

Should we screen for breast cancer?

Polarised arguments about the benefits and harms of breast screening are not helping women to make an informed decision. **Klim McPherson** looks at the evidence and calls for dispassionate analysis of all available data



MAURO FERMARELLO/SPL

The burden of breast cancer is unremitting and we must do anything we can to contain it. It seems obvious that detecting tumours before they are clinically apparent is a good idea. But screening all women aged 50-70 every three years is but one way of containing the disease, and its appreciable financial costs need to be borne in mind (£75m (€90m; \$110m) a year in the UK: £37.50 per woman invited and £45.50 per woman screened).

Screening for a progressive disease is justified only if earlier diagnosis and treatment improve disease progression. Since all healthy women aged 50-70 are called for breast screening, the benefits (reduced mortality) ought to be unambiguous and considerable and the risks of harm small. Since this is not always the case, national screening programmes rely primarily on faith, supported by large randomised studies. They are essentially based on their face validity and the apparent benefit among women with screen detected minute tumours. Scientific evidence plays too little a part, not least because it is much disputed.

The recent guidelines by the US Preventive Services Task Force on screening for breast cancer reviewed the evidence, most of which comes from eight large trials.¹ The task force recommended changing current practice in the US to biannual mammography rather than annual screening and only for women aged 50-74, excluding those aged 40-49, provoking an almost hysterical response from screening enthusiasts.² The recommendations were rejected.³

The US report pools results from all randomised studies and estimates that the mortality reductions attributable to breast screening are 15% for women aged 39-49, 14% for those aged 50-59,

and 32% for those aged 60-69.⁴ The effect in women over 70 is estimated from only one trial with too little precision. These effects might seem adequate, but under age 60 the pooled effects are of marginal statistical significance, while some large trials show no benefit. Worse still, estimated numbers of women needed to be invited to a US screening programme in order to save one life are high. For the younger group it is nearly 2000 while in those aged 60-69 it is still nearly 400. The box gives the figures for the UK.

Individual benefit from mammography is thus very small, but this is not widely understood. In part this is due to obfuscation from organisers of mammography services assuming that a positive emphasis is needed to ensure reasonable compliance.

The Nordic Cochrane Centre has published analyses questioning the value of screening. Their 2002 review assessed the seven trials then available,

which they rated for quality, and found no reliable evidence of reduced mortality.⁵ The most recent update⁶ includes eight trials and estimates a 15% mortality across age groups. They are rightly concerned that women are not provided with adequate information about these uncertainties and that the extent of overdiagnosis is underplayed in screening publicity.⁷

Whatever we believe about the science, there is no doubt that screening for breast cancer has limited benefit and some possibility of harm for an individual woman and marginal cost effectiveness for a community. Has the time come for a serious scientific rethink of the benefits of the NHS screening programme in the context of cost effective care?

Reducing breast cancer burden

The incidence of breast cancer has risen by 50% since 1980, although age standardised mortality is thankfully falling (by 35%).⁸ The reduced mortality is mainly down to a valiant concentration on therapeutics, research, and practice and a strong commitment to rationalise cancer services in the NHS. But screening has had its effect too: both on the rise in incidence and on the fall in mortality. Quantifying these relative effects when much else is changing simultaneously is where the problems arise. Reducing incidence must be the primary goal, with reducing mortality an important but secondary end point. Mammography unavoidably increases incidence considerably by bringing the diagnosis forward but also by diagnosing cancer that would never become apparent and by false positive results. The last two result in overdiagnosis and overtreatment and have no effect on mortality. Thus real mortality reductions attributable to mammography must be large and secure to justify this possible harm.

Early breast cancer and DCIS

Enthusiasm for breast screening has uncovered a disease hitherto much less common and correspondingly poorly studied: ductal carcinoma in situ (DCIS). Mammography may also detect invasive tumours that are not going to cause any trouble during the rest of a woman's life. And mammography itself is minimally carcinogenic. We do not know enough about DCIS to know how to treat it optimally, and we have imprecise measures of the extent of overtreatment.

DCIS represents about 20% of screen detected cancers.⁸ The population incidence shoots up above the age of 50. Half of women with DCIS detected will develop invasive disease,⁹ as far as we know. Previously most were treated with mastectomy; nowadays more are treated with lumpectomy and radiotherapy. Around 15% of women with treated DCIS, have a recurrence within 10 years; hence treatment can be withheld only when the prognosis is more specific than it can be now—for example, by more reliably taking account of grade or necrotic tendency. This is a dilemma for

Reductions in mortality and number needed to screen in UK

- If mortality among women aged 40-55 in the UK is 0.41%, then with a 15% reduction from screening the risk would become 0.35%
- The difference attributed to screening is $0.0041 - 0.0035 = 0.00062$ (0.062%)
- The number of women who need to be screened to avoid one death is $1/0.00062 = 1610$
- Even at age 60, when risk of death from breast cancer over the next 15 years is 1.2%, a 32% mortality reduction still requires 259 women to be screened to avoid one death.

Clearly the number needed to be invited will be higher. The calculations assume that results from randomised controlled trials can be applied to the NHS and a 15 year effect (maximum follow-up in modern trials is 13 years).

women,¹⁰ both in deciding whether to go for routine mammography in the first place and in what to do when confronted with a diagnosis of DCIS.

Overdiagnosis

If we assume the US estimates of effect are correct and apply them to UK death rates, then over 15 years the rate of prevented deaths in a screened population ranges from 0.06% of women at age 40 to 0.40% at age 60. We need then to be able compare these estimates reliably with risks of overdiagnosis, including some DCIS. An overdiagnosed case is one in which a cancer is detected and treated that would never have presented clinically. Clearly this is an unobservable entity in individual women. The only way to estimate the extent of overdiagnosis, just as for reductions in mortality, is to compare the number of accumulated cases of breast cancer in a screened population with that in a comparable unscreened population over a long period. The difference between the two groups after a long period would provide an estimate of the extent of overdiagnosis.

Evidence from randomised trials is essential here for many reasons. Use of hormone replacement therapy has changed greatly since the 1980s, mainly because it was shown to increase the risk of breast cancer.¹¹ It also affects the sensitivity and specificity of mammography.¹² Obesity has also increased, and that too increases risk of breast cancer among postmenopausal women. If use of hormone replacement therapy and obesity are associated with the uptake of screening, which is highly likely,¹³ then unscrambling that confounding will be impossible. Sophisticated modelling will not help either.

A recent analysis from the Nordic Cochrane Centre claims that overdiagnosis constitutes one third of breast cancers detected in screening programmes.¹⁴ And in 2006 Zackrisson et al, who followed up 43 283 women aged 45-69 in the Malmö randomised trial of mammographic screening, estimated that 10% of cancers in the screening group were overdiagnosed after 15 years of follow-up.¹⁵ This is 0.93% of screened women. There were 150 more women diagnosed with cancer in the screened group than in the randomised control group at the end of the trial's original 10 year follow-up, and 115 after 25 years.¹⁶

More recently, Duffy and colleagues have responded to these varying estimates of overdiagnosis in screening programmes by analysing the Swedish Two County Trial of 55 985 women aged 40-74.¹⁷ They argue that the lives saved by screening greatly outnumber overdiagnosed cases, estimating overdiagnosis to be 0.43% and lives saved as 0.88% at 20 years (table). However, the US task force estimates give a figure for lives saved of about 0.4% for women over 60 with optimistic assumptions. The 0.88% is likely to be a "random high" because the two counties trial

Approximate estimated effects of a mammography screening programme (% of women screened) after 15 years

| Source | Data source | Reduction in mortality | Overdiagnosis |
|--|-------------------------|------------------------|-----------------|
| US Preventive Services Task Force ⁴ : | | | |
| Age 40-9 | | 0.06 | |
| Age 50-9 | | 0.11 | Range 0.10-0.86 |
| Age 60-9 | | 0.40 | |
| Nordic Cochrane Centre ⁶ | 8 randomised trials | 0.11 | 2.8 |
| Zackrisson et al (Malmö) ¹⁵ | 43 283 women aged 40-69 | 0.18 | 0.93 |
| Duffy et al (Two counties) ¹⁷ | 55 875 women aged 40-74 | 0.88 | 0.30 |

observed an unusually large reduction in mortality. The estimate of overdiagnosis from this trial is very low (0.43% is roughly approximately 0.30% at 15 years) compared with that of Zackrisson et al (0.30% v 0.93%). But it is higher than the authors have previously estimated from multistage modelling, which is important because these original modelling analyses were crucial in designing the NHS screening programme.

But it is misleading to accept Duffy et al's estimate as if it comes from a randomised trial. They estimate overdiagnosis by comparing incidence observed with regression estimates of the trend in incidence before the control group was offered screening and give no direct comparison of the observed diagnosed cases over time in the two groups. Other things also need clarifying: why did they not discuss Zackrisson et al's¹⁵ estimate of overdiagnosis in their report when they did criticise the Nordic Cochrane Centre, and why did they not discuss DCIS?

The Nordic Cochrane Centre has responded at considerable length,¹⁸ providing substantive justification and pointing to possible conflicts of interest in Duffy et al's publication. Duffy et al's response dismisses this as "off topic" and comments that they "do not have the leisure" to fully enumerate the centre's inaccuracies. And still Zackrisson is not discussed.

The US task force estimates, from five trials in which the control group did not receive mammography, that the excess incidence of invasive and ductal carcinoma in situ is between 0.007% and 0.073% a year.¹⁹ These are approximately equivalent to 0.10% and 0.73% at 15 years, with UK incidence data. The low figure comes from a 1960s trial that is probably less relevant now, and the high figure is from Zackrisson et al. Apart from this two Canadian trials are the most informative, being recent and with 13 years' follow-up. Among women aged 50-59, the cumulative excess risk (and hence overdiagnosis) in the screened arm of these trials was 0.34%. Hence the estimate from Zackrisson does look high. But the Canadian trial used here compared annual mammography and physical examination with annual physical examination, not a control group and found no mortality reduction associated with mammography.²⁰

Physical examination in both groups could therefore attenuate the effect on overdiagnosis of mammography.

Sorting out the evidence

Mammography does save lives, more effectively among older women, but does cause some harm. Do the benefits justify the risks? The misplaced propaganda battle seems to now rest on the ratio of the risks of saving a life compared with the risk of overdiagnosis, two very low percentages that are imprecisely estimated and depend on age and length of follow-up. Duffy et al say this ratio is about 2 (0.88%/0.43%) whereas Zackrisson et al put it at about 0.2 (0.18/0.93). So are women more likely to be overdiagnosed than to have their life saved by screening mammography? Even if women knew the answer balancing the two would be deeply problematic, but such differences in estimates for this ratio are unacceptable (table).

Estimates of rates of overdiagnosis are less secure than estimates of mortality because they were not the primary outcome measure, and consequently estimation is possible in only three randomised studies. The US task force, using estimated averages, says that the studies are too heterogeneous to combine statistically.

Arguments that polarise are unhelpful and

"The misplaced propaganda battle seems to now rest on the ratio of the risks of saving a life compared with the risk of overdiagnosis"

render women, many with strong preferences, more helpless. For too long they have been misled and confused by too much agenda driven analyses of these data.

What is required now is a full examination of all the data, preferably individual patient data, from all the recent studies, by dispassionate epidemiologists to get the best estimates in the UK screening setting. Synthesis of observed data from trials and the screening service could surely now be put to better use by, for example, the National Institute for Health and Clinical Excellence, with collaboration, in six months.

Meanwhile the NHS screening programme needs to be really clear about these uncertainties when communicating with women, and organisers of current trials need to be clear about how much of this uncertainty will be addressed, with what precision, and by when. More importantly, we all need to understand better how a national

programme of such importance could exist for so long with so many unanswered questions. Maybe the breast screening service exists on its face validity plus concomitant scientific polarisation in the face of uncertainty? That would be irresponsible.

Klim McPherson visiting professor of public health epidemiology, University of Oxford

klm.mcpherson@obs-gyn.ox.ac.uk

Competing interests: The author has completed the unified competing interest form at www.icmje.org/coi_disclosure.pdf (available on request from him) and declares (1) no financial support for the submitted work from anyone other than their employer; (2) no financial relationships with commercial entities that might have an interest in the submitted work; (3) no spouses, partners, or children with relationships with commercial entities that might have an interest in the submitted work; and (4) no non-financial interests that may be relevant to the submitted work.

Provenance and peer review: Commissioned; externally peer reviewed.

- 1 US Preventive Services Task Force. Screening for breast cancer: recommendation statement. *Ann Int Med* 2009;151:716-26.
- 2 Electronic responses. Screening for breast cancer: recommendation statement *Ann Int Med* 2009. www.annals.org/content/151/10/716.full/reply.
- 3 Kamerow D. Mammograms, poor communication, politics. *BMJ* 2009;339:b5175.
- 4 Nelson HD, Tyne K, Nalk A, Bougatsos K, Chan BK, Humphrey L. Screening for breast cancer: an update for the US Preventive Services Task Force. *Ann Int Med* 2009;151:727-37.
- 5 Olsen O, Gotzsche PC. Cochrane review on screening for breast cancer with mammography. *Lancet* 2002;358:1340-2.
- 6 Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2009;4:CD001877.
- 7 Gotzsche P, Hartling OJ, Nielsen M, Brodersen J, Jørgensen KJ. Breast screening the facts—or maybe not. *BMJ* 2009;338:b86.
- 8 Cancer Research UK. Breast cancer—the facts. <http://info.cancerresearchuk.org/cancerstats/types/breast>.
- 9 Thompson A, Brennan K, Cox A, Gee J, Harcourt D, Harris A, et al. Evaluation of the current knowledge limitations in breast cancer research: a gap analysis. *Breast Cancer Res* 2008;10:R26.
- 10 Flanders J. Reality of ductal carcinoma in situ. *BMJ* 2009;338:b958.
- 11 Rossouw JE, Anderson GL, Prentice RL, LaCroix AZ, Kooperberg C, Stefanick ML, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women. *JAMA* 2002;288:321-68.
- 12 Banks E. Hormone replacement therapy and the sensitivity and specificity of breast cancer screening: a review. *J Med Screen* 2001;8:29-34.
- 13 Moser K, Patrick J, Beral V. Inequalities in the reported use of breast and cervical screening in Great Britain: analysis of cross sectional study. *BMJ* 2009;338:b2025.
- 14 Jørgensen KJ, Gotzsche P. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. *BMJ* 2009;339:b2587.
- 15 Zackrisson S, Andersson I, Janzon L, Manjer J, Game JP. Rate of over-diagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study. *BMJ* 2006;332:689-92.
- 16 Welch HG. Overdiagnosis and mammography screening. *BMJ* 2009;339:b1425.
- 17 Duffy SW, Tabar L, Olsen AH, Vitak B, Allgood PC, Chen TH, 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.
- 18 Electronic responses to Jørgensen KJ et al. Breast cancer mortality in organised mammography screening in Denmark: comparative study. *BMJ* 2010. www.bmj.com/cgi/eletters/340/mar23_1/c1241.
- 19 Moss S. Overdiagnosis in randomised controlled trials of breast cancer screening. *Breast Cancer Res* 2005;7:230-4.
- 20 Miller AB, To T, Baines C, Wall C. Canadian national breast screening study-2: 13-year results of a randomized trial in women aged 50-59. *J Can Canc Inst* 2000;92:1490-9.

Cite this as: *BMJ* 2010;340:c3106

doc2doc White paper, white elephant?



According to 62% of healthcare professionals on doc2doc, BMJ Group's global clinical online community, general practitioners are not the right people to control England's £80bn NHS commissioning budget. The community discusses whether the new NHS white paper is worth the paper it's written on. . .

tnolan: "I don't see how having 500 groups nationwide deciding on services, rather than 150 primary care trusts, will be any more GP driven than it is at the moment."

DaVinci: "Is this just a way of telling current managers that they aren't doing a good job? Or are we just shaking things loose for the sake of it?"

Max Allen: "It comes as no surprise that there is disparity between what was said during the election and what is done: that's politics."

It is undoubtedly misplaced optimism. The same reason we thought England might have won the World Cup. I hoped for better performances from Rooney and Cameron. I hoped the government we chose would be the government we would get."

Hugh Rogers: "We will now spend three years scrapping old organisations and recreating them under new banners. This is a brilliant exercise in how to occupy the workforce with rearranging deckchairs, while distracting the whole NHS from the real task of improving or defending quality in the face of a real terms financial downturn."

Saleem Khwaja: "If this is not backdoor privatisation, what is? Any change carried out in a hurry or merely for ideological reasons, and without staged testing is courting disaster."

Ed Davies: "The white paper *Equity and excellence: liberating the NHS* comes to us in association with 2008's *High quality care for all*, 2006's *Our health our care our say*, 2004's *Trainspotting* inspired *Choosing health*, 2002's *Securing our future health* (featuring exclusive 2004 Wanless remix), 2000's *NHS plan*, 1997's *The new NHS*, the letters Y O Y, and the number 666."

Follow the debate at <http://bit.ly/blXBhw>

FROM BMJ.COM

The sick and famous, and other blogs

"Singer Cheryl Cole may turn out to be the most famous person to get malaria in 2010, but of course she's not the only one. And more importantly, she's hardly typical of malaria patients who are likely to be impoverished residents with poor access to prophylaxis living in areas with run-down public health infrastructures. While we all wish her a full recovery, has the publicity generated been useful to malaria control? And how can people working in health and medicine do more than stand back as spectators on the massive coverage that celebrity illness, injury, and death can create?" asks Simon Chapman.

Jeremy Sare reports on NRG1 (naphyrone), which the home secretary has recently agreed to bring under the control of the Misuse of Drugs Act 1971. Reportedly it has

"approximately ten times the potency of cocaine." For drug users, that is not the good selling point ministers may think it is. Ten times the strength means simply too strong. "The Advisory Council on the Misuse of Drugs is desperate to re-establish its relevance after haemorrhaging much of its scientific expertise in recent months. But, here, they have elected to attack a straw man. The drug may be highly potent but if no one is taking it, then where is the harm and why the urgency?" writes Jeremy.

Regular blogger Emily Spry is back in London after a year spent working in Freetown, Sierra Leone. She writes about the readjustment she has had to make. "It has been amazing to come back to the luxuries of reliable water and power, safe and efficient transport,



and a seemingly unlimited procession of utterly fabulous clothes, perfumes, shops, magazines, interiors, and many other material things that I normally have precisely no interest in," she says. She has also found it hard to explain to everyone what her time abroad was really like.

Taking over from Emily in the Pikin Hospital, Freetown, is Sandra Lako. In her first two blog posts she introduces herself and tells us about what led her to working in Sierra Leone.

Read these blogs and others at www.bmj.com/blogs