

## RESEARCH

# Possible net harms of breast cancer screening: updated modelling of Forrest report

 OPEN ACCESS

James Raftery *professor of health technology assessment*, Maria Chorozioglou *research fellow*

Wessex Institute, Faculty of Medicine, University of Southampton, Southampton SO16 7PX, UK

## Abstract

**Objective** To assess the claim in a Cochrane review that mammographic breast cancer screening could be doing more harm than good by updating the analysis in the Forrest report, which led to screening in the United Kingdom.

**Design** Development of a life table model, which replicated Forrest's results before updating and extending them with data from relevant systematic reviews, trials, and other models based on purposive literature searches.

**Participants** Women aged 50 and over invited for breast cancer screening.

**Main outcome measures** Quality adjusted life years (QALYs), combining life years gained from screening with losses of quality of life from false positive diagnoses and surgery.

**Results** Inclusion of the effects of harms reduced the updated estimate of net cumulative QALYs gained after 20 years from 3301 to 1536 or by more than half. The best estimates from the Cochrane review generated negative QALYs for the first seven years of screening, 70 QALYs after 10 years, and 834 QALYs after 20 years. Sensitivity analysis showed these results were robust to a range of assumptions, particularly up to 10 years. It also indicated the importance of the level and duration of harms from surgery.

**Conclusions** This analysis supports the claim that the introduction of breast cancer screening might have caused net harm for up to 10 years after the start of screening.

## Introduction

The Forrest report in 1986,<sup>1</sup> which led to the introduction of mammographic breast screening in the United Kingdom, analysed the costs and benefits in terms of quality adjusted life years (QALYs). One of the earliest uses of QALYs to guide policy, it suggested that screening would reduce the death rate from breast cancer by almost one third with few harms and at low cost (for details see appendix on bmj.com).

The key data used in the Forrest report were drawn from two randomised trials, the Swedish two counties trial<sup>2</sup> and the Health Insurance Plan (HIP) New York trial.<sup>3</sup> The Forrest report claimed that overdiagnosis was not a problem, based on the New York trial, but noted that the Swedish trial found possible overdiagnosis of 20%. It stated that "further follow up is required to find out whether this excess persisted." We have updated the Forrest report's estimates for mortality and extended them to include the effects of false positives and overdiagnosis. Since the Forrest report, the harms of mammographic breast cancer screening have been acknowledged. A WHO report defined false positives and overdiagnosis:

"The term false positive refers to an abnormal mammogram (one requiring further assessment) in a woman ultimately found to have no evidence of cancer. Overdiagnosis refers to the diagnosis and treatment of cancer that would never have caused symptoms. Thus a false positive result can be found only in a woman without cancer, while overdiagnosis can only be made for women with cancer."<sup>4</sup>

It went on to note that "overdiagnosis is a foreign concept to most prospective screenees (and many clinicians)."

The WHO report noted that a considerable part of overdiagnosis involved ductal carcinoma in situ, which accounts for around a fifth of mammographically detected cancers. While this is a risk factor for breast cancer, only a minority of these develop into breast cancer. Indeed the inclusion of the term "carcinoma" in ductal carcinoma in situ has been questioned.<sup>5</sup>

The WHO report claimed that the success of breast cancer screening programmes should be assessed only in terms of mortality: "Screening programmes should ultimately be monitored in terms of deaths, the measure directly related to the purpose of screening." A focus solely on deaths, however, implies ignoring harms to the living.

Correspondence to: J P Raftery [j.p.raftery@soton.ac.uk](mailto:j.p.raftery@soton.ac.uk)

Extra material supplied by the author (see <http://www.bmj.com/content/343/bmj.d7627?tab=related#webextra>)

Appendix: Details of the original Forrest report, the Southampton model, and extra tables and figures

Gøtzsche and Nielsen's Cochrane review<sup>6</sup> raised the disturbing possibility that mammographic breast cancer screening could be doing more harm than good. This was because of their lower estimate of the reduction in mortality from breast cancer and their inclusion of the harms from overtreatment. They said that "this means that for every 2000 women invited for screening throughout 10 years, one will have her life prolonged, and 10 healthy women, who would not have been diagnosed if there had not been screening, will be diagnosed as breast cancer patients and will be treated unnecessarily. Furthermore, more than 200 women will experience important psychological distress for many months because of false positive findings. It is thus not clear whether screening does more good than harm."<sup>6</sup>

Their meta-analysis included eight randomised trials, three of which they considered adequately randomised and five suboptimally randomised. Only the suboptimally randomised trials found a significant effect of screening on deaths ascribed to breast cancer. For all the eight trials taken together the relative risk reduction for mortality from breast cancer was 19% (95% confidence interval 26% to 13%) after 13 years. Given the quality of the evidence, Gøtzsche and Nielsen's best estimate of the effect of screening was a 15% decline in mortality.

The increased risk of surgery was the basis of Gøtzsche and Nielsen's estimate of unnecessary treatment. Four trials provided data on breast operations (mastectomies and lumpectomies), with more performed in the screened groups than in the control groups: the relative risk increase was 31% (22% to 42%) for the two adequately randomised trials and 35% (26% to 44%) for all four trials. For false positive results, Gøtzsche and Nielsen stated "it seems that screening inflicts important psychological distress for many months on more than a 10th of the healthy population of women who attend a screening program."

A systematic review and meta-analysis by Nelson and colleagues for the US Preventive Services Task Force independently analysed the same eight clinical trials in the Cochrane review but by age group.<sup>7 8</sup> This put the reduction in mortality from breast cancer at 15% for those aged 39-49, 14% for those aged 50-59, and 32% for those aged 60-69. It used US registry data to suggest that about 10% of those screened would have a false positive result requiring further investigation.<sup>7</sup> It differed from the Cochrane review in relation to overdiagnosis. "Rates of overdiagnosis vary from less than 1% to 30% with most from 1% to 10%. Estimates differ by outcome (invasive vs in situ breast cancer), by whether cases are incident or prevalent, and by age. The studies are too heterogeneous to combine statistically."<sup>7</sup> These studies, it should be noted, included both randomised trials and observational studies.

Thus the two systematic reviews agreed that screening reduced mortality from breast cancer but differed in how much. Nelson and colleagues estimated a false positive rate around 10% per round of screening, while Gøtzsche and Nielsen put it at around 10% over 10 years. Only Gøtzsche and Nielsen provided data on the increased relative risk of surgery with screening, with two estimates: 31% based on the better quality trials and 35% based on all trials reporting this outcome.

We assessed the claim of Gøtzsche and Nielsen by updating the Forrest report framework, extended to include harms. The Forrest report used life tables to estimate the number of women surviving by year up to 15 years in two cohorts aged 50, only one of which was screened. Deaths could be from breast cancer or all other causes. Baseline mortality and the reduction from triennial breast cancer screening were based on the two randomised controlled trials then available. The difference in life years between the two cohorts after 15 years was expressed

in QALYs by reducing their quality of life by 8% to reflect the effects of treatment.

## Methods

The Southampton model used the same life table approach as Forrest to estimate life years. To ensure that the Southampton model was fully compatible, we confirmed that use of Forrest inputs generated the same number of deaths in our model.

Forrest took baseline mortality from breast cancer from the two trials then available but acknowledged that as this was below the English mortality rate from breast cancer, his results were underestimates. In updating Forrest, we corrected this by using the mortality rate from breast cancer for England.<sup>9</sup> We also took the baseline risk of surgery for breast cancer from the English NHS.<sup>10 11</sup> Data for both these baselines were for 1985, the latest year before screening for which we could locate data. These changes meant more favourable results for screening than if we used the control arms of trials as baselines.

We drew parameter inputs for the Southampton model (table 1) from the published literature, giving priority to systematic reviews, followed by randomised clinical trials and other published models, and then observational data supplemented by clearly stated assumptions when necessary. Sensitivity analysis varied mean estimates to their 95% confidence intervals and other inputs by  $\pm 33\%$ . The results of individual sensitivity analyses are reported in the appendix on bmj.com. Probabilistic sensitivity analysis varied key inputs simultaneously by sampling from their probability distributions for 10 000 iterations.

All the input values are listed in table 1 with sources and discussed more fully in the appendix on bmj.com. In brief, the changed relative risks for breast cancer mortality and surgery from screening were based on the meta-analyses of the relevant trials.<sup>6-8</sup> The losses of quality of life from false positive results and surgery were based on a systematic review, supplemented by relevant randomised trials and values used in previous models. The extent and duration of the loss of quality of life from surgery have been least researched. We assumed a 6% permanent loss from surgery, less than in the Forrest report but informed by recent randomised trials.<sup>12 13</sup> Sensitivity analyses explored changing the extent and duration of these and other values. Figure 1 illustrates the modelling approach.

## Setting, participants, and outcome measures

The setting was England. The outcomes of 100 000 women aged 50 were modelled in two cohorts, one screened the other not. The outcome measures were deaths from breast cancer, deaths from all other causes, and the number of women having false positive diagnoses and surgery, which we combined into the main outcome—quality adjusted life years (QALYs).

## Results

Figure 2 graphically presents the five scenarios, and table 2 summarises the results. Scenario 1 shows the QALY gains that Forrest would have got if he had used English breast cancer mortality rates as baseline with his risk reductions. Scenario 2 updates this with the reduction in mortality from breast cancer for all ages from all eight trials. The losses of quality of life from surgery and false positive diagnoses were added in scenario 3. Scenario 4 used the reduction in mortality from breast cancer suggested by Gøtzsche and Nielsen.<sup>6</sup> In scenario 5, we used the reductions in mortality from breast cancer by age group from Nelson et al.<sup>7 8</sup> The results are based on 100 000 women being

invited for mammographic screening, with 73% attending, and are presented for each year up to 20 years after the entry to the screening programme.

Scenario 1 accumulated just over 3300 net QALYs after 20 years. This is what Forrest would have got had he used as baseline the breast cancer mortality rate for England and the mortality reduction from the two trials. When we updated the estimate for reduction in breast cancer mortality for all ages, with the meta-analysis of the eight trials (scenario 2), the net cumulative QALY gain at 20 years fell to around 3100 QALYs or by about 6%. When we added harms in scenario 3, this was reduced to just over 1500 QALYs or by half. When we changed the reduction in mortality from breast cancer to that suggested by Gøtzsche and Nielsen,<sup>6</sup> the net QALYs at year 20 fell to 834 (scenario 4). Scenario 5, based on the reductions in mortality from breast cancer by age group suggested by Nelson et al,<sup>7,8</sup> generated 1685 QALYs by year 20.

Scenarios 3, 4, and 5 had negative cumulative QALY values for the first four, seven, and eight years, respectively, but had positive values after 10 years. The harms from surgery and false positive diagnoses impacted from the start because they were linked to each round of screening. Mortality from breast cancer, however, was reduced only after several years but accumulated over time so that positive net QALY applied by 10 and especially by 20 years.

## Sensitivity analyses

Sensitivity analyses explored the effects of varying input values independently (see appendix on [bmj.com](http://bmj.com)). When we combined four key parameters (reduced mortality from breast cancer, increased surgery for breast cancer, and losses of quality of life from false positive results and from surgery) in a probabilistic sensitivity analysis for scenario 3 (Forrest updated including harms), the net cumulative gain in QALYs after 20 years was between 771 and 2136 (mean 1532), with lower values in the earlier years (fig 3). The mean number of years with negative QALYs was four, with a range of two to nine.

We did not include the duration of harms from surgery in the probabilistic sensitivity analysis because of uncertainty about the appropriate distribution. Instead we used deterministic sensitivity analyses to explore reducing the duration of harms from surgery, assumed as permanent in the base case, to five and 10 years (see appendix on [bmj.com](http://bmj.com)). This led to unchanged net QALYs up to five and 10 years but with more QALYs over longer periods.

## Discussion

Assessment of the effects of mammographic breast screening in terms of mortality or life years inevitably shows positive benefits because of the omission of harms. Despite its espousal of a QALY framework, the Forrest report focused mainly on life years gained, which it adjusted for quality of life only from necessary surgery and ignored all other harms. Our analysis shows that inclusion of the harms from false positive results and unnecessary surgery reduced the benefits of screening by about half with negative net QALYs in the early years after the introduction of screening.

We assumed that the loss of quality of life in women who had unnecessary surgery was the same as for those who had had “necessary” surgery. A key feature of overtreatment is that individuals affected cannot be identified. Of Gøtzsche and Nielsen’s 10 women who had unnecessary surgery, all believed that it was necessary.<sup>6</sup> This has been dubbed the paradox of

overtreatment—“overdiagnosis and overtreatment create a paradoxical popularity because each individual justifies their experience by believing they have had a dramatic benefit.”<sup>20</sup> The more people are (over)treated, the more people think screening saved their lives.

Would knowing whether or not treatment was necessary affect quality of life? None of the surveys of quality of life included overtreatment, implicitly assuming all surgery was necessary. To answer this question surveys would have to ask each woman whether her quality of life would be affected if it could be shown that her surgery had been unnecessary. While the methodological problems of measuring quality of life in cancer screening are considerable,<sup>21</sup> ignoring overtreatment is inexcusable.

Ways of reducing the harms from screening might include less frequent screens, particularly for younger women. While further modelling might explore the clinical and cost effectiveness of various options, conclusions will inevitably be limited without better estimates of the level and impact of overtreatment.

## Strengths and limitations

Our analysis does have limitations. Following the Forrest report, it relies heavily on clinical trials, most of which were completed in other countries several decades ago. As mortality from breast cancer in 1985 was higher in England than in those trials, we took as baseline the rate for England for 1985, before screening was introduced. We have assumed that the risk reductions shown in the trials apply to this higher baseline rate. An assessment of the value added by screening today might require disentangling the effects of screening from the effects of improved treatments, which is difficult.<sup>22-24</sup> It would also require consideration of how screening methods have changed. Double view mammography and improved imaging might reduce the false positive rate but could have increased overtreatment by creating more (harmful) true positive diagnoses.<sup>25</sup>

As with breast cancer mortality, our baseline risk of breast cancer surgery was that for England in 1985. We assumed that the risk increase shown in the trials applied to this risk. Observational studies, including those summarised in the US systematic review, provided a wide range of estimates of overtreatment from 1%<sup>7</sup> to 52%<sup>6,7,26-28</sup> but have also been criticised for poor quality.<sup>29</sup> We assigned a single loss of quality of life to all forms of surgery but acknowledge that lesser harms are likely with lumpectomy than with mastectomy. Against this, we have not included the harms from radiotherapy and chemotherapy. Future studies on quality of life might usefully distinguish between the effects of different treatments.

How plausible are the losses in QALYs from surgery? The Forrest report estimate of 8% seems to have been based on a single small study from which it took the lowest estimate<sup>30</sup> (see appendix on [bmj.com](http://bmj.com)). A 2010 systematic review of health state utilities in breast cancer<sup>15</sup> found only two relevant studies. One put the loss in utility at 8% in year one, 4% in intervening years, and 11% in the last year of life. The other put the loss at 38% in year one after diagnosis, 31% in years one to five, and 29% after five years. The 2010 UK COMICE trial put the loss in quality of life from surgery in 1625 women with a low risk breast cancer at 5% after 12 months.<sup>12</sup> The five year follow-up to the PRIME trial showed that the quality of life losses after surgery were unchanged after five years.<sup>13</sup> Overall, our assumption of a permanent 6% loss in quality of life from surgery does not seem unreasonable, but more robust estimates are needed.

We assumed the base case to be a loss in quality of life from false positive results of 5% of full health for 0.2 years. This is

lower than that of a relevant US model<sup>16</sup> but similar to the Dutch model.<sup>17</sup> The 2010 systematic review of the utility losses from breast cancer included estimates of this loss of between 11% and 34%<sup>15</sup> but warned that the studies could not be synthesised. The time frame of up to 20 years is long relative to the duration of the trials, results of which have been synthesised up to 13 years. Extrapolation required an assumption of constant benefits and harms from additional rounds of screening. Longer time frames generate greater net QALYs but rely on increasingly strong assumptions, both to do with the rate of survival from breast cancer and the pattern of losses in quality of life over time.

We assumed no recurrence of cancer, despite the 10 year survival rate for breast cancer being 72% in the UK.<sup>30</sup> We also assumed no re-operations, even though 17% of women with tumours detected at screening in the UK had more than one therapeutic operation in 2006–7.<sup>31</sup> While it is possible that some cancers that were detected early by screening might have progressed in longer time frames, a recent analysis has shown no decline in the incidence of advanced breast cancer.<sup>32</sup> Modelling the longer term effects of breast cancer screening should include these factors.

Finally, our list of benefits and harms excluded the potential reassurance from a negative result on mammography. As a negative mammogram has little predictive value, any reassurance is limited to relief at not having cancer at that time.<sup>33</sup> The 2010 systematic review of utility states in breast cancer<sup>15</sup> found no evidence of improved quality of life from negative results.

## Comparison with other studies

Our results can be compared with attempts to model the effectiveness of mammographic screening in terms of cost per QALY. Although Stout et al<sup>16</sup> included only losses of quality of life in a sensitivity analysis, their inclusion roughly doubled the cost per QALY. The Dutch MISCAN study concluded that including the effects on quality of life of both treatment and false positive results had little consequence,<sup>17</sup> but this seems to be because of the relatively low level of surgery assumed in that model. In a review of the cost effectiveness of extending the age range for the UK breast screening programme, Madan et al<sup>18</sup> showed that inclusion of losses of quality of life from false positive results considerably increased the cost per QALY.

## Conclusions and policy implications

Overall, our study supports the suggestion by Gøtzsche and Nielsen that mammographic breast cancer screening could be causing more harm than good after 10 years.<sup>6</sup> Scenario 4, based on Gøtzsche and Nielsen's best estimate, had negative QALYs for the first seven years after screening and minimal gains of 70 QALYs after 10 years. Thereafter, net QALYs accumulate but much less than would be expected by our updating of the Forrest report. The uncertainty around this result, explored in scenarios 3 and 5 and in greater detail in other scenarios, applies more to the longer than the shorter term. Harms largely offset the gains up to 10 years, after which the gains accumulate at an increasing rate.

More research is required on the extent of unnecessary treatment and its impact on quality of life. Most of the observational studies of overtreatment have focused on the relation between the incidence of breast cancer and mortality rather than on the levels of treatment, especially surgery. The effects of treatment on quality of life could be established observationally or in longer follow-up studies of trials.<sup>13</sup> Improved ways of identifying those most likely to benefit from surgery and for measuring the

levels and duration of the harms from surgery should be research priorities.

As randomised trials might be the only way to resolve the extent of overtreatment, researchers in countries that have not yet implemented breast cancer screening should consider trials that include the harms of screening. There have been suggestions for more sophisticated approaches to the prevention and treatment of breast cancer.<sup>33, 34</sup> From a public perspective, the meaning and implications of overdiagnosis and overtreatment need to be much better explained and communicated to any woman considering screening.

We thank Ruairidh Milne for noting that the Forrest report used QALYs, the authors of both the Cochrane review (Peter Gøtzsche) and US Preventive Task Force review (Heidi Nelson) for help with queries, David Turner, Jason Madan and Lily Yao for advice on modelling, and Maryrose Tarpey for ongoing insights. Responsibility for errors remains with the authors.

Contributors: JR was responsible for the idea, early modelling, writing up of successive drafts, and is guarantor. MC was responsible for the detailed modelling. Both authors were involved in the final drafts of both the paper and the appendix.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Competing interests: All authors have completed the ICMJE uniform disclosure form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) (available on request from the corresponding author) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: Not required.

Data sharing: No additional data available.

- 1 Forrest P. Breast cancer screening. Report to the Health Ministers of England Wales Scotland and N Ireland by a working group chaired by Professor Sir Patrick Forrest. HMSO, 1986. [www.cancerscreening.nhs.uk/breastscreen/publications/forrest-report.html](http://www.cancerscreening.nhs.uk/breastscreen/publications/forrest-report.html).
- 2 Tabar L, Fagerberg CJ, Gad A, Baldetorp L, Holmberg LH, Grontoft O, et al. Reduction in mortality from breast cancer after mass screening with mammography. Randomised trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health and Welfare. *Lancet* 1985;ii:829-32.
- 3 Shapiro S. Ten to fourteen year effect of screening on breast cancer mortality. *J Natl Cancer Inst* 1982;69:349-55.
- 4 WHO IARC. Breast cancer screening. International Agency for Research on Cancer/WHO, 2002.
- 5 Welch GH, Woloshin S, Schwartz LM. The sea of uncertainty surrounding ductal carcinoma in situ—the price of screening mammography. *J Natl Cancer Inst* 2008;100:228-9.
- 6 Gøtzsche PC, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2009;4:CD001877.
- 7 Nelson H, Tyne K, Naik A, Bougatsos C, Chan BK, Humphrey L, et al. Screening for breast cancer: an update for the US Preventive Services Task Force. *Ann Intern Med* 2009;151:727-37.
- 8 Nelson H, Tyne K, Naik A, Bougatsos BS, Chan B, Nygren P, et al. Screening for breast cancer: an update for the US Preventive Services Task Force. Agency for Healthcare Research and Quality, 2009. [www.uspreventiveservicestaskforce.org/uspstf09/breastcancer/brcanes.pdf](http://www.uspreventiveservicestaskforce.org/uspstf09/breastcancer/brcanes.pdf).
- 9 Office for National Statistics. Mortality statistics. Cause. Review of the Registrar General on deaths by cause, sex and age in England and Wales, 1985. ONS, 1986 (series DH2, No 12).
- 10 OPCS, Department of Health. Hospital inpatient Enquiry. Inpatient and day case trends 1979-1985: England. HMSO, 1987 (MB4 87/1 No 29).
- 11 OPCS, Department of Health. Hospital inpatient Enquiry. Summary table: 1985 England. HMSO, 1987 (MB4 No 26).
- 12 Turnbull LW, Brown SR, Olivier C, Harvey I, Brown J, and the COMICE Trial Group. Multicentre randomised controlled trial examining the cost-effectiveness of contrast-enhanced high field magnetic resonance imaging in women with primary breast cancer scheduled for wide local excision (COMICE). *Health Technol Assess* 2010;14:1-182. [www.hta.ac.uk/fullmono/mon1401.pdf](http://www.hta.ac.uk/fullmono/mon1401.pdf).
- 13 Williams LJ, Kunkler IH, King CC, Jack W, van der Pol M. A randomised controlled trial of post operative radiotherapy following breast conserving surgery in a minimum risk population. Quality of life at 5 years in the PRIME trial. *Health Technol Assess* 2011;15:1-57. [www.hta.ac.uk/project/1697.asp](http://www.hta.ac.uk/project/1697.asp).
- 14 Smith Bindman R, Chu PW, Miglioretti DL, Sickels EA, Blanks R, Ballard-Barbash, et al. Comparison of screening mammography in the United States and the United Kingdom. *JAMA* 2003;290:2129-37.
- 15 Peasgood T, Ward S, Brazier J. A review and analysis of health state utility values in breast cancer. SCHARR, University of Sheffield, 2010 (HEDS Discussion Paper 10/15). [www.shef.ac.uk/content/1/c6/01/87/47/HEDS%20DP%2010-15.pdf](http://www.shef.ac.uk/content/1/c6/01/87/47/HEDS%20DP%2010-15.pdf).

**What is already known on this topic**

Mammographic screening for breast cancer saves lives but also imposes losses in quality of life from false positive results and unnecessary treatment

It has been suggested that the harms outweigh the benefits, but this has not been quantified

**What this study adds**

By combining the life years saved with the quality of life losses in quality adjusted life years (QALYs), this study combined the benefits and harms into a single measure

The net QALYs from screening were negative for the early years after the introduction of screening, after which net positive QALYs accumulated but by much less than predicted by the Forrest report

- 16 Stout NK, Rosenberg MA, Trentham-Dietz A, Smith MA, Robinson SM, Fryback DG. Retrospective cost effectiveness analysis of screening mammography. *J Natl Cancer Inst* 2006;98:77482.
- 17 De Haes JC, de Koning HJ, van Oortmarssen GJ, van Agt HME, de Bruyn H, van der Maas P. The impact of a breast cancer screening programme on quality adjusted life years. *Int J Cancer* 1991;49:538-44.
- 18 Madan J, Rawdin A, Stenvenson M, Tappenden P. A rapid response economic evaluation of the UK cancer reform strategy breast cancer screening program extension via a plausible bounds approach. *Value in Health* 2008;13:215-21.
- 19 Advisory Committee on Breast Cancer Screening. Screening for breast cancer in England. Past and future. NHS Breast Screening Programme, 2006 (Publication 61)
- 20 Raffle A, Gray M. Screening evidence and practice. Oxford University Press, 2007:82
- 21 Brodersen J, McKenna S, Doward L, Thorsen H. Measuring the psychosocial consequences of screening. *Health Qual Life Outcomes* 2007;5:3.
- 22 Blanks RG, Moss SM, McGahan CE, Quinn MJ, Babb PJ. The effects of the NHS breast screening programme on mortality from breast cancer in England 1990-8. *BMJ* 2000;321:665-9.
- 23 Quinn M, Babb P, Brock A, Kirby L, Jones J. Cancer trends in England 1950-1999. Stationery Office, 2008 (SDMPS 66).
- 24 Pisani P, Forman D. Declining mortality from breast cancer in Yorkshire 1983-98. *Br J Cancer* 2004;90:652-6.
- 25 Jorgensen KJ, Keen JD, Gotzsche PC. Is mammographic screening justifiable considering its substantial overdiagnosis rate and minor effect on mortality? *Radiology* 2011;3:621-7.
- 26 Duffy SW, Tabar L, Olson AH, Vitak B, Allgood P, Chen T, et al. Absolute numbers of lives saved and overdiagnosis by breast cancer screening, from a randomised trial and from the Breast Cancer Screening Programme in England. *J Med Screen* 2010;17:25-30.
- 27 Jorgenson KJ, Gotzsche P. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. *BMJ* 2009;339:b2587.
- 28 Welch HG, Frankel BA. Likelihood that a woman with screen detected breast cancer had her "life saved" by screening. *Arch Intern Med* 2011 Oct 24, epub ahead of print.
- 29 Jørgensen KJ, Gotzsche PC. The background review for the USPSTF recommendation on screening for breast cancer. *Ann Intern Med* 2010;152:538-9.
- 30 Buxton M, Ashby J. The time trade off approach to health state valuation. In: Teeling Smith G, ed. *Measuring health: a prospective approach*. John Wiley, 1988.
- 31 Cancer Statistics. Breast Cancer UK/Cancer Research UK, 2009:9.
- 32 Autier P, Boniol M, Middleton R, Doré JF, Héry C, Sheng T, et al. Advanced breast cancer incidence flowing population based mammographic screening. *Ann Oncol* 2011;22:1726-35.
- 33 Welch HG, Schwartz SL, Woloshin S. Overdiagnosed. Making people sick in pursuit of health. Beacon, 2011.
- 34 Baum M. *Breast beating. A personal odyssey in the quest for an understanding of breast cancer, the meaning of life and other easy questions*. Ansham, 2010.

**Accepted:** 18 November 2011

Cite this as: *BMJ* 2011;343:d7627-

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited, the use is non commercial and is otherwise in compliance with the license. See: <http://creativecommons.org/licenses/by-nc/2.0/> and <http://creativecommons.org/licenses/by-nc/2.0/legalcode>.

## Tables

**Table 1 | Data used to estimate QALYs for mammographic breast cancer screening of women aged 50, by scenario**

| Parameter and scenario                                   | Value   | Sources  |
|--|---|--|
| Relative risk reduction in mortality from breast cancer: |   |  |
| Scenario 1   | -30% for 10 years, 0% thereafter  | Forrest report <sup>1</sup>  |
| Scenarios 2-3  | -19%  | Gøtzsche and Nielsen Cochrane review <sup>6</sup>  |
| Scenario 4   | -15%  | Gøtzsche and Nielsen Cochrane Review <sup>6</sup>  |
| Scenario 5   | -14% at age 50-9 over 10 years, -32% at age 60-9 over 10 years                                      | Nelson et al, US systematic review <sup>7,8</sup>  |
| False positive rate:                                     |   |  |
| All scenarios  | 6.39% at 1st invitation; 3.06% at 2nd and subsequent invitation                                     | Smith-Bindman et al <sup>14</sup>  |
| Loss of quality of life                                  | -5%   | Best estimates based on Peasgood, <sup>15</sup> Stout, <sup>16</sup> De Haes, <sup>17</sup> Madan <sup>18</sup>                      |
| Duration of loss (years)                                 | 0.2   |  |
| Breast cancer surgery:                                   |   |  |
| Relative risk  | 35% (26% to 44%)  | Gøtzsche and Nielsen Cochrane review <sup>6</sup>  |
| Loss of quality of life:                                 |   |  |
| Scenarios 1-2  | -6% for lives saved only  | Forrest report <sup>1</sup>  |
| Scenarios 3, 4, 5  | -6% for all who had surgery   | Peasgood, <sup>15</sup> Stout, <sup>16</sup> De Haes <sup>17</sup>   |
| Sensitivity analysis                                     | ±2%   | COMICE trial <sup>13</sup>   |
| Duration of loss   | Permanent in base case <sup>6</sup>   | Forrest report, <sup>1</sup> Peasgood, <sup>15</sup> Stout, <sup>16</sup> De Haes, <sup>17</sup> PRIME trial follow-up <sup>14</sup> |
| Sensitivity analyses                                     | 5 and 10 years' duration  |  |
| Baseline mortality from breast cancer                    | Rates per 100 000: 73.65 at age 50-54, 97.55 at age 55-59, 117.47 at age 60-64, 123.03 at age 65-69 | NHS mortality statistics for 1985, England <sup>9</sup>  |
| Baseline risk of breast cancer surgery                   | Rate per 100 000: 438.01 aged 45-64   | HIPE for England 1985 <sup>10,11</sup>   |
| Screening attendance rate                                | 73.2%   | Advisory Committee on Breast Cancer Screening <sup>19</sup>  |
| Time frame   | 0-20 years, scenarios (except scenarios with 5 and 10 year duration)                                | —  |

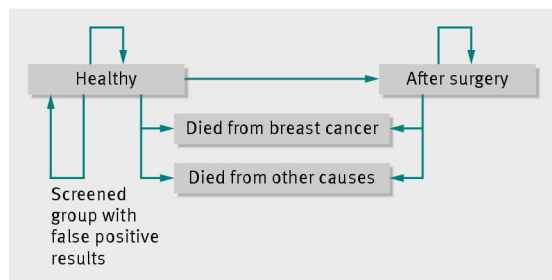
QALY=quality adjusted life year.

Table 2| Net QALYs gained over time in women undergoing breast screening by scenario

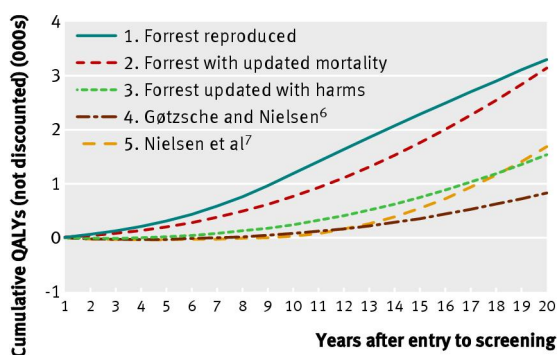
| Scenario  | At 5 years | At 10 years | At 20 years |
|---|------------|-------------|-------------|
| 1. Original Forrest report  | 304        | 1189        | 3301        |
| 2. As 1 updated for breast cancer mortality from eight trials                   | 195        | 764         | 3145        |
| 3. As 2 with harms added  | 12         | 240         | 1536        |
| 4. As 3 but with mortality suggested by Gøtzsche and Nielsen <sup>6</sup>       | -31        | 70          | 834         |
| 5. As 3 but baseline mortality and reductions as in Nelson et al <sup>7,8</sup> | -42        | 27          | 1685        |

QALY=quality adjusted life year.

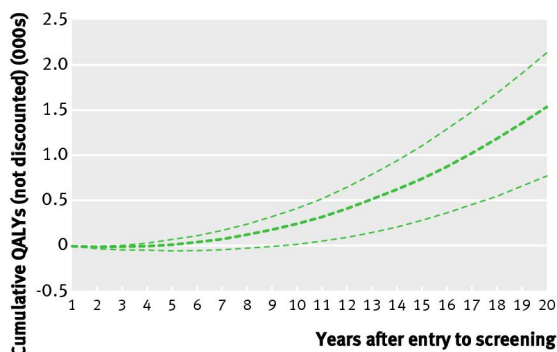
## Figures



**Fig 1** Outline of Southampton breast screening model: this applies to two cohorts of women aged 50, one screened the other not



**Fig 2** Breast cancer screening over 20 years: net QALYs by year after start of screening according to different scenarios



**Fig 3** Probabilistic sensitivity analysis, including reduced mortality from breast cancer, increased surgery for breast cancer, and losses of quality of life from false positive results and from surgery, showing cumulative QALYs for 100 000 iterations, scenario 3 (Forrest updated with harms)