Estimating breast cancer mortality reduction and overdiagnosis due to screening for different strategies in the United Kingdom

N B Gunsoy*,[1], M Garcia-Closas[1] and S M Moss[2]

[1]Division of Genetics and Epidemiology, Institute of Cancer Research, Sutton, Surrey, UK and [2]Centre for Cancer Prevention, Queen Mary University of London, Wolfson Institute of Preventive Medicine, London, UK

Background: The benefits and harms of population-wide mammography screening have been long debated. This study evaluated the impact of screening frequency and age range on breast cancer mortality reduction and overdiagnosis.

Methods: We developed a Markov simulation model for the evaluation of mammography screening in a cohort of British women born in 1935–40.

Results: For triennial screening in women aged 47–73, breast cancer mortality reduction and overdiagnosis was 18.1% (95% confidence interval: 17.3%, 19.0%) and 5.6% (5.1%, 6.1%), of all breast cancer deaths and diagnoses, respectively, from age 40 to 85 years. For annual screening in the same age range, estimates for both outcomes increased considerably to 35.0% (34.2%, 35.7%) and 7.6% (7.1%, 8.1%), respectively. For the age extension of triennial screening from 50–70 to 47–73, we estimated 5 (3, 7) incremental breast cancer deaths avoided and 14 (9, 19) incremental cases overdiagnosed per 10 000 women invited for screening.

Conclusions: Estimates of mortality reduction and overdiagnosis were highly dependent on screening frequency, age range, and uptake, which may explain differences between some previous estimates obtained from randomised trials and from service screening.

Mammography screening was introduced in the UK from 1988 after evidence from screening trials showed considerable reductions in breast cancer mortality. Initially, triennial screening was offered to women aged 50–64; the age range was extended to include women up to 69 years of age in 2002, and is currently being extended further to include women aged 47–73 (NHS Breast Screening Programme, 2010). The impact of population-wide screening on breast cancer mortality is still under debate. A disadvantage of screening is overdiagnosis, defined as a screen-detected breast cancer that would never have presented clinically in a woman's lifetime in the absence of screening. Estimates of overdiagnosis due to screening vary considerably, with some discrepancies attributable to methodological differences (de Gelder *et al*, 2011; Puliti *et al*, 2011). Benefits and harms of screening were recently assessed in two major reviews (Paci and EUROSCREEN Working Group, 2012; Marmot *et al*, 2013).

The first review, carried out by the Independent UK Panel on Breast Cancer Screening, concluded that 43 breast cancer deaths were avoided and 129 cases of breast cancer were overdiagnosed for every 10 000 women invited to screening every 3 years from age 50 to 70 years in the UK. In a meta-analysis of nine randomised trials conducted from 1963 to 1991 with at least 13 years of follow-up, the authors estimated a breast cancer mortality reduction due to screening of 20% (95% confidence interval: 11%, 27%). Using data on excess incidence from three trials in which women in the control arm were never invited to screening, overdiagnosis was estimated at 11% (9%, 12%) of all cancers diagnosed in invited women during screening and subsequent follow-up, equivalent to 19% (15%, 23%) of all cancers detected in the screening period. However, the authors acknowledged the uncertainty around these estimates (Marmot *et al*, 2013).

The second review, carried out by the EUROSCREEN working group, combined evidence from incidence-based mortality and case-control studies to estimate breast cancer mortality reduction due to population screening in Europe. The group also reviewed breast cancer overdiagnosis studies based on European data. The group concluded that for biennial screening for 20 years starting at age 50 in a typical European country, and with follow-up to age 79, a 38–48% reduction in breast cancer mortality and overdiagnosis of 6.5% of all breast cancer would be expected in screened women (Paci and EUROSCREEN Working Group, 2012). Differences in estimates of benefits and harms of mammography screening may be due to differences in delivered interventions, data sources used, the choice of denominators, and length of follow-up (Beral et al, 2011; Paci and EUROSCREEN Working Group, 2012).

The emergence of new treatments, roughly concurrent with the introduction of screening, also improved survival from breast cancer. A comparison of breast cancer-specific survival by Nottingham Prognostic Index (NPI) showed a substantial improvement in survival across all prognostic categories between cancers diagnosed in 1980–86 and 1990–99 (Blamey et al, 2007). The reduction in breast cancer mortality observed in the UK after the introduction of mammography screening is therefore likely to represent a combination of improved treatment and early detection due to screening. However, the contribution of each factor is not easily distinguishable (Blanks et al, 2000), and may result in inflated estimates of the impact of screening.

Estimates from previous trials and European studies may not reflect the outcomes expected from screening in the UK, potentially due to differences in screening delivery (e.g., triennial screening provided in the UK vs biennial in most other European countries), healthcare provision, and underlying risk. The objective of this study was to use a Markov simulation model to estimate the impact of mammography screening on breast cancer mortality and overdiagnosis under various screening scenarios with varying screening frequency and age at invitation, in order to provide insight on (1) the benefits and harms of mammography screening in the UK, (2) the likely impact of the most recent age extension to the NHSBSP and (3) the potential for further changes in age range and screening frequency. Our results were also compared and contrasted with those reported in the two major reviews (Paci and EUROSCREEN Working Group, 2012; Marmot et al, 2013).

MATERIALS AND METHODS

Model structure. We constructed a Markov model to compare the number of breast cancer diagnoses, breast cancer deaths, and life-years accumulated under various screening strategies. The model included 13 states: healthy, preclinical non-progressive in situ, preclinical progressive in situ, preclinical invasive, diagnosed in situ, and diagnosed invasive breast cancer by NPI category (excellent, good, moderate 1, moderate 2, poor, and very poor), death from BC, and death from other causes (Figure 1). The division of in situ breast cancers into progressive and non-progressive was aimed at facilitating the interpretation of mean sojourn time estimates for in situ breast cancers that eventually progress into invasive breast cancers and those that do not. At the start of each simulation, women could be healthy or in any preclinical cancer state. At the end of each monthly cycle, women in these states could stay in their current state or progress. Once diagnosed with breast cancer, women could only die from breast cancer, die from competing causes, or survive.

Model parameters. Details are provided in Supplementary Materials. In summary, parameters used for calculating transition probabilities between states within monthly cycles were extracted directly or estimated from various sources (Moss et al, 1993;

Blamey et al, 2007; Gunsoy et al, 2012; Office for National Statistics, http://www.ons.gov.uk; NHS Breast Screening Programme, 2012) and included parameters for screening, breast cancer incidence and progression, breast cancer prognosis, and mortality (Supplementary Tables 1, 2, and 3).

Screening parameters included the sensitivity of mammography screening and the uptake at first invitation, at subsequent invitation for previous attenders, and at subsequent invitation for previous non-attenders (Moss et al, 1993; Gunsoy et al, 2012; NHS Breast Screening Programme, 2012). These parameters were defined in relation to age and were assumed to be independent of screening frequency (Supplementary Table 1). The relative risk of radiation-induced breast cancer was calculated using an excess relative risk model (Preston et al, 2002) with assumed exposure of 4.5 mGy per two-view mammogram across all ages (Young, 2002; Young et al, 2005), and a lag period of 10 years between exposure and induction of breast cancer.

The prognosis of invasive breast cancer was estimated using (1) the distribution of NPI at diagnosis for screen-detected and interval cancers in a population-offered screening and for cancers diagnosed in a population not offered screening (Breast Screening Frequency Trial Group, 2002; West Midlands Cancer Intelligence Unit, 2007; Wishart et al, 2008; Allgood et al, 2011) (personal communication (Moss (1999))), (2) 10-year breast cancer-specific survival by NPI obtained for cancers detected between 1990–99 (Blamey et al, 2007), and (3) long-term breast cancer relative survival estimates (Taylor et al, 2003) (Supplementary Table 3). In situ breast cancer prognosis was determined using 10-year breast cancer-specific survival of 98.2% reported in an overview of randomised trials of radiotherapy of ductal carcinoma in situ (Early Breast Cancer Trialists' Collaborative Group et al, 2010). Survival was further adjusted using the hazard ratio for screen-detection estimated after adjusting for nodal status, tumour size, tumour grade, chemotherapy, and hormone therapy use (Wishart et al, 2010). Age-specific mortality rates from competing causes for 2008 were obtained from ONS mortality statistics (Office for National Statistics, 2008).

Breast cancer progression was modelled through three screen-detectable states. We assumed that progressive in situ breast cancers could not be diagnosed in the absence of screening, that non-progressive in situ breast cancers were detectable clinically in the absence of screening, and that all preclinical invasive breast cancers had a mandatory progressive in situ precursor. Over-diagnosed cancers were not represented explicitly in the model, and thus were not assigned specific breast cancer survival parameters. The mean sojourn time for women aged 40–49 years was obtained from a UK-based study (Gunsoy et al, 2012) and was estimated in other age groups using a cohort simulation model: the age-specific mean sojourn time in each breast cancer state was obtained by calibration of the age-specific incidence of invasive and in situ breast cancer predicted by our model to that observed in the cohort between 1971 and 2010, as calculated from ONS cancer registrations and population estimates (Office for National Statistics, www.ons.gov.uk). Model predictions were validated against historical overall and screen-detected incidence of breast cancer (details provided in Supplementary Materials). The age-specific mean sojourn time in preclinical breast cancer states increased with age. The mean sojourn time was longest in the non-progressive in situ state, increasing from 1.8 years in women aged 40–49 to 13 years in women aged 60 and older. In contrast, the mean sojourn time in the progressive in situ state was short (Table 1). The probability of being screened in each 5-year period and the incidence of screen-detected invasive and in situ breast cancer were calculated using data from the NHSBSP annual statistics bulletins (NHS Breast Screening Programme, 2012). The underlying incidence of breast cancer was obtained by simulating an identical cohort in the absence of screening (Supplementary Table 2). This was performed in OpenBUGS 3.2 (Lunn et al, 2009).
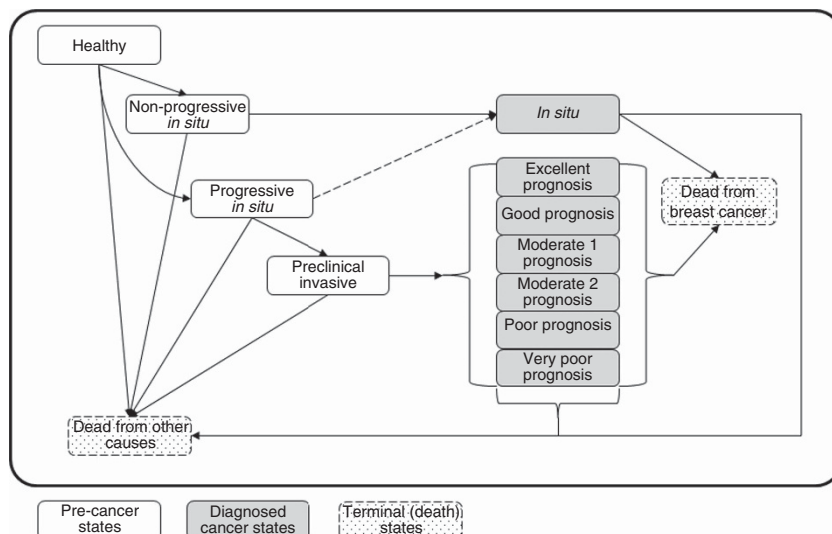
Figure 1. The 13-state Markov simulation model for the evaluation of mammography screening. Boxes indicate health states. White boxes are pre-cancer states including healthy and three preclinical breast cancer states: preclinical non-progressive *in situ*, preclinical progressive *in situ*, and preclinical invasive. Grey boxes group seven diagnosed breast cancer states that were diagnosed *in situ*, and diagnosed invasive breast cancer by six NPI categories. Arrows indicate possible transitions between states. The dotted arrow indicates a transition possible only through screen-detection. Dotted boxes are terminal states for death from breast cancer and from other causes.

Table 1. Estimates of age-specific mean sojourn time: results of a cohort simulation model using data on breast cancer incidence, screening, and competing mortality

Age group	Mean sojourn time, years (95% Bayesian credible interval)		
	Non-progressive *in situ*	Progressive *in situ*	Invasive
40–49	1.8 (0.7, 4.2)	0.11 (0.06, 0.19)	0.9 (0.7, 1.3)
50–59	4.6 (2.2, 11.4)	0.18 (0.08, 0.32)	2.2 (1.7, 2.8)
60+	13.1 (5.7, 34.1)	0.27 (0.11, 0.43)	5.3 (4.2, 6.9)

Outcomes. The number of breast cancer deaths predicted from age 40 to 85 years for each screening scenario was compared with that predicted in the absence of screening. We defined an overdiagnosed cancer as a screen-detected breast cancer, which would not have been diagnosed in the absence of screening by age 85 years. Breast cancer overdiagnosis in screening scenarios was estimated as the difference between the cumulative incidence of invasive and *in situ* breast cancer in each scenario and that in the absence of screening, using three different denominators (as in (Biesheuvel *et al*, 2007)); (1) all breast cancers diagnosed in the absence of screening from age 40 to 85 years, (2) cancers diagnosed in the screening period, and (3) all screen-detected breast cancers. For validation purposes, the proportion of screen-detected cancers among women eligible for screening was calculated for triennial screening between ages 50–70 and compared with the proportion reported by the NHSBSP (NHS Breast Screening Programme, 2012).

Scenarios. The following screening strategies in women followed from age 40 to 85 years were considered as follows:

1. No screening,
2. Triennial screening in women aged 50–70,
3. Triennial screening in women aged 47–73,
4. Triennial screening in women aged 40–73,
5. A combination of annual screening from age 40 to 47 years followed by triennial screening to age 73,
6. Annual screening in women aged 47–73,
7. And annual screening in women aged 40–73.

The duration of follow-up, from age 40 to 85 years, was selected to balance (1) adequate length of follow-up to allow for a compensatory drop in incidence (i.e., the fall in breast cancer incidence following the end of screening), (2) dilution of breast cancer mortality effects due to extended follow-up, and (3) reliability in input parameters. Estimates of population mortality and incidence were unreliable for women over the age of 85 years, and data from trials suggest that the effect of screening on breast cancer mortality lasts for around 10 years (Marmot *et al*, 2013).

Sensitivity analysis. We performed a series of sensitivity analyses to assess the impact of higher and lower uptake of screening, and of uncertainty in estimates of sensitivity and mean sojourn time. Scenarios included: (1) 100% uptake, (2) 15% reduced overall uptake, (3) 20% increased sensitivity, (4) 20% reduced sensitivity, (5) long mean sojourn times, and (6) short mean sojourn times. Sensitivity analyses were performed for triennial screening in women aged 47–73 only.

Calculations. In both the main and sensitivity analyses, we performed 200 sets of 10 000 iterations in each scenario. Breast cancer mortality reduction and overdiagnosis due to screening was then calculated in each set, for which we report the mean and 95% confidence interval. These simulations were performed in TreeAge Pro 2011 (TreeAge Software, Williamstown, MA, USA).

RESULTS

The predicted breast cancer mortality reduction due to screening ranged from 15.9% to 36.7% of all breast cancer deaths from age 40 to 85 years for different scenarios (Table 2). Estimates of overdiagnosis ranged from 4.3% to 8.9% of all invasive and *in situ* breast cancers diagnosed from age 40 to 85 years (Table 3). More frequent screening was associated with higher reductions in breast cancer mortality and higher risk of overdiagnosis with around a two-fold increase in both when moving from triennial to annual screening.

Breast cancer overdiagnosis varied according to the chosen denominator. In relation to screen-detected cancers, overdiagnosis was similar across scenarios. For the two remaining denominators, triennial screening from 50–70 years was consistently estimated to have the lowest risk of overdiagnosis, 4.3% of all breast cancers diagnosed from age 40 to 85 and 6.7% of cancers detected within the screening period. Annual screening from 40–73 years yielded the highest estimated risk of overdiagnosis, 8.9% and 10.1% relative to all cancers diagnosed from age 40 to 85 and to cancers diagnosed within the screening period, respectively. Compared with triennial screening from 47–73 years, the risk of overdiagnosis in relation to all breast cancers from age 40 to 85 years was significantly higher in annual screening and when annual screening from 40–47 years was added to annual or triennial screening from 47–73 years.

Absolute numbers of breast cancer deaths avoided, breast cancer cases overdiagnosed, and life-years gained are presented in Table 4. More frequent screening consistently resulted in more breast cancer deaths avoided and more breast cancer cases overdiagnosed. The mean ratio of breast cancer deaths avoided and cases overdiagnosed ranged from 0.8 to 1.1 across interventions, and was highest for annual screening from 47–73 years and lowest for annual screening from age 40 followed by triennial screening from 47–73 years. Significant differences in the ratio of breast cancer

deaths avoided per case overdiagnosed were not observed in most cases. More frequent screening or an increase in age range was also associated with increases in overall life-years gained per 10 000 women invited for screening, although average life-years gained per screening round decreased.

The incremental gain in breast cancer mortality reduction and increase in overdiagnosis from age 40 to 85 years estimated from an age extension from 50–70 to 47–73 years were significant, 2.3% (1.3%, 3.4%) and 1.3% (0.8%, 1.8%), respectively (Table 5). In absolute terms, this corresponded to five (3, 7) extra breast cancer deaths avoided and 14 (9, 19) extra cases overdiagnosed, a ratio of 0.4 (0.1, 0.8). In comparison, the predicted incremental reduction in breast cancer mortality from an increase in screening frequency from triennial to annual was 20.3% (19.3%, 21.3%), with an incremental increase in overdiagnosis of 2.0% (1.5%, 2.5%), corresponding to a ratio of 1.9 breast cancer deaths avoided per case overdiagnosed. The addition of triennial screening from 40–47 years to 47–73 years was predicted to lead to an increment in breast cancer mortality reduction of 2.2% (1.0%, 3.4%) and in overdiagnosis of 0.3% (− 0.3%, 0.8%) between the ages of 40 to 85 years, corresponding to a mean ratio of 1.6 breast cancer deaths avoided per case overdiagnosed compared with triennial screening between ages 47–73. For the addition of annual screening from 40–47 years, increments were 3.5% (2.3%, 4.7%) for breast cancer mortality reduction and 1.3% (0.8%, 1.8%) for overdiagnosis, corresponding to a ratio of 0.5 breast cancer deaths avoided per case overdiagnosed compared with triennial screening between ages 47–73.

Sensitivity analysis. Full compliance to screening was associated with significant increases in breast cancer mortality reduction and overdiagnosis, whereas reduced uptake was associated with significant decreases in both outcome measures. High sensitivity was also associated with small increases in both outcome measures, but to a much smaller extent than full compliance. Longer sojourn times (upper limits reported in Table 1) did not modify mortality reduction, but was associated with an increase in overdiagnosis of nearly two-fold. In contrast, shorter mean sojourn times (lower limits reported in Table 1) were associated with a significant decrease in both outcomes (Table 6). In terms of absolute number, modifying uptake or sensitivity resulted in similar ratios of breast cancer deaths avoided per cases overdiagnosed than in the base case, whereas longer mean sojourn times resulted in a reduction of this ratio and shorted mean sojourn times resulted in an increase of this ratio (Supplementary Table 5).

Table 2. Breast cancer mortality reduction due to screening compared with no screening in a cohort of British women followed up from age 40 to 85 years for various screening age ranges and frequencies

Scenario (frequency and age range)	Breast cancer mortality reduction
	Mean (95% CI)
Triennial, 50–70	15.9 (15.0, 16.8)
Triennial, 47–73	18.1 (17.3, 19.0)
Triennial, 40–73	20.1 (19.2, 21.1)
Combination[a]	21.3 (20.5, 22.1)
Annual, 47–73	35.0 (34.2, 35.7)
Annual, 40–73	36.7 (35.9, 37.5)

Abbreviation: CI = confidence interval.
[a]Annual screening from age 40 followed by triennial screening from age 47 to 73.

Table 3. Breast cancer overdiagnosis due to screening compared with no screening in a cohort of British women followed up from age 40 to 85 years for various screening age ranges and frequencies

	Breast cancer overdiagnosis due to screening as percent of		
	All breast cancers[a] diagnosed from age 40 to 85	Breast cancers[a] diagnosed in screening period	Screen-detected breast cancers[a]
Scenario (frequency and age range)	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)
Triennial, 50–70	4.3 (3.8, 4.8)	6.7 (6.0, 7.4)	11.8 (10.5, 13.0)
Triennial, 47–73	5.6 (5.1, 6.1)	7.1 (6.5, 7.7)	12.5 (11.4, 13.6)
Triennial, 40–73	5.8 (5.2, 6.3)	6.8 (6.2, 7.5)	12.2 (11.1, 13.3)
Combination[b]	6.9 (6.4, 7.4)	8.0 (7.4, 8.6)	13.5 (12.6, 14.5)
Annual, 47–73	7.6 (7.1, 8.1)	9.6 (9.0, 10.1)	11.7 (11.0, 12.4)
Annual, 40–73	8.9 (8.3, 9.5)	10.1 (9.5, 10.7)	12.7 (12.0, 13.5)

Abbreviation: CI = confidence interval.
[a]Both *in situ* and invasive breast cancers were included in estimates.
[b]Annual screening from age 40 followed by triennial screening from age 47 to 73.

Table 4. Breast cancer deaths avoided and breast cancer cases overdiagnosed in a cohort of British women followed up from age 40 to 85 years per 10 000 women invited to screening, for selected-screening schedules

Scenario (frequency and age range)	Breast cancer deaths avoided[a] (per 10 000)	Breast cancer cases overdiagnosed[b] (per 10 000)	Ratio (per case overdiagnosed)	Incremental life-years gained (per 10 000)	Incremental life-years gained (per screening round)
	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)	Mean	Mean
Triennial, 50–70	41 (39, 44)	45 (40, 50)	0.9 (0.8, 1.1)	476	59.5
Triennial, 47–73	47 (44, 49)	58 (53, 63)	0.8 (0.7, 0.9)	498	49.8
Triennial, 40–73	52 (50, 55)	60 (54, 66)	0.9 (0.8, 1.0)	600	50.0
Combination[c]	55 (53, 58)	72 (67, 78)	0.8 (0.7, 0.9)	691	40.7
Annual, 47–73	90 (88, 92)	80 (75, 85)	1.1 (1.0, 1.2)	1021	37.8
Annual, 40–73	95 (92, 97)	93 (87, 99)	1.0 (0.9, 1.1)	1152	33.9

Abbreviation: CI = confidence interval.
[a]Baseline in the absence of screening: 256 (254–259) deaths per 10 000.
[b]Baseline in the absence of screening: 1054 (1050–1058) diagnoses per 10 000.
[c]Annual screening from age 40 followed by triennial screening from age 47 to 73.

Table 5. Estimated relative and absolute incremental gains for selected extensions of the UK breast screening programme.

Outcome	Relative increment (%) Mean (95% CI)	Absolute increment (cases per 10 000) Mean (95% CI)
Extending age range from 50–70 to 47–73		
Breast cancer mortality reduction	2.3 (1.3, 3.4)	14 (9, 19)
Breast cancer overdiagnosis	1.3 (0.8, 1.8)	5 (3, 7)
Ratio (per case overdiagnosed)		0.4 (0.1, 0.8)
Extending from triennial 47–73 to triennial 40–73		
Breast cancer mortality reduction	2.2 (1.0, 3.4)	5 (2, 7)
Breast cancer overdiagnosis	0.3 (−0.3, 0.8)	3 (−3, 9)
Ratio (per case overdiagnosed)		1.6 (0.2, not applicable)
Extending from triennial 47–73 to annual 40–47 plus triennial 47–73		
Breast cancer mortality reduction	3.5 (2.3, 4.7)	7 (5, 10)
Breast cancer overdiagnosis	1.3 (0.8, 1.8)	15 (9, 20)
Ratio (per case overdiagnosed)		0.5 (0.2, 1.1)
Increasing frequency from triennial to annual 47–73		
Breast cancer mortality reduction	20.3 (19.3, 21.3)	43 (40, 45)
Breast cancer overdiagnosis	2.0 (1.5, 2.5)	22 (17, 28)
Ratio (per case overdiagnosed)		1.9 (1.5, 2.6)

Abbreviation: CI = confidence interval. Numbers in this table do not match increments in Tables 3, 4, and 5 as denominators are different (value in the comparator rather than value for no screening).

Table 6. Breast cancer mortality reduction and overdiagnosis due to screening women every three years in women aged 47–73 in a cohort of British women followed up from age 40 to 85 years. Results of sensitivity analyses.

	Breast cancer mortality reduction % Mean (95% CI)	Overdiagnosis of all breast cancer % Mean (95% CI)
Base case	18.1 (17.3, 19.0)	5.6 (5.1, 6.1)
Modified uptake		
100%	24.3 (23.3, 25.2)	7.5 (6.9, 8.1)
15% decreased	12.2 (11.1, 13.3)	3.9 (3.3, 4.5)
Modified sensitivity		
100%	20.0 (19.0, 21.0)	6.2 (5.6, 6.8)
Modified mean sojourn time		
Long sojourn times[a]	18.0 (17.1, 18.9)	9.7 (9.1, 10.3)
Short sojourn time[b]	14.8 (13.9, 15.7)	3.1 (2.6, 3.5)

Abbreviation: CI = confidence interval.
[a]Upper limits reported in Table 1.
[b]Lower limits reported in Table 1.

DISCUSSION

Our model predicted that triennial screening in women aged 47–73 years resulted in an 18.1% (17.3%, 19.0%) reduction in breast cancer mortality and a 5.6% (5.1%, 6.1%) increase in breast cancer incidence between the ages of 40 and 85 years. The probability of a screen-detected cancer being overdiagnosed was 12.5% (11.4%, 13.6%).

The age range in the NHSBSP was recently extended from 50–70 to 47–73 years; this extension is currently being introduced on a trial basis over two or three screening cycles (NHS Breast Screening Programme, 2010). The predicted absolute impact of this extension was five extra breast cancer deaths avoided and 14 extra cases overdiagnosed per 10 000 women invited. Although our predictions for this extension point toward more breast cancer deaths avoided, these need to be considered carefully against the excess harms, as the ratio of 0.4 incremental breast cancer deaths avoided per incremental case overdiagnosed from our model was low.

Our study estimated a 20.3% incremental breast cancer mortality reduction when increasing screening frequency from triennial to annual in women aged 47–73. This was in the same order of magnitude as an estimate from a model-based prediction study using data from the UK Breast Screening Frequency Trial, which estimated a 17% reduction in the risk of breast cancer death in annual compared with triennial screening in women aged 50–62

years at entry with lifetime follow-up (van Ravesteyn et al, 2011). However, this increase in frequency was also associated with a substantial increase in overdiagnosis.

The incremental gain in breast cancer mortality reduction from further extending the screening age range to include women from age 40 years was 2.2% for triennial screening and 3.5% for annual screening between the ages of 40 and 47 years. This increased benefit was lower than the two-fold increase (35.0% related to 18.1% (Table 2)) observed in women aged 47–73 when increasing the frequency of screening from triennial to annual, and may be due to lower breast cancer incidence and shorter sojourn times in younger women. However, it may be that much higher gains in life expectancy are achieved when screening women between the ages of 40–47 years. In our model, extending the age range for triennial screening to include women aged 47–73 resulted in an increment of 22 life-years gained compared with an age range of 50–70 (calculated from Table 4). This would equate to 11 life-years per extra screening round (two screening rounds added). In comparison, a further extension of the age range to include women aged 40–73 yielded an estimated gain of 102 life-years compared with an age range of 47–73, which is equivalent to 51 life-years per extra screen (two extra rounds added). Although our estimates relating to the inclusion of younger women in screening programmes are mostly applicable to the UK, the implications of these findings may also be relevant to other European countries.

A recent high-profile independent review of benefits and harms of breast cancer screening in the UK provided estimates of breast cancer mortality of 20% (11%, 27%) based on screening trials with an average follow-up of 13 years, and overdiagnosis as a proportion of cancers detected in the screening period of 19% (15%, 23%), although a high degree of uncertainty in the latter was highlighted by Marmot et al (2013). The estimate of mortality reduction was assumed to be applicable to breast cancer deaths between the ages of 55–79 years. An equivalent estimate from our model of mortality reduction from age 55–79 years from triennial screening in women aged 50–70 was 18.8% (17.7%, 20.0%), similar to the estimate suggested by the review. In absolute terms, our estimate of 41 breast cancer deaths avoided per 10 000 invited for screening was similar to the estimate of 44 estimated in the review. There are also possible reasons for a lower mortality reduction due to current population screening compared with that observed in the randomised trials including changes in the prognosis of all breast cancers due to both improved treatment and increased awareness. In addition, the majority of the trials had shorter screening intervals and on average higher uptake (81% weighted average (Marmot et al, 2013)), than currently observed in the NHSBSP (73% (NHS Breast Screening Programme, 2012)).

The estimate of overdiagnosis of 19% in the review was expressed in relation to cancers detected in the screening period, and hence is comparable with our estimate of 6.7% for triennial screening from age 50–70 years. In absolute terms, the estimate in the review of 129 cases overdiagnosed per 10 000 women invited for screening was much higher than our estimate of 45 cases overdiagnosed. The estimate in the review was based on data from three trials with screening intervals of 12 months or 18–24 months with very high uptake (weighted average 84% (Marmot et al, 2013)), and one including only women aged 40–49 years. Our model predicted that higher screening frequency and higher uptake and wider age range was associated with higher overdiagnosis. Other possible reasons for the difference include an increase in opportunistic screening in women who are not invited for or do not attend screening, increasing the incidence of breast cancer in the absence of organised screening estimated in our model, leading to lower breast cancer mortality reduction and overdiagnosis attributable to screening invitation, and differences in breast cancer risk and progression between the UK population and those included in trials. For the latter, previous studies have reported

lower estimates of mean sojourn times when using UK data compared with other countries (Yen et al, 2003; Gunsoy et al, 2012), also observed in the age-specific mean sojourn time estimates in our model compared with those obtained from Swedish data (Duffy et al, 1997): shorter mean sojourn times would ultimately result in lower estimates of overdiagnosis, as observed in our sensitivity analysis.

The EUROSCREEN working group, which reviewed evidence for population screening in Europe, estimated that for every 10 000 women screened biennially for 20 years from age 50–70 years, 90 breast cancer deaths were avoided and 40 cases overdiagnosed, a ratio of 2.3 breast cancer deaths avoided per case overdiagnosed (Paci and EUROSCREEN Working Group, 2012). This estimate was based on assuming 100% uptake, and exposure to a total of 10 biennial screens between the ages of 50 and 70 years. Using these assumptions in our model, and assuming a similar distribution of cancer prognosis as in annual screening (Wai et al, 2005), for screened women followed up from age 50 to 80, estimates were 34.7% (33.8%, 35.7%) for breast cancer mortality reduction, slightly below the range of 38–48% estimated in the review, and 5.4% (4.9%, 5.9%) for overdiagnosis relative to all breast cancers diagnosed from age 50 to 80 years, within the range of 1–10% reported in the review, but slightly lower than the reported mean of 6.5%. In absolute terms, this corresponded to 75 cases of breast cancer deaths avoided and 51 cases overdiagnosed between the ages of 50 and 80 years per 10 000 screened, compared with 90 and 40, respectively, reported in the review.

Our study design had several strengths for estimating the impact of screening. Firstly, the difference in prognosis in screen-detected cases was intentionally not deduced from screening trials or observational studies: instead, we used the distribution of NPI at diagnosis for screen-detected, interval, and unscreened cases and modelled survival from diagnosis using data on 10-year and long-term breast cancer-specific survival: this made possible the evaluation of various screening frequencies and age ranges. Also, our model did not rely on the comparison of data from two potentially different populations. Instead, predicted mortality and incidence estimates with and without screening were modelled in the same cohort using the same parameters and assumptions.

Despite this, important limitations need to be considered. The natural history of breast cancer was modelled through three preclinical states where in situ breast cancers were divided into progressive and non-progressive cancers, and all invasive breast cancers were assumed to have a mandatory in situ precursor. These assumptions, although shown to have a good fit to screening data (Yen et al, 2003; Gunsoy et al, 2012), may not adequately portray the progression of breast cancer. The use of historic incidence and survival data may not reflect current observed mortality rates. The choice of cohort may have altered absolute estimates of the impact of screening. To address this, we also calculated the absolute impact of screening using breast cancer incidence and mortality estimates for 2010 in the UK (see Supplementary Materials for details). Although absolute numbers were higher due to a higher underlying incidence of breast cancer, the ratio of breast cancer deaths avoided per case overdiagnosed were equal to those observed in our selected cohort (Supplementary Table 4). Simulated women were followed up until age 85 rather than until death: this was due to high variability in breast cancer incidence and all-cause mortality estimates in women beyond the age of 85 years in the UK. This, however, gives a minimum of 12 years follow-up after the end of screening, which is above the minimum of 5–10 years quoted by Marmot et al (2013), and comparable with the 13-year follow-up accrued in most studies included in meta-analyses of randomised screening trials. Uncertainty in key parameters including mean sojourn times and sensitivity may affect predictions of overdiagnosis and mortality reduction. We addressed this by performing a sensitivity analysis to evaluate the impact of

changes in these parameters: only extreme changes in mean sojourn time resulted in meaningful changes in outcomes. Estimates of breast cancer mortality reduction and overdiagnosis were 18% and 10% for long sojourn times, and 15% and 3% for short sojourn times, respectively. In addition, the estimation of breast cancer mortality reduction based on surrogate measures, such as the distribution of prognosis at detection, may have led to inaccurate estimates. However, this was a necessary feature for investigating the impact of screening frequency, age range, and uptake. In addition, due to a lack of reliable estimates of incidence and mortality risk differences between screening attenders and non-attenders, it was assumed that women who do and do not attend screening have the same risk of developing and dying of breast cancer. However, screen-detected cancers did have slightly better NPI-adjusted survival in our model, following estimates reported in the PREDICT study (Wishart et al, 2010), which may have accounted for some differences in the risk of dying from breast cancer between screening attenders and non-attenders.

CONCLUSIONS

Our model predicted that breast cancer mortality reduction and overdiagnosis due to screening from age 47–73 years in the UK were 18% (17%, 19%) and 5.6% (5.1%, 6.1%), respectively based on mortality and incidence between ages 40 and 85. The probability of overdiagnosis for a screen-detected case was 12.5% (11.4%, 13.6%), meaning that one in eight screen-detected cancers were over-diagnosed. Some of the discrepancies from other studies could be explained by differences in denominators or length of follow-up, and by more frequent screening and higher uptake in previous trials and studies, associated with increased breast cancer mortality reduction and increased overdiagnosis in our simulations. When screening was extended to women aged 40–47 years, the increment in life-years per added screening round was much higher than in other age groups, whereas only small incremental increases in breast cancer mortality reduction were observed. Future evaluative studies of screening should consider the impact of frequency and age range in obtained estimates. Further estimates of benefits, harms, and cost-effectiveness of screening in the general UK population need to be undertaken to evaluate the impact of recent and further extensions in the UK national screening programme.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

NBG participated in the conception and design of the study, the development of the methodology and the interpretation of results, performed the analyses, and drafted the manuscript. MG-C participated in the conception of the study, the interpretation of results, and reviewed the manuscript. SMM participated in the conception and design of the study, the interpretation of results, revised and reviewed the manuscript, and supervised the study. All authors read and approved the final manuscript.

REFERENCES

Allgood PC, Duffy SW, Kearins O, O'Sullivan E, Tappenden N, Wallis MG, Lawrence G (2011) Explaining the difference in prognosis between screen-detected and symptomatic breast cancers. Br J Cancer 104: 1680–1685.

Beral V, Alexander M, Duffy S, Ellis IO, Given-Wilson R, Holmberg L, Moss SM, Ramirez A, Reed MW, Rubin C, Whelehan P, Wilson R, Young KC (2011) The number of women who would need to be screened regularly by mammography to prevent one death from breast cancer. J Med Screen 18: 210–212.

Biesheuvel C, Barratt A, Howard K, Houssami N, Irwig L (2007) Effects of study methods and biases on estimates of invasive breast cancer overdetection with mammography screening: a systematic review. Lancet Oncol 8: 1129–1138.

Blamey RW, Ellis IO, Pinder SE, Lee AHS, Macmillan RD, Morgan DAL, Robertson JFR, Mitchell MJ, Ball GR, Haybittle JL, Elston CW (2007) Survival of invasive breast cancer according to the Nottingham Prognostic Index in cases diagnosed in 1990–1999. Eur J Cancer 43: 1548–1555.

Blanks RG, Moss SM, McGahan CE, Quinn MJ, Babb PJ (2000) Effect of NHS breast screening programme on mortality from breast cancer in England and Wales, 1990–8: comparison of observed with predicted mortality. BMJ 321: 665–669.

Breast Screening Frequency Trial Group (2002) The frequency of breast cancer screening: results from the UKCCCR randomised trial. United Kingdom Co-ordinating Committee on Cancer Research. Eur J Cancer 38: 1458–1464.

de Gelder R, Heijnsdijk EAM, van Ravesteyn NT, Fracheboud J, Draisma G, de Koning HJ (2011) Interpreting overdiagnosis estimates in population-based mammography screening. Epidemiol Rev 33: 111–121.

Duffy SW, Day NE, Tabar L, Chen H-H, Smith TC (1997) Markov models of breast tumor progression: some age-specific results. J Natl Cancer Inst Monogr 22: 93–97.

Early Breast Cancer Trialists' Collaborative GroupCorrea C, McGale P, Taylor C, Wang Y, Clarke M, Davies C, Peto R, Bijker N, Solin L, Darby S (2010) Overview of the randomized trials of radiotherapy in ductal carcinoma in situ of the breast. J Natl Cancer Inst Monogr 2010: 162–177.

Gunsoy NB, Garcia-Closas M, Moss SM (2012) Modelling the overdiagnosis of breast cancer due to mammography screening in women aged 40 to 49 in the United Kingdom. Breast Cancer Res 14: R152.

Lunn D, Spiegelhalter D, Thomas A, Best N (2009) The BUGS project: Evolution, critique and future directions. Stat Med 28: 3049–3067.

Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M (2013) The benefits and harms of breast cancer screening: an independent review. Br J Cancer 108: 2205–2240.

Moss S (1999) A trial to study the effect on breast cancer mortality of annual mammographic screening in women starting at age 40. Trial Steering Group. J Med Screen 6: 144–148.

Moss SM, Coleman DA, Ellman R, Chamberlain J, Forrest AP, Kirkpatrick AE, Thomas BA, Price JL (1993) Interval cancers and sensitivity in the screening centres of the UK trial of early detection of breast cancer. Eur J Cancer 29A: 255–258.

NHS Breast Screening Programme (2010) Age extension full randomised control trial. http://www.clinicaltrials.gov/ct2/show/NCT01081288 (accessed 20 November 2013).

NHS Breast Screening Programme (2012) Statistics bulletin: Breast screening programme, England. http://www.cancerscreening.nhs.uk/ (accessed 20 November 2013).

Office for National Statistics (2008) Mortality statistics: Deaths registrations (Series DR). http://www.ons.gov.uk/ (accessed 20 November 2013).

Paci E. EUROSCREEN Working Group (2012) Summary of the evidence of breast cancer service screening outcomes in Europe and first estimate of the benefit and harm balance sheet. J Med Screen 19(Suppl 1): 5–13.

Preston DL, Mattsson A, Holmberg E, Shore R, Hildreth NG, Boice JD (2002) Radiation effects on breast cancer risk: a pooled analysis of eight cohorts. Radiat Res 158: 220–235.

Puliti D, Miccinesi G, Paci E (2011) Overdiagnosis in breast cancer: design and methods of estimation in observational studies. Prev Med 53: 131–133.

Taylor R, Davis P, Boyages J (2003) Long-term survival of women with breast cancer in New South Wales. Eur J Cancer 39: 215–222.

van Ravesteyn NT, Heijnsdijk EAM, Draisma G, de Koning HJ (2011) Prediction of higher mortality reduction for the UK breast screening frequency trial: a model-based approach on screening intervals. *Br J Cancer* **105**: 1082–1088.

Wai ES, D'Yachkova Y, Olivotto IA, Tyldesley S, Phillips N, Warren LJ, Coldman AJ (2005) Comparison of 1- and 2-year screening intervals for women undergoing screening mammography. *Br J Cancer* **92**: 961–966.

West Midlands Cancer Intelligence Unit (2007) Breast Cancer Clinical Outcome Measures (BCCOM) Project; analysis of the management of symptomatic breast cancers diagnosed in 2004; 3rd year report. http://www.associationofbreastsurgery.org.uk/media/2327/bccom_ar09.pdf (accessed 20 November 2013).

Wishart GC, Greenberg DC, Britton PD, Chou P, Brown CH, Purushotham AD, Duffy SW (2008) Screen-detected vs symptomatic breast cancer: is improved survival due to stage migration alone? *Br J Cancer* **98**: 1741–1744.

Wishart GC, Azzato EM, Greenberg DC, Rashbass J, Kearins O, Lawrence G, Caldas C, Pharoah PD (2010) PREDICT: a new UK prognostic model that predicts survival following surgery for invasive breast cancer. *Breast Cancer Res* **12**: R1.

Yen M-F, Tabár L, Vitak B, Smith RA, Chen H-H, Duffy SW (2003) Quantifying the potential problem of overdiagnosis of ductal carcinoma in situ in breast cancer screening. *Eur J Cancer* **39**: 1746–1754.

Young KC (2002) Radiation doses in the UK trial of breast screening in women aged 40–48 years. *Br J Radiol* **75**: 362–370.

Young KC, Burch A, Oduko JM (2005) Radiation doses received in the UK breast screening programme in 2001 and 2002. *Br J Radiol* **78**: 207–218.

Supplementary Information accompanies this paper on British Journal of Cancer website (http://www.nature.com/bjc)