The benefits and harms of breast cancer screening: an independent review

*Independent UK Panel on Breast Cancer Screening**

Lancet 2012; 380: 1778–86
Published Online
October 30, 2012
http://dx.doi.org/10.1016/S0140-6736(12)61611-0

See **Editorial** page 1714

*Members listed at end of paper

Correspondence to:
Prof Sir Michael Marmot, UCL Department of Epidemiology and Public Health, UCL, London, WC1E 7HB, UK
m.marmot@ucl.ac.uk

Whether breast cancer screening does more harm than good has been debated extensively. The main questions are how large the benefit of screening is in terms of reduced breast cancer mortality and how substantial the harm is in terms of overdiagnosis, which is defined as cancers detected at screening that would not have otherwise become clinically apparent in the woman's lifetime. An independent Panel was convened to reach conclusions about the benefits and harms of breast screening on the basis of a review of published work and oral and written evidence presented by experts in the subject. To provide estimates of the level of benefits and harms, the Panel relied mainly on findings from randomised trials of breast cancer screening that compared women invited to screening with controls not invited, but also reviewed evidence from observational studies. The Panel focused on the UK setting, where women aged 50–70 years are invited to screening every 3 years. In this Review, we provide a summary of the full report on the Panel's findings and conclusions. In a meta-analysis of 11 randomised trials, the relative risk of breast cancer mortality for women invited to screening compared with controls was 0·80 (95% CI 0·73–0·89), which is a relative risk reduction of 20%. The Panel considered the internal biases in the trials and whether these trials, which were done a long time ago, were still relevant; they concluded that 20% was still a reasonable estimate of the relative risk reduction. The more reliable and recent observational studies generally produced larger estimates of benefit, but these studies might be biased. The best estimates of overdiagnosis are from three trials in which women in the control group were not invited to be screened at the end of the active trial period. In a meta-analysis, estimates of the excess incidence were 11% (95% CI 9–12) when expressed as a proportion of cancers diagnosed in the invited group in the long term, and 19% (15–23) when expressed as a proportion of the cancers diagnosed during the active screening period. Results from observational studies support the occurrence of overdiagnosis, but estimates of its magnitude are unreliable. The Panel concludes that screening reduces breast cancer mortality but that some overdiagnosis occurs. Since the estimates provided are from studies with many limitations and whose relevance to present-day screening programmes can be questioned, they have substantial uncertainty and should be regarded only as an approximate guide. If these figures are used directly, for every 10 000 UK women aged 50 years invited to screening for the next 20 years, 43 deaths from breast cancer would be prevented and 129 cases of breast cancer, invasive and non-invasive, would be overdiagnosed; that is one breast cancer death prevented for about every three overdiagnosed cases identified and treated. Of the roughly 307 000 women aged 50–52 years who are invited to begin screening every year, just over 1% would have an overdiagnosed cancer in the next 20 years. Evidence from a focus group organised by Cancer Research UK and attended by some members of the Panel showed that many women feel that accepting the offer of breast screening is worthwhile, which agrees with the results of previous similar studies. Information should be made available in a transparent and objective way to women invited to screening so that they can make informed decisions.

Introduction

After the recommendations made by Professor Sir Patrick Forrest in his report on breast screening in 1986,[1] women have been invited to screening through the NHS Breast Cancer Screening Programme since 1988. Since screening was established in the UK, additional follow-up data have become available from the trials on which the Forrest Report recommendations were based and from other randomised trials. Moreover, many observational studies have assessed existing population screening programmes.

This additional information has stimulated a continuing debate about the potential benefits and harms of population breast screening. The debate has focused on the reduction in mortality attributable to screening, the numbers of women overdiagnosed, and the way that the risks and benefits are communicated to women invited for screening. The arguments have become polarised between those who believe that the benefit of a decrease in mortality outweighs the harms and those who believe the opposite. These contrasting views of the evidence have arisen partly from disagreements about the validity and applicability of the available randomised controlled trials of breast screening, and partly from questions about the usefulness and interpretation of observational data for breast cancer incidence and mortality.

The debate about the benefits and harms of breast screening is not unique to the UK and its breast cancer screening programmes. In 2002, the International Agency for Research on Cancer[2] reviewed the evidence for breast screening and proposed recommendations for further research and the implementation of screening programmes. In 2009, the US Preventive Services Task Force re-examined the efficacy of various screening modalities. They recommended that women younger than 50 years do not need to be screened routinely and women aged 50–74 years should have biennial rather than annual screens.[3] The Canadian Taskforce on Preventative Health Care updated their guidelines for breast screening in 2011, and concluded that the reduction

in mortality associated with screening mammography is small for women aged 40–74 years at average risk of breast cancer.[4] They also noted a greater reduction in mortality for women 50 years and older than for those younger than 50 years, and that the risk of overdiagnosis and biopsy might be greater for younger than for older women. They recommended that women aged 50–74 years be screened routinely, but stated that substantial uncertainty exists around the evidence for this recommendation.[4] The Cochrane Review[5] concluded that, despite their substantial methodological limitations, trials showed that screening reduced breast cancer mortality, but at the cost of substantial harm from overdiagnosis.

As a result of this debate, Professor Sir Mike Richards (National Cancer Director, England) and Dr Harpal Kumar (Chief Executive Officer of Cancer Research UK) asked Professor Sir Michael Marmot to assemble and chair an independent Panel to review the evidence for the benefits and harms of breast screening in the UK. The terms of reference of the Panel and the details of the membership and support for the Panel are available in appendix 1 of the full report.[6] Although the Panel has substantial expertise in epidemiology and medical statistics and in current breast cancer diagnosis and treatment practices, no Panel member has previously published work on breast screening, which helps to ensure an objective and independent assessment of the evidence. A patient advocate was an integral member of the Panel.

The Panel reviewed the extensive published work and listened to testimonies from experts in the specialty who were the main contributors to the debate.[6] The Panel focused on the effects of screening on mortality and overdiagnosis in the context of the UK breast screening programmes, which currently invite women aged 50–70 years for a screening mammography every 3 years. The Panel's full report[6] provides substantial background information and references to the evidence and recommendations summarised in this Review.

The effect of breast screening on mortality

Measurement of benefit

The aim of screening is to advance the time of diagnosis so that prognosis can be improved by earlier intervention. Earlier diagnosis leads to an increase in the apparent incidence of breast cancer and extends the time between diagnosis and death, even if screening does not confer any benefit. The appropriate measure of benefit, therefore, is the reduction in mortality from breast cancer in women offered screening compared with those who are not.

Randomised controlled trials

Randomised controlled trials potentially provide the most reliable information about the effects of breast screening. High-quality randomised controlled trials are prone to fewer distorting effects, or biases, than are observational studies. For this reason, the Panel's quantitative estimate of the benefits of breast screening is based on randomised trials of breast screening. A specific issue raised by some commentators is that most of the randomised trials of breast screening date from the 1980s or earlier. Since

	New York HIP	Malmö I and II	Swedish Two County	Canada I and II	Stockholm	Göteborg	UK Age trial	Edinburgh
Start date	1963	1976	1977	1980	1981	1982	1991	1978
Randomisation method	Individual	Individual	Cluster	Individual	Day of birth	Day of birth*	Individual	Cluster
Population of women								
Source	IC	P	P	Various†	P	P	PC	PC
Number of women‡ (clusters)	62 000	60 076	133 065 (45)	89 835	60 800	52 222	160 921	54 654 (87)
Age group (years)	40–64	45–69 and 43–49	38–75	40–49 and 50–59	39–65	39–59	39–41	45–64
Invited group intervention	M+PE	M	M+SE	M+PE+SE	M	M	M	M+PE
Mammography								
Number of views	2	2 then 1 or 2	1	2	1	2 then 1	2 then 1	2 then 1
Screening interval (months)	12	18–24	24–33	12	24–28	18	12	24
Number of screening rounds	4	6–8	2–4	4–5	2	4–5	8–10	2–4
Duration of screening (years)	3	12	7	5	4	7	8	6
Attendance	65%	74%	85%	88%	82%	84%	81%	65%
Control group intervention	None	None	None	PE+SE§	None	None	None	None
Follow-up								
Controls invited for screening¶	Not known	Never||	After 7 years	Never	After 4 years	After 7 years	After 10 years	After 10 years
Cause of death determination	L	IEC, NS	L, IEC, NS	IEC, NS	IEC, NS	NS	NS	NS

Information is taken from various publications, but mainly the Cochrane Review,[5] Nyström and colleagues,[7] and Tabar and colleagues.[8] These summaries are sometimes simplifications of characteristics that differ between subtrials or subgroups. Some discrepancies also exist between different publications. IC=insurance company register. P=population register. PC=primary care register. M=mammography. PE=physical examination. SE=self-examination. L=local. IEC=independent endpoint committee. NS=national statistics or register. *Day of birth, and later individual. †Includes P, IC, employee recruitment, and general publicity. Women were randomly assigned after initial PE, and evidence suggests that the women attending screening had a higher rate of breast cancer at that initial attendance than was expected from an age-matched population. ‡Some of these numbers are approximate, because the numbers vary in different publications. §After the initial assessment, only the women in Canada II underwent systematic PE during the screening period—in Canada I, they were taught how to do a physical examination. ¶Systematic invitation of all controls. ||Applies to Malmö I ages 55–69 years.

Table 1: **Characteristics of the randomised trials of breast cancer screening**

then, treatment and management of breast cancer have improved, and changes in the incidence of breast cancer and in mammographic techniques have occurred. Are the trials still relevant? Such a question can be asked of any area of medical investigation or treatment, because trials refer to the past and our use of interventions relates to the future, but the question is particularly relevant for breast cancer screening.

11 relevant randomised trials[5] have been undertaken and results reported (New York HIP, Malmö I and II, Swedish Two County [Kopparberg and Östergötland], Canada I and II, Stockholm, Göteborg, UK Age trial, and Edinburgh); the three trials with two parts have sometimes, but not always, been reported separately in publications. All the trials compared women invited to screening against a control group not invited, but varied greatly in their methodology (table 1).

The Panel considered the published concerns about the randomisation process in some trials, but felt that only in the case of the Edinburgh trial were the concerns sufficiently problematic to merit exclusion. All age groups were considered. Although adjudication of the cause of death has been one of the major criticisms of some of the trials, the Panel does not think that it would exaggerate the estimates of relative risk obtained from individual trials, or from a meta-analysis of trials. Little effect on mortality is noted within the first 5 years of screening, and most of the trials started systematic screening of the control group after 4–10 years (table 1), so the Panel believes that a follow-up of about 10–15 years after randomisation provides the most reliable estimate of the relative risk.

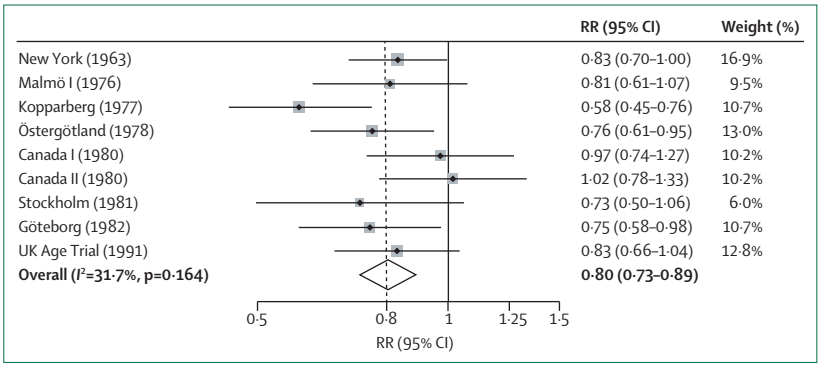

Figure 1: **Meta-analysis of breast cancer mortality after 13 years of follow-up in breast cancer screening trials**
Adapted from the Cochrane Review.[5] RR=relative risk. Malmö II is excluded because follow-up of about 13 years was not available; the Swedish Two County (Kopparberg and Östergötland) and Canada I and II trials are split into their component parts; the Edinburgh trial is excluded because of severe imbalances between randomised groups. Weights are from random-effects analysis.

	Overall RR (95% CI)
This review	
13-year follow-up in trials reported in the Cochrane Review[5] random-effects meta-analysis	0·80 (0·73–0·89)
Cochrane review[5]	
Fixed-effect meta-analysis of the above trials	0·81 (0·74–0·87)
As above, but excluding women <50 years	0·77 (0·69–0·86)
Trials considered adequately randomised (Canada, Malmö, and UK Age trial) had RR 0·90 (95% CI 0·79–1·02); trials deemed suboptimally randomised gave RR 0·75 (0·67–0·83). As a compromise between these two estimates, the authors concluded that an RR of 0·85 was plausible	0·85
US Task Force[9]	
RR 0·86 (95% CI 0·75–0·99) for women aged 50–59 years, and RR 0·68 (0·54– 0·87) for those aged 60–69 years. These estimates have an inverse-variance weighted average RR of 0·81	0·81
Canadian Task Force[4]	
Routinely screening for breast cancer with mammography every 2–3 years for women aged 50–69 years was rated as a weak recommendation based on moderate-quality evidence according to GRADE criteria[11]	0·79 (0·68–0·90)
Duffy et al, 2012[10]	
Review of all trials and age groups	0·79 (0·73–0·86)

RR=relative risk.

Table 2: **Estimates of RR in a comparison of invited women versus control women in the trials of breast cancer screening**

Meta-analysis of relative risk of dying from breast cancer in the trials

The Panel's primary conclusions about breast cancer mortality are based on data reported in the Cochrane Review,[5] which provided results for 13 years of follow-up of the groups as randomised, excluding the Edinburgh trial (analysis 1·2 in that report). The Panel did not distinguish between the trials labelled adequately randomised and suboptimally randomised in the Cochrane Review, but considered the evidence from all the trials. Random-effects meta-analysis, rather than fixed-effect meta-analysis, was used to estimate an average effect across the trials. Use of random effects acknowledges that the trials could estimate different quantities, which is likely because of their clinical heterogeneity, whereas a fixed-effect analysis estimates an assumed common effect across all the trials. Figure 1 shows the results, with the relative risks of breast cancer mortality. The overall relative risk, comparing invited versus control women, is 0·80 (95% CI 0·73–0·89). There was some heterogeneity in the relative risks from different trials, but it was not statistically significant (figure 1). Thus, the relative risk reduction in breast cancer mortality in the groups invited to screening is estimated to be 20% (95% CI 11–27).

Other meta-analyses of the breast cancer screening trials have provided different estimates of the relative risk (table 2).[5,9,10,12,13] Different meta-analyses are based on different trials, durations of follow-up, definitions of outcome, and methods of analysis. Nevertheless, there is general agreement of about a 20% relative risk reduction in breast cancer mortality resulting from invitation to screening. The Panel noted that all-cause mortality is not an appropriate outcome for trials of breast screening because the trials were not designed with sufficient power for this outcome.

Estimation of relative risk from observational studies

The randomised controlled trials were all undertaken at least 20–30 years ago. More contemporary estimates

of the benefit of breast cancer screening are from observational studies.

The Panel reviewed three types of observational studies. The first were ecological studies comparing areas or periods when screening programmes were and were not in place. These studies have generated remarkably diverse findings. Major advances in the treatment of breast cancer, which have the largest effect on mortality trends, outweigh any smaller effect of screening. Therefore, the Panel did not consider these studies to be helpful in estimation of the effect of screening on mortality, both because of the changes over time in the use of more effective treatments and because of the difficulty in exclusion of imbalances in other factors that could affect breast cancer mortality. The other two types of study, case-control studies and incidence-based mortality studies, showed breast screening to confer a greater benefit than did the trials. Although these studies generally attempted to control for non-comparability of screened and unscreened women, the Panel was concerned that residual bias could inflate the estimate of benefit of screening. However, the Panel notes that the findings of these studies are in the same direction as those of the trials.

Conclusion about relative risk reduction

In the Panel's opinion, the best evidence for the relative benefit of screening on mortality reduction comes from randomised controlled trials of breast screening. Meta-analysis showed a 20% reduction in mortality in women invited for screening. However, three types of uncertainty surround this estimate. The first is statistical: the 95% CI around the relative risk of 0·80 was 0·73–0·89. The second is bias. During the decades since these trials were reported, there has been much discussion of their internal validity. The third is the relevance of these old trials to current screening programmes. Do the improvements in treatment for breast cancer mean that screening is no longer relevant? In the absence of data to the contrary, the Panel concluded that the benefits of screening and those of better treatments are likely to be independent, and thus that the estimates of the relative reduction in breast cancer mortality achieved with screening are similar now to when the trials were undertaken. The Panel did not feel able to quantify the additional uncertainty around its estimate of benefit, but notes that other views of the randomised controlled trials have yielded similar estimates of benefit (table 2).

Absolute benefit

Estimates of the absolute benefit of screening have varied from one breast cancer death prevented for about 100 women screened to one such death prevented for 2000 women invited to screening (table 3). The major determinants of this large variation are the age of the women screened, the duration of screening, and the length of follow-up.[13] The age of the women invited is important, since mortality from breast cancer increases substantially with age. Rather than use data for absolute risk reduction directly from the trials, which have limited follow-up and shorter active screening periods than the UK programmes, the Panel applied a relative reduction of 20% to achieve the observed cumulative absolute risk of breast cancer mortality within the ages 55–79 years in the UK. This approach assumes that women who are first invited to screening at age 50 years, and continued to be invited for 20 years (as in the UK), would gain no benefit in the first 5 years, but that the mortality reduction would continue for 10 years after screening ended. This assumption yielded the estimate that for every 235 women invited to screening, one breast cancer death would be prevented, representing 43 breast cancer deaths prevented per 10 000 women invited to screening. The absolute benefit for women actually attending screening would be higher: 180 women would need to be screened to prevent one breast cancer death. These estimates carry the same uncertainties as mentioned previously.

	Description	Number of women
This review	Based on an RR reduction of 20% for women aged 55–79 years in the UK	235 women invited, 180 women screened
Cochrane review[5]	Absolute risk reduction based on the 13-year follow-up in the trials considered adequately randomised	2000 women invited
US Task Force[9]	Based on 7 years of screening and 13 years of follow-up	1339 women invited aged 50–59 years, and 377 invited aged 60–69 years
Canadian Task Force[4]	Women aged 50–69 years screened every 2–3 years for about 11 years	720 women screened
Duffy et al, 2010[12]	Based on 22-year follow-up of women aged 50–69 years in the Swedish Two-County trial, assuming that the absolute risk reduction for the 7 years of screening can be multiplied up to reflect 20 years in the UK screening programmes	113 women screened
Beral et al, 2011[13]	Women aged 50–70 years regularly screened for 10 years, based on summary of published evidence	400 women screened

RR=relative risk.

Table 3: **Absolute risk reduction, expressed as number of women who need to be invited or screened to prevent one breast cancer death, in the trials of breast cancer screening**

Overdiagnosis

The major harm of screening considered by the Panel was that of overdiagnosis. Because cancers are detected earlier with screening, the cancer incidence is expected to be higher in screened women during the screening period. The period between detection of a cancer at screening and when it would have presented clinically is the lead time and is an inevitable part of screening. In theory, when screening stops, the cancer incidence should fall, so that by the end of the screening period plus lead time, the cumulative incidence in the screened and control populations should be the same.[14]

Some screen-detected cancers, however, might never have progressed to become symptomatic in the absence of screening, and some women would die from another cause before the cancer became evident. These cancers are nonetheless treated. This adverse consequence (harm) of screening is called overdiagnosis or overdetection, and is defined as the "detection of cancers that would never have been found were it not for the screening test".[2] It refers to all cancers, invasive or in situ, because both are actively treated.

Whether a particular woman has had an overdiagnosed cancer cannot be judged. The best estimate of the extent of overdiagnosis relevant to the UK screening programmes would come from a comparison of the number of cancers in women screened for 20 years with that in an unscreened comparable population—similar in terms of age, exposure to breast cancer risk factors, and availability of treatment—and followed up until death. An excess of cancers in the screened group would represent the amount of overdiagnosis. Unfortunately, such a study does not exist, and therefore estimations of overdiagnosis need some indirect inference from available studies.

Estimation of overdiagnosis from randomised trials

Overdiagnosis can be estimated from randomised controlled trials or observational studies. Valid estimates depend on similar underlying breast cancer risks in screened and unscreened women and the fact that the effect of lead time has been accounted for.[15] Randomised controlled trials have the huge advantage that, by design, they compare groups of women with the same average prognosis, although the screening phase of these trials varied in length and was always shorter than that used in the UK national screening programmes.

Because screening advances detection of breast cancer, follow-up should continue after the screening period to allow catch-up of diagnoses in the unscreened group. In principle, the extended period of follow-up should correspond to the lead time, but the average lead time is debated and is not equal for all cancers. As the follow-up is extended to after the screening period, many new cancers in both the screened and unscreened groups will be included irrespective of screening, and the ratio of total numbers of diagnosed cancers will converge towards 1.[15] An adequate follow-up is a minimum of 5–10 years after the end of the intervention period.[16,17] Although the definition of an overdiagnosed case, and thus the numerator in a ratio, is clear, the choice of denominator has been the source of further variability in reported estimates. For example, de Gelder and colleagues[18] described seven different approaches.

The most reliable estimates of overdiagnosis come from three randomised controlled trials, in which women in the control group were not offered screening at the end of the trial (in the other randomised controlled trials, all the women in the control group were offered screening at the end of the active period of the trial, which itself might be expected to lead to some overdiagnosis and thus to an overall underestimate of overdiagnosis). These three trials are Malmö I,[19] for women aged 55–69 years, and the two Canadian trials,[20,21] which screened women for 5 years and reported follow-up data at 11 years—ie, about 6 years after the end of screening. Table 4 shows estimates of overdiagnosis from these three trials using four methods. The estimates from the three randomised controlled trials are similar, but there are substantial differences according to the denominator used. The HIP study was excluded from the estimation of overdiagnosis because the Panel had great difficulty obtaining a consistent set of figures for incidence on which to base any such assessment.[22] Moreover, in the first report from the HIP trial, lobular carcinoma in situ cases are included, but would now not be, and whether the subsequent, longer-term cancer data in the post-trial period did or did not include lobular carcinoma in situ (or ductal carcinoma in situ) is not clear.[22]

The Panel believes that there is no single optimum way to estimate overdiagnosis but that the two most useful estimates are: from the population perspective, the proportion of all cancers ever diagnosed in women invited to screening that are overdiagnosed (method B in table 4); and from the perspective of a woman invited to screening, the probability that a cancer diagnosed during the screening period represents overdiagnosis.[23,24] Because the number of interval cancers depends on screening frequency,[25] interval cancers should be included in the denominator (method C in table 4).

	A	B	C	D
Malmö I ages 55–69 years	11·7% (82/698)	10·5% (82/780)	18·7% (82/438)	29·1% (82/282)
Canada I	14·1% (82/581)	12·4% (82/663)	22·7% (82/361)	29·4% (82/279)
Canada II	10·7% (67/626)	9·7% (67/693)	16·0% (67/420)	19·8% (67/338)

Numbers of excess cancers are expressed as a percentage of different denominators. A=excess cancers as a proportion of cancers diagnosed over whole follow-up period in unscreened women. B=excess cancers as a proportion of cancers diagnosed over whole follow-up period in women invited for screening. C=excess cancers as a proportion of cancers diagnosed during screening period in women invited for screening. D=excess cancers as a proportion of cancers detected by screening in women invited for screening.

Table 4: **Estimates of overdiagnosis in randomised trials without systematic end-of-trial screening of the control group, according to four calculation methods**

Figure 2 shows meta-analyses of these overdiagnosis proportions from the three trials. The frequency of overdiagnosis was roughly 11% from a population perspective and about 19% from the perspective of a woman invited to screening. These estimates are subject to the same sources of uncertainty as noted for the estimates of mortality from the randomised controlled trials. Estimates of overdiagnosis have additional uncertainties about which estimate to use, and the data are not available for all studies to calculate overdiagnosis with the suggested methods. Additionally, the estimates are not tailored to the UK screening programme or a 20-year screening period.

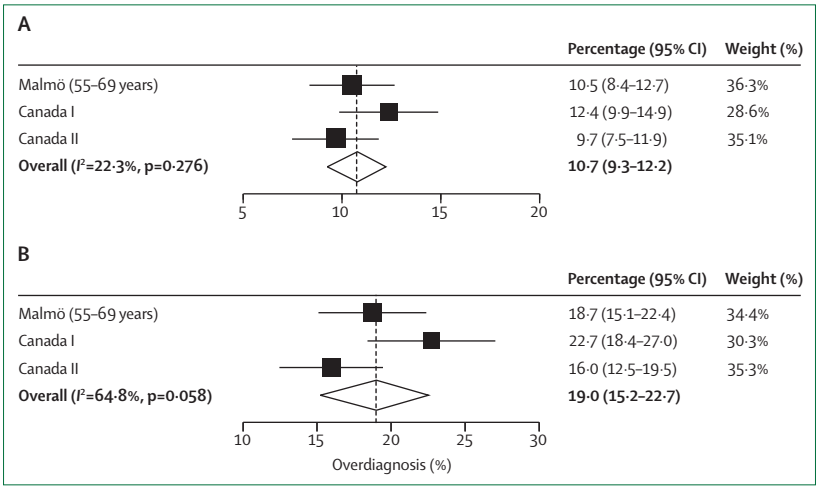

Figure 2: **Meta-analysis of estimates of overdiagnosis from trials without systematic end-of-study screening of the control group**

(A) Excess cancers as a proportion of cancers diagnosed over long-term follow-up in women invited for screening. (B) Excess cancers as a proportion of cancers diagnosed during the screening period in women invited for screening. Weights are from random-effects analysis.

Estimation of overdiagnosis from observational studies

Many studies have examined the effects of screening in populations and assessed the degree of overdiagnosis. Breast cancer rates fluctuate over time in populations and the large differences in estimates of overdiagnosis reflect this variation, as well as different lengths of follow-up, statistical assumptions, and ways to account for lead time.

Some studies have compared the post-screening incidence of breast cancer with a projection of previous incidence trends in the screened population, and have resulted in very different estimates of overdiagnosis. The Panel concluded that by changing the assumptions underlying these studies, a vast range of estimates of overdiagnosis could be obtained. Since there seems to be no a priori reason to favour one set of assumptions over another, the Panel do not think that approaches based on extrapolation offer a robust method to estimate overdiagnosis.

Several groups have compared breast cancer incidence trends over time in screened and unscreened countries or regions within the same time period.[26] The estimates of risk of overdiagnosis from these studies vary greatly. The investigators of studies in Denmark and Norway have produced estimates of overdiagnosis that are greater than those that the Panel calculated from the randomised controlled trials. Key issues are comparability of populations and distinction of lead time from overdiagnosis. The investigators of one review concluded that, taking these effects into account, overdiagnosis risks were lower than those from the randomised controlled trials.[14]

Conclusion

The Panel believes that overdiagnosis occurs. The consequence of overdiagnosis is that women have their cancer treated by surgery, and in many cases radiotherapy and medication, but neither the woman nor her doctor can know whether this particular cancer would be one that would have become apparent without screening and could possibly lead to death, or one that would have remained undetected for the rest of the woman's life. The Panel believes that data from three of the randomised controlled trials without end-of-trial screening of controls provide the most reliable estimates of the extent of overdiagnosis, but notes that there is a rather small amount of data, and numerical estimates are subject to several uncertainties in common with estimates of mortality benefit. Results of observational studies support the existence of overdiagnosis, but estimates of its magnitude are unreliable. The overdiagnosis risks estimated from old randomised controlled trials might not reflect those in current screening programmes. However, no clear evidence exists to suggest that the current risk of overdiagnosis would be lower or higher than in the original trials.

Although longer follow-up is needed, the Panel thinks that the best estimate of overdiagnosis for a population invited to be screened is roughly 11%, defined as the excess incidence in the screening population as a proportion of the long-term expected incidence. An alternative definition addresses the answer to the question "if I am invited to enter into the screening programme and am given a cancer diagnosis during the screening period, what is the likelihood of overdiagnosis?" The Panel believes the evidence suggests that this probability is about 19%. Application of this risk of 19% overdiagnosis to the present cumulative incidence of cancer (invasive and in situ) in women aged 50–69 years in the UK suggests that one in 77 women aged 50 years invited to screening for 20 years will have an overdiagnosed cancer; this is a rate of 129 per 10 000 women invited to screening.

Ductal carcinoma in situ

Ductal carcinoma in situ is a malignant tumour that arises from the epithelial tissues of the breast and consists of neoplastic cells. However, these cells do not infiltrate beyond the limiting basement membrane and therefore remain within the ducts where they arose. Ductal carcinoma in situ is usually grouped by grade into high, intermediate, or low grade. Although the cells have the appearance of malignancy,

they do not show invasiveness, so carcinoma in situ is not in itself a life-threatening disease.

The importance of ductal carcinoma in situ in breast screening is that it occurs more frequently in screen-detected than in symptomatic cancers (about 20% *vs* 5%). Although ductal carcinoma in situ can be associated with invasive cancer, and therefore can be a marker of malignancy, it can also relapse. For example, in the UK, Australia, and New Zealand trial,[27] after wide local excision of screen-detected ductal carcinoma in situ, without any further treatment, relapse in the breast occurred in about 19% of cases, and was invasive in half of these cases. This frequency of relapse in ductal carcinoma in situ that has been treated shows that although it may contribute to overdiagnosis, it is wrong to assume that all ductal carcinoma in situ represents overdiagnosis.

The relevant question is therefore not whether ductal carcinoma in situ progresses to invasive cancer (which it can) but whether it might progress to an invasive cancer that causes symptoms within the lifetime of the woman concerned. Progression will depend mainly on the age of the woman, her life expectancy at the point of diagnosis, and perhaps other factors, such as hormonal exposure and obesity. Studies do not show any significant effect of ductal carcinoma in situ on survival, even at 20 years of follow-up, but the increasing survival rate might mean that, for women in their 50s and even 60s, a diagnosis of ductal carcinoma in situ could affect their longer-term survival.[28]

Thus, in the diagnosis of ductal carcinoma in situ with a screening programme, a balance has to be struck between the potential benefits for some women of the identification and treatment of a precancerous lesion and the risks for others of the treatment of something that would never have affected the woman in her lifetime.

Other harms

The Panel considered other harms but in less detail, not because they are unimportant, but because there is more agreement about their nature and magnitude.

Some women have pain with mammography, which is sometimes enough to deter them from further attendances. About 4% of women attending for screening are recalled for repeat mammography and possible biopsy.[29] Of these women, nearly one in five will have cancer; of the remainder, nearly 70% will need only further imaging and 30% a biopsy (under local anaesthetic in >90% of cases). These procedures can cause psychological distress.[30]

Most patients (99%) diagnosed with a screen-detected breast cancer will undergo surgery, and about 70% radiotherapy, 70% adjuvant endocrine therapy, and 25% adjuvant chemotherapy. All these procedures are associated with well recognised morbidities but a low mortality rate (<0·15% of women treated). Additionally, the adverse psychological results of a breast cancer diagnosis and subsequent treatment have been well documented.[31]

For women who enter the screening programme, these harms are mostly foreseeable and quantifiable. The main concern is for women whose cancer is overdiagnosed (although such cases cannot be individually identified), for whom the morbidities are not offset by any potential gain from reduced mortality.

Conclusions

Breast screening extends lives. The Panel's review of the evidence on benefit suggests a 20% reduction in mortality in women invited to participate in a 20-year screening programme. A lot of uncertainty surrounds this estimate, but it represents the Panel's overview of the evidence. This reduction corresponds to one breast cancer death prevented for every 235 women invited to screening, and one death averted for every 180 women who attend screening. The breast screening programmes in the UK, which invite women aged 50–70 years to screening every 3 years, probably prevent about 1300 breast cancer deaths every year—equivalent to about 22 000 life-years saved.

However, there is a cost to women's wellbeing; mammographic screening detects cancers that would not have come to clinical attention in the woman's lifetime were it not for screening (overdiagnosis). The Panel estimated that in women invited to screening, about 11% of the cancers diagnosed in their lifetime constitute overdiagnosis, and about 19% of the cancers diagnosed during the period that women are actually in the screening programme. However, the Panel emphasises that these figures are best estimates from inadequate data. The Panel considers any excess mortality that arises from investigation and treatment of breast cancer to be minimal and substantially outweighed by the benefits of treatment.

When these data are combined to assess benefit and overdiagnosis, the Panel estimates that for 10 000 UK women invited to screening from the age of 50 years for 20 years, about 681 cancers will be discovered, of which 129 will represent overdiagnosis, and 43 deaths from breast cancer will be prevented. Therefore, for every breast cancer death prevented, about three overdiagnosed cases will be identified and treated. Of the approximately 307 000 women aged 50–52 years who are invited to screening every year, just over 1% would have an overdiagnosed cancer during the next 20 years. In view of the uncertainties that surround the estimates, the figures cited give a false impression of accuracy.

Policy recommendations

The Panel concludes that the UK breast screening programmes confer significant benefit and should continue. For each woman, the choice is clear. On the positive side, screening confers a reduction in the risk of mortality from breast cancer because of early detection and treatment. On the negative side, is the knowledge that she has perhaps a 1% chance of having a cancer diagnosed and treated that would never have caused problems if she had not been screened.

Clear communication of these harms and benefits to women is essential and is the core of how a modern

health system should function. A body of knowledge exists about how women want information to be presented, which should affect how information is presented to the public.

Research recommendations

In view of the uncertainties in the published results of the randomised trials, the Panel supports the ongoing meta-analysis of centrally collated individual patient data from all the randomised trials. The Panel also supports the use of randomised trials to investigate the balance of benefit to harm of breast cancer screening to women younger than 50 years and those older than 70 years. In view of the problems with estimation of the degree of overdiagnosis from existing published work, the Panel encourages concerted attempts to provide more reliable and up-to-date estimates. Further research is needed to improve the precision of screening to better distinguish between breast cancers that will or will not cause harm during a woman's lifetime. Randomised trials that elucidate the appropriate treatment of screen-detected ductal carcinoma in situ of different grades are also encouraged, as is the Sloane Project to gain a better understanding of the natural history of ductal carcinoma in situ.[32] The Panel also believes that the overall cost-effectiveness of the UK breast cancer screening programmes needs to be reassessed in view of this report.

Contributors

All members of the Panel contributed substantially to the full report and to the drafting of this Review and this article.

Independent UK Panel on Breast Cancer Screening

UK M G Marmot (Chair, UCL Department of Epidemiology and Public Health, UCL, London), D G Altman (Centre for Statistics in Medicine, University of Oxford), D A Cameron (Edinburgh Cancer Research Centre, University of Edinburgh), J A Dewar (Department of Surgery and Oncology, Ninewells Medical School, Dundee), S G Thompson (Department of Public Health and Primary Care, University of Cambridge), Maggie Wilcox (lay member).

Secretariat

Martine Bomb, Samantha Harrison, Sara Hiom, Nick Ormiston-Smith (Cancer Research UK); Dulcie McBride (Department of Health).

Conflicts of interest

All members of the Panel declare that they have no conflicts of interest, and have not previously worked or published on the topic of breast cancer screening. The Secretariat provided administrative and editorial support to the Panel, but played no part either in directing its scientific work or in the conclusions drawn.

Acknowledgments

The Review was commissioned by Professor Sir Mike Richards (National Cancer Director, England) and Dr Harpal Kumar (Chief Executive, Cancer Research UK). The Department of Health provided funding to the Independent Breast Screening Review, which supported the additional resources needed to undertake the Review. The Panel is very grateful to the experts, listed in the full report, who provided verbal or written evidence, or both, for the Review.

References

1 Forrest APM. Breast cancer screening: report to the Health Ministers of England, Wales, Scotland and Northern Ireland by a working group. London: HM Stationery Office, 1987.

2 International Agency for Research on Cancer. IARC handbooks of cancer prevention (volume 7): breast cancer screening. Lyon: IARC, 2002.

3 Woolf SH. The 2009 Breast Cancer Screening Recommendations of the US Preventive Services Task Force. *JAMA* 2010; **303:** 162–63.

4 Canadian Task Force on Preventive Health Care. Recommendations on screening for breast cancer in average-risk women aged 40–74 years. *CMAJ* 2011; **183:** 1991–2001.

5 Gøtzsche PC, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2011; **1:** CD001877.

6 Cancer Research UK. Breast Screening Review. www.cruk.org.uk/breastscreeningreview (accessed Oct 30, 2012).

7 Nyström L, Andersson I, Bjurstam N, Frisell J, Nordenskjold B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet* 2002; **359:** 909–19.

8 Tabar L, Vitak B, Chen TH, et al. Swedish two-county trial: impact of mammographic screening on breast cancer mortality during 3 decades. *Radiology* 2011; **260:** 658–63.

9 Nelson HD, Tyne K, Naik A, Bougatsos C, Chan BK, Humphrey L. Screening for breast cancer: an update for the U.S. Preventive Services Task Force. *Ann Intern Med* 2009; **151:** 727–37, W237–42.

10 Duffy SW, Ming-Fang Yen A, Hsiu-Hsi Chen T, et al. Long-term benefits of breast screening. *Breast Cancer Management* 2012; **1:** 31–38.

11 Schünemann H, Brozek J, Oxman A, eds. GRADE handbook for grading the quality of evidence and the strength of recommendations. The GRADE Working Group, 2009.

12 Duffy SW, Tabar L, Olsen AH, et al. Absolute numbers of lives saved and overdiagnosis in breast cancer screening, from a randomized trial and from the Breast Screening Programme in England. *J Med Screen* 2010; **17:** 25–30.

13 Beral V, Alexander M, Duffy S, et al. The number of women who would need to be screened regularly by mammography to prevent one death from breast cancer. *J Med Screen* 2011; **18:** 210–12.

14 Welch HG. How much overdiagnosis? 10 March 2006. http://www.bmj.com/rapid-response/2011/10/31/how-much-overdiagnosis (accessed 14 Oct, 2012).

15 Puliti D, Duffy S, Miccinesi G, et al. The EUROSCREEN Working Group. Methodology and estimate of overdiagnosis in breast cancer service screening: a review of the European studies. *J Med Screen* 2012; **19** (suppl 1)**:** 42–56.

16 Biesheuvel C, Barratt A, Howard K, Houssami N, Irwig L. Effects of study methods and biases on estimates of invasive breast cancer overdetection with mammography screening: a systematic review. *Lancet Oncol* 2007; **8:** 1129–38.

17 Puliti D, Miccinesi G, Paci E. Overdiagnosis in breast cancer: design and methods of estimation in observational studies. *Prev Med* 2011; **53:** 131–33.

18 de Gelder R, Heijnsdijk EA, van Ravesteyn NT, Fracheboud J, Draisma G, de Koning HJ. Interpreting overdiagnosis estimates in population-based mammography screening. *Epidemiol Rev* 2011; **33:** 111–21.

19 Zackrisson S, Andersson I, Janzon L, Manjer J, Garne JP. Rate of over-diagnosis of breast cancer 15 years after end of Malmo mammographic screening trial: follow-up study. *BMJ* 2006; **332:** 689–92.

20 Miller AB, To T, Baines CJ, Wall C. Canadian National Breast Screening Study-2: 13-year results of a randomized trial in women aged 50–59 years. *J Natl Cancer Inst* 2000; **92:** 1490–99.

21 Miller AB, To T, Baines CJ, Wall C. The Canadian National Breast Screening Study-1: breast cancer mortality after 11 to 16 years of follow-up. A randomized screening trial of mammography in women age 40 to 49 years. *Ann Intern Med* 2002; **137** (5 part 1)**:** 305–12.

22 Shapiro S. Evidence on screening for breast cancer from a randomized trial. *Cancer* 1977; **39:** 2777–82.

23 Welch HG, Schwartz LM, Woloshin S. Ramifications of screening for breast cancer: 1 in 4 cancers detected by mammography are pseudocancers. *BMJ* 2006; **332:** 727.

24 Harris R, Yeatts J, Kinsinger L. Breast cancer screening for women ages 50 to 69 years a systematic review of observational evidence. *Prev Med* 2011; **53:** 108–14.

25 The Breast Screening Frequency Trial Group. The frequency of breast cancer screening: results from the UKCCCR Randomised Trial. *Eur J Cancer* 2002; **38:** 1458–64.

26 Jorgensen KJ, Gotzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. *BMJ* 2009; **339:** b2587.

27 Cuzick J, Sestak I, Pinder SE, et al. Effect of tamoxifen and radiotherapy in women with locally excised ductal carcinoma in situ: long-term results from the UK/ANZ DCIS trial. *Lancet Oncol* 2011; **12:** 21–29.

28 National Health Service Breast Screening Programme and Association of Breast Surgery. An audit of screen detected breast cancers for the year of screening April 2012 to March 2011. 2012. http://www.cancerscreening.nhs.uk/breastscreen/publications/baso2010-2011.pdf (accessed Sept 13, 2012).

29 NHS Breast Screening Programme. Annual Review 2011. http://www.cancerscreening.nhs.uk/breastscreen/publications/nhsbsp-annualreview2011.pdf (accessed Sept 13, 2012).

30 Bond M, Pavey T, Welch K, et al. Psychological consequences of false-positive screening mammograms in the UK. *Evid Based Med* 2012; published online Aug 2. DOI:10.1136/eb-2012-100608.

31 Fallowfield LJ, Hall A. Psychosocial and sexual impact of diagnosis and treatment of breast cancer. *Br Med Bull* 1991, **47:** 388–99.

32 The Sloane Project. UK prospective audit of screen detected non-invasive carcinomas and atypical hyperplasias of the breast: Progress Report 2008/09 and 2009/10. http://www.sloaneproject.co.uk/documents/Sloane%20Annual%20Report%202010%20web.pdf (accessed Oct 12, 2012).