The benefits and harms of breast cancer screening

The Independent UK Panel on Breast Cancer Screening (Nov 17, p 1778) estimates that one woman will avoid dying from breast cancer for every three women who are overdiagnosed. This estimate is far too positive, and the panel did not pay attention to important issues in the testimony we submitted to it in February, 2012.

The panel used the data from our Cochrane review, but did not consider it important that some trials are more reliable than others, and estimated a 20% effect of screening. On the basis of assumptions that this effect would be the same today and would exist undiminished up to age 79 years, 10 years after screening stopped, the panel estimated that screening prevents about 1300 breast cancer deaths every year in the UK.

These are serious errors. Contrary to the opinion of the panel, the important advances in treatment that have occurred since the trials were done will reduce the effect of screening substantially. A woman who would have died without screening in the past might now live so much longer, because of better treatment, that she dies of a heart attack. Screening can have no effect for such women. Breast cancer awareness has also reduced the effect of screening.

The panel does not think that adjudication of the cause of death was a problem in the trials. We have documented at length, in our Cochrane review and in our report to the panel, that this is a huge problem, which inevitably biases the trials in favour of screening even when blinded endpoint committees have been used.

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. Whether an outcome is appropriate or not has nothing to do with power. What matters is whether the outcome is reliable, and since mortality from breast cancer is not reliable, we need to look at other mortality outcomes. Screening did not reduce total mortality, nor mortality from cancer, including breast cancer (relative risk 1.02, 95% CI 0.95–1.01). It is also important to be aware that some of the healthy overdiagnosed women will die from their treatment. For example, radiotherapy increases deaths from heart disease by 27%.

The panel based their 19% estimate of overdiagnosis on trials that ran for only 7–9 years, although screening is offered for 20 years. The estimate of overdiagnosis in the Cochrane review was 29%, and observational studies have found 33% overdiagnosis in Denmark (which has an ideal control group because 80% of the country was not screened for 17 years) and 52% in a systematic review of countries with organised screening programmes.

Is it acceptable that a public health initiative each year converts thousands of healthy women into cancer patients unnecessarily, which is fatal for some of them?

We declare that we have no conflicts of interest.

Peter C Gøtzsche, Karsten Juhl Jørgensen_npc@cochrane.dk
Nordic Cochrane Centre, Righospitalet, 2100 Copenhagen, Denmark


The Independent UK Panel on Breast Cancer Screening concludes that information should be made available transparently to women invited to screening, to allow informed decisions.

The conclusion: “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 overdiagnosed” is midway between the estimates of the Cochrane review (ten cases of overdiagnosis for each breast cancer death avoided) and of the Euroscreen Working Group (seven to nine lives saved for every four cases overdiagnosed).

But to say that “screening extends lives”, with about 22 000 life-years saved in the UK, is not evidence-based. For “one life saved”, women read “one death avoided”, not “one breast cancer death avoided”, only to die from some other cancer or from something else, without any undisputed survival advantage. Indeed, data show that trials with adequate randomisation did not find an effect of screening on cancer mortality, including breast cancer, after 10 years: relative risk 1.02 (95% CI 0.95–1.1); the same was true for the trials suboptimally randomised: 0.99 (95% CI 0.93–1.06). Moreover, the trials with adequate randomisation did not find an effect on all-cause mortality after 13 years (0.99, 0.95–1.03); neither did the trials suboptimally randomised (0.99, 0.97–1.01).

This finding might be due to insufficient power to assess all-cause mortality, but nobody can guarantee lives saved or extended because all-cause mortality at present does not differ, and because the net cumulative effect of repeated mammographies, and unnecessary surgery, radiotherapy, adjuvant endocrine therapy, and chemotherapy cannot be assumed to be harmless.

I declare that I have no conflicts of interest.

Alberto Donzelli
adonzelli@asl.milano.it
The report of the Independent UK Panel on Breast Cancer Screening is undermined by contradictions and omissions. From 1989 until 2008, breast cancer mortality declined by about 30% in the UK female population aged 50 years and older. The panel skipped over observational studies on mortality trends, done according to methods recommended by the International Agency for Research on Cancer handbook on breast cancer screening, because “major advances in the treatment of breast cancer, which have the largest effect on mortality trends, outweigh any smaller effect of screening”. However, the panel considered that, ultimately, the 20% reduction in breast cancer mortality found in randomised trials should be found back in the UK. We do not know how alleged benefits of screening and treatment can be reconciled other than by concluding that screening has been much less effective than suggested by randomised trials.

The panel did not allude to data showing no reduction in the incidence of advanced breast cancer in the UK (and elsewhere) after screening introduction. Because screening is about detecting cancer when at an early stage before it has evolved into an advanced, less curable stage, incidence rates of advanced cancer should decrease in populations where screening has been widespread for a long time. Monitoring of advanced cancer is independent of the effects of treatments. We would appreciate understanding why the panel did not pay attention to these data.

We declare that we have no conflicts of interest.

*Philippe Autier, Mathieu Boniol, Peter Boyle philippe.autier@i-pri.org
International Prevention Research Institute (IPRI), 69006 Lyon, France

In their review, the Independent UK Panel on Breast Cancer Screening calculate a 3:1 ratio of 129 over-diagnosed cases to 43 breast cancer deaths prevented. However, the 43 is an underestimate by a factor of two.

The panel correctly recognise that mortality reductions are delayed—if women aged 50 years were invited to screening for the next 20 years, reductions would start at about age 55 years (5 years after the first screen) and end at 79 years (10 years after the last screen). However, the panel’s meta-analysis, like the others cited, ignored this delay. It included years in which little effect would be expected: the years immediately after the first screen, and the years well after the last screen. In the appropriate age window, the sustained reduction is at least 40%, not the 20% the panel obtained.

This common underestimation becomes evident when the yearly mortality data are examined. Thus, for every 10,000 women invited, 86 (not 43) deaths from breast cancer would be prevented: a ratio of 1:5:1, not 3:1. We declare that we have no conflicts of interest.

*James A Hanley, Zhihui Liu, Maurice McGregor james.hanley@mcgill.ca
McGill University, Montreal, QC H3A 1A2, Canada

The publication by the Independent UK Panel on Breast Cancer Screening is a milestone in the breast cancer screening debate. We hope it will start a new era of dispassionate scientific discussion about outcome research in screening. There are two major issues for future discussion: the facts and the communication of those facts.

The panel estimated a 20% reduction in breast cancer mortality for invited women from the results of randomised controlled trials (RCT). As part of the Euroscreen project, we reviewed European incidence-based mortality studies and case-control studies. Our meta-analyses showed a 25–31% reduction in mortality for women invited, and a 38–48% reduction for women screened, taking self-selection bias into consideration. Thus, particularly in terms of intention to treat, there is some consistency between the summary estimates of the RCTs and the European observational studies.

The Euroscreen overview further found two constellations of estimates for overdiagnosis. The first consisted of studies taking adequate account of trends in underlying risk and lead time. These estimated overdiagnosis as lower than 10% (average 6·5% in our estimate for screened women). The second was made up of studies not adequately adjusting for these complexities. This group obtained
higher estimates, with a maximum of 54%. Given the limitations and methodological problems of the studies, we concluded that the second constellation was systematically overestimating overdiagnosis.

We have reservations about the methods adopted by the panel regarding overdiagnosis, although we agree that this an area of uncertainty. To limit the excess of cases to breast cancers diagnosed in the study period is likely to have inflated the overdiagnosis estimates: first, because it is not clear that the invited women in the Malmö trial all stopped being screened at the age specified in the protocol; and second, because the Canada trial excesses were noted relatively soon after the screening period. Consequently, excesses in the Canada trial might still have been affected by lead time. Additionally, these trials were dominated by the prevalence screen, at which overdiagnosis is greater than for the incidence screen, at which these trials were dominated by a Canada trial might still have been affected by lead time. Consequently, excesses in the Canada trial might still have been affected by lead time. Additionally, these trials were dominated by the prevalence screen, at which overdiagnosis is greater than for the incidence screen. This problem would not apply to a screening programme, in which a participant is expected to have one prevalence screen and seven to nine incidence screens between age 50 and 70 years. Finally, the treatment of the overdiagnosed proportions as binomial, with the denominator of total number of cancers in invited or screened women, gives a false impression of precision of the estimates of overdiagnosis. For example, the excess incidence in the Malmö trial study group was not significant, whereas the CIs in the panel’s review suggest that it was.

Given these facts, with the strengths and weaknesses of the data available, there are limitations when we use the data for communication, which calls for absolute numbers. In balance sheet estimates, many assumptions are needed and the reader should be clearly informed of the ingredients used. It is easy to overlook the series of steps taken, and assumptions required, to arrive at the estimates of benefit and harm.

On the basis of two 30–35-year-old randomised studies, the Independent UK Panel on Breast Cancer Screening concludes that there is 19% over-diagnosis when screening with mammography. This estimate is diluted and biased because of extensive screening in the control groups.

Furthermore, detection rates and the level of overdiagnosis have increased 100% or more as the sensitivity of the technique has improved (ultrasound, double view, computer-assisted reading, and MRI). When screening was introduced in Sweden (1986–89) and in Norway (1996–97), we noted a 50% increase in invasive breast cancer.

Additionally, the increase in ductal carcinoma in situ was much larger than in the old trials. The total increase in diagnosis in Norway was 75%, and when women were no longer invited to screening, there was no incidence decline (figure). If only half of this 75% increase is overdiagnosis (ie, 37.5% overdiagnosis), then there should be no cancer after age 69 years. When comparing age-matched cohorts in Sweden and in Norway, we found that almost all of the incidence increase when screening was due to detection of lesions that normally would go into spontaneous regression.

Before screening started in Sweden, there was no incidence increase, and Norwegian breast cancer rates in unscreened age groups have been constant during

---

**Correspondence**

[Figure: Incidence of invasive breast cancer and ductal carcinoma in situ in women aged 50–69 years in Norway invited to prevalence screening in 1996–97 and women older than 69 years who are no longer invited to screening. Vertical lines indicate years with prevalence screening.]

We declare we have no conflicts of interest.

Eugenia Paci, *Mireille Broeders, Solveig Hofvind, Stephen W Duffy, on behalf of the Euroscreen working group m.broeders@ebh.umcn.nl


1991–2009. There is no underlying incidence increase. The increase in Sweden came before women started using hormone-replacement therapy (HRT) and the increase was also 50% among women younger than 50 years who were not using HRT. Furthermore, after 2002, HRT use dropped 80% in Norway without any decline in the breast cancer incidence. HRT cannot explain the increase.

I declare that I have no conflicts of interest.

Per-Henrik Zahl
per-henrik.zahl@fhi.no
National Institute of Public Health, PO Box 4404
Nydalen, 0403 Oslo, Norway


The review of breast screening by the Independent UK Panel on Breast Cancer Screening is welcome. There is much to agree with in the report, but the overdiagnosis estimates are likely to be overestimates. The panel estimates that around 11% of cancers in women invited to screening are overdiagnosed. The true figure is likely to be half this or even smaller.

The estimates of overdiagnosis are necessarily indirect and are based on excess cumulative breast cancer incidence assessed some years after screening has stopped in three of the trials. If the comparison is made too soon, the excess incidence will reflect the early diagnosis of potentially fatal cancers, which is an essential element of the screening programme. If follow-up of the Canadian trials in particular had been longer, the excess incidence would have been smaller. This effect can be seen from the Canadian National Breast Screening Study, in which the incidence of invasive cancer in the two groups becomes closer with increasing follow-up.

In the Health Insurance Plan of Greater New York study, the panel excluded for reasons of uncertainty about the data, in particular about the inclusion or not of lobular carcinoma in situ, the control group was not systematically screened. 11 years of observation after screening were available. The excess incidence in the study group is 38% at the end of year 1, 15% at year 4, 5% at year 12, and 3.5% at the end of year 15. The 15-year figure pertains to 1873 cancers, and the inclusion or not of a dozen or so cases of lobular carcinoma in situ would not materially affect the result. These figures are equally relevant to those of the panel, and their inclusion would give a substantially lower estimate of overdiagnosis.

I declare that I have no conflicts of interest.

Stephen W Duffy
s.w.duffy@qmul.ac.uk
Cancer Research UK Centre for Epidemiology, Mathematics and Statistics, Wolfson Institute of Preventive Medicine, Bart’s and The London School of Medicine and Dentistry, Queen Mary University of London, London E1 2QM, UK


The Independent UK Panel on Breast Cancer Screening concludes that those invited for breast cancer screening had a 20% relative reduction in breast cancer mortality after 13 years. Since the 1980s, there have been changes in mammography, which could affect overdiagnosis, and advances in systemic treatment of breast cancer.

The balance of benefit and harm was as follows. For 10,000 women invited to screening from age 50 years for 20 years, the panel estimated that 681 cancers will be diagnosed, of which 129 (19%) will represent overdiagnosis, and 43 deaths from breast cancer will be prevented. Overdiagnosed women (one in 60 actually screened) undergo unnecessary surgery.

A crucial judgment by the panel was: “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.” But what evidence on statistical independence did the panel seek? Whereas large tumours are unlikely to include over-diagnosed breast cancer, screening-detected small tumours could be a mixture of those that have, and have not, the potential to progress in a woman’s lifetime, so that randomisation to systemic therapy might not benefit those without such potential. Randomised systemic therapy for similarly sized, non-screen tumours has the potential to benefit all such tumours, with an anticipated larger treatment effect. In summary, what evidence might there be for, or against, differential treatment effects according to mode of breast cancer diagnosis when tumour characteristics are suitably accounted for?

Doctors cannot identify over-diagnosed cases, and so a woman’s decision to screen or not could be made differently by age group (50–59 years vs 60–69 years), or idiosyncratically: no change in the UK’s breast screening uptake would be remarkable.

I am a woman within the breast cancer screening age range.

Sheila M Bird
sheila.bird@mrc-bsu.cam.ac.uk
MRC Biostatistics Unit, Cambridge CB2 0SR, UK
Trustworthy assessment of mammography screening necessitates complete and balanced reporting of benefit and harm. However, the Independent UK Panel on Breast Cancer Screening is framing the data.\(^1\) Reporting of harm is incomplete and bias is in favour of possible benefit.

In the Summary and Conclusions, the relative risk reduction of breast cancer mortality is reported as 20%, whereas the possible harm of over-diagnosis is reported as an absolute percentage increase: “just over 1% would have an overdiagnosed cancer in the next 20 years”. To use equivalent modes of presentation, the benefit should read something like “just over 0.4% would be prevented from dying of breast cancer in the next 20 years”.

Additionally, in the Conclusions, the panel estimates that in the UK about 1300 breast cancer deaths would be prevented every year. But the corresponding figures for over-diagnosis of breast cancer and recalls for additional diagnostic tests, including 25% of recalled women getting a biopsy, are not reported. On the basis of the information in the review,\(^1\) I calculated that there would be around 4300 women with overdiagnosis of breast cancer every year, and about 100,000 recalls and 25,000 biopsies every year (although it is difficult to estimate the recall and biopsy figures from the information provided in the review).

Finally, important data on interval cancers are missing, and the costs of screening, including costs of over-diagnosis and overtreatment, should be provided. I declare that I have no conflicts of interest.

Ingrid Muhlhauser

Ingrid_muehlhauser@uni-hamburg.de

University Hamburg, Faculty of Mathematics, Informatics, and Natural Sciences (MIN), Health Sciences and Education, 20146 Hamburg, Germany


In the mammography screening trials, the best outcomes in the screened groups were evident in the Swedish Two County and Göteborg trials, which included a large number of premenopausal women, and offered screening to women in the control groups after 7 years.\(^1\) In only three trials (Malmö I, Canada I, and Canada II) were women in the control groups not offered screening, and these trials showed that the screened and control groups had similar outcomes.\(^3\)

Several investigators have pointed out that, in the mammography screening trials, premenopausal women had a transient increase in breast cancer mortality during the initial years after the start of these studies.\(^2,3\) Also, we have reported that, after initiation of mammography screening in the USA, there was a transient excess of breast cancer mortality in African American women (who are more likely than white women to develop breast cancer during their premenopausal years).\(^4\) The reason for the transient excess mortality in premenopausal women invited to screening is not clear, but it has been suggested that the detection of occult cancers and surgery might potentially perturb the natural history of breast cancer in these women.\(^3,4\)

In both the Swedish Two County and Göteborg trials, the initiation of screening in the control groups might have transiently increased breast cancer mortality in the controls, and thereby made outcomes in the screened groups seem better. Thus, the “improved” long-term outcomes in women in the screened groups might simply be attributable to the transient worse outcomes in the controls. I declare that I have no conflicts of interest.

Ismail Jatoi

jatoi@uthscsa.edu

Division of Surgical Oncology, University of Texas Health Science Center, San Antonio, TX 78229, USA


Authors’ reply

These letters variously suggest that the Independent UK Panel on Breast Cancer Screening either understated or overstated the benefits of breast screening and either underestimated or overestimated the risk of over-diagnosis. It was just such divergent views that led to the convening of the panel. The panel heard from expert witnesses who put most of the points contained in these letters to us.

Our responses are set out in our full report\(^1\) and we give the relevant page numbers below.

The panel was aware of the concern about bias in the ascertainment of endpoints, but also noted that bias could diminish the apparent benefit of screening as well as enhance it (p 23). Hence, the panel judged that the relevant outcome measure for breast screening was breast cancer mortality, best estimated from all the trials excluding the Edinburgh trial. A 20% reduction in breast cancer mortality would yield only 3.0% and 1.2% relative risk reductions in all-cancer and all-cause mortality. Even an overview of all the trials would be underpowered to show effects on...


