Annual mammographic screening to reduce breast cancer mortality in women from age 40 years

Duffy, Stephen; Vulkan, Daniel; Cuckle, Howard; Parmar, Dharmishta; Sheikh, Shama; Smith, Robert

Published in:
Health technology assessment (Winchester, England)

DOI:
10.3310/hta24550

Publication date:
2020

Document Version
Publisher's PDF, also known as Version of record

NIHR | National Institute for Health Research

Journals Library

Check for updates

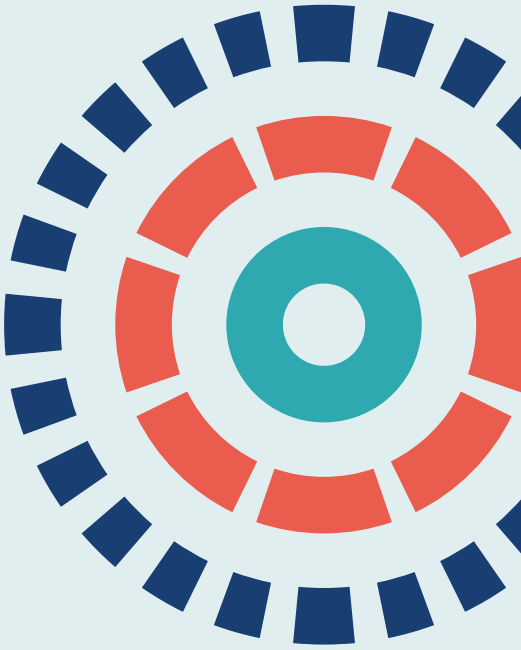
Health Technology Assessment

Volume 24 • Issue 55 • October 2020

ISSN 1366-5278

Annual mammographic screening to reduce breast cancer mortality in women from age 40 years: long-term follow-up of the UK Age RCT

Stephen Duffy, Daniel Vulkan, Howard Cuckle, Dharmishta Parmar, Shama Sheikh, Robert Smith, Andrew Evans, Oleg Blyuss, Louise Johns, Ian Ellis, Peter Sasieni, Chris Wale, Jonathan Myles and Sue Moss



DOI 10.3310/hta24550

Annual mammographic screening to reduce breast cancer mortality in women from age 40 years: long-term follow-up of the UK Age RCT

Stephen Duffy,[1*] Daniel Vulkan,[1] Howard Cuckle,[2] Dharmishta Parmar,[1] Shama Sheikh,[3] Robert Smith,[4] Andrew Evans,[5] Oleg Blyuss,[1] Louise Johns,[3] Ian Ellis,[6] Peter Sasieni,[3] Chris Wale,[1] Jonathan Myles[1] and Sue Moss[1]

[1]Centre for Cancer Prevention, Queen Mary University of London, London, UK
[2]Department of Obstetrics and Gynaecology, Tel Aviv University, Tel Aviv, Israel
[3]Cancer Prevention Group, King's College London, London, UK
[4]Cancer Screening, American Cancer Society, Atlanta, GA, USA
[5]Division of Cancer Research, University of Dundee, Dundee, UK
[6]Faculty of Medicine and Health Sciences, University of Nottingham, UK

*Corresponding author

Declared competing interests of authors: Peter Sasieni is a member of the National Institute for Health Research (NIHR) Clinical Trials Unit and Policy Research Unit and reports grants from the NIHR. Sue Moss reports grants from the Department of Health and Social Care, Cancer Research UK, the Medical Research Council, the Health Technology Assessment programme, the National Cancer Institute and the American Cancer Society during the conduct of the study.

Published October 2020
DOI: 10.3310/hta24550

This report should be referenced as follows:

Duffy S, Vulkan D, Cuckle H, Parmar D, Sheikh S, Smith R, *et al*. Annual mammographic screening to reduce breast cancer mortality in women from age 40 years: long-term follow-up of the UK Age RCT. *Health Technol Assess* 2020;**24**(55).

Health Technology Assessment is indexed and abstracted in *Index Medicus*/MEDLINE, *Excerpta Medica*/EMBASE, *Science Citation Index Expanded* (SciSearch®) and *Current Contents®*/ Clinical Medicine.

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 3.370

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, the Cochrane Library and Clarivate Analytics Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 10/57/55. The contractual start date was in July 2012. The draft report began editorial review in August 2019 and was accepted for publication in May 2020. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Abstract

Annual mammographic screening to reduce breast cancer mortality in women from age 40 years: long-term follow-up of the UK Age RCT

Stephen Duffy,[1*] Daniel Vulkan,[1] Howard Cuckle,[2] Dharmishta Parmar,[1] Shama Sheikh,[3] Robert Smith,[4] Andrew Evans,[5] Oleg Blyuss,[1] Louise Johns,[3] Ian Ellis,[6] Peter Sasieni,[3] Chris Wale,[1] Jonathan Myles[1] and Sue Moss[1]

[1]Centre for Cancer Prevention, Queen Mary University of London, London, UK
[2]Department of Obstetrics and Gynaecology, Tel Aviv University, Tel Aviv, Israel
[3]Cancer Prevention Group, King's College London, London, UK
[4]Cancer Screening, American Cancer Society, Atlanta, GA, USA
[5]Division of Cancer Research, University of Dundee, Dundee, UK
[6]Faculty of Medicine and Health Sciences, University of Nottingham, UK

*Corresponding author s.w.duffy@qmul.ac.uk

Background: There remains disagreement on the long-term effect of mammographic screening in women aged 40–49 years.

Objectives: The long-term follow-up of a randomised controlled trial that offered annual mammography to women aged 40–49 years. The estimation of the effect of these mammograms on breast cancer and other-cause mortality, and the effect on incidence, with implications for overdiagnosis.

Design: An individually randomised controlled trial comparing offering annual mammography with offering usual care in those aged 40–48 years, and thus evaluating the effect of annual screening entirely taking place before the age of 50 years. There was follow-up for an average of 23 years for breast cancer incidence, breast cancer death and death from other causes. We analysed the mortality and incidence data by Poisson regression and estimated overdiagnosis formally using Markov process models.

Setting: Twenty-three screening units in England, Wales and Scotland within the NHS Breast Screening Programme.

Participants: Women aged 39–41 years were recruited between 1990 and 1997. After exclusions, a total of 53,883 women were randomised to undergo screening (the intervention group) and 106,953 women were randomised to have usual care (the control group).

Interventions: The intervention group was invited to an annual breast screen with film mammography, two view at first screen and single view thereafter, up to and including the calendar year of their 48th birthday. The control group received no intervention. Both groups were invited to the National Programme from the age of 50 years, when screening is offered to all women in the UK.

Main outcome measures: The main outcome measures were mortality from breast cancers diagnosed during the intervention phase of the trial (i.e. before the first National Programme screen at 50 years), mortality from all breast cancers diagnosed after randomisation, all-cause mortality, mortality from causes other than breast cancer, and the incidence of breast cancer.

Results: There was a statistically significant 25% reduction in mortality from breast cancers diagnosed during the intervention phase at 10 years' follow-up (relative rate 0.75, 95% confidence interval 0.58 to 0.97; $p = 0.03$). No reduction was observed thereafter (relative rate 0.98, 95% confidence interval 0.79 to 1.22). Overall, there was a statistically non-significant 12% reduction (relative rate 0.88, 95% confidence interval 0.74 to 1.03; $p = 0.1$). The absolute benefit remained approximately constant over time, at one death prevented per 1000 women screened. There was no effect of intervention on other-cause mortality (relative rate 1.02, 95% confidence interval 0.97 to 1.07; $p = 0.4$). The intervention group had a higher incidence of breast cancer than the control group during the intervention phase of the trial, but incidence equalised immediately on the first National Programme screen at the age of 50–52 years.

Limitations: There was 31% average non-compliance with screening and three centres had to cease screening for resource and capacity reasons.

Conclusions: Annual mammographic screening at the age of 40–49 years resulted in a relative reduction in mortality, which was attenuated after 10 years. It is likely that digital mammography with two views at all screens, as practised now, could improve this further. There was no evidence of overdiagnosis in addition to that which already results from the National Programme carried out at later ages.

Future work: There is a need for research on the effects of modern mammographic protocols and additional imaging in this age group.

Trial registration: Current Controlled Trials ISRCTN24647151.

Funding: This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 24, No. 55. See the NIHR Journals Library website for further project information. Other funding in the past has been received from the Medical Research Council, Cancer Research UK, the Department of Health and Social Care, the US National Cancer Institute and the American Cancer Society.

Contents

List of tables

List of figures

List of abbreviations

AgeX	Age Extension	RCT	ransomised controlled trial
CI	confidence interval	RR	relative rate
GP	general practitioner	USPSTF	US Preventive Services Task Force
NHSBSP	NHS Breast Screening Programme		

Plain English summary

It is known that breast cancer screening with mammography (i.e. X-ray of the breasts) in women aged ≥ 50 years leads to a reduction in the number of deaths from breast cancer. In the UK, the NHS Breast Screening Programme offers regular screening to women aged 50–70 years. There is still some disagreement about the effect of such screening on the risk of death from breast cancer for those aged 40–49 years. There is also concern about overdiagnosis, that is, the finding of breast cancer that would not have been diagnosed in a woman's lifetime if she had not been screened.

This study recruited 160,921 women aged 39–41 years and randomly assigned one in three of the women to be offered annual mammographic screening from age 40 to 48 years. The women were followed up for occurrence of breast cancer, death from breast cancer and death from all other causes.

We found that the women who were offered the screening were 25% less likely to die of breast cancer in the first 10 years in the trial. This mortality reduction was reduced with later follow-up, with a 12% reduction after an average of 23 years. There was no effect of offering screening on death from other causes.

During the early years of the trial, the women offered screening had larger numbers of breast cancers diagnosed, but this excess disappeared after the first National Programme screen. This suggests that there is no overdiagnosis from screening those aged 40–49 years over and above that which already results from screening those aged ≥ 50 years.

Scientific summary

Background

The effect of mammographic screening on breast cancer mortality in women aged < 50 years has been a matter for discussion for several decades. A lesser effect of screening on breast cancer mortality has been observed for women aged < 50 years in randomised controlled trials, partly because this age group has more radiologically dense breast tissue than those aged ≥ 50 years, and possibly also because of the more rapid progression of cancers diagnosed in younger women.

There is continuing disagreement and uncertainty on the magnitude of the major desirable effect of screening in this age group, the reduction in breast cancer mortality and on some major adverse effects, notably overdiagnosis of breast cancer. In this context, overdiagnosis is the diagnosis of breast cancer as a result of screening that would not have occurred in the person's lifetime if they had not been screened.

Objectives

The primary objective was to determine the effect of annual mammographic screening on breast cancer mortality for those aged 40–49 years. We also aimed to estimate the effect on other-cause and all-cause mortality, and the effect on breast cancer incidence, to assess the implications for overdiagnosis of breast cancer.

Methods

A total of 160,921 women were randomised in a 1 : 2 ratio to the intervention group or the control group. After exclusions, the trial included 160,836 women who had data available for analysis. Recruitment took place between October 1990 and September 1997. Individual randomisation was performed, stratified by general practice so that one-third of the women in any practice were allocated to the intervention group. Women were aged 39–41 years at time of entry to the trial. The trial was conducted in 23 NHS Breast Screening Programme units in England, Wales and Scotland. Women in the intervention group were sent a letter of invitation and an information leaflet that clearly stated that the woman was being asked to participate in a research trial, and that her acceptance of the invitation was taken to be her informed consent to participate. Women in the intervention group were invited for annual mammography screening until the calendar year of their 48th birthday. At 50 years, both they and the women in the control group became eligible for 3-yearly invitation to screening as part of the NHS Breast Screening Programme, and received their first invitation between the ages of 50 and 52 years.

Screening in the trial was by two-view mammography at the first screen, with single-view mammography thereafter unless otherwise indicated. Mammograms were double-read. All women, including non-attenders, were reinvited annually unless they requested otherwise. Women who moved to areas that were not covered by the trial were not reinvited for screening as part of the trial, but were able to self-refer to either their previous or their nearest participating screening centre. Screening in three centres ceased prematurely (after four, five and six rounds) because of the inability of the centres to manage the additional workload with the available resources.

Women were flagged for follow-up by the NHS Central Register (a responsibility now belonging to the Office for National Statistics with data collation by NHS Digital), and the triallists were notified of all breast cancers, breast cancer deaths and deaths from all other causes up to 28 February 2017.

Mortality data from breast cancers, other causes and all causes were analysed by Poisson regression for significance testing between the intervention and the control groups, and for the estimation of relative rates and confidence intervals on these. In addition, we calculated Nelson–Aalen estimates of cumulative hazard.

The primary end point was mortality from breast cancers diagnosed in the intervention phase of the trial, before the first National Programme invitation. In estimating the effect on mortality from cancers diagnosed in the intervention period of the trial, there is a potential bias against the intervention because the intervention group will include deaths from cancers diagnosed at screening whose time of diagnosis would have been at or after the first NHS Programme invitation, and which would therefore not be included in the control group. This bias can be minimised by including cancers diagnosed at a contemporaneous screen at the end of the intervention period in both groups. We therefore performed a secondary analysis redefining the intervention period cancers as those diagnosed up to and including the first NHS Programme screen in both groups.

We compared incidence between the intervention and control groups before the first National Programme screen, up to and including the first National Programme screen and up to the final follow-up at 28 February 2017. We also derived tentative estimates of overdiagnosis using Markov process models.

Results

At 10 years, there was a statistically significant 25% reduction in mortality (relative rate 0.75, 95% confidence interval 0.58 to 0.97; $p = 0.03$). For $\geq$ 10 years, there was no reduction observed (relative rate 0.98, 95% confidence interval 0.79 to 1.22; $p = 0.9$). Overall, there was a 12% reduction in breast cancer mortality, which was not statistically significant (relative rate 0.88, 95% confidence interval 0.74 to 1.03; $p = 0.1$).

For the corresponding breast cancer mortality figures for the secondary analysis of cancers diagnosed up to and including the first NHS Programme screen in both groups, the 10-year results were identical to the primary analysis: a statistically significant 25% reduction in mortality (relative rate 0.75, 95% confidence interval 0.58 to 0.97; $p = 0.03$). For $\geq$ 10 years, a small, statistically non-significant reduction was observed (relative rate 0.95, 95% confidence interval 0.77 to 1.17; $p = 0.6$). At complete follow-up, there was a 14% reduction in breast cancer mortality that was of borderline statistical significance (relative rate 0.86, 95% confidence interval 0.73 to 1.01; $p = 0.07$).

After adjustment for selection bias, the effect of actually being screened was estimated as a statistically significant 34% reduction in breast cancer mortality up to 10 years after randomisation (relative rate 0.66, 95% confidence interval 0.46 to 0.95; $p = 0.02$), a statistically non-significant 2% reduction after 10 years (relative rate 0.98, 95% confidence interval 0.75 to 1.27; $p = 0.9$) and a statistically non-significant 16% reduction overall (relative rate 0.84, 95% confidence interval 0.68 to 1.04; $p = 0.1$).

There was no difference between intervention and control groups in mortality from other causes than breast cancer (relative rate 1.02, 95% confidence interval 0.97 to 1.07; $p = 0.4$) or from all-cause mortality (relative rate 1.01, 95% confidence interval 0.96 to 1.05; $p = 0.8$).

There was an excess of cancers (total invasive and in situ) up to the time of the first National Programme screen, which was not present thereafter. Tentative formal estimation of overdiagnosis

suggested that 80 breast cancers were overdiagnosed in the intervention phase of the trial, 8.5% of cancers diagnosed in this period in the intervention group, and an absolute rate of 0.2% over eight annual screens. However, the equalisation of incidence at the time of the first National Programme screen indicates that these would have been diagnosed by screening after the age of 50 years in any case.

Conclusions

Annual mammographic screening at 40–49 years conferred a reduction in breast cancer mortality. The relative reduction is attenuated after 10 years, possibly because of a lesser effect of screening in some aggressive grade 3 tumours in this age group. There was no evidence of overdiagnosis in addition to that which already results from the National Programme carried out at later ages. These results pertain to the epoch before digital mammography and universal two-view imaging, so the effectiveness nowadays may be greater than that observed here.

Trial registration

This trial is registered as ISRCTN24647151.

Funding

This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 24, No. 55. See the NIHR Journals Library website for further project information. Other funding in the past has been received from the Medical Research Council, Cancer Research UK, the Department of Health and Social Care, the US National Cancer Institute and the American Cancer Society.

Chapter 1 Introduction

In the UK, as in many other countries, following unequivocal evidence from randomised controlled trials (RCTs),[1] breast cancer screening with X-ray mammography is offered to women aged 50–70 years.[2] The interscreening interval is 3 years. The effect of mammographic screening on breast cancer mortality in women aged < 50 years has been a matter for discussion for several decades.[3–6] A lesser effect of screening on breast cancer mortality has been observed for women aged < 50 years in the RCTs, partly because this age group has more radiologically dense breast tissue than those aged ≥ 50 years, and possibly also because of the more rapid progression of cancers diagnosed in younger women.[7] We do not propose to report on a full systematic review of the evidence here. However, to put this report in context, we provide a brief narrative overview of the major international reviews, evidence summaries and consequent recommendations in recent years regarding the age at which to start breast screening. What follows is not an exhaustive list of organisations making recommendations on breast cancer screening for this age group, but it is representative and demonstrates the variation in conclusions from major reviews of the subject.

Following an independent systematic review, the American Cancer Society made a strong recommendation for screening commencing at the age of 45 years and a qualified recommendation that women aged 40–44 years have the opportunity to start screening at any point during ages 40–44 years.[8] The US Preventive Services Task Force (USPSTF) recommends screening from the age of 50 years and notes that screening between the ages of 40 and 49 years is a matter of individual choice.[9] The American College of Radiology, on the other hand, recommends annual mammographic screening from the age of 40 years.[10] The International Agency for Research on Cancer concluded that there was limited evidence that mammographic screening reduced mortality from breast cancer in women aged 40–49 years.[11] The European Guidelines for Breast Cancer Screening and Diagnosis include a conditional recommendation for screening at ages 45–49 years, but not at ages 40–44 years.[12]

In 2007, the UK's Cancer Reform Strategy announced a policy to extend breast screening to nine screening rounds between the ages of 47 and 73 years, presumably because of considerations of similarity between the ages of 47–49 years and 50 years with respect to screening effects and breast cancer.[13] However, this policy was never fully implemented and remains under research.

Although recommendations vary with respect to mammographic screening in women aged < 50 years, there is a general conclusion, with few dissenters, for example Miller,[4] that screening in this age group does prevent deaths from breast cancer. However, the absolute magnitude of the estimated mortality benefit varies substantially between reviews. The American Cancer Society's review, published in 2015, estimates that between 753 and 1770 persons will need to be screened in the age group 40–49 years to prevent one breast cancer death.[8] In 2016, the USPSTF quotes a figure of 3333 persons needing to screen in the age group 39–49 years to prevent one breast cancer death.[14] The European Guidelines, published online in 2019, quote that between 1299 and 2273 people aged between 45 and 49 years need to be screened to prevent one breast cancer death.[12] The differences in estimates of absolute benefit are because of the variation in a number of inputs, principally the follow-up period for cancer deaths (because screening in the present prevents deaths 10, 15 or 20 years in the future) and the relative risk estimate used.

In view of the continuing uncertainties and variability of estimates of benefit, it is important to exploit to the full all UK research resources that address this issue. In the UK, there are two RCTs specifically aimed at this age group: the UK Breast Screening Age trial (comparing usual care with annual screening for 7 years from age 39 to 41 years, with follow-up continuing thereafter)[15] and the ongoing Age Extension (AgeX) trial [evaluating extending the age range of the NHS Breast Screening Programme (NHSBSP) from 50–70 to 47–73 years].[16] The AgeX trial is essentially two trials, one offering screening at ages 47–49 years and another offering screening at ages 71–73 years.

The UK Breast Screening Age trial was initiated in 1990 and now has an average of > 23 years' follow-up. After exclusions, this trial randomised 106,953 women aged 39–41 years to the usual care group and 53,883 women to the intervention group (i.e. they were offered annual mammography for 7 years). The 17-year results were published in 2015 and showed a 25% reduction in breast cancer mortality with screening at 10 years, although this was attenuated with further follow-up.[17] We now have follow-up data to an average of 23 years. The follow-up phase is a National Institute for Health Research (NIHR) Health Technology Assessment (HTA) programme-funded project and is the subject of this report.

The aims of the follow-up project as set out in the protocol were:

- Analysis comparing breast cancer, other cause and all-cause mortality between intervention and control groups to the end of 2010.
- The same analysis to the end of 2017. The primary end point will be mortality from breast cancers diagnosed during the intervention phase of the trial, when the intervention group was invited to screening and the control group was not, that is, for all women in intervention and control groups, up to immediately before their first NHSBSP invitation.
- Analyses of cumulative mortality from breast cancer over the total follow-up period and at 5, 10 and 15 years from randomisation.
- Analysis of cumulative mortality from breast cancers diagnosed before first National Programme screen, for the most part diagnosed at ages 40–49 years.
- Estimation of the effect of screening in those women attending for screening ('per-protocol' analysis). This will use an established method[18] to adjust for selection bias, which arises because non-attenders for screening are likely to be at different level of risk of breast cancer mortality than those attending.
- Estimation of the absolute long-term benefit of screening in terms of the number needed to screen in this age group to prevent one breast cancer death.
- Analyses of the cumulative incidence of breast cancer, invasive and in situ, by trial from date of trial entry will be conducted. We will look specifically at incidence prior to the first NHSBSP invitation, incidence up to and including the first NHS National Programme screen-detected cancers and incidence up to the end of follow-up. We shall estimate overdiagnosis overall, and the additional overdiagnosis from starting screening at 40 years instead of 50 years, as in the National Programme.

The first aim has already been achieved and published, with results reported to 2011.[17] For completeness, the results will be briefly summarised in this report. In addition to the analyses above, we shall estimate overdiagnosis with formal adjustment for lead time, that is, separating excess incidence because of earlier diagnosis from excess incidence because of overdiagnosis.[19] We shall also conduct 'diagnostic' analysis to clarify the reasons for specific results, such as the early mortality benefit that is later diluted.[17]

Chapter 2 Methods

Trial design

The trial profile is summarised in *Figure 1*. A total of 160,921 women were randomised in a 1 : 2 ratio to the intervention or the control group, and after exclusions the trial included 160,836 women who had data available for analysis. Recruitment took place between October 1990 and September 1997. Individual randomisation was performed and stratified by general practice so that one-third of the women in any practice were allocated to the intervention group. Women were identified from general practitioner (GP) lists then held by Family Heath Services Authorities. Randomisation was at the individual level, stratified by general practice. GPs were given prior sight of lists and could remove women whom they considered unsuitable for invitation, for example those already under care for breast cancer from 1992 onwards, randomisation and allocation to a trial group were carried out on the health authorities' computer system using ad hoc software. Prior to this, in three centres that started the trial early, random numbers generated by the trial co-ordinator were applied to GP lists from the Family Health Services Authorities. It was not possible to blind the screening centres to trial group allocation.

Women were aged 39–41 years at time of entry to the trial. The trial was conducted in 23 NHSBSP screening units in England, Wales and Scotland. Participating units are listed in the appendix to Moss *et al.*[17] Women in the intervention group were sent a letter of invitation and information leaflet that clearly stated that the woman was being asked to participate in a research trial, and her acceptance of the invitation was

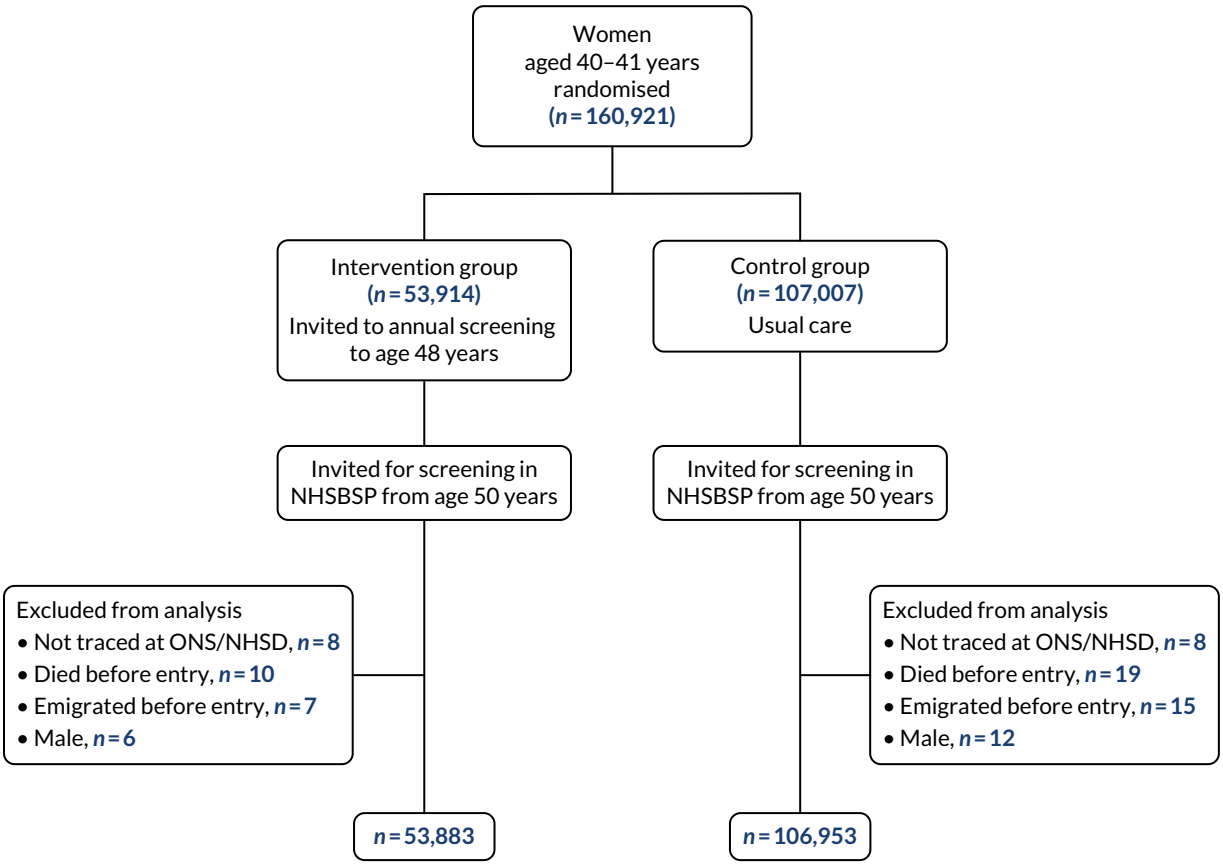


FIGURE 1 Flow diagram of trial profile. NHSD, NHS Digital; ONS, Office for National Statistics.

taken to be her informed consent to participate. Women in the intervention group were invited to eight annual mammographic screens, until the calendar year of their 48th birthday. Thus, the trial was designed to evaluate screening, all of which took place before the age of 50 years. At 50 years, both they and the women in the control group became eligible for 3-yearly invitation to screening as part of the NHSBSP, and received their first invitation between the ages of 50 and 52 years.

Screening in the trial was by two-view mammography at the first screen, with single-view mammography thereafter unless otherwise indicated. Mammograms were double-read. All women, including non-attenders, were reinvited annually unless they requested otherwise. Women who moved to areas that were not covered by the trial were not reinvited for screening as part of the trial, but were able to self-refer to either their previous or their nearest participating screening centre. Screening in three centres ceased prematurely (after four, five and six rounds) because of the inability of the centres to manage the additional workload with the available resources.

The trial database contains information on all screening as part of the trial in women in the intervention group. It also contains data on the first screening invitation and attendance at ages 50–52 years as part of the NHSBSP in women in both the intervention and the control groups. These data on the first screening invitation and attendance at ages 50–52 years have been collected not only from the 23 centres participating in the trial but also from all NHSBSP screening units in England, Wales and Scotland, thus providing information on screening in women who have moved away from their original trial centre. Data on this first NHSBSP screen were estimated to be 93% complete, with similar percentages in the two trial groups. Information on screening includes attendance, outcome of initial mammogram (i.e. whether or not the woman was recalled for further assessment) and final outcome of the screening episode.

Up to the end of 2009, pathological information including in situ/invasive status, invasive tumour size, lymph node status and histological grade was routinely supplied to the triallists. In addition, prior to 2009, pathology was reviewed and the pathological variables reclassified by a panel of three expert breast pathologists, using pathology slides where available.[20,21] Of the 7890 breast cancers diagnosed between the start of the trial and February 2017, 3641 (46%) underwent full pathological review. This included all cancers diagnosed in the intervention phase of the trial. In cases undergoing review, there was good agreement between original and reviewed classifications with respect to invasive status, tumour size, histological grade and node status.[20] In the analyses below, reviewed pathology classifications were used where available; otherwise, classifications from the original pathology reports were used.

The primary outcome measure of the trial is mortality from breast cancer. As noted above, the primary end point was mortality from breast cancers diagnosed during the intervention phase of the trial, when the intervention group was invited to screening and the control group was not (i.e. before the first NHSBSP invitation, from which point both groups were invited to screening). It was decided from the outset of the trial to use underlying cause of death from the death certificate rather than undertake a verification exercise. All women in this trial were flagged at the NHS Central Register, now controlled by NHS Digital, and > 99.9% were successfully traced. This register provides data on all cancer registrations and deaths, including data on underlying coded cause of death. We have notifications to 28 February 2017, an average follow-up of 23 years. We had originally planned to have data to the end of 2017 but because of changes in information governance policies on the part of the data custodians, notifications ceased 10 months early.

Statistical methods

The trial was originally designed to recruit 190,000 women to have 80% power for a 20% reduction in breast cancer mortality at 10 years' follow-up, at 5% significance level. However, financial and workload constraints on NHS breast screening units slowed recruitment and no new centres entered after 1996. In 1999, the Data Monitoring Committee recommended that, as further accrual would

result in only marginal gains in power and would delay achievement of mean follow-up times, recruitment should cease.[15] The revised power, based on the original estimates of breast cancer mortality in the control group of 3.3 per 1000, was 72% at 10 years' follow-up and 80% at 14 years' follow-up. See the protocol of this follow-up project.[22] Mortality data over time were analysed by Poisson regression for the purpose of significance testing between the intervention and the control groups, and for the estimation of relative rates (RRs) and confidence intervals (CIs) on these. This is the recommended method of analysis for data on counts of events in a given period.[23] In addition, we calculated Nelson–Aalen estimates of cumulative hazard.[24] For the primary end point, in addition to estimating the intention-to-treat effect, we also estimated the effect of actually being screened using the method described in Cuzick *et al.*[18]

In estimating the effect on mortality from cancers diagnosed in the intervention period of the trial, there is a potential bias against the intervention because the intervention group will include deaths from cancers diagnosed at screening whose time of diagnosis would have been at or after first NHS Programme invitation, and which would therefore not be included in the control group. Duffy and Smith[25] describe how the bias can be minimised by including cancers diagnosed at a contemporaneous screen at the end of the intervention period in both groups. Therefore, we performed a secondary analysis redefining the intervention period cancers as those diagnosed up to and including the first NHS Programme screen in both groups.

In addition to mortality from breast cancer, we also estimated the effect of intervention on deaths from all causes and deaths from causes other than breast cancer. In particular, we considered deaths from causes other than breast cancer in the breast cancer patients. If there was any systematic bias in classification of death from breast cancer with respect to screening status, we would expect to see different rates of death from other causes in the breast cancer patients between the intervention and the control group.

We estimated incidence of breast cancer, in situ and invasive, in the intervention and the control groups, up to just before first NHSBSP screen, up to and including first NHSBSP screen and up to end of follow-up (i.e. February 2017). For estimation of overdiagnosis, we used the customary definition of breast cancers diagnosed as a result of screening that would not have been diagnosed in the person's lifetime if screening had not taken place. Formal estimation of overdiagnosis therefore implies quantitative estimates of the size of this subgroup of cancers and is not simply a comparison of incidence between screened and unscreened groups, which is affected by lead time. However, such estimation is tentative (i.e. it has associated uncertainties in addition to simple statistical variation) because, for a given cancer that is detected by screening and treated, we can never know what would have happened if it had not been diagnosed and treated at that point. We derived estimates of overdiagnosis by estimating the expected numbers of cancers detected at each screen if there were no overdiagnosis, and subtracting these from the observed numbers. The expected numbers were derived as a function of estimates of screening test sensitivity and mean sojourn time (the duration of the preclinical screen-detectable period); the latter was estimated by maximum likelihood from interval cancer rates, which, by definition, do not include overdiagnosed cancers.[26]

To estimate the expected number of cancers at incident (second and subsequent) screens, we used the estimate of programme sensitivity (the proportion of cancers detected at screening in a population attending screening). This was derived by Launoy *et al.*[27] as:

$$P_s = \frac{s(1-e^{-\lambda t})}{\lambda t(1-(1-s)e^{-\lambda t})}. \quad (1)$$

The expected number of screen-detected cancers at a given screen E should satisfy:

$$P_s = \frac{E}{E + C}, \tag{2}$$

where C is the number of interval cancers corresponding to that screening round. Thus, we calculate E as:

$$E = \frac{P_s C}{1 - P_s}. \tag{3}$$

We derived independent estimates of S for individual screens at 40–47 years by interpolation and extrapolation of the estimates of Carney *et al.*,[28] who calculated sensitivity as 68.6% at 40–44 years and 72.5% at 45–49 years. We estimated λ conditional on these sensitivity estimates from the incidence of interval cancers, modelling progression from preclinical cancer to clinical symptomatic disease as a Markov process. Further details are given by Michalopoulos and Duffy.[26]

This method cannot be used for first screens. However, the expected number of cancers diagnosed at a prevalent (first) screen is:

$$\frac{N_p I S}{\lambda}, \tag{4}$$

where S and λ are as defined above, N_p is the number of women screened for the first time and I is the underlying incidence. We estimated the underlying incidence among screening attenders as 0.573 per 1000 per year, by adjusting the control group incidence for selection bias (those who choose to be screened have potentially different risk of breast cancer from those who do not) using the method described in Moss *et al.*[29]

Chapter 3 Mortality results

Reporting policy

The results are reported in line with Consolidated Standards of Reporting Trials (CONSORT) guidelines.

Recap on 2011 results

At the time of analysis of the 2011 mortality results, we had a median follow-up time of 17 years. *Table 1* shows the cumulative person-years, breast cancer deaths (from cancers diagnosed up to the end of the intervention period, as noted in *Chapter 2*) and RRs with 95% CIs on these, for deaths prior to 10 years from randomisation, deaths occurring ≥ 10 years after randomisation and all deaths from these breast cancers. Average ages during the two periods (i.e. prior to 10 years and ≥ 10 years) were 45 and 54 years, respectively. There was a statistically significant reduction in mortality in the first 10 years (RR 0.75, 95% CI 0.58 to 0.97), but no reduction thereafter (RR 1.02, 95% CI 0.80 to 1.30).

Table 2 shows cumulative deaths from all breast cancers diagnosed since randomisation by trial group and 5-year period, not only those diagnosed during the intervention period as in *Table 1*. Thus, even in the first 10 years, numbers are slightly larger than those in *Table 1*. Note that these are cumulative rather than period specific as in Moss *et al.*[17] Again, there was an early reduction in mortality that was later diluted. It should be noted that the later period in *Table 2* includes deaths from cancers diagnosed after both groups of the trial were subject to screening in the National Programme, so dilution is expected.

TABLE 1 Mortality from breast cancers diagnosed in the intervention period, by trial group and period of observation, to 31 December 2011

	Intervention group		Control RR		
Period of observation (years)	Deaths	Person-years	Deaths	Person-years	RR (95% CI)
0 to < 10	83	532,747	219	1,058,322	0.75 (0.58 to 0.97)
≥ 10	99	408,221	193	810,395	1.02 (0.80 to 1.30)
Total	182	940,968	412	1,868,717	0.88 (0.74 to 1.04)

TABLE 2 Cumulative mortality from all breast cancers diagnosed since randomisation to 31 December 2011

	Intervention group		Control group		
Period of observation (years)	Deaths	Person-years	Deaths	Person-years	RR (95% CI)
0 to < 5	27	267,864	69	532,104	0.78 (0.50 to 1.21)
0 to < 10	83	532,748	221	1,058,324	0.75 (0.58 to 0.97)
0 to < 15	181	793,911	406	1,576,547	0.89 (0.74 to 1.07)
Total	242	940,969	515	1,876,717	0.93 (0.80 to 1.09)

At follow-up to the end of 2011, there were 2127 deaths from all causes in the intervention group (940,969 person-years), and 4320 in the control group (1,868,717 person-years). This gave a RR of 0.98 (95% CI 0.93 to 1.03).

When interpreting the 2011 results, it was noted that both short- and long-term survival in the intervention group were improved for those with cancers of histological grades 1 and 2, but that improved survival was observed only in the first 10 years for those with grade 3 cancers.[17] This suggested that, for these more aggressive tumours, early detection was delaying but not preventing death from breast cancer. This would go some way to explaining the attenuation of the mortality benefit after 10 years.

Breast cancer mortality to February 2017

Table 3 shows breast cancer mortality to February 2017 from the cancers diagnosed during the intervention phase of the trial, up to 10 years after randomisation and from 10 years onwards, by group. The corresponding Nelson–Aalen cumulative mortality graphs are given in *Figure 2*. At 10 years there was a statistically significant 25% reduction in mortality (RR 0.75, 95% CI 0.58 to 0.97; $p = 0.03$). For ≥ 10 years, no reduction was observed (RR 0.98, 95% CI 0.79 to 1.22; $p = 0.9$). Overall, there was a 12% reduction in breast cancer mortality, which was not statistically significant (RR 0.88, 95% CI 0.74 to 1.03; $p = 0.1$). The absolute benefit, however, remains similar regardless of follow-up period (see *Figure 2*).

TABLE 3 Mortality from breast cancers diagnosed in the intervention period, by trial group and period of observation, to 28 February 2017

	Intervention group		Control group		
Period of observation (years)	Deaths	Person-years	Deaths	Person-years	RR (95% CI)
0 to < 10	83	532,729	219	1,058,236	0.75 (0.58 to 0.97)
≥ 10	126	668,281	255	1,326,770	0.98 (0.79 to 1.22)
Total	209	1,201,010	474	2,385,006	0.88 (0.74 to 1.03)

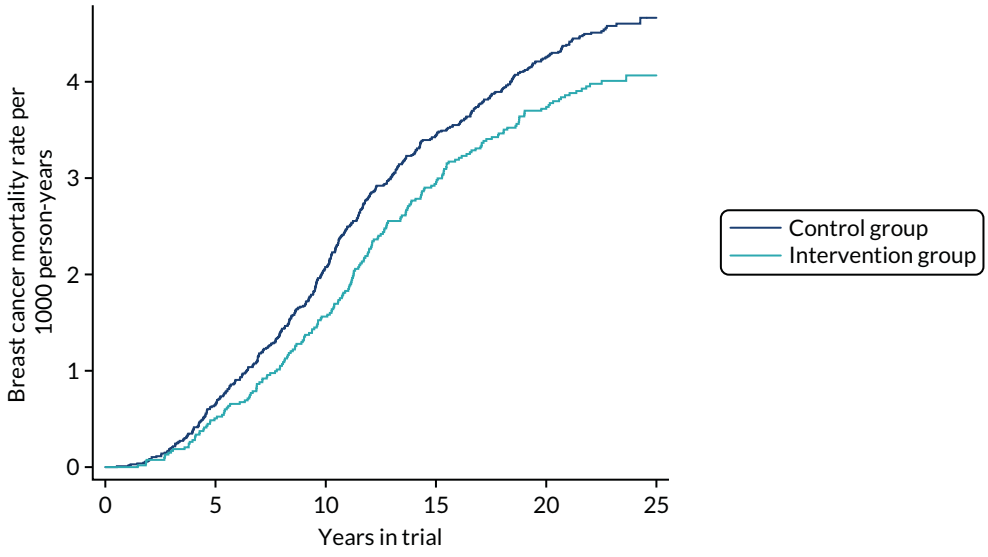


FIGURE 2 Cumulative breast cancer mortality to February 2017 from cancers diagnosed during the intervention phase of the trial.

There were 55 deaths in non-attenders, lapsed attenders or those who had ceased screening, 32 deaths before 10 years after randomisation and 23 thereafter. This gives the selection bias-adjusted per-protocol effect of being screened as a statistically significant 34% reduction in breast cancer mortality up to 10 years after randomisation (RR 0.66, 95% CI 0.46 to 0.95; $p = 0.02$), a statistically non-significant 2% reduction after 10 years (RR 0.98, 95% CI 0.75 to 1.27; $p = 0.9$) and a statistically non-significant 16% reduction overall (RR 0.84, 95% CI 0.68 to 1.04; $p = 0.1$).[18]

Table 4 shows the corresponding breast cancer mortality figures for the secondary analysis of cancers diagnosed up to and including the first NHS Programme screen in both groups. Numbers are larger than in *Table 3* because of the additional inclusion of cancers diagnosed at the first NHS Programme screen. At 10 years, the results were identical to the primary analysis: a statistically significant 25% reduction in mortality (RR 0.75, 95% CI 0.58 to 0.97; $p = 0.03$). For $\geq$ 10 years, a small, statistically non-significant, reduction was observed (RR 0.95, 95% CI 0.77 to 1.17; $p = 0.6$). Overall, there was a 14% reduction in breast cancer mortality, which was of borderline statistical significance (RR 0.86, 95% CI 0.73 to 1.01; $p = 0.07$). There was no statistically significant effect of the intervention on mortality from breast cancers diagnosed after the intervention period (RR 0.99, 95% CI 0.79 to 1.24; $p = 0.9$).

Table 5 shows cumulative deaths from all breast cancers diagnosed since randomisation by trial group and 5-year period, up to 28 February 2017. The corresponding Nelson–Aalen plots of cumulative mortality by continuous time are shown in *Figure 3*. Again, there was an early reduction in mortality that was later diluted, as expected.

The absolute reduction in breast cancer mortality was 0.6 per 1000 women invited. This corresponds to needing to invite 1667 women and, given the 65% average participation rate, needing to screen 1083 women in this age group to prevent one breast cancer death. This finding is relatively stable over time and end point. At 10 years, the number of women who would need to be screened is 1300. At the final follow-up, using only deaths from cancers diagnosed during the intervention phase, the figure is 1300, and for the secondary end point including the cancers diagnosed at first NHSBSP screen the number needed to screen is 1050.

TABLE 4 Mortality from breast cancers diagnosed up to and including the first NHSBSP screen, by trial group and period of observation, to 28 February 2017

	Intervention group		Control group		
Period of observation (years)	Deaths	Person-years	Deaths	Person-years	RR (95% CI)
0 to < 10	83	532,729	219	1,058,236	0.75 (0.58 to 0.97)
≥ 10	133	668,281	279	1,326,770	0.95 (0.77 to 1.17)
Total	216	1,201,010	498	2,385,006	0.86 (0.73 to 1.01)

TABLE 5 Cumulative mortality from all breast cancers diagnosed since randomisation to 28 February 2017

	Intervention group		Control group		
Period of observation (years)	Deaths	Person-years	Deaths	Person-years	RR (95% CI)
0 to < 5	27	267,852	69	532,066	0.78 (0.50 to 1.21)
0 to < 10	83	532,729	221	1,058,236	0.75 (0.58 to 0.96)
0 to < 15	182	793,852	406	1,576,346	0.89 (0.75 to 1.06)
Total	338	1,201,010	743	2,385,006	0.90 (0.79 to 1.03)

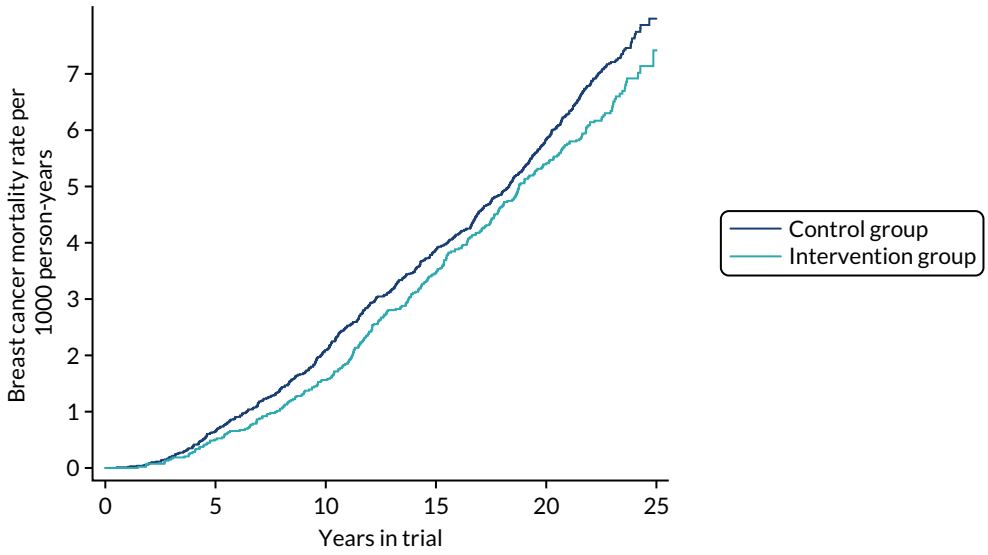


FIGURE 3 Cumulative breast cancer mortality, regardless of time of diagnosis, from randomisation to 28 February 2017.

All-cause and other-cause mortality

Figure 4 shows the Nelson–Aalen cumulative mortality from all causes over time. There was no difference in all-cause mortality to the end of follow-up between the two groups (RR 1.01, 95% CI 0.96 to 1.05; $p = 0.8$). For mortality from causes other than breast cancer, the result was very similar (RR 1.02, 95% CI 0.97 to 1.07; $p = 0.4$).

We also investigated two further specific causes of death, all cancers (including breast cancer) and ischaemic heart disease. There was no statistically significant effect on deaths from all cancers (RR 0.98, 95% CI 0.93 to 1.05; $p = 0.6$). Similarly, no statistically significant effect was observed on deaths from ischaemic heart disease (RR 1.03, 95% CI 0.87 to 1.20; $p = 0.7$).

In considering deaths from all causes and causes other than breast cancer, it is worth investigating the deaths of the cancer patients, in case treatment of cancers has a differential effect on mortality between the study and control groups. Accordingly, we compared the intervention and control groups with respect to fatality from all causes and all causes except breast cancer in the breast cancer patients.

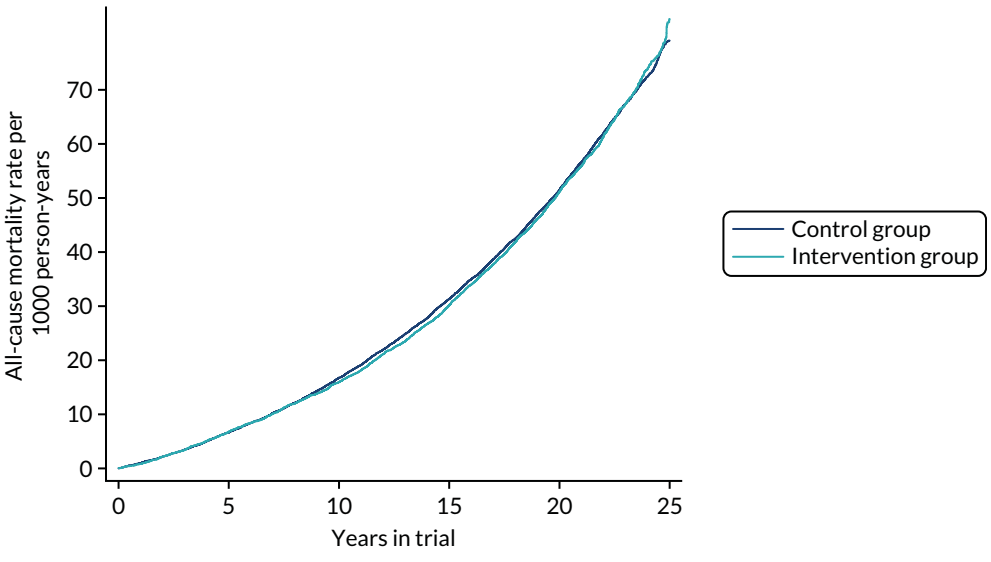


FIGURE 4 Cumulative mortality from all causes from randomisation to 28 February 2017.

The results showed a statistically significant reduction in fatality from any cause in breast cancer patients in the intervention group (RR 0.87, 95% CI 0.77 to 0.98; $p = 0.02$) and a statistically non-significant reduction in the number of deaths from causes other than breast cancer in the intervention group (RR 0.86, 95% CI 0.67 to 1.11; $p = 0.2$). This suggests that there is no bias in favour of screening from using death from breast cancer as the end point.

Further investigation of the reduction in effect over time

As noted above, it was suggested in the report on the 2011 results that one possible reason for the reduction in the effect on breast cancer mortality, even using only those cancers diagnosed in the intervention period, may have been a reduced effect of early detection on invasive cancers of histological grade 3. The suggestion was that, for these more aggressive tumours, the early detection may have, on average, only postponed death from breast cancer rather than preventing it altogether. Although the relatively stable absolute effect suggests that the theory of postponement rather than prevention of death is not the issue, results from this further follow-up are consistent with a lesser effect of screening on grade 3 cancers. *Figure 5* shows survival by group for invasive tumours of grades 1 and 2 combined (this shows survival from randomisation rather than diagnosis to avoid issues of lead time bias). There is a survival advantage for the intervention group, which persists over the long term. *Figure 6* shows the corresponding survival curves for invasive grade 3 cancers. Note that rather larger numbers in the intervention group had usable pathology data as more of these were diagnosed in the intervention period when supply of pathology data to the triallists was routine.

Population mortality results, which avoid potential biases from length bias and overdiagnosis, are similarly consistent with this. *Table 6* gives the mortality from breast cancer by grade and group for 5, 10, 15 and 20 years of follow-up. For tumours of grades 1 and 2, there is an approximate 30% reduction in mortality that is maintained at long-term follow-up. For grade 3 cancers, there is a lesser mortality reduction that is no longer present by 15 years of follow-up.

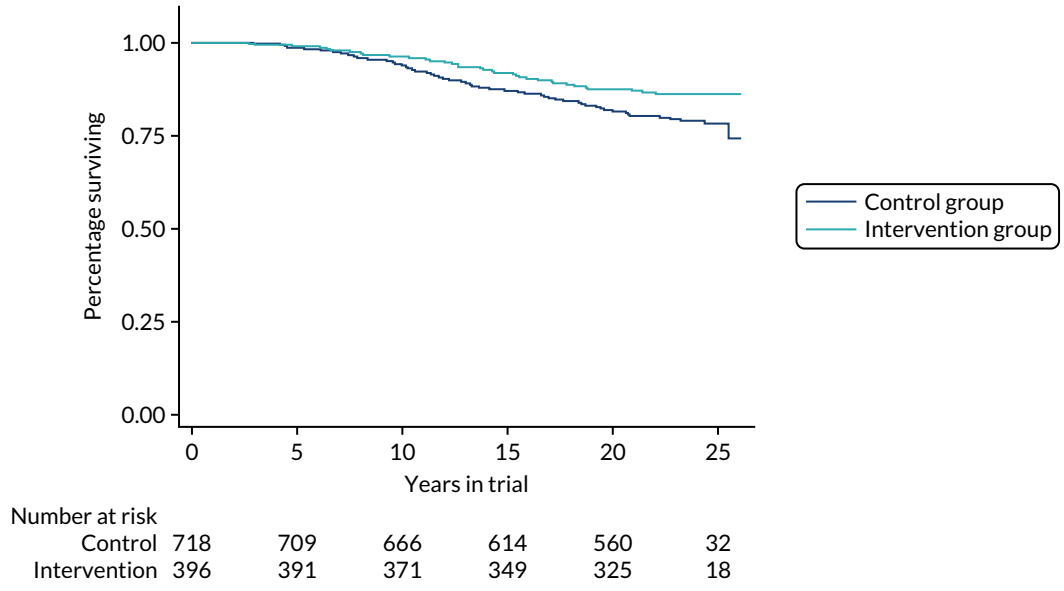


FIGURE 5 Survival to death from breast cancer of invasive cancers of grades 1 or 2, diagnosed during the intervention period by trial group.

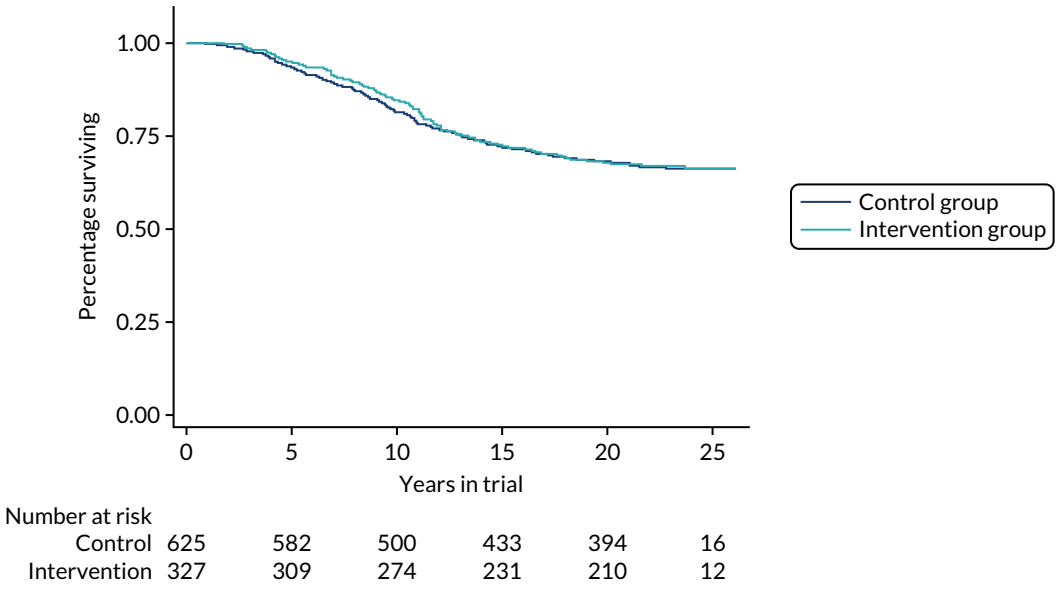


FIGURE 6 Survival to death from breast cancer of invasive cancers of grade 3, diagnosed during the intervention period by trial group.

TABLE 6 Cumulative breast cancer mortality by grade of tumour, period of follow-up and trial group

Grade	Follow-up (years)	Intervention group		Control group		RR (95% CI)
		Breast cancer deaths (*n*)	Rate/ 1000 women	Breast cancer deaths (*n*)	Rate/ 1000 women	
1 and 2	5	3	0.0001	9	0.0001	0.66 (0.18 to 2.44)
	10	16	0.0003	47	0.0004	0.68 (0.38 to 1.19)
	15	33	0.0006	93	0.0009	0.70 (0.47 to 1.05)
	20	43	0.0008	112	0.0010	0.76 (0.45 to 1.08)
3	5	17	0.0003	42	0.0004	0.80 (0.46 to 1.41)
	10	50	0.0009	114	0.0011	0.87 (0.62 to 1.21)
	15	89	0.0017	173	0.0016	1.02 (0.79 to 1.32)
	20	99	0.0018	190	0.0018	1.03 (0.81 to 1.32)

Chapter 4 Breast cancer incidence and estimation of overdiagnosis

Cumulative incidence of breast cancer over time

Figure 7 shows the cumulative incidence of breast cancer, invasive and in situ, from randomisation to 28 February 2017 in the intervention and control groups. *Table 7* gives the number of cancers, person-years and RRs (and 95% CIs) for incidence up to immediately before the first invitation to NHSBSP screen, up to and including first NHSBSP screen and at complete follow-up to the end of February 2017. The RR of breast cancer incidence just before the first NHSBSP screen was 1.09 (95% CI 1.00 to 1.19; $p = 0.03$). The RR up to and including first NHSBSP screen was 0.99 (95% CI 0.92 to 1.07; $p = 0.7$). At the end of follow-up, the RR was 0.99 (95% CI 0.94 to 1.04; $p = 0.6$). This indicates that there was no excess in the intervention group in addition to cancers diagnosed in the NHSBSP from the age of 50 years onwards. Thus, any overdiagnosis conferred by screening at 40–49 years would occur in any case if the population were screened from ≥ 50 years.

Table 8 shows the corresponding results for invasive cancers only. The results show a very small excess in the intervention group in the intervention phase and a very small deficit thereafter. Neither of these differences was statistically significant.

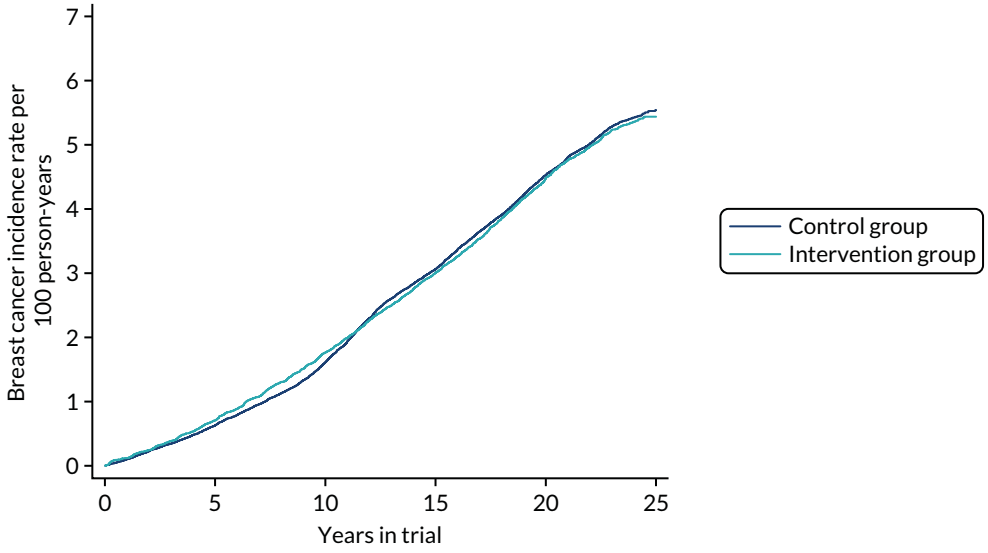


FIGURE 7 Cumulative incidence of breast cancer, invasive and in situ, by period of follow-up and trial group.

TABLE 7 Numbers of breast cancers, invasive and in situ, and person-years of observation, by period of follow-up and group

	Intervention group		Control group		
Period of follow-up	Breast cancers	Person-years	Breast cancers	Person-years	RR (95% CI)
Just before first NHSBSP screen	953	569,016	1731	1,129,491	1.09 (1.00 to 1.19)
Up to and including first NHSBSP screen	1132	569,016	2278	1,129,491	0.99 (0.92 to 1.07)
Up to 28 February 2017	2617	1,174,649	5260	2,334,516	0.99 (0.94 to 1.04)

TABLE 8 Cumulative incidence of breast cancer, invasive cancers only, by period of follow-up and trial group

Period of follow-up	Intervention group		Control group		RR (95% CI)
	Breast cancers	Person-years	Breast cancers	Person-years	
Just before first NHSBSP screen	835	569,016	1628	1,129,491	1.02 (0.94 to 1.11)
Up to and including first NHSBSP screen	970	569,016	2021	1,129,491	0.95 (0.88 to 1.04)
Up to 28 February 2017	2288	1,177,990	4640	2,339,852	0.98 (0.93 to 1.03)

Pathological attributes of cancers diagnosed

Owing to changes in information governance and data ownership, we do not have detailed pathology data on more recent cancers diagnosed since the trial groups entered the National Programme. However, from *Tables 7* and *8*, we can calculate that, in the intervention phase of the trial prior to the first National Programme screen, 6% of control group cancers and 12% of intervention group cancers were in situ. After all subjects had entered the National Programme, the corresponding figures were close to 15% in the control group and 13% in the intervention group.

For invasive cancers diagnosed during the intervention phase, *Table 9* shows the size, node status, grade, vascular invasion status and tumour type, by trial group, giving results in the intervention group

TABLE 9 Pathological factors by trial group and detection mode in the intervention group

Factor	Category	Intervention group, *n* (%)		Control group, *n* (%)
		Screen detected	Symptomatic	
Total invasive		256	579	1628
Node status	Negative	207 (81)	328 (61)	932 (62)
	Positive	49 (19)	213 (39)	575 (38)
	NK	0	38	121
Size	≤ 20 mm	196 (81)	265 (56)	788 (58)
	> 20 mm	47 (19)	206 (44)	562 (42)
	NK	13	108	278
Grade	1	48 (20)	50 (10)	122 (9)
	2	114 (48)	184 (38)	596 (44)
	3	74 (32)	253 (52)	625 (47)
	NK	20	92	285
Vascular invasion	Yes	53 (22)	169 (33)	439 (31)
	No	189 (78)	335 (67)	989 (69)
	NK	14	75	200
Tumour type	Ductal	179 (72)	399 (76)	1127 (77)
	Lobular	19 (8)	67 (13)	175 (12)
	Medullary	3 (1)	8 (2)	17 (1)
	Mucinous	4 (2)	9 (2)	32 (2)
	Tubular	40 (16)	39 (7)	105 (7)
	Other	1 (< 1)	2 (< 1)	15 (1)
	NK	10	55	157

NK, not known.

for screen-detected and symptomatic cancers (interval cancers and cancers diagnosed in non-attenders for screening) separately. The screen-detected cancers tended to be smaller (as expected) and were less likely to be node positive, to be grade 3 or to have vascular invasion than the cancers in the control group and the symptomatic cancers in the intervention group. The symptomatic cancers in the intervention group had very similar attributes to those in the control group. There was a higher proportion of tubular carcinoma cases in the screen-detected cancers in the intervention group, suggesting a measure of length bias in screen detection.

Tentative estimation of overdiagnosis

Table 10 shows the numbers of screen-detected and interval cancers by screening round over the eight screening rounds in the intervention phase of the trial, with the sensitivity estimates derived from Carney *et al.*[28] From the interval cancers and using maximum likelihood similarly to Michalopoulos and Duffy,[26] we estimated λ, the inverse of the mean sojourn time, as 0.92 (95% CI 0.54 to 1.49). *Table 10* also shows the expected numbers of screen-detected cancers as estimated from λ and sensitivity. The total number of expected screen-detected cancers was 164, compared with 244 observed. This suggests that 80 cancers were overdiagnosed, 8.5% of the breast cancers diagnosed in the intervention phase of the trial, or 32.8% of screen-detected cancers, and an absolute rate of 0.2% of women being screened over 8 years.

In addition, the equalisation of incidence of breast cancer between the control and intervention groups following the first NHSBSP screen indicates that these would have been diagnosed by screening after the age of 50 years in any case. Therefore, starting screening at 40 years instead of 50 years did not lead to any additional overdiagnosis.

TABLE 10 Women screened, observed and expected screen-detected cancers, interval cancers and sensitivity estimates, by screening round in the intervention group

Screening round	Women screened (*n*)	Screen-detected cancers (*n*)	Interval cancers (*n*)	Estimated screening sensitivity (%)	Expected screen-detected cancers (*n*)
1	35,582	31	10	67.0	15
2	33,547	21	18	68.6	19
3	31,753	23	22	69.3	23
4	31,117	22	21	70.1	22
5	31,169	34	21	70.9	23
6	29,695	34	20	71.7	22
7	28,452	37	15	72.5	17
8	24,904	42	20	73.3	23

Chapter 5 Discussion and implications

This trial found a statistically significant 25% reduction in breast cancer mortality with the offer of annual screening for breast cancer in the first 10 years following randomisation. The reduction was diluted thereafter. There is some evidence that the dilution was due to a number of aggressive grade 3 cancers for which the effect of screening was small and not apparent at all after 10 years. This is in contrast to results of the Swedish Two-County Trial on which the NHS Programme is based. In the Two-County Trial, of women aged 40–74 years (73% were aged ≥ 50 years), most of the breast cancer mortality reduction was in grade 3 cancers.[30] The magnitude of the observed benefit is consistent with a recent meta-analysis of the RCTs of mammography.[31] However, it should be noted that our trial specifically recruited subjects who were aged 39–41 years so that all of the trial screening would take place before the age of 50 years. The other trials did not have this design feature.

There was no indication of an effect of the intervention on deaths from other causes, and no effect on all-cause mortality. The latter is to be expected, as the effect on all-cause mortality is overwhelmingly driven by causes of death on which the intervention has no effect.[32] In this trial, breast cancer deaths from cancers diagnosed in the intervention phase constituted only 7% of all deaths.

Interestingly, there was no statistically significant difference between the trial groups with respect to ischaemic heart disease. It has been suggested that early detection implies more frequent local excision, which in turn requires more radiotherapy, which increases the risk of ischaemic heart disease.[33] This does not seem to be the case in this trial, although the lack of an effect may be due to lack of statistical power because there were only 17 ischaemic heart disease deaths in the breast cancer patients. It should also be noted that, with modern techniques, risks to the heart from radiotherapy are considerably reduced.[34,35] Similarly, there was no evidence of an increase in risk of death from all causes in the intervention group or in death from other causes among breast cancer patients, suggesting that the concerns about excess deaths from the treatment of screen-detected cancers are unfounded. It has to be admitted, however, that there are potential inaccuracies in using cause of death as given on the death certificate. Screening makes it more likely that a breast cancer is diagnosed and, if a subject dies with a previous diagnosis of breast cancer, there may be a tendency for that breast cancer to be considered as either causing or contributing to death. Some breast cancer patients may die of metastatic disease from another, occult primary tumour and their death may then be classified as a breast cancer death.

The results translate to an absolute benefit of one breast cancer death prevented per 1000 women regularly screened in this age group, which is considerably smaller than that observed at 50–69 years (and requiring more frequent screening).[1] However, it is a larger benefit than most of the review findings quoted in *Chapter 1*, almost certainly because of the long-term follow-up reported here. In addition, it is worth considering that technological changes that have taken place since the trial's intervention phase might modify this. Digital mammography rather than film is the standard, as is two-view mammography at all screens. There is clear evidence that both substantially improve screening sensitivity, with digital mammography specifically showing improved sensitivity in women < 50 years.[36,37] It may be, however, that the improved sensitivity of state-of-the-art screening may also result in increased overdiagnosis.

It should also be noted that therapies have changed dramatically since the inception of this trial. We did not have access to treatment data, but, in the intervention period of this trial, the earliest diagnosis of a breast cancer was in 1991 and the latest in 2008. It is fair to say that those cancers diagnosed more recently in this period will have had greater access to effective and potentially tumour-targeted systemic therapies [e.g. hormone therapies for oestrogen receptor-positive cancers, and trastuzumab (Herceptin®; Roche Diagnostics, Hertford, UK) for human epidermal growth factor receptor 2 (HER2)-positive cancers] than those diagnosed early in the trial period. In addition, it is not clear whether or not the combination of modern screening and modern systemic therapies would be synergistic.

However, it is certainly possible that earlier detection in combination with more effective treatments may have benefits that are considerably greater than the effects observed here.

The results of analysis of breast cancer incidence rates suggest relatively low absolute levels of overdiagnosis, and no additional overdiagnosis above that arising from screening women aged 50–70 years, as the incidence equalised with the first NHSBSP screen at or shortly after 50 years. That is, any cancers overdiagnosed at 40–49 years in the intervention group were balanced by their equivalent cancers being overdiagnosed at ≥ 50 years in the control group. In addition, invasive cancers showed a small, statistically non-significant deficit in the intervention group following the first NHSBSP screen. This is consistent with findings in the NHSBSP, the Swedish Two-County Trial[38] and the Gothenburg Trial[39] that diagnosis and treatment of ductal carcinoma in situ was followed by a reduction in subsequent invasive breast cancer incidence.

Overdiagnosis is not the only adverse effect of screening. Another important human cost is false-positive recall for assessment of suspicious lesions that transpire not to be cancer. In this trial, the rate of false positives was 4.9% at first screening and 3.2% at subsequent screens.[40] The first is considerably smaller than the corresponding first-round false-positive rate in the NHS Programme in women aged 50–70 years, and the second is comparable to the corresponding rate observed at subsequent rounds in the NHS Programme.[41] It is not clear whether or not these rates would change with more modern screening methods.

It is planned that results reported in *Chapters 3* and *4* will be expanded on and submitted separately to peer-reviewed medical journals. In addition, further methodological work on estimation of screening sensitivity and overdiagnosis in this population is at an advanced stage, and will be submitted in the near future.

Limitations of the trial include the 31% average non-compliance with screening and the fact that three centres had to cease screening for resource and capacity reasons. These would tend to bias the results against screening. The issues are further discussed by Moss *et al.*[17] Other limitations include the fact that the technical aspects of the screening were considerably different from the state-of-the-art screening today and the range of systemic treatments that are available now, which were not available in the 1990s when much of the diagnostic and treatment activity of the trial was carried out.

What are the implications for clinical practice and future research? The results indicate that there is a reduction in breast cancer mortality associated with the offer of screening in women aged 40–49 years. There is no evidence of overdiagnosis in addition to that which is accrued in screening women aged 50–70 years. It is also likely that the use of digital mammography and universal two-view examination could lead to a greater benefit. Policy-makers may usefully consider this potential improvement in addition to the mortality benefit observed in this trial in deciding lower age limits for population screening. The finding of a mortality reduction is likely to be generalisable from this population-based trial, but, because of the changes in screening technology and practice, and the limitations with respect to compliance, resource and capacity noted above, the actual size of the reduction in practice is likely to be larger than observed here.

In terms of research, two major questions should be addressed. First, can digital mammography and universal double-reading improve on the effects observed here, in particular the long-term effects? In this respect, an update of the Swedish Mammography Screening in Young Women Cohort study[42] to the epoch of digital mammography would be useful. The second question is whether or not alternative or additional imaging modalities might be needed to improve the effectiveness of screening in this age group because of the higher mammographic density in premenopausal women. There is already ample evidence of increased detection from digital breast tomosynthesis, and from adding magnetic resonance imaging or ultrasound to mammography.[43,44] There is a need for further research to determine to what extent this increased detection will be reflected in a greater effect on mortality from breast cancer.

Acknowledgements

The investigators are grateful to a large number of radiologists, pathologists, other clinicians and screening office staff who contributed to this study over the years, and to the 160,921 women who participated in the trial. We thank Dr Matejka Rebolj for helpful discussion and Ms Raissa Frank for diligent data management.

Patient and public involvement

Because this project comprised only follow-up of a trial that had already been completed, there was no patient and public involvement.

Contributions of authors

Stephen Duffy (https://orcid.org/0000-0003-4901-7922) took over as chief investigator after the retirement of Sue Moss, and was responsible for supervising statistical analysis and drafting the report.

Daniel Vulkan (https://orcid.org/0000-0003-4738-9378) was responsible for the primary data analyses.

Howard Cuckle (https://orcid.org/0000-0003-2450-2403) and **Sue Moss (https://orcid.org/0000-0001-8463-3160)** were jointly responsible for the study concept, design, initiation, management and conduct.

Dharmishta Parmar (https://orcid.org/0000-0001-8767-8364) was responsible for the study informatics and manuscript organisation.

Shama Sheikh (https://orcid.org/0000-0002-7498-4538), **Oleg Blyuss (https://orcid.org/0000-0002-0194-6389)**, **Chris Wale** and **Jonathan Myles (https://orcid.org/0000-0002-1198-1209)** contributed to the statistical analysis.

Robert Smith (https://orcid.org/0000-0003-3344-2238) and **Peter Sasieni (https://orcid.org/0000-0003-1509-8744)** contributed to the oversight of statistical analysis and the interpretation of results.

Andrew Evans (https://orcid.org/0000-0002-3320-0215) was responsible for the radiological review.

Louise Johns (https://orcid.org/0000-0003-2837-1682) contributed to the study informatics and statistical analysis.

Ian Ellis (https://orcid.org/0000-0001-5292-8474) was responsible for the pathology review.

Publication

Duffy SW, Vulkan D, Cuckle H, Parmar D, Sheikh S, Smith RA, *et al.* Effect of mammographic screening from age 40 years on breast cancer mortality (UK Age trial): final results of a randomised, controlled trial. *Lancet Oncol* 2020;**21**:1165–72.

Data-sharing statement

Requests for data sharing should be sent to the corresponding author. Requests involving outcomes will be referred to NHS Digital, which provided the outcomes. Only anonymised data will be shared.

Patient data

This work uses data provided by patients and collected by the NHS as part of their care and support. Using patient data is vital to improve health and care for everyone. There is huge potential to make better use of information from people's patient records, to understand more about disease, develop new treatments, monitor safety, and plan NHS services. Patient data should be kept safe and secure, to protect everyone's privacy, and it's important that there are safeguards to make sure that it is stored and used responsibly. Everyone should be able to find out about how patient data are used. #datasaveslives You can find out more about the background to this citation here: https://understandingpatientdata.org.uk/data-citation.

References

1. Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. *Lancet* 2012;**380**:1778–86. https://doi.org/10.1016/S0140-6736(12)61611-0
2. Paci E, Broeders M, Hofvind S, Puliti D, Duffy SW, EUROSCREEN Working Group. European breast cancer service screening outcomes: a first balance sheet of the benefits and harms. *Cancer Epidemiol Biomarkers Prev* 2014;**23**:1159–63. https://doi.org/10.1158/1055-9965.EPI-13-0320
3. Moskowitz M. Breast cancer: age-specific growth rates and screening strategies. *Radiology* 1986;**161**:37–41. https://doi.org/10.1148/radiology.161.1.3532183
4. Miller A. Is routine mammography screening appropriate for women 40-49 years of age? *Am J Prevent Med* 1991;**7**:55–62. https://doi.org/10.1016/S0749-3797(18)30967-X
5. van Schoor G, Moss SM, Otten JD, Donders R, Paap E, den Heeten GJ, *et al.* Effective biennial mammographic screening in women aged 40–49. *Eur J Cancer* 2010;**46**:3137–40. https://doi.org/10.1016/j.ejca.2010.09.041
6. Ray KM, Joe BN, Freimanis RI, Sickles EA, Hendrick RE. Screening mammography in women 40-49 years old: current evidence. *AJR Am J Roentgenol* 2018;**210**:264–70. https://doi.org/10.2214/AJR.17.18707
7. Swedish Cancer Society and the Swedish National Board of Health and Welfare. Breast-cancer screening with mammography in women aged 40–49 years. *Int J Cancer* 1996;**68**:693–9. https://doi.org/10.1002/(SICI)1097-0215(19961211)68:6<693::AID-IJC1>3.0.CO;2-Z
8. Oeffinger KC, Fontham ET, Etzioni R, Herzig A, Michaelson JS, Shih YC, *et al.* Breast cancer screening for women at average risk: 2015 guideline update from the American Cancer Society. *JAMA* 2015;**314**:1599–614. https://doi.org/10.1001/jama.2015.12783
9. Siu AL, Force USPST. Screening for breast cancer: U.S. Preventive Services Task Force Recommendation Statement. *Ann Int Med* 2016;**164**:279–96. https://doi.org/10.7326/M15-2886
10. Monticciolo DL, Newell MS, Hendrick RE, Helvie MA, Moy L, Monsees B, *et al.* Breast cancer screening for average-risk women: recommendations from the ACR Commission on breast imaging. *J Am Coll Radiol* 2017;**14**:1137–43. https://doi.org/10.1016/j.jacr.2017.06.001
11. IARC Working Group on the Evaluation of Cancer-Preventive Strategies. Breast cancer screening. *IARC Handbooks Cancer Prev* 2016;**15**:1–469.
12. European Commission. *European Guidelines on Breast Cancer Screening and Diagnosis.* URL: https://ecibc.jrc.ec.europa.eu/recommendations/ (accessed 5 January 2020).
13. Sasieni P, Cuzick J. The UK breast-screening programme should start at age 47 years. *Lancet* 2003;**362**:246–7. https://doi.org/10.1016/S0140-6736(03)13922-0
14. Nelson HD, Fu R, Cantor A, Pappas M, Daeges M, Humphrey L. Effectiveness of breast cancer screening: systematic review and meta-analysis to update the 2009 U.S. Preventive Services Task Force Recommendation. *Ann Intern Med* 2016;**164**:244–55. https://doi.org/10.7326/M15-0969
15. Moss SM, Cuckle H, Evans A, Johns L, Waller M, Bobrow L, Trial Management Group. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. *Lancet* 2006;**368**:2053–60. https://doi.org/10.1016/S0140-6736(06)69834-6

16. Moser K, Sellars S, Wheaton M, Cooke J, Duncan A, Maxwell A, *et al.* Extending the age range for breast screening in England: pilot study to assess the feasibility and acceptability of randomization. *J Med Screen* 2011;**18**:96–102. https://doi.org/10.1258/jms.2011.011065

17. Moss SM, Wale C, Smith R, Evans A, Cuckle H, Duffy SW. Effect of mammographic screening from age 40 years on breast cancer mortality in the UK Age trial at 17 years' follow-up: a randomised controlled trial. *Lancet Oncol* 2015;**16**:1123–32. https://doi.org/10.1016/S1470-2045(15)00128-X

18. Cuzick J, Edwards R, Segnan N. Adjusting for non-compliance and contamination in randomized clinical trials. *Stat Med* 1997;**16**:1017–29. https://doi.org/10.1002/(SICI)1097-0258(19970515)16:9<1017::AID-SIM508>3.0.CO;2-V

19. Duffy SW, Lynge E, Jonsson H, Ayyaz S, Olsen AH. Complexities in the estimation of overdiagnosis in breast cancer screening. *Br J Cancer* 2008;**99**:1176–8. https://doi.org/10.1038/sj.bjc.6604638

20. Anderson TJ, Waller M, Ellis IO, Bobrow L, Moss S. Influence of annual mammography from age 40 on breast cancer pathology. *Hum Pathol* 2004;**35**:1252–9. https://doi.org/10.1016/j.humpath.2004.07.011

21. Anderson TJ, Sufi F, Ellis IO, Sloane JP, Moss S. Implications of pathologist concordance for breast cancer assessments in mammography screening from age 40 years. *Hum Pathol* 2002;**33**:365–71. https://doi.org/10.1053/hupa.2002.32222

22. Duffy SW, Vulkan D, Cuckle H, Parmar D, Sheikh S, Smith RA, *et al.* Effect of mammographic screening from age 40 years on breast cancer mortality (UK Age trial): final results of a randomised, controlled trial. *Lancet Oncol* 2020;**21**:1165–72.

23. Breslow NE, Day NE. Statistical methods in cancer research. Volume II – the design and analysis of cohort studies. *IARC Sci Publ* 1987;**82**:1–406.

24. Aalen O. Nonparametric inference for a family of counting processes. *Ann Stat* 1978;**6**:701–26. https://doi.org/10.1214/aos/1176344247

25. Duffy SW, Smith RA. A note on the design of cancer screening trials. *J Med Screen* 2015;**22**:65–8. https://doi.org/10.1177/0969141315577847

26. Michalopoulos D, Duffy SW. Estimation of overdiagnosis using short-term trends and lead time estimates uncontaminated by overdiagnosed cases: results from the Norwegian Breast Screening Programme. *J Med Screen* 2016;**23**:192–202. https://doi.org/10.1177/0969141315623980

27. Launoy G, Duffy SW, Prevost TC, Bouvier V. [Detection of cancer, sensitivity of the test and sensitivity of the screening program.] *Rev Epidemiol Sante Publique* 1998;**46**:420–6.

28. Carney PA, Miglioretti DL, Yankaskas BC, Kerlikowske K, Rosenberg R, Rutter CM, *et al.* Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. *Ann Intern Med* 2003;**138**:168–75. https://doi.org/10.7326/0003-4819-138-3-200302040-00008

29. Moss S, Thomas I, Evans A, Thomas B, Johns L, Trial Management Group. Randomised controlled trial of mammographic screening in women from age 40: results of screening in the first 10 years. *Br J Cancer* 2005;**92**:949–54. https://doi.org/10.1038/sj.bjc.6602396

30. Tabar L, Chen TH, Yen AM, Chen SL, Fann JC, Chiu SY, *et al.* Effect of mammography screening on mortality by histological grade. *Cancer Epidemiol Biomarkers Prev* 2018;**27**:154–7. https://doi.org/10.1158/1055-9965.EPI-17-0487

31. Bastardis-Zakas K, Iatrakis G, Navrozoglou I, Peitsidis P, Salakos N, Malakassis P, Zervoudis S. Maximizing the benefits of screening mammography for women 40–49 years old. *Clin Exp Obstet Gynecol* 2010;**37**:278–82.

32. Sasieni PD, Wald NJ. Should a reduction in all-cause mortality be the goal when assessing preventive medical therapies? *Circulation* 2017;**135**:1985–7. https://doi.org/10.1161/CIRCULATIONAHA.116.023359

33. Gøtzsche PC. The debate on breast cancer screening with mammography is important. *J Am Coll Radiol* 2004;**1**:8–14. https://doi.org/10.1016/S1546-1440(03)00017-6

34. Simonetto C, Eidemüller M, Gaasch A, Pazos M, Schönecker S, Reitz D, *et al.* Does deep inspiration breath-hold prolong life? Individual risk estimates of ischaemic heart disease after breast cancer radiotherapy. *Radiother Oncol* 2019;**131**:202–7. https://doi.org/10.1016/j.radonc.2018.07.024

35. Corradini S, Ballhausen H, Weingandt H, Freislederer P, Schönecker S, Niyazi M, *et al.* Left-sided breast cancer and risks of secondary lung cancer and ischemic heart disease: Effects of modern radiotherapy techniques. *Strahlenther Onkol* 2018;**194**:196–205. https://doi.org/10.1007/s00066-017-1213-y

36. Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S, *et al.* Diagnostic performance of digital versus film mammography for breast-cancer screening. *N Engl J Med* 2005;**353**:1773–83. https://doi.org/10.1056/NEJMoa052911

37. Wald NJ, Murphy P, Major P, Parkes C, Townsend J, Frost C. UKCCCR multicentre randomised controlled trial of one and two view mammography in breast cancer screening. *BMJ* 1995;**311**:1189–93. https://doi.org/10.1136/bmj.311.7014.1189

38. Duffy SW, Agbaje O, Tabar L, Vitak B, Bjurstam N, Björneld L, *et al.* Overdiagnosis and overtreatment of breast cancer: estimates of overdiagnosis from two trials of mammographic screening for breast cancer. *Breast Cancer Res* 2005;**7**:258–65. https://doi.org/10.1186/bcr1354

39. Duffy SW, Dibden A, Michalopoulos D, Offman J, Parmar D, Jenkins J, *et al.* Screen detection of ductal carcinoma in situ and subsequent incidence of invasive interval breast cancers: a retrospective population-based study. *Lancet Oncol* 2016;**17**:109–14. https://doi.org/10.1016/S1470-2045(15)00446-5

40. Johns LE, Moss SM, Age Trial Management Group. False-positive results in the randomized controlled trial of mammographic screening from age 40 ('Age' trial). *Cancer Epidemiol Biomarkers Prev* 2010;**19**:2758–64. https://doi.org/10.1158/1055-9965.EPI-10-0623

41. NHS Digital. *Breast Screening Programme England, Statistics for 2014–15*. NHS Digital; 2016. URL: www.hscic.gov.uk/pubs/brstscreen1415 (accessed 13 January 2020).

42. Hellquist BN, Duffy SW, Abdsaleh S, Björneld L, Bordás P, Tabár L, *et al.* Effectiveness of population-based service screening with mammography for women ages 40 to 49 years: evaluation of the Swedish Mammography Screening in Young Women (SCRY) cohort. *Cancer* 2011;**117**:714–22. https://doi.org/10.1002/cncr.25650

43. Cho N, Han W, Han BK, Bae MS, Ko ES, Nam SJ, *et al.* Breast cancer screening with mammography plus ultrasonography or magnetic resonance imaging in women 50 years or younger at diagnosis and treated with breast conservation therapy. *JAMA Oncol* 2017;**3**:1495–502. https://doi.org/10.1001/jamaoncol.2017.1256

44. Ciatto S, Houssami N, Bernardi D, Caumo F, Pellegrini M, Brunelli S, *et al.* Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study. *Lancet Oncol* 2013;**14**:583–9. https://doi.org/10.1016/S1470-2045(13)70134-7

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library