in women aged 40-49 years and 20.5% of breast cancer detected in women aged 50-59 years.³⁰

Harm: false negative mammograms

The sensitivity of screening mammography is lower in women under 50 years of age compared with older women. Sensitivity for the detection of invasive cancer at first screen has been estimated at 75% (95% CI 52.9-89.4) for women aged 40-49 and at 92.3% (95% CI 78.0-98.0) for women aged 50-59 years.²⁶ This means that about one quarter of women with invasive cancer in this age group will be falsely reassured. It is possible that women may delay investigation of breast symptoms if they are participating in a screening program, believing either that their last mammogram was normal and that the symptom is not due to cancer, or that investigation can wait until their next screening mammogram is due. Any effect on mortality is already included in the mortality estimates from trial data, however we are unaware of any data looking at the behavioural or psychological consequences of false negative mammograms. We do know that most women do not appreciate that screening is intended for asymptomatic people.³⁶ It is thus important that women are advised that not all cancers will be detected by screening mammography and that they should not ignore symptoms which develop between screens.

The sensitivity of screening mammography is related to breast density, which is in turn related to age and menopausal status, although these relations are complex. Kerlikowske et al found that the sensitivity of screening mammography was 98% and 84% for women aged 50 years or more with predominantly fatty and dense breast tissue respectively ($p < 0.01$).²⁶ However, for women aged less than 50 years, the sensitivity was similar for women with both fatty (82%) and dense breasts (86%) ($p = 0.79$). Looking at menopausal status, among women less than 50 years of age, sensitivity was 81% and 100% for pre- and post-menopausal women respectively, although this difference was not statistically significantly ($p = 0.22$).

Further evidence that hormonal status is important to the sensitivity of screening mammography comes from a study of the effect of estrogen replacement therapy (ERT) on the sensitivity of screening mammography.³⁷ This study found the sensitivity was significantly lower among women currently taking ERT (69%) compared with both former (94%) and never users (94%) of ERT.

Thus, although there is very little evidence available to date, and no direct evidence from the trials, it does appear that hormonal or menopausal status affects the sensitivity, and thus the benefit, of screening mammography.

Tabar and others have suggested that the screening interval should be reduced to compensate for the relative insensitivity of screening mammography in younger women.^{9,11} Screening at more frequent intervals would increase the number of cancers detected, although the improvement with annual screening cannot be estimated directly at present as there are insufficient published empirical data available to allow this. Further discussion of this issue is presented in Section 5.

Harm: radiation

The amount of radiation exposure which results from screening mammography is small and it has been estimated that the adverse effect, if any, is greatly outweighed by the mortality benefit.³⁸ For example, backward linear models have been developed from observational studies of high dose radiation. This suggests that if ten annual mammograms were done in 100,000 women this would induce less than eight breast cancers and these would occur at least a decade later.³⁹ Half of those cases (4) might eventually die from their breast cancer. If we compare this to the, at least, forty deaths prevented (as predicted by trial data) then the benefit to risk ratio is at least 10:1. It should be noted that these might already have occurred in the randomised trials and hence the benefit seen already takes into account the adverse effect of radiation. The fact that mortality benefit is seen in trials after up to eighteen years of follow-up and appears to continue with longer duration of follow-up also suggests that the benefits of screening mammography outweigh the harms of the extra radiation.

5 Weighing up benefits and harms: a balance sheet

In order to illustrate and compare these benefits and harms we have taken a hypothetical cohort of 10,000 women invited for screening from 40 years of age and developed a ‘balance sheet’ of the estimated number of women within the cohort who would be screened, investigated and diagnosed with cancer (see Table 4). Data for benefits are based on our meta-analyses of mortality reduction presented above and have been modelled to simulate the benefits that would accrue from commencing screening at 40 rather than 50 years of age. Data for the harms are from data published by the National Program, BreastScreen Australia.⁴⁰ We have assumed 80% compliance for first screens, as compliance rates in the trials varied from 65% to 100% with an average of 80% for the first screen (or at least one screen) and 75% compliance for subsequent screens. These compliance rates are somewhat higher than Australian rates,⁴⁰ but the balance sheet is nevertheless illustrative of the outcomes that could be expected if the mortality benefits observed in the trials are applied to the Australian situation.

Table 4
Balance sheet - anticipated outcomes of ten years of screening offered to 10,000 Australian women from 40 years of age

	First screen	Each of screens two to five	Total attributable to ten years of screening
No. of screens	8,000	7,500	38,000
No. of assessments	528	405	2,148
No. of biopsies	48	45	228
No. of invasive cancers	22	19	98
No. of DCIS	5	4	21
No. of deaths prevented after thirteen years			6.7 95% CL -1.4, 14.7

Table 4 shows that over thirteen years approximately seven deaths would be prevented per 10,000 women regularly screened from 40 years of age rather than from 50 years of age. Note the 95% confidence limits (CL) around this estimate are wide and include a small adverse effect. The balance sheet also shows that of the 10,000 women in the cohort, about 2,000 will be recalled for assessment over the ten years of screening.

Annual versus biennial screening

Tabar et al have reported that breast cancer progresses more rapidly in younger women and that to obtain substantial benefit from screening in younger women screening should be conducted every twelve to eighteen months.^{9,41} We would like to make the following comments with respect to this suggestion.

- 1 There are no empirical data to show that there is an increased benefit to be obtained by screening more frequently. However, Tabar et al have published estimates of the expected relative mortality for third yearly, second yearly and annual screening based on Markov chain and generalised linear modelling techniques.⁹ The results are reported as 4%, 18% and 36% relative reductions in mortality for third yearly, second yearly and annual screening respectively. While it is very likely that annual screening would result in an incremental benefit over second yearly screening, it is logically impossible to assert, as the Tabar et al model does, that by halving the screening interval the benefit can be doubled. It seems that currently, there is insufficient information available to be able to estimate the effect of annual screening with any certainty.
- 2 The frequency with which harms occur will be increased if the screening interval is reduced. If we were to move from second-yearly to annual screening there would be the potential to double the number of women who are recalled for further assessment. Few of these women will have cancer and many will be subjected to unnecessary investigation, inconvenience, cost and anxiety.

Improvements in mammography

It has been suggested that the detection of early breast cancer by screening mammography has improved since the randomised trials were conducted and that consequently screening mammography is now as effective in women aged 40-49 years as it is for women aged over 50 years.^{42,43} Improvements include the use of two view mammography⁴⁴ and double reading of films.⁴⁵ Double reading and two views were used by about half of the trials (see Table 2). It is

unclear whether screening mammography in Australia is achieving better detection of small cancers than in the randomised trials and therefore whether we would expect greater mortality benefits now because of improvements in mammographic technology. Data from Victoria show that among women 40-49 years, 27.5% of all cancers detected were 10mm or less in size.³⁰ This percentage is based on small numbers of cancers (forty cancers in the age group in total) and therefore has very wide CIs. It is difficult to compare this percentage with the percentage of small cancers detected in the randomised trials as the relevant data are either not published or are reported using different classifications of cancer size (for example some papers report the number of cancers 'less than 10mm' and others report cancers '10mm or less'). One way of exploring the possible effect of the standard use of double reading and two view mammography on mortality is to assume a direct relationship between increased detection and mortality reduction. Then we could adjust our estimate of the number of deaths prevented by the percentage improvement in detection attributable to double reading and two view mammography (a total of approximately 30% improvement in detection).^{44,45} However, as noted above about half the trials used double reading and two views. Therefore it seems reasonable to only increase the estimate of the number of deaths prevented by half the anticipated improvement (that is 15%). This adjustment increases the estimated number of deaths prevented to 7.7 deaths prevented per 10,000 women screened from age 40 for ten years over thirteen years.

Weighing up benefits and harms of screening among high risk women

Risk factors other than age generally do not provide sufficient risk differentiation to greatly alter the benefit achieved by screening. However, women who are at high risk of breast cancer can potentially be identified by family history or genetic testing and these women could gain much greater absolute benefits by beginning screening in their forties.

Currently there are no trials which have tested whether screening is beneficial for women at high risk of breast cancer so we have no empirical evidence available to us. Nevertheless, we know that we need to screen about 2,600 women at population risk to prevent one death from breast cancer; and therefore we can estimate the number needed to screen among higher risk women. Women at moderate and high risk (as defined by family history) account for about 4% and 1% of the population respectively.⁴⁶ Among women at moderately increased risk, that is whose risk is approximately double that of the general population, we could expect the benefit of screening to be approximately doubled and we would therefore need to screen only

about 1,300 moderate risk women to prevent one death. Among women who are at approximately four times higher risk than the general population (that is women who have a very strong family history with up to a one in two risk of developing breast cancer) we should only have to screen about 650 such women to prevent one death from breast cancer. Thus we would anticipate that screening women at increased risk of breast cancer from 40 years of age should result in substantial benefits; among the high risk women we would expect to prevent about thirty deaths per 10,000 women invited to screening. Consequently the cost-effectiveness figures would also be substantially better than in the general population.

Table 5
Estimated benefits of mammographic screening among higher risk women

	General population	Moderately increased risk	High risk
Estimated risk of breast cancer	1 in 13 to 1 in 8	1 in 8 to 1 in 4	1 in 4 to 1 in 2
Approximate number needed to screen to prevent one death	2,600	1,300	650
Approximate number of deaths prevented per 10,000 women invited to screening	7	14	28

It should be noted that the above estimates are based on screening at the standard two-yearly interval. However, women who test positive for genetic mutations associated with breast cancer, for example the BRCA and BRCA2 genes, may be advised to have a screening mammogram every year or even every six months.² There are no data to suggest that screening at this frequency would be more beneficial. There is also the possibility that women with inherited or acquired defects in DNA repair mechanisms may be more susceptible to radiation-induced cancer and thus that the benefits may be offset to an unknown extent by very frequent screening.²¹

6 Cost-effectiveness estimates

The cost-effectiveness estimates for extending mammographic screening to women aged 40-49 have been calculated using two methods. The first method uses the original MISCAN modelling data (contained in the report of the national evaluation for screening women aged 50-69 years presented to the Australian Health Ministers' Advisory Council, AHMAC),⁴⁸ but with updated costs and effects based on the most recent screening costs and trial-based measures of effect. This method is called the 'adjusted AHMAC' model throughout this section. The second method estimates cost-effectiveness directly; using the effect estimates presented in the balance sheet (see Table 4) and the most recent Australian screening costs.

Costs

Costs for the cost-effectiveness analyses presented below are derived from an adjustment of the costs used in the national evaluation of mammographic screening for women aged 50-69 years. This approach calls for an assessment of the validity of the original screening cost estimates derived in 1990 for the AHMAC.⁴⁸ Pilot project cost data were used to estimate the cost of screening during the phase-in period for a national program of five years, and the cost of screening for the steady-state period beyond five years. Cost estimates (in 1990 Australian dollars) on a per screen basis for recruitment, screening and assessment were \$120 in the first year of a national screening project through to \$80 per screen for a steady-state program. Carter and Cheok⁴⁹ conducted a breast cancer screening costing study to see whether the actual costs of the national program were consistent with the national program's funding formula of \$120 per woman screened during the phase-in period, and \$80 per woman screened once the program reached its steady-state level. The results of these studies are summarised in Table 6 and are expressed in 1996 prices. The AHMAC 1996 prices are the original 1990 AHMAC costs updated to 1996 using health price inflators (HPI).⁵⁰ It should be noted that the Carter and Cheok 1996 prices represent the actual costs of screening during the phase-in period of the National Program for the Early Detection of Breast Cancer (NPEDBC).

Table 6
Cost estimates on a per screen basis for recruitment, screening and assessment

Cost estimate	Year 1	Year 2	Year 3	Year 4	Year 5
	\$	\$	\$	\$	\$
AHMAC 1990 prices	120	110	100	90	80
AHMAC 1996 prices	135	124	112	101	90
Carter and Cheok 1996 prices	121	134	n/a	97	92

Whilst there are some variations in cost during the phase-in period, the Year 5 steady-state cost is very similar; \$90 versus \$92. We have used the steady-state cost of screening and assessment of \$90 per woman screened to estimate the cost-effectiveness of screening women aged 40-49.

In using these adjusted average costs we have not accounted for the current capacity of the national screening program to accommodate screening women aged 40-49. Ideally, it is the marginal cost per screening women aged 40-49 that should be used to analyse the impact of expanding the current screening program. However, given the scale of expanding the program to include women aged 40-49 years and the lengthy time horizon, it is reasonable to assume that the marginal cost per woman screened would in any case approximate the average cost per woman screened.

Cost-effectiveness analysis

Adjusted AHMAC model

The AHMAC analysis found that the marginal cost of moving from a policy of second yearly screening in the age group 50-69 years to a policy of second yearly screening in the age group 40-69 years was \$49,500 per life year gained. Further it showed that moving from second yearly to yearly screening in the 40-49 age group had a marginal cost-effectiveness of \$187,000 per life year gained.^{48, 51} However this modelling was based on an assumption of a 7% reduction in breast cancer mortality due to screening in the 40-49 age group. The most recent evidence suggests that the benefit is larger though more delayed than that used in the model; there is about a 15% relative reduction in breast cancer mortality at thirteen years of follow-up but this does not appear for at least seven to ten years. A complete reanalysis would be time consuming and expensive, and we

have therefore recalculated the AHMAC analysis adjusting for the increase in benefit from 7% to 15% and the 12.5% estimated increase in costs between 1990 and 1996. The results are shown in Table 7.

We would like to emphasise that the cost-effectiveness of screening in the 40-49 age group is particularly sensitive to the size of the mortality benefit, which is difficult to estimate for two reasons. First, the confidence limits around the benefit are still quite wide. Second, it is uncertain what proportion of the benefit seen in the randomised trials occurs because of screening beyond the age of 50. Simulations by de Koning and colleagues suggest that about 40% of the benefit seen in trials was due to screening beyond 50 years of age.²¹ Table 7 therefore presents two point estimates of cost-effectiveness (with CL derived from the CL around the RRR) assuming either (a) 40% of the benefit or (b) 0% of the benefit that was seen in the trials comes from screening beyond the age of 50.

Table 7
Cost per life year saved for (a) 40% and (b) 0% effect seen in trials is due to screening after 50 years of age, assuming relative risk reduction of 0.15 (95% CL 0.01, 0.29)

	Estimated cost per life year saved (15% relative risk reduction)	95% CL (29% relative risk reduction)	95% CL (1% relative risk increase)
(a)	\$40,617	22,409	infinite
(b)	\$24,370	13,445	infinite

From Table 7, we can say that the marginal cost of introducing screening from 40 years of age, based on this adjustment to the AHMAC analysis, ranges from \$24,370-\$40,617 per life year saved based on a relative reduction in mortality of 15%. The range between these two point estimates reflects the proportion of benefit from screening that occurs beyond the age of 50. *The figure of \$40,617 is likely to be more appropriate because it takes account of the fact that part of the mortality benefit is due to screening after 50 years of age.* The CL around these estimates (that is the range of plausible values) are very wide. The marginal cost per life year gained could be as little as \$13,000 assuming a 29% RRR and 0% of effect due to screening beyond the age of 50, or it could be infinite assuming 1% RR increase.

Direct estimate of cost-effectiveness from trial data

Using the data from Table 4 it is possible to directly estimate the costs that would be incurred by introducing screening for women from

40 years of age. We did this by using the cost estimate from Carter and Cheek of \$92 per screen.⁴⁹ This estimate includes the costs of subsequent investigations for women whose mammograms are reported as abnormal. By this method, the total cost for 38,000 screens over ten years is \$3,496,000 and the cost per death prevented is approximately \$526,000. The gain is 6.7 deaths prevented between seven and thirteen years later which results in a gain in discounted life years of 44.9 (discounting both costs and life years at 5% per annum) or approximately \$64,800 per discounted life year gained. Thus, the direct, but approximate measure of cost-effectiveness is somewhat greater than that obtained using the adjusted model described above.

Comment on the cost-effectiveness analysis

To summarise, the cost-effectiveness analysis suggests the marginal cost of screening women from 40 years of age is probably between \$24,000 and \$65,000 per life year gained. From an economic point of view, the central question is whether extending mammographic screening represents value for money. This is a value judgement based on relative rather than absolute cost of achieving health gain. One approach to answering the question is to compare screening women aged 40-49 with other possible uses of the health care resources involved. The cost per life year gained for the national mammographic screening program (screening women aged 50-69 every two years) is estimated at \$17,000 per life year saved, \$37,500 for cervical cancer screening⁴⁸ and \$25,000 for annual faecal occult blood testing for colorectal cancer.⁵² Both the mammographic and cervical cancer screening programs were regarded as acceptable value for money by the Commonwealth Government. The evidence for colorectal cancer screening is currently being reviewed by the Australian Health Technology Advisory Committee (AHTAC). Our estimate of \$24,000 per life year gained for mammographic screening of women aged 40-49 (assumption b) clearly lies within the range deemed acceptable value for money by the Commonwealth. On the other hand, our estimate of \$41,000 (assumption a), which is likely to be more accurate as it takes account of the fact that some of the mortality benefit is due to screening after 50 years of age, lies just outside this range. The estimate from the direct approximation method of \$65,000 lies well outside the acceptable range.

As noted earlier, the cost-effectiveness estimates need to be adjusted proportionately to estimate the cost-effectiveness of screening among higher risk women. For example, among women at moderately increased risk the cost-effectiveness estimate would be about half that in average risk women.

7 Conclusions

We conclude that the evidence available to date shows a modest benefit of commencing screening from 40 rather than 50 years of age. This benefit is not apparent until about eight years after the commencement of screening. Our estimate is that seven deaths will be prevented by thirteen years after beginning screening for every 10,000 women invited to second yearly screening from 40 years of age.

The benefit is likely to depend on the hormonal status of women in their forties, as the sensitivity of screening mammography appears to be dependent on hormonal status and breast density. Greater benefit will be obtained from each screening examination for women in their late forties compared to women in their early forties because of the increasing incidence of breast cancer and the improved sensitivity of screening mammography as women become menopausal.

The benefit is likely to be proportionally greater among women at increased risk of breast cancer. For women at moderate risk (based on family history) about fifteen deaths per 10,000 women could be prevented and for women at high risk up to thirty deaths per 10,000 women could be prevented.

To be considered against these benefits are the increased rates of false positives and false negatives that will occur if screening is offered to younger women. About 5% of women will be reported as abnormal at each screening and will require further investigation; ultimately about 95% of these women will not have breast cancer but they will have experienced anxiety, inconvenience, cost and discomfort to varying degrees. Also unknown is the extent to which it may reduce attendance for screening in older age groups when the benefits of screening are greater. About 25% of women aged 40-49 (compared with 7-8% of women aged 50-69 years) who do have invasive breast cancer will be incorrectly classified as normal. Thus it is critical women appreciate the relative inaccuracy of screening mammography in this age group. The acceptability of this level of misclassification to women is unknown.

The costs of introducing second yearly screening for women from 40 years of age cannot be estimated accurately at present, due to the wide CI around the estimate of the relative mortality reduction observed in trials of mammographic screening of women aged 40-49 years. We estimate it will cost about \$40,000-\$65,000 per life year gained to start screening at age 40, although it is possible the cost could be as little as \$13,000 per life year gained or infinitely large.

Costs can be expected to be proportionately lower for women at higher risk.

To summarise, the benefit of second yearly screening of women from 40 years of age is modest and the harms and costs are not inconsequential. We strongly suggest that the NHMRC National Breast Cancer Centre needs to be guided by studies looking at how women weigh the benefits and harms of screening, and whether women, having been given full information about screening, would choose to be screened. This should be done before the cost-effectiveness issues are considered in determining future policy. If women perceive the benefit of introducing screening earlier outweighs the harms we suggest the total financial and resource implications of expanding the screening program to cater for women aged 40-49 years are estimated to ensure that there are adequate facilities to handle the increased numbers of screens and investigations.

References

- 1 Kricger A, Jelfs P. Breast cancer in Australian women 1921-1994. Australian Institute of Health and Welfare, Canberra; 1996.
- 2 Coates M, McCredie M, Armstrong B. Cancer in New South Wales incidences and mortality 1993. Sydney: Cancer Control Information Centre, NSW Cancer Council; 1996.
- 3 Kerlikowske K, Grady D, Rubin SM, Sandrock C, Ernster VL. Efficacy of screening mammography: a meta-analysis. *JAMA* 1995;273:149-54.
- 4 Elwood JM, Cox B, Richardson AK. The effectiveness of breast cancer screening by mammography in younger women. Report prepared for the Department of Health, New Zealand. Hugh Adam Cancer Epidemiology Unit and Department of Health, New Zealand; 1992.
- 5 Nystrom L, Rutqvist L, Wall S et al. Breast cancer screening with mammography: overview of Swedish randomised trials. *Lancet* 1993;341:973-8.
- 6 Wald N, Chamberlain J, Hackshaw A, on behalf of the Evaluation Committee. Consensus conference on breast cancer screening Paris, 1993. Report of the Evaluation Committee. *Oncology* 1994;51:380-9.
- 7 Glasziou P, Woodward AJ, Mahon CM. Mammographic screening trials for women aged under 50: a quality assessment and meta-analyses. *Med J Aust* 1995;162:625-9.
- 8 Smart CR, Hendrick RE, Rutledge JH et al. Benefit of mammography screening in women ages 40 to 49 years: current evidence from randomised controlled trials. *Cancer* 1995;75:1619-26.
- 9 Breast-cancer screening with mammography in women aged 40-49 years. Report of the organising committee and collaborators, Falun Meeting, Falun, Sweden 1996. *Int J Cancer* 1996;68:693-9.
- 10 Andersson I, Aspegren K, Janzo L et al. Mammographic screening and mortality from breast cancer: the Malmo mammographic screening trial. *BMJ* 1988;297:943-8.

- 11 Tabar L, Fagerberg G, Hsis-Hsi Chen et al. Efficacy of breast cancer screening by age. *Cancer* 1995;75:2507-17.
- 12 Frisell J, Eklund G, Hellstrom L et al. Randomised study of mammographic screening - preliminary report on mortality in the Stockholm trial. *Breast Cancer Res Treat* 1991;18:49-56.
- 13 Nystrom L, Rutqvist L, Wall S et al. Breast cancer screening with mammography: overview of Swedish randomised trials. *Lancet* 1993;341:973-8.
- 14 Roberts M, Alexander F, Anderson T et al. Edinburgh trial of screening for breast cancer: mortality at seven years. *Lancet* 1990;335:241-6.
- 15 Shapiro S, Venet W, Strax P et al. Ten-to-fourteen year effect of screening on breast cancer mortality. *J Natl Cancer Inst* 1982;69:349-55.
- 16 Miller A, Baines C To T, Wall C. Canadian national breast screening study: 1. Breast cancer detection and death rates among women aged 40-49 years. *Can Med Assoc J* 1992;147:1459-88.
- 17 Hakama M, Pukkala E, Kallio M et al. Effectiveness of screening for breast cancer in women under 50 years at entry: The Kotka Pilot Project in Finland. *Int J Cancer* 1995;63:55-7.
- 18 Fletcher S, Black W, Harris R et al. Report of the International workshop on screening for breast cancer. *J Natl Cancer Inst* 1993;85:1644-56.
- 19 de Koning HJ, Boer R, Warmerdam PG et al. Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer screening trials. *J Natl Cancer Inst* 1995;87:1217-23.
- 20 Tabar L. Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer screening trials. *J Natl Cancer Inst* 1996;88:52-5.
- 21 National Institutes of Health. Consensus development consensus statement: Breast cancer screening for women ages 40-49. National Institutes of Health; January 1997.
- 22 Boyd NF, Jong RA, Yaffe MJ et al. A critical appraisal of the Canadian breast cancer screening study. *Radiology* 1993;189:661-3.

- 23 Bailar JC, MacMahon B. Randomisation in the Canadian National Breast Screening Study: a review for evidence of subversion. *Can Med Assoc J* 1997;156:193-9.
- 24 Cohen MM, Kaufert PA, MacWilliam L, Tate RB. Using an alternative data source to examine randomisation in the Canadian national breast screening study. *J Clin Epidemiol* 1996;49:1039-44.
- 25 Fahey T, Griffiths S, Peters TJ. Evidence based purchasing: understanding results of clinical trials and systematic reviews. *BMJ* 1995;311:1056-60.
- 26 Kerlikowske K, Grady D, Barclay J et al. Effect of age, breast density, and family history on the sensitivity of first screening mammography. *JAMA* 1996;276:33-8.
- 27 Newcombe RG. Explanatory and pragmatic estimates of the treatment effect when deviations from allocated treatment occur. *State Med* 1988;7:1179-87.
- 28 Coates M, Day P, McCredie M, Taylor R. Cancer in New South Wales incidence and mortality 1992. NSW Central Cancer Registry, NSW Cancer Council; 1995.
- 29 World Health Statistics Annual 1995. World Health Organisation Geneva; 1996.
- 30 BreastScreen Victoria. 1995 Annual Statistical Report. BreastScreen Victoria, Melbourne; 1996.
- 31 Cockburn J, Staples M, Hurley SF, De Luise T. Psychological consequences of screening mammography. *J Med Screening* 1994;1:7-12.
- 32 Bull AR, Campbell JM. Assessment of the psychological impact of a breast screening programme. *Br J Radiol* 1991;64:510-5.
- 33 Kerlikowske K, Grady D, Barclay J et al. Positive predictive value of screening mammography by age and family history of breast cancer. *JAMA* 1993;270:2444-50.
- 34 Emster VL, Barclay J, Kerlikowske K et al. Incidence of and treatment for ductal carcinoma in situ of the breast. *JAMA* 1996;275:913-8.
- 35 Eusebi V, Foschini MP, Cook MG et al. Long-term follow-up of in situ carcinoma of the breast with special emphasis on clinging carcinoma. *Semin Diagn Pathol* 1989;6:165-73

- 36 NHMRC National Breast Cancer Centre. Report on the 1996 Breast Health Survey. NHMRC National Breast Cancer Centre, Woolloomooloo (NSW); 1997.
- 37 Laya MB, Larson EB, Taplin SH, White E. Effect of estrogen replacement therapy on the specificity and sensitivity of screening mammography. *J Natl Cancer Inst* 1996;88:643-9.
- 38 Law J. Risk and benefit associated with radiation dose in breast screening programmes - an update. *Br J Radiol* 1995;68:870-6.
- 39 Feig SA. Mammographic screening of women aged 40-49 years. Benefit, risk and cost considerations. *Cancer* 1995;76:2097-106.
- 40 National program for the early detection of breast cancer: evaluation of phase 1, 1991-1994. Commonwealth Department of Human Services and Health, Canberra; 1995.
- 41 Tabar L, Fagerberg G, Chen H et al. Screening for breast cancer in women aged under 50: mode of detection, incidence, fatality and histology. *J Med Screening* 1995;2:94-8.
- 42 Thurfjell L, Lindgren JA. Breast cancer survival rates with mammographic screening: similar favourable survival rates for women younger and those older than 50 years. *Radiology* 1996;201:421-6.
- 43 Burhenne LJ, Burhenne HJ, Kan L. Quality oriented mass mammography screening. *Radiology* 1995;194:185-8.
- 44 Wald NJ, Murphy P, Major P et al. UKCCCR multicentre randomised controlled trial of one and two view mammography in breast cancer screening. *BMJ* 1995;311:1189-93.
- 45 Ciatto S, Rosselli Del Turco M, Morrone D et al. Independent double reading of screening mammograms. *J Med Screening* 1995;2:99-101.
- 46 Hopper J. Genetic testing: review of the scientific literature. NHMRC National Breast Cancer Centre, Woolloomooloo (NSW); 1995.
- 47 Hoskins KF, Stopher Je, Calzone KA et al. Assessment and counselling for women with a family history of breast cancer: a guide for clinicians. *JAMA* 1995;273:577-85.
- 48 Australian Institute of Health Prevention Program Evaluation Series No 1. Breast cancer screening in Australia: future directions. Breast Cancer Screening Evaluation Steering Committee. AGPS, Canberra; 1990.

- 49 Carter R, Cheok F. National report: breast cancer screening costing study. National Program for the Early Detection of Breast Cancer, Commonwealth Department of Human Services and Health; 1994.
- 50 Australian Institute of Health and Welfare. Health Expenditure. Health Expenditure Bulletin: number 12. December 1996.
- 51 Carter R, Glasziou P, van Oortmarssen G et al. Cost-effectiveness of mammographic screening in Australia. *Aust J Public Health* 1993;17:42-50.
- 52 Salkeld G, Young G, Irwig L et al. Cost-effectiveness analysis of screening by faecal occult blood testing for colorectal cancer in Australia. *Aust J Public Health* 1996;20:138-43.

Appendix 1

Table A1
Elwood, November 1992

Trial	Yrs of follow-up	Result	95% CI
Malmo	7	1.33	0.64-2.73
Two county	7	1.14	0.63-2.08
Stockholm	7	1.10	0.47-2.57
Edinburgh	7	0.98	0.45-2.11
HIP	7	0.78	0.48-1.25
NBSS Canada	7	1.36	0.84-2.21
All	7	1.08	0.85-1.39

Table A2
Wald, 1994

Trial	Yrs of follow-up	Result	95% CI
Malmo	9	0.51	0.22-1.19
Two county	11	1.04	0.68-1.59
Stockholm	7	1.00	0.49-2.04
Göteborg	5	0.72	0.26-1.99
Edinburgh	10	0.87	0.46-1.65
HIP	18	0.78	0.52-1.18
NBSS Canada	8	1.36	0.84-2.21
All	9.71	0.93	0.76-1.15

Table A3
Kerlikowske, 1995

Trial	Yrs of follow-up	Result	95% CI
Malmo	12	0.51	0.22-1.17
Two county - W (K)	12	0.75	0.41-1.36
Two county - E (O)	12	1.28	0.76-2.33
Stockholm	8	1.04	0.53-2.05
Göteborg	7	0.73	0.27-1.97
Edinburgh	10	0.78	0.46-1.51
HIP	10	0.77	0.50-1.16
NBSS Canada	7	1.36	0.84-2.21
All	9.75	0.92	0.75-1.13

Table A4
Glasziou, 1995

Trial	Yrs of follow-up	Result	95% CI
Malmo	12	0.52	0.22-1.20
Two county	12	1.06	0.70-1.62
Stockholm	8	0.82	0.40-1.68
Gothenburg	7	0.72	0.26-1.99
Edinburgh	10	0.85	0.41-1.79
HIP	10	0.82	0.54-1.25
NBSS Canada	10	1.36	0.83-2.21
All	9.9	0.95	0.76-1.18

Table A5
Nystrom (overview of Swedish studies), 1993 (with individual study results as presented in Fletcher 1993)

Trial	Yrs of follow-up	Result	95% CI
Malmo	12	0.51	0.22-1.17
Two county - W (K)	12	0.75	0.41-1.35
Two county - E (O)	12	1.28	0.70-2.33
Stockholm	8	1.04	0.53-2.05
Gothenburg	7	0.73	0.27-1.92
All	10.2	0.87	0.63-1.20

Note: the individual results in Table A2 are from the evaluation model; consequently the combined estimate is also the evaluation model estimate. In the body of our report, we have used the combined estimate for the follow-up model which is methodologically more appropriate but the individual trial data are not available for the follow-up model. The follow-up model gives a combined estimate of 0.90, 95% CI 0.65-1.24.

Table A6
Smart, 1995

Trial	Yrs of follow-up	Result	95% CI
Malmo	12	0.51	0.22-1.17
Two county - W (K)	13	0.73	0.37-1.41
Two county - E (O)	13	1.02	0.52-1.99
Stockholm	8	1.04	0.53-2.05
Gothenburg	7	0.73	0.27-1.97
Edinburgh	10	0.78	0.46-1.51
HIP	18	0.77	0.53-1.11
NBSS Canada	7	1.36	0.84-2.21
All	11	0.84	0.69-1.02

Table A7
Falun, 1996

Trial	Yrs of follow-up	Result	95% CI
Malmo	15	0.67	0.35-1.27
Two county - W (K)	15	0.67	0.37-1.22
Two county - E (O)	15	1.02	0.59-1.77
Stockholm	12	1.08	0.54-2.17
Gothenburg	10	0.59	0.33-1.06
Edinburgh	10	0.73	0.43-1.25
HIP	18	0.77	0.53-1.11
NBSS Canada	10	1.10	0.79-1.54
All	13.13	0.85	0.71-1.01

Addendum

Some common questions about breast cancer screening in 40-49 year old women.

Prepared by Professor Les Irwig, Associate Professor Paul Glasziou and Dr Alexandra Barratt.

The NHMRC National Breast Cancer Centre (NBCC) report *Review of the evidence about the value of mammographic screening in 40-49 year old women* has prompted some common questions which we outline and discuss below.

Q1. Is there really any significant effect of mammographic screening in 40-49 year old women?

It is true that the 15% reduction in mortality is only at the threshold of the usual criterion of statistical significance of $p=0.05$. However, we know that mammographic screening is effective in women over 50 years of age, so there is additional information to support the view that the effect is real. Anyway, a more useful approach is to examine the likely range of effect. Based on the 95% confidence intervals, the likely impact of mammographic screening in 40-49 year old women on breast cancer mortality is somewhere in the range between no effect through to a 29% reduction in mortality. Similarly, while the point estimate for number needed to screen is 2,638 (ie 2,638 women aged 40-49 years must be screened to prevent one death from breast cancer), the 95% confidence interval is 1,190 to infinite. The width of these confidence intervals emphasises the uncertainty in our estimates of the effectiveness of mammographic screening in this age group.

Q2. Would it not be better to estimate the effect of screening separately in women aged 40-44 and those aged 45-49?

Doing a subgroup analysis for women in different age groups would be very sensible. Unfortunately, the data available to us from the trials are not sufficiently detailed to allow this analysis. However, we estimate that of the total number of deaths prevented by screening women aged 40-49 years, about two-thirds are prevented by screening women aged 45-49 years and about one third by screening women aged 40-44 years. This estimate is based on both the known increase

in incidence of breast cancer with increasing age and the expected increase in the relative risk reduction achievable by screening mammography during the decade (as shown by the diamonds in Figure 3, page 15 of the report). Cost-effectiveness will therefore also be better in women aged 45-49 years than in women aged 40-44 years.

Q3. The recently published Gothenburg study shows a 44% reduction in mortality from breast cancer after 11 years of follow-up.¹ Does this not change the conclusions of the report?

The NBCC report included similar Gothenburg data which was presented at the Falun meeting.² To clarify exactly what data were used in the NBCC report, we give full details in the Table on the following page. As can be seen from the Table, the Gothenburg data used in the NBCC report were 19 deaths in the study (screening) group and 37 in the control group. The data presented in the formal publication of the Gothenburg data were 18 deaths in the study group and 40 in the control group.¹ These minor differences only alter the meta-analytic estimates of effect minimally: the relative risk changes from 0.85 (95% confidence interval 0.71,1.01) to 0.84 (95% confidence interval 0.71, 1.00) and the number needed to screen to delay one death changes from 2,638 to 2,334.

The effect of screening in the Gothenburg study cannot be regarded as truly larger than all other studies. Tests for heterogeneity show that all the observed differences between study results can easily be explained by chance (chi square test for heterogeneity = 5.20 for our original meta-analysis and 6.39 with the new Gothenburg data, p for both tests >0.2).

Table A:
Data for meta-analysis of the effect of mammographic screening on risk of death from breast cancer, women aged 40-49 years

Study	Study group deaths	Study group	Control group deaths	Control group	Odds ratio
Two County	45	19844	39	15604	0.91
Malmo	15	3795	23	3769	0.65
Stockholm	25	14842	12	7103	1.00
Göteborg	19	10821	37	13101	0.62
Edinburgh	25	11370	31	10269	0.73
HIP	49	14432	65	14701	0.77
NBSS	73	25214	66	25216	1.11
Total	251	100318	273	89763	0.85

Q4. Isn't modern mammography more effective than the mammographic techniques which were used in the randomised trials?

To address this question we examined the likely effect of known improvements in sensitivity due to double reading and two-view mammography. The results of this analysis can be found in the section 'Improvements in mammography' (page 20-21) of the report. We also did a meta-analysis of effectiveness by date of trial commencement which showed no clear trend in either direction (ie improving effect for earlier or later trials). However, it is difficult to interpret this analysis because the effect of newer techniques is confounded by length of follow-up (earlier trials show a greater benefit because they have more years of follow-up data available).

¹ Bjurstam N, Bjorneld L, Duffy SW. The Göteborg Breast Screening Trial. First results on mortality, incidence and mode of detection for women aged 39-49 years at randomization. *Cancer* 1997;80:2091-9.

² Breast-cancer screening with mammography in women aged 40-49 years. Report of the organising committee and collaborators, Falun meeting, Falun, Sweden 1996. *Int J Cancer* 1996;68:693-9.