BREAST CANCER

THE COST OF SCREENING A NATION

A Polygeia Research Paper
November 2016

Editors: Alex Astley, Sophie Caseby
Researchers: Yasmine Bark, Kritica Dwivedi, Aishah Prastowo, James Whitehouse

1 in 8 women will be diagnosed with breast cancer during their lifetime

POLYGEIA
STUDENTS SHAPING GLOBAL HEALTH POLICY

Abstract

The national breast cancer screening programme, introduced in 1988, gave promise in reducing mortality of the most common cancer in the UK. Inviting women aged 50-70 to triennial mammographic screening aims to detect cancers at the earliest stage possible – improving prognosis and ultimately seeking to reduce the number of breast cancer deaths. The programme has led to an estimated 20% reduction in breast cancer mortality, saving 1320 lives each year, however it is continually debated whether breast screening does more harm than good. The controversies, which will be discussed in detail in this report, include inadequate cost-effectiveness, overdiagnosis and overtreatment, and drawbacks in clinical trial design. Our policy recommendations aim to address some of these criticisms, and focus mainly on a proposed restriction of screening to high-risk subpopulations of women, which could greatly overcome both the clinical and economic costs of breast cancer screening.

Abbreviations

BRCA1/2: Breast cancer 1/2 gene; BSE: Breast self-examination; BTWSP: Breast Test Wales Screening Programme; CBE: Clinical breast examination; CE: Cost-effectiveness; DCIS: Ductal carcinoma *in situ*; GDP: Gross domestic profit; ICER: Incremental cost-effectiveness ratio; LORIS: Low risk DCIS trial; LYS: Life year saved; MRI: Magnetic resonance imaging; NICE: National Institute for Health and Care Excellence; NHS: National Health Service; NHSBSP: NHS Breast Screening Programme; NYHIP: New York Health Insurance Plan; OECD: Organisation of Economic Co-operation and Development; PSA: Prostate-specific antigen; QALY: Quality-adjusted life year; RCT: Randomised controlled trial; *TP53*: Tumour protein p53 gene; UK: United Kingdom; WHO: World Health Organisation

Editor Profiles

Alex Astley

Alex is an MSc student at SOAS who specialises in public health of Sub-Saharan Africa. He has spent time on different public health projects including three months on a rural sexual health and hygiene project in Ghana. Upon returning from Ghana he has been interested in applying the skills he learnt to a UK scenario. He hopes that the project will lead to a widening understanding of the controversies surrounding breast cancer.

Sophie Caseby

Sophie became involved with Polygeia whilst studying for an MSc in Integrated Immunology at University of Oxford. Her initial interest in global health focused on control of infectious diseases, developed through research on parasitology at University of São Paulo, and vaccinology at Oxford's Jenner Institute. Sophie has enjoyed expanding her interests to UK public health, and is now working as an Analyst at one of Polygeia's collaborators, Costello Medical Consulting.

Researcher Profiles

Yasmine Bark

Yasmine is a European Research Council PhD scholar. She is a pharmacist and has recently graduated with an MSc in Experimental Therapeutics from University of Oxford. Yasmine is interested in public and global health. She is involved in the project addressing the effectiveness of the breast cancer screening programme.

Kritica Dwivedi

A third year medical student at Exeter College, Oxford, Kritica developed her interest in academic research and international health policy during her final honour's school project researching tuberculosis vaccine strategies. She is now extremely excited to be part of Polygeia, researching the current breast cancer screening programme.

Aishah Prastowo

Aishah is a third year DPhil student in Engineering Science at Brasenose College, Oxford. Her thesis is on novel microfluidic techniques for biotechnology applications. She is interested in the interface between engineering and biosciences, as previously she did a Masters at Paris Descartes University (France), studying interdisciplinary approaches to life sciences, and an undergraduate degree in engineering at Universitas Gadjah Mada (Indonesia).

James Whitehouse

James is currently a fourth year student studying History and Philosophy of Science at Homerton College, Cambridge. He is particularly interested in evidence based medicine and biomedical ethics. He has previously written a dissertation on the ethics of breast cancer screening, and joined the Polygeia team working on this issue.

Table of Contents

Introduction

The national breast cancer screening programme, introduced in 1988, gave promise in reducing mortality of the most common cancer in the UK, which one in eight women are diagnosed with during their lifetime. The scheme invites women aged 50-70 to mammographic breast screening every three years, allowing tumours to be detected at the earliest stage possible. This facilitates less aggressive treatment, aiming to improve prognosis and ultimately reduce number of breast cancer deaths (1).

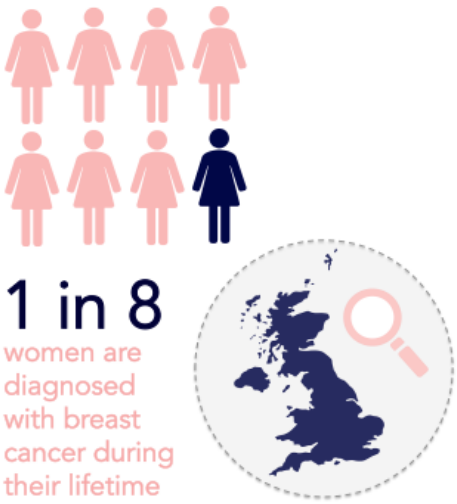

Figure 1. Breast Cancer across the UK

The programme has proven success, having detected 13,000 invasive breast cancers in one year, with an estimated 20% reduction in breast cancer mortality (2). However, it is still continually debated whether screening causes more harm than good, and the Independent UK Panel on Breast Cancer Screening were commissioned in 2012 to review both the costs and benefits of the programme. The panel, which consisted of oncology and epidemiology experts, as well as a patient advocate, concluded that 'breast cancer screening reduces mortality at a cost of overdiagnosis'. They also highlighted how transparent and objective information should be made available, allowing women to make informed decisions on their attendance to screening (2). As well as overdiagnosis and associated overtreatment, criticisms have also targeted the inadequate cost-effectiveness, since several estimates have put the cost of screening at far higher than £30,000 per quality-adjusted life year (QALY) – the threshold used by the National Institute of Health and Care Excellence (NICE) to determine whether interventions should be funded by the NHS (3). Lastly, concerns also surround clinical trial design, since their relevance to present-day screening is continually questioned.

This report will discuss the economic and clinical issues of mammographic screening and associated overdiagnosis. To address these issues within the wider context of screening, we will investigate the success of mammographic programmes in other countries, and evidence from the clinical trials which led to recommendation and roll-out of the UK programme. The risk of radiation-induced cancers within the target population for screening, which is due to be extended to women aged 47-73, will also be explored as well as the role of breast-self examination. Lastly, we will present our own sustainable policy to address some of the criticisms the screening programme has faced.

Introductory Discussion on the Cost-Effectiveness of the Screening Programme: the NICE Threshold

Yasmine Bark

The breast cancer screening programme was established in 1988, subsequent to publication of the Forrest Report in 1986 (4). Additional data from observational studies and randomised trials have since become available, fuelling an ongoing discussion debating about the benefit vs. harm of the screening programme (5). The Independent UK Panel on Breast Cancer Screening's 2012 review revealed the main harm as overdiagnosis (2). This is the otherwise-not-detected diagnosis of breast cancer resulting from screening. An estimated 9% to 12% of excess breast cancer incidence was seen coupled with long term screening, and 15% to 23% of incidence of the disease associated with active screening in the invited groups studied by the panel. The conclusion was that though screening contributes to reduced deaths from breast cancer, the costs of overdiagnosis and overtreatment emerge (3).

In 1999, the National Institute for Health and Care Excellence (NICE, formerly known as National Institute for Clinical Excellence) was established by the Secretary of State for Health. The initial purpose of this institute was to make sure that the NHS in England and Wales are on hold of the most clinically and cost-effective drugs and treatments. At the government level, NICE creation is expected to improve access to the best value treatments and guidance across the NHS. NICE guidance for breast cancer assessment includes an invitation to all women between 47 and 73 years old for screening mammography every three years through the NHS Breast Screening Programme (NHSBSP) in England or the Breast Test Wales Screening Programme (BTWSP) in Wales (2,6). A report was submitted to the NHS in 2011 about the cost-effectiveness of screening policies in elder women. It was suggesting that extending the upper age limit for the group of women invited to the screening programme from 70 to 78 years would bring cost-effective use of the available resources. However, the report presented several limitations, since it is usually harder to calculate the quality of life in the elderly population mainly due to the shortness of life expectancy. However, for most patients, diagnosis in the breast clinic should be based on a triple assessment: 'clinical assessment, mammography and/or ultrasound imaging, and core biopsy and/or fine needle aspiration cytology'. On follow-up, a copy of the imaging recommendations as stated by NICE is as follows (7):

Follow-up imaging

- *Offer annual mammography to all patients with early breast cancer, including DCIS, until they enter the NHSBSP/BTWSP. Patients diagnosed with early breast cancer who are already eligible for screening should have annual mammography for 5 years.*

- *On reaching the NHSBSP/BTWSP screening age or after 5 years of annual mammography follow-up we recommend the NHSBSP/BTWSP stratify screening frequency in line with patient risk category.*
- *Do not offer mammography of the ipsilateral soft tissues after mastectomy.*
- *Do not offer ultrasound or MRI for routine post-treatment surveillance in patients who have been treated for early invasive breast cancer or DCIS.*

NICE is also in charge of evaluating the clinical and cost-effectiveness of the treatments. A recommendation in regards to this task then follows within the UK NHS. An analysis of the cost-effectiveness takes into consideration at least two courses of action and appraises their benefits to their costs. The expected incremental cost-effectiveness ratio (ICER) is used to summarise the comparison (8). Results are usually described with cost per QALY gained. NICE's approach adopts a tool to allow the consideration of the different courses of action by decision makers whenever there are multiple options (9). According to NICE appraisal process, to establish that a course of action, which is in this case the screening programme, is likely recommended for use in the NHS, the threshold should typically be between £20,000 and £30,000 per QALY (10). Assumptions used by economic models include health state utilities in six conditions starting with: disease-free, and stages 0 (cancer *in situ*), I, II, III, and IV of breast cancer. These last five stages are identified after diagnosis. Diagnosed women are assumed to be in one of the five health states of the disease over a predefined period as determined by their prognostic profile. There are no sufficient data on the duration of the diagnosis process on quality of life hence the need for the assumptions. Patients in stages 0, I/II, III, and IV are assumed to have a reduced quality of life for one, two, three years, and lifetime respectively. At the end of the state duration, they are assumed to return to disease-free (11). The latest economic evaluation submitted to the NHS suggested that for the breast cancer screening to be cost-effective under a willingness-to-pay of £20,000 per QALY is only moderately likely and that there are still uncertainties in the effects (3).

Randomised Controlled Trials for Breast Cancer Screening

James Whitehouse

This section looks at the early evidence for screening, which was used to bring in the NHS screening programme. It considers what the Forrest Report did with the RCT and cost data in calculating expected QALY outcomes. It then looks at problems with the evidence used, and makes suggestions for how better evidence could be obtained.

The Forrest Report

The introduction of the NHS breast screening programme in 1988 followed the recommendations of a government committee on screening, which were published in 1986 as the Forrest Report. This section looks at the evidence from the randomised controlled trials (RCTs) which were used to make the recommendations in the Forrest Report. As of 1986, there were only two RCTS which had delivered any meaningful results, the New York Health Insurance Plan trial (hereafter NYHIP), and the Swedish Two County Trial. There are two main reasons for this scarcity of trials. Firstly, because cancer screening is for women who have no symptoms, any trial requires a very large sample size to find any statistically significant benefit, usually over 100,000 women in total. Secondly, because cancer is a relatively long-term disease, trials must be run for many years before any results can be drawn. Initial results of the Two County trial were published nine years after the trial began, and the NYHIP trial ran for 18 years (4). The headline results of both of these trials, which the Forrest Report uses, is that the rate of breast cancer mortality was significantly reduced in both trials, NYHIP found a 23% reduction after 18 years, the Two County trial found a 31% reduction after 7 years (4). [This interim result was available to the Forrest Committee; the final results of the trial were not published until 1995 (12).] These results were considered sufficient to justify the NHS screening program, combined with the assumption that screening was economically beneficial, due to the lives it could save (4).

Expected QALY Outcomes

The Forrest Report focuses on QALY measurements as the basis for judging the effectiveness of screening, and extrapolates results from the two RCTs to make economic predictions. The problem here is that the extrapolation is done directly, and because of the relatively few number of deaths per year in the two RCTs, there is a lot of noise in the data. For example between years 5-9 the expected number of breast cancer deaths in the screened group is between 27 and 35, except for in year 7 where there were 50 deaths. Most annual results show approximately a 50% reduction in breast cancer mortality, but year 7 shows less than a 20% reduction. A much larger sample size of women with breast cancer would give a much better signal to noise ratio, giving smoother data, as well as reducing the risk that the positive result for screening is due to chance. Smoothed data would produce a clearer target figure for the expected reduction in breast cancer mortality; i.e. the 20-30% figures generated from the two RCTs. The conclusions reached in the report are based on the economic analysis of the data from the two RCTs.

Combined with evidence gathered from an Edinburgh study on breast cancer screening, the Forrest Report calculates a cost of £11.66 per individual screening (4). Using the results from the RCTs a total cost per QALY is generated, with a baseline assumption of 0.5% breast cancer incidence per year and a 15 year increase in life expectancy for those who are diagnosed via screening (4). From these assumptions, the cost per QALY was deemed to be £3,044 per QALY in 1986 value, which accounting for inflation corresponds to £8,300 in 2016. As of 2016, the NHS considers treatments that cost less than £20,000 per QALY to be cost-effective, and so from this data the screening programme is seen to be cost-effective. However, whilst the economic analysis is highly detailed, including taking into account the travel costs for the women being screened, the clinical outcomes, which affect the cost-effectiveness the most, are not as definitive as the RCT results claim. One such clinical outcome is false positive diagnoses.

Ignoring False Positives

A problem with any medical test is the risk of false positives. Mammography does not directly detect cancer, but looks for different density of breast tissue. Different tissue density can be due to cancer, but also many other factors such as cysts, benign lumps or just natural variation in tissue density. False positives occur when an individual is incorrectly diagnosed with having a disease when they actually do not, and is inevitable as no test is perfect. However, false positives are especially relevant in the case of breast cancer screening because it is a test on asymptomatic people. The rate of breast cancer is low in the screened population of women aged 50-70, at around 0.5% per year. Because of this low incidence, the base rate fallacy becomes relevant. In short, the problem is that even a very accurate test used on a low incidence disease produces many false positives compared to actual positives. The accuracy of a test is measured by its specificity and sensitivity. Sensitivity measures the percentage of people who do have the disease and will be correctly diagnosed as such by the test. Specificity measures the rate of true negatives, which controls the rate of false positives. If a test has 95% specificity then 95% of the time it is used on those who do not have the disease, it will give a negative result. A test with 95% sensitivity and specificity appears very accurate, however it becomes less useful when applied to a population with a low incidence of a disease. Sensitivity and specificity for screening have been measured in the region of 85-95% (13). Consider the outcomes for a test that has 95% specificity and sensitivity. If 1000 women are screened, then we would expect five cases of breast cancer to be correctly diagnosed, 50 women to have false positive results, and 945 women to correctly be given a negative result. This means that less than 10% of those who have a positive screening result actually have breast cancer. In reality, this is why second screenings are done, as well as biopsies and other procedures to try to reduce the number of false positive diagnoses that lead to treatment. The only real solution to the base rate problem is to find a way to raise the rate of breast cancer in the screened group. This can be done by only offer screening to those at an elevated risk of cancer, due to genetic or environmental factors. For example, those who have a mutated BRCA1 gene have a 60-90% lifetime risk of breast cancer, and so screening would be more beneficial for them than the general population (14). This is also an argument against extending the age range of current screening programmes as women under the age of 50 have an even lower risk of developing breast cancer (15).

Screening can detect cancer long before symptoms occur; in many cases it can detect what is termed 'stage 0' cancer, or ductal carcinoma *in situ* (DCIS). Around 30% of DCIS cases become invasive cancer, so pre-emptive treatment can be recommended (16). However this will lead to 70% of women with DCIS being unnecessarily treated. Because screening can detect many asymptomatic cancers as well as DCIS, it is expected that there should be a higher rate of cancers detected in a screened group than the control. However, in the NYHIP trial, after 7 years there was no difference in cancer detection rates between the two groups. The Forrest Report stated that 'overdiagnosis was not a problem' in the case of NYHIP, and that the 20% overdiagnosis seen in the Two County trial may have been due to an improvement in the sensitivity of mammography over time (4). However, the negligible difference between breast cancer rates in the NYHIP trial groups suggests that the sensitivity of early mammography was very low.

Criticisms of RCT Results

The results from the NYHIP trial may favour screening, but should not be seen as an unqualified recommendation. In the study group there were 304 diagnoses of breast cancer in the first five years of the study. The study group was made up of 31,000 women who were offered four screenings at yearly intervals (4). 79 of the 304 cases were in women who were in the screening group, but who chose not to be screened, and so can be disregarded. Of the cases where the women actually underwent screening, 132 were diagnosed by screening, and 93 were diagnosed after a negative screening test (4). This low detection rate is further compounded because 'screening' in the NYHIP trial involved both mammography and a clinical breast exam (4). Of those 132 diagnoses via screening, only 55% were present in the mammography, the other 45% were diagnosed via clinical exam after mammography gave a negative result. If the sensitivity of mammography is only 55%, this might be enough to consider the NYHIP results to be invalid, considering that all of the later RCTs claim to have around a 90% sensitivity (4). Without these results, the only useful data at the time of the Forrest Report came from the Two County Trial , which did not have the same low sensitivity issues.

However, the results of the Two County Trial have also been questioned in a 2000 meta-analysis by Gøtzsche and Olsen, which claimed that there was insufficient randomisation in many of the RCTs, including the NYHIP and Two Country Trial, which had led to the effectiveness of screening being overstated (17). The Two County Trial was criticised for its method of cluster randomisation, where groups of the population instead of individuals were the units of randomisation. Because of the clustering there is an age imbalance in both counties. In Kopparberg the study women were 0.45 years older than the control group ($P<0.0001$), and in Östergötland the study women were 0.27 years older than the control group ($P<0.0001$) (17). Therefore, even though the age discrepancy would favour the control group, there are clear issues with randomisation in this trial. If the two groups are not sufficiently similar in age, then it is possible that they had other underlying differences with regards to health, and so the results of the trial may be biased. Hence, Gøtzsche and Olsen argued that the results from both of the early trials should be ignored in their meta-analysis.

Ideal Trials

We now consider how an ideal trial into the efficacy of screening should be run, which would overcome some of the issues with the early RCTs. It should have the components of any good RCT, such as sufficient randomisation (on an individual basis), even though double blinding is not possible. The clinical endpoint, i.e. what outcome is being measured is important. For example, the average length of time patients live with cancer for is not a good endpoint, because screening detects cancer earlier than other methods, and so artificially inflates this figure. Both the NYHIP and Two County Trial measured the rates of breast cancer mortality in the two groups. This is not subject to the same bias as the previous suggestion; however it does still favour a positive result for screening because it ignores other negative outcomes. Such negative outcomes include complications from the screening procedure or biopsy as well as overdiagnosis and overtreatment, which will negatively impact the overall benefits of screening. Ideally the measurement should be the average lifespan in two otherwise identical groups, because this accounts for all outcomes of screening, both positive and negative. This will not produce such a simple headline figure as '30% reduction in breast cancer mortality', but it is more truthful. It is necessary to know all the outcomes to calculate a true QALY measurement. If those who are screened have even a slightly elevated risk of dying due to the screening procedure, then this needs to be considered as a cost of screening and included in any cost-benefit analysis.

Policy Recommendations

- Consider offering screening to only high-risk populations.
- Use overall lifespan as the measurement in future RCTs.
- Collect more data on overdiagnosis from screening to use for future cost-benefit analysis.
- Do not extend the screening program to include younger women until there is sufficient evidence to justify the increased risk of overdiagnosis.

Cost-Effectiveness of Breast Cancer Screening Programmes in Different Countries

Aishah Prastowo

In the UK, breast screening is currently offered to women aged 50-70 by the NHS (18) and currently – in some areas – is at the process of extension to ages 47-73. Similar programmes have been running in other countries, with various age ranges and screening intervals. Here we compare breast cancer screening programmes in different countries to the one offered by the NHS, in order to explore the possibility of adopting a more effective screening approach.

In 2013, Yoo et al. systematically reviewed and compared the cost-effectiveness of breast cancer screening programmes in different countries (19). They collected cost-effectiveness (CE) data from previous studies which define CE as cost per either life year saved (LYS) or quality adjusted life year (QALY). LYS and QALY are both used to evaluate health interventions; LYS is a more straightforward approach based only on the living years, while QALY takes into account the health quality (20). For comparison, the cost per LYS/QALY is normalised to each country's per-capita GDP as of the reference year (denoted here as CE/per-capita GDP). The figures accumulated in this study are reproduced in **Table 1**.

Table 1. Cost-effectiveness of breast cancer screening programmes in different countries. If there are several studies on the same country, only the latest study is presented here. Adapted from (19).

Country	Reference Year of Cost	Cost (US$) per LYS/QALY	Per-capita GDP (US$)	CE/Per-capita GDP Ratio
India	2001	3308	460	7.19
Japan	1986	14300	16882	0.85
China	2008	64400	1731	37.20
South Korea	2009	29964	16959	1.77
Spain	2005	4691	26056	0.18
UK	1985-1986	3730	13009	0.37
Finland	1995	18955	25609	0.74
Norway	1996	14554	36555	0.40
The Netherlands	1985-1990	3235	16116	0.20
Switzerland	2007	15468	57490	0.27
US	2006	48884	46760	1.05

Based on the CE/per-capita GDP, among 11 countries, the UK has one of the most cost-effective programmes, ranking only below Spain, The Netherlands and Switzerland. Asian countries tend to have high CE/per-capital GDP ratio, due to low cancer incidence rates and less sensitive screening because of racial characteristics. Because of these inherent differences, it is more relevant to compare the UK screening programme with other Western countries than Asian countries.

The programme details of the EU countries included in Yoo et al.'s study (19) are listed in Table 2 (21). The UK has the longest interval between screenings, i.e. three years compared to two years for Spain, Finland and The Netherlands. The screening coverage is also lower than in Finland and the Netherlands. For the target population screened, the lower age limit of 50 is used in all these countries, except for in several regions in Spain where screening is offered to women from age 45. The upper age limits are 69 (Spain, Finland), 70 (UK) and the highest is 74 (the Netherlands).

Table 2. Population-based nationwide breast cancer screening programmes in some EU countries as of March 2014. Adapted from (21).

Country	Programme Start Date	Nationwide Coverage	Screening Interval (years)	Age of Target Population	% Coverage in 2010
Spain	1990	2009	2	(45) 50-69	67
UK	1988	1995	3	50-70	73
Finland	1987	1989	2	50-69	85
The Netherlands	1988	1997	2	50-74	80

Alternative screening strategies – i.e. alternative target population age range and screening interval – have been proposed and simulated in different studies for different scenarios, which resulted in different argumentation regarding screening of women aged younger than 50. Whilst some showed potential improvement in CE with the extended lower age limit, it is still a debatable recommendation to reduce the lower screening age limit, as the risk factors in screening younger women need to be included, e.g. overdiagnosis. For example, in the study by Sankatsing et al. for the Netherland's case, although extending the screening age range for women aged 40-49 is cost-effective, an increase of overdiagnosis by 74 per 1000 women was estimated for this proposed extended range (compared to the current lower age limit of 50) (22). Carles et al.'s study on Catalonia also supports screening for younger women, and through using LYS or QALY as an output, effective screening could be achieved when the starting screening age is either 40 or 45, with 69 as the upper limit; however the overdiagnosis consequence was not calculated in their model (23). In contrast to those two studies, a simulation performed by Gocgun et al. for Canada's case revealed the ineffectiveness of screening women aged 40-49 based on the LYS (24) – this contradicting evidence further highlights the complications in calculating cost-effectiveness of population-based screening programmes.

Radiation Risk: Should We Screen Women Under the Age of 50?

Alex Astley

In most countries belonging to the Organisation of Economic Co-operation and Development (OECD), population breast cancer screening is available for women from the ages 50-70, however there has always been significant debate about the relative advantages of expanding the age range to include women from aged 40-50. In recent years the UK government has repeatedly stated it's intention that it wants to widen the screening age, to include women aged 47 (and up to 73), once again bringing the debate to the forefront of public attention. Although most of the wider debate surrounding the issue of breast cancer screening (as fore mentioned in this paper) has focused on false negatives and overdiagnoses, the contentious issue here is the risk of radiation-induced cancers. By focusing solely on the relative risk of radiation-induced cancers (and largely ignoring the other issues) we will show that the risk of radiation-induced cancers is not strong enough to warrant excluding women under the age of 50 from screening programmes.

Whilst there is lot of disagreement over the harms of screening women under the age of 50, there is general agreement on the benefits (25). In Hellquist et al.'s 2011 study, of the screening programme in Sweden, they found that the increase in age range led to a 26% reduction in mortality (26). Their finding does have some limitations due to the non-random allocation of women into control and trial groups; this is due to the peculiarities of how the study began. In 1987/88 the guidelines, in Sweden, for who was to be screened were changed allowing councils who had a lack of resources to restrict access for screening to women aged 50-70. This restriction therefore allowed Moss and her team to compare the results between the councils. Although this does affect randomisation it is our belief that this doesn't inherently damage their finding due to the fact it was a population study and that there is no evidence to suggest there was a correlation between councils who restricted access to the screening programme and prior breast cancer prevalence.

Indeed Hellquist et al's findings are supported by many other studies including most notably Moss' 17 yearlong study in the UK. Moss et al. found that there was a significant reduction in mortality rates for those in the treatment group compared with those in the control group in the first 10 years (25). However, unlike the Swedish study, Moss' study did not find a difference in mortality between the control and treatment groups after 10 years (25). There are several potential reasons why we see this convergence in mortality but the most compelling of the theories is that whilst the screening programme reduces mortality for those with grade one or grade two tumours it only delays mortality for those with grade three tumours. This postponement of mortality will therefore cause a spike in mortality after a period of time; whilst this may appear to reduce the effectiveness of the screening programme, it is important to note that although it has not reduced mortality it has led to an increase in life expectancy for women affected. Increasing life expectancy by 10 years in itself is a significant positive impact, even if

there so far has been no gain in 'curing' breast cancer for women with grade three tumours.

The increase in life expectancy, alongside the reduction in mortality, has meant that screening women aged 40-49 has been calculated as cost-effective at £18,838 a year (27). This clearly puts it within NICE's guideline for treatment it will support. However this figure is calculated using the current methodology regarding overdiagnosis and false positives, which we have previously criticised. We therefore believe this figure is an overestimation of the cost-effectiveness, and that in reality it will have a much higher cost per life saved. Nonetheless this is a criticism in the way risks and negatives are calculated, it does not discredit the ample evidence that screening women aged 40-49 reduces mortality.

Although there is a reduction in mortality there is significant controversy over the associated costs of widening the age range for the screening programme. Of particular focus is the risk of radiation-induced cancers (25). As part of all screening programmes women are offered a mammography, which works in a similar way to an X-ray in that it uses radiation to detect tumours. A side effect of using radiation, as a method of detection is that it can also induce cancers, the more exposure women are given to radiation the higher the chance of a radiation-induced cancer (28). Thus if the age range is widened, particularly if it is lowered, then women will receive a higher exposure to harmful radiation. This higher exposure to radiation and the ensuing higher risk of an induced cancer, potentially even higher mortality risk, has led many to argue that the screening programme should not be extended (27).

However Gelder et al claim that the arguments against expanding the screening programme due to high risk of radiation-induced cancers are based on outdated evidence (28). They find that previous studies had overestimated the level of radiation a standard mammography test emitted; the latest figures find that the average screening produces just 1.3 MGy compared with the previous average of 2.38 MGy. Only 1% of the mammographies tested produced over 5MGy or higher. Secondly they found that radiation causes fewer induced cancers than previously thought. At 1.3 MGy Gelder et al. found that there would be an increase in the mortality rate by 3.7 per 100,000 women, or 349 prevented deaths for every induced death. At 5 MGy, which we established as extremely rare, there would be an increase in the mortality rate by 6.3 per 100,000 women, or 178 prevented deaths for every induced death. Whilst it is always regrettable that medical interventions can have negative impacts it is important to note that no procedure is without controversies. Using tools of analysis similar to those used by national health agencies, Gelder et al. find that even if their calculation of risk was wrong by a factor of three the benefits would still outweigh costs.

Thus the overwhelming evidence suggests that the costs of radiation-induced cancers are far outweighed by the benefits. This means that we have reached the conclusion that the screening of women aged 40-49 should not be restricted because of the risk of radiation-induced cancers.

Of course this is based upon the current method of cost analysis, which as we have pointed out grossly underestimates the effect of false positives and overdiagnosis which

both affect women aged 40-49 in a far greater propensity. Moreover the values placed on lives lost due to side effects are in essence entirely arbitrary. Whilst we will go on and try and provide a working solution to the former problem we have not and cannot, in such a short paper, propose a philosophical answer to the second issue. We do however note that this is an area of the literature that has been largely ignored but due to the limitations of our study have largely restricted ourselves to an analysis of clinical and economic costs of the screening programme.

Policy Recommendations (with the caveats discussed)

- Radiation risk is low-negligible meaning it cannot be used as grounds to not widen the age range of the screening programme.
- Further, more public, research is needed on the ethical assumptions that we make when valuing lives.
- Although low-risk there needs to be more investment into screening technology that uses lower doses of radiation.
- Although low-risk unnecessary screenings should be avoided.

Overdiagnosis and Overtreatment

Kritica Dwivedi

One of the most common issues raised against the breast cancer screening programme introduced in the UK has been that of overdiagnosis. Overdiagnosis is when a breast cancer is diagnosed that would not have caused a problem to the individual in their lifetime (2,29–32). This is due to the nature of breast cancers - as women age, a large majority would go on to develop some form of cancerous growth in their breasts; however, due to the slow progressing nature of these cancers, many would not have ever caused any symptoms or issues for the women in their lifetime, until they died of a different cause (33). It is based on the principle that, if you autopsy elderly women who died of other causes, some form of cancerous lump that was undetectable and asymptomatic in their lifetime would often be found in the breast tissue (29,31,34,35). For example, around 10% of invasive breast cancers are not symptomatic during lifetime, but detectable post mortem (36). This is distinct from, and should not be confused with, false positives. A false positive is when a cancer is falsely diagnosed where none exists, whereas overdiagnosis is when a 'cancer' is truly present, but would not progress to be dangerous to the health of the individual in their lifetime (30).

Although it may not seem like a substantial problem, overdiagnosis of cancers that pose no threat to the patient in their lifetime is a significant issue. This is because overdiagnosis leads to overtreatment: unnecessary treatment of these cancers, which can have significant physical and psychological consequences (30,33). Moreover, overdiagnosis can also lead to further testing to confirm the diagnosis, such as biopsies and fine needle aspirations. These not only have negative health impacts for the patient, but also financial implications for the NHS which ultimately reduces the cost-effectiveness of the screening programme on a national scale (30). Therefore, we must carefully consider the issue of overdiagnosis when making policy recommendations for the NHS as a whole.

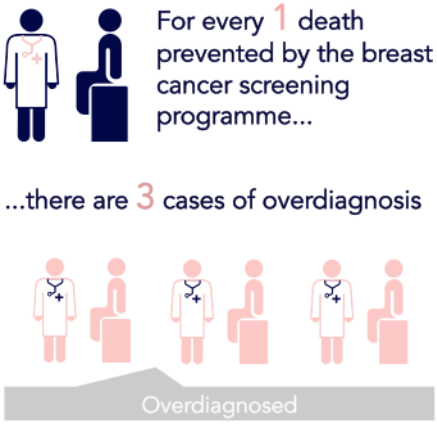

Figure 2. Overdiagnosis Caused by Mammographic Screening

Which cancers are overdiagnosed?

Certain cancers are more likely to be overdiagnosed than others. There is a higher risk of overdiagnosis of ductal carcinoma *in situ* (DCIS) than invasive breast cancers (37). However, any form of breast cancer – DCIS or invasive – can be overdiagnosed, either because the tumour lacks the potential to progress to clinical stage or might regress on its own, or because the woman dies of other causes before tumour surfaces clinically due to its slow progressing nature (30). Rapidly progressing, invasive cancers are detected by screening and lead to efficient treatment of the cancer and early detection. However, DCIS rather than invasive breast cancer is more likely to be overdiagnosed

(38–41). It is often considered an early stage of breast cancer and is often more slow progressing: DCIS is the most common type of non-invasive breast cancer, found inside the milk ducts of breast tissue that has not invaded into surrounding tissue (therefore, termed *in situ*). DCIS itself is not life-threatening and tends not to produce symptoms, with most diagnoses occurring due to screening mammograms (40). However, having it does increase the risk of developing an invasive breast cancer later on, as the cancer cells could eventually invade the surrounding breast tissue and metastasise to other organs (31). The biology of DCIS and its potential to progress into invasive cancer has yet to be formally established (38–42).

It is worth noting, however, that studies show that only a subset of DCIS progress into a clinically significant invasive cancer in a patient's lifetime (38,43,44). However, the natural history of DCIS, and the lifetime risk of it developing into a clinically significant invasive breast cancer is still unknown and debated (31,42,45). A study estimated that <5% of women diagnosed with DCIS will die of breast cancer within 30 years after their diagnosis (46); however, another study claimed that, if undetected, 84% of high-grade DCIS would progress to invasive disease in 5 years, most intermediate grade DCIS would progress to invasive disease in 10 years, and low-grade DCIS could become invasive in 15 years or longer (35). Although the true life time risk of development of DCIS to invasive cancer is unknown, it is largely estimated than less than 50% of DCIS would progress to invasive cancers if left untreated (30,41). DCIS is a common diagnosis out of the screening programme, and cannot be ignored (31): approximately 60,000 cases of DCIS are diagnosed in the United States (US) each year, accounting for about one out of every five new breast cancer cases, and about one in four breast cancers diagnosed by the screening programme in the UK (30).

The increasing evidence for overdiagnosis

Overdiagnosis has come a long way from first being denied when screening programmes were introduced, to becoming accepted as a common problem of such a screening programme. In 2012, the US National Cancer Institute agreed that overdiagnosis and overtreatment need to be recognised as a common problem of cancer screening programmes (38,39). The issue is now supported by various randomised control trials as well as observational studies. For example, early cancer detection increased when the screening programme was introduced, as would be expected by such a programme with the ability to detect cancers before they were symptomatic; however, no significant decrease in advanced cancer detection, as would occur if cancers were being detected earlier before they reached an advanced stage, was observed (31,38,41). For example, in Norway, a 58% increase in stage 1 (localised cancer) and 22% increase in stage 2 (regional cancer with or without lymph node involvement) was observed, but no subsequent decrease in advanced stages (stage III and IV) was observed (30). This suggests that although more cancers were being detected early, it was not early detection of invasive cancers, but that of cancers that did not present as advanced cancers in the unscreened population that were largely being picked up by the screening programme.

Estimating overdiagnosis – why is there no agreement?

Estimates of overdiagnosis have been extremely varied. There have been many problems estimating the frequency of overdiagnosis, with as varied estimates as 0-50%

from the studies discussed below (33). In a review of over 50 individual studies, the estimates ranged from anywhere between 0-35% of all diagnosed cancers (2). Excess incidence studies in the US found at least 20% screen-detected breast cancers are overdiagnosed. In contrast, a European review from seven countries estimated 1-10% overdiagnosis, adjusting for breast cancer risk and lead time (33). On the other hand, a study in Norway and Sweden concluded up to 54% cancers diagnosed through the screening programme were overdiagnosed (30,32).

A major cause of divergence in estimates between studies occurs because of the differences in study design (30,33) There have been two main approaches to determining overdiagnosis: an excess incidence observed approach, which looks at the differences in overall incidence of cancer in a screened versus unscreened population; and a lead-time approach which makes inferences about lead-time or natural history of breast cancer and estimates overdiagnosis based on these assumptions (29,47). Lead-time is the estimated amount of time that a diagnosis is brought forward due to early detection by screening, so the time by which screening advances the detection of a cancer (29,35). Lead-time approaches adjust for this because, upon first initiation of a screening programme, a sudden increase in diagnosis of cancer would be observed as asymptomatic cancers which would develop into clinically significant cancers are detected earlier than before, and should not be included in an estimate of overdiagnosis (33,35). However, estimates for lead-times make assumptions about the data (41), and there has been a suggestion that lead time is often substantially overestimated when overdiagnosis is not taken into account, and subsequently vastly underestimates overdiagnosis in these model-based approaches (37,48). This is a possible explanation for why statistical modelling studies which use a lead-time based approach often report a lower than 5% estimate of overdiagnosis, whereas observational studies which use an observed excess incidence approach range from about 22-54% (30,47).

Further to study design, there are many other sources of variation when determining an estimate for overdiagnosis that may be contributing to the discordance amongst studies (37,49). Many of the studies differ in patient populations with different underlying cancer incidence trends due to differences in risk factors and age ranges screened, as well as differences in screening and follow-up times along with screening policies, uptake and intensity (30,33,41,47). Moreover, when calculating overdiagnosis estimates, studies often differed in the assumptions used for lead-time and progression of DCIS into invasive cancers, and whether they used both DCIS and invasive cancers in the calculations when determining overdiagnosis (37,47,50,51). A Cochrane Review in 2009, which accounted for bias and differences in study design and the heterogeneity in data, estimated overdiagnosis to be about 30% upon meta-analysis of seven randomised controlled trials (32). The review, however, did comment on the fact that it did not include DCIS in its estimates, and therefore the true estimate of overdiagnosis may be even higher (30,32). Upon review of the 11 randomised controlled trials identified in the Cochrane Review, a provisional estimate of 19% of all diagnosed cancers in women who were screened was agreed upon by the Independent UK panel led by Marmot in 2012 (2).

Moreover, a recent study recently commented on the fact that randomised controlled trials often look at data on an intention-to-treat basis (52). This means that they include all participants in their originally allocated group (i.e. screened vs. unscreened), even if the participants do not adhere to the group they were allocated or withdraw early from the trial. Although this is good practice as it reduces selection bias and confounding, when assessing the effect of interventions, such as the risk of overdiagnosis to an individual patient, it dilutes the significant measure of effect. This is because estimates made for all women invited to screening would be attenuated compared to those who actually attend screening and adhered to the protocol of the study (52). In response, a meta-analysis of the same randomised controlled trials analysed in the Cochrane and Marmot reviews found that the de-attenuated overdiagnosis estimate (estimating only for women who actually adhered to the group they were allocated) was 29.7% of diagnosed cancers (52). Due to the increasing evidence for overdiagnosis and unnecessary treatment of many DCIS cases, a low-risk DCIS trial (LORIS) was recently registered, to compare the effects of immediate treatment of DCIS compared with no treatment/active monitoring in the UK, and should provide an interesting addition to the discussion when its results are published years down the line (43)(44). The state of the field at the moment requires a clear accord to be reached with regards to the standardisation of a method for estimation of overdiagnosis, as the diversity of the estimates currently leaves it challenging to interpret the results or make policy decisions based on such heterogeneous data (2,33,47).

Why is overdiagnosis a problem?

There are many consequences of overdiagnosis that must be considered to fully understand why it is a significant problem with the current screening programme. Overdiagnosis often leads to overtreatment, which has many effects – both immediate and late on the health of individual women. As currently, it is not possible to differentiate between cancers that are likely to become symptomatic and pose a threat to an individual's health with those which will not, treatment is recommended in most cases (35). There are many possible options of treatment once diagnosed with breast cancer: from hormone, radiation, and chemotherapy to surgical therapy – all of which can have potentially unsavoury side effects. Out of all the women who have breast cancer found through screening, more than 99% have surgery, 87% have hormone therapy, 80% have radiotherapy and about 26% have chemotherapy. In the UK, 30% of women diagnosed with screen-detected DCIS are treated by mastectomy and 70% by breast-conservation surgery (43,44).

Immediate risks of therapy include surgical deformity - both from minor lumpectomies that precisely incise the tumour and some surrounding tissue, to major radical mastectomies that remove the entire breast; as well as toxicity that can occur from hormone, radiation and chemotherapy (34). Late effects of therapy also include late sequelae such as lymphoedema: an accumulation of lymph in the limbs resulting in persistent swelling of the limb, which occurs in about one in five people treated for breast cancer (34). Therapeutic radiation can lead to cardiac toxicity through damage to coronary arteries or development of heart failure, as well as scarring and even the development of new radiotherapy-induced cancers (31,32). A meta-analysis of radiotherapy showed that there was a 27% excess mortality from heart disease and a 78% excess mortality from lung cancer (32). The radiation dose per typical two-view

mammogram is approximately 4 mSv, annual mammograms at which may theoretically cause one breast cancer per 1,000 women screened aged 40-80.

Overdiagnosis and cost-effectiveness

The issue of overdiagnosis not only has a significant impact on the women that are diagnosed with cancer and potentially go through unnecessary treatment, but also on the cost-effectiveness of the screening programme as a whole. Currently, NICE has a £20,000-£30,000 per QALY threshold for new therapies (3). QALYs take into account not only the extension of life, but also the quality of the extended survival. One QALY equates to one year of life in perfect health. The Marmot review in 2012, upon meta-analysis of 11 randomised controlled trials, determined the relative risk reduction of screening to be 20% compared to the unscreened population (31). The absolute mortality benefit of breast cancer screening has been estimated to be approximately 21-28% with a 56% attendance in Australia, for screening between the ages of 50-69, which is similar to the screening programme in the UK (although we boasted a higher attendance, approximately 70% in 2009/2010) (37). It is estimated that participation in the screening programme would result in 8 deaths prevented for every 1,000 women screened every two years (compared to the every three years invited in the UK) from 50-74 (37). The Cochrane Review in 2009 estimated that screening reduces breast cancer mortality by 15%; however, given that overdiagnosis and overtreatment was estimated to be 30%, it meant that for every 2000 women invited for screening throughout 10 years, one would avoid dying of breast cancer whilst 10 healthy women would be overdiagnosed with a cancer that would never affect them in their lifetime, and be treated unnecessarily for it (32).

A modelling study in 2013 looked at the cost-effectiveness of the breast cancer screening programme, and found that only 45% of the systems that were modelled based on the figures concluded by the Marmot review for overdiagnosis and mortality benefit (19% and 20% respectively) were under the NICE £20,000 threshold for interventions (3). The cost-effectiveness estimates were particularly sensitive to the values used for the reduction in deaths from breast cancer and for overdiagnosis. This is because a balance has to be obtained between the absolute mortality benefit achieved, corresponding to every gain in QALYs due to early detection of cancers through screening, with the loss caused in QALYs of patients due to unnecessary invasive treatment and testing along with the anxiety caused by overdiagnosis. It is worth noting that the study mentioned that if the figures concluded by the 2009 Cochrane review were used, i.e. that there was 15% reduction in breast cancer mortality, the cost-effectiveness ratio increased to just over £70,007 per QALY, which is significantly higher than the NICE threshold (3).

For every 10,000 women invited to screening in the UK, it is estimated that about 43 deaths would be prevented, whilst 129 cases of overdiagnosis would occur, which gives an estimate of one death prevented for every three overdiagnosed under the screening programme (2). It is also worth noting that, just because cancers are more likely to be diagnosed early, the screening programme would appear to increase the years survived after diagnosis, even if the woman died at the same time in the screened population compared to the unscreened population. Therefore, this should also be taken into account when the benefit of the screening programme is assessed.

Overdiagnosis affects the QALY estimate as it leads to unnecessary diagnosis and also often treatment of slow progressing cancer. Further diagnostic tests such as biopsies and fine needle aspirations can be recommended, which increases the cost spent on these overdiagnosed cancers, also increasing the absolute cost of the screening programme. The cost to treat one overdiagnosed cancer was estimated to be around £1,800 by one study (3). Moreover, undergoing treatment can reduce the quality of life of patients, due to the many negative side effects of surgical therapy, such as radiotherapy and chemotherapy. This further reduces the cost per QALY: by reducing the quality of life of patients, and therefore increasing the threshold needed to achieve a full quality of life year. However, there is no agreed figure for the effect of overdiagnosis on the QALY threshold, largely due to the very varied estimates of overdiagnosis as a whole and the problems associated with estimating the effects of overdiagnosis. Therefore, it is difficult to conclude an exact figure for how overdiagnosis affects the QALY threshold, although it seems likely that it would decrease the cost-effectiveness of the programme to beyond the NICE recommended threshold (3,53).

How can we optimise the breast cancer screening programme to reduce overdiagnosis?

In light of the evidence, the issue of overdiagnosis and the deficit it causes to the cost-effectiveness of the screening programme can be reduced. In contrast with plans to increase the age range of the screened population, selectively targeting screening to only high-risk populations would reduce overdiagnosis (3). This is because it would be more likely that breast cancers detected in high-risk women would be fast-developing and would have progressed to be symptomatic and proven problematic to their health in their lifetime; it would also reduce the amount of overdiagnosed breast cancers that are needlessly detected in other women (38,39). For example, one study estimated that the planned age extension would cost £27,400 per QALY gained, with only a 29% probability of cost-effectiveness at a threshold of £20 000 per QALY (3,5).

Therefore, many studies looking at breast cancer screening across the globe have suggested that maybe a one-size fits all, blanket model of screening the population is no longer the most appropriate or cost-effective method of screening (38,39,46). Perhaps a more personalised approach, tailored to individual women's risk of developing breast cancer, with a goal of maximising benefits and minimising harms would be a better approach going forward. A stratified approach, in which high-risk women are selectively targeted for more frequent screening/monitoring, whereas low-risk women could have less frequent or no screening at different age ranges has been shown to a more cost-effective approach by modelling studies, and would also minimise the issue of overdiagnosis (54). A suggestion has been made that perhaps individual screening should begin at the age when the risk of developing breast cancer for a woman is equal to that for an average risk 50-year-old woman, which is approximately 2% in the next 10 years, and should stop when the risk of co-mortality from other diseases overtakes that from the risk of breast cancer mortality (54).

This would, however, require development of criteria for high-risk patient screening. For example, recently NICE determined it cost-effective to genetically test and screen those with high risk through family history or BRCA1/2 mutations with MRI or mammography (10). Currently, risk is estimated based on family history in the following stratification:

- High risk (with less than 1% of women) is defined as an estimated risk of greater than 8% between age 40 and 50 years, or a lifetime risk of 30% or greater, or a 20% or greater chance of a faulty BRCA1, BRCA2 or TP53 gene.
- Moderate risk (which accounts for about 20% of all cases) is defined as 3-8% estimated risk between age 40 and 50 years, or a lifetime risk of 17% or greater but less than 30%.
- Low risk is any risk below the above mentioned thresholds (10)

Therefore, those with known genetic mutations that increase the risk for breast cancer such as BRCA1, BRCA2 and TP3 mutations, and with a familial history of breast cancer could be selectively screened. Furthermore, women with exposure to other risk factors which are associated with increased breast cancer risk such as obesity, frequent alcohol use, oestrogen replacement therapy or oestrogen contraceptive pill use, could have their risk assessed, and those of an elevated risk could be invited as eligible under a targeted screening programme (31). A list of the known risk factors for breast cancer is included in Table 3.

Table 3. List of known risk factors for developing breast cancer.

Risk Factors for Breast Cancer	
- Age	
- Genetic Mutations	BRCA1 and BRCA2 CHEK2 ATM FGFR2 MAP3KI
- Early menarche	
- Late menopause	
- First pregnancy after 30	
- Family history	
- History of breast surgery	
- High breast density	
- Radiation exposure	
- Post-menopause hormone therapy	
- Post-menopausal obesity	
- Alcohol misuse	

Breast cancer screening could still be offered for women who ask for it if they were particularly concerned of their risk; however, the entire population could have less frequent, or no screening every three years (54). This could be compared to how prostate cancer is not screened for in the UK, despite being the most common cancer in men, just as breast cancer is the most common cancer in women. Although prostate-specific antigen (PSA) levels have been shown to correlate with prostate cancer, due to the unreliability of results to distinguish successfully between cancers and a high rate of false positives and negatives, along with the high rate of overdiagnosis for every life saved, a screening programme using PSA testing is not yet established in the UK,

although men are encouraged to seek testing if they face any symptoms or are concerned (38,39). Even though a European study showed a 20% reduction in deaths from prostate cancer if there was a screening programme (similar to the absolute mortality benefit in the breast cancer screening programme), when balanced with the overdiagnosis and overtreatment that would result from the establishment of such a programme, it was decided against (30).

Therefore, it is worth considering that, upon review of breast cancer screening programmes, the Cochrane review in 2009 stated that they found 'no convincing evidence that UK breast cancer screening service has saved lives, but solid evidence of serious and common harms' (32,43,44). Furthermore, on the basis of further review of the evidence, the Swiss Medical Board concluded in 2014 that 'there was no clear evidence of survival benefit, but clear evidence of overtreatment, with up to 14 patients getting an unnecessary diagnosis for one who would possibly benefit' for its own breast cancer screening programme, and consequently decided against the start of any new screening programmes whilst placing a time limit on the existing ones (30,43,44). In light of the increasing concerns, the cost that overdiagnosis poses to both the cost-effectiveness of the programme, and the quality of life of patients must thoroughly be considered when assessing the success of the breast cancer screening programme.

Policy Recommendations

- Advise against expanding the age range that is being trialled as this would likely further increase the issue of overdiagnosis and reduce cost-effectiveness, especially when extending to an even older population who are more susceptible to developing cancers that are often overdiagnosed. Moreover, increased time in a screening population increases cumulative radiation dose, and increases the chance of women being diagnosed with false positives.
- Consider targeting screening programme on an individual risk-basis, rather than a blanket population level. Targeting high-risk patients specifically would not only reduce the overall cost of the screening programme but also minimise the issue of overdiagnosis.
- Encourage patients to be mindful of symptoms, whilst still providing mammography on the NHS to those who request it, with the aim to still detect cancers early.

The Role of Breast Self-Examination within the Mammographic Screening Programme

Sophie Caseby

Breast self-examination (BSE) has long been advocated as an inexpensive, non-invasive technique to detect pre-symptomatic breast cancer. Systematic inspection and palpation of the breast by the individual would ideally detect tumours at smaller, earlier stages which are likely to be more receptive to available treatments, and allow a more breast conserving approach to therapy (55)(56). However similarly to mammographic screening, BSE has been subject to numerous criticisms, with known limitations and contradicting evidence on overall effectiveness (57). Previous clinical trials and publications investigating the value of BSE are discussed below, whilst considering whether investment in BSE education programmes and breast awareness campaigns could assist in the proposed restriction of mammographic screening to high-risk subpopulations.

Early studies show benefit of BSE on tumour prognosis

The World Health Organisation (WHO) recognised in 1982, prior to the introduction of the NHS breast cancer screening programme, that the potential efficacy of BSE should be considered in greater detail (58). Studies in this pre-mammographic era gave initial evidence for effectiveness of BSE. For example, Hill et al.'s (1988) meta-analysis of 12 studies including 8118 breast cancer patients demonstrated that women reporting BSE performance had significantly fewer cases of lymph node involvement than women who had never practised the examination, with their tumours more likely to measure less than 2cm in diameter (59). A closer look at the individual case-control studies supports the hypothesis that BSE would lead to tumours being detected with fewer axillary-lymph node metastases (60), smaller pathologic size (61), and at an earlier clinical stage (62). However these retrospective studies also highlight particular features of BSE that would be required for more favourable prognosis. Mant et al. (1987) only observed significant differences between the non-practising group and the women who had both practised and been taught BSE, rather than the cohort practising BSE without direct teaching (57). Further case-control studies observed an increased risk of breast cancer mortality when omitting two or three of specific BSE components, including visual examination of whole breast and axillae, use of finger pads for circular palpations, and examination with three middle fingers (63). These requisite skills of BSE were also noted in Newcomb et al.'s (1991) study, with decreased risk of advanced-stage breast cancer being limited to women with more thorough BSE technique (64). The importance of being highly proficient in BSE therefore emphasises the need for high-quality, structured education programmes if self-examination is to be recommended. Overall these earlier studies may imply an advantage of BSE in terms of favourable prognosis, yet they are limited by small sample sizes, selection bias (65), and lack of evidence for reduced breast cancer mortality rates – the ultimate aim of any screening programme. Furthermore, patient

self-reporting is an unreliable method to measure BSE activity, due to difficulties in measuring adherence and competence (56), therefore limiting the applicability of non-randomised, retrospective studies.

Larger population-based trials for BSE report increased rates of invasive biopsies, whilst giving no evidence for reduced breast cancer mortality

Miller et al. (1985) proposed that the only truly valid criterion of BSE effectiveness would be a reduction in breast cancer mortality rate in a study population compared with a suitable control population (58). This called for larger randomised controlled trials (RCTs) to be conducted, which with appropriate randomisation, would distribute confounding variables equally between study and control groups. A 2003 Cochrane Review by Kösters and Gøtzsche (66) found only two RCTs investigating BSE compared with no screening intervention, based in Shanghai (67) and Russia (68), with a combined trial population of 388,535 women. Both studies came to the agreed conclusion that BSE has little benefit, with no statistically significant difference in breast cancer mortality between intervention and control group. Therefore, BSE cannot have resulted in breast cancer being diagnosed at a sufficiently less advanced stage for therapy to have altered course of disease. Moreover, there was good evidence for harm from BSE, with significantly more invasive diagnostic procedures performed with benign results, such as needle and excision biopsies, which could lead to breast deformity and scars (66).

The Shanghai trial is thought to be well-designed, and of higher quality than the Russia trial, with better compliance rates and an 11 year follow-up period (69). The lack of evidence for any survival advantage in the intervention group therefore cannot be attributed to inadequate quality or length of the study, and further RCTs are unlikely to uncover any additional benefits of BSE (66). This also highlights how BSE is inferior to mammographic screening, which shows a reduction in breast cancer mortality after just five years of trial follow-up. The women in the Shanghai intervention group received intensive BSE instruction from medical personnel with multiple reinforcement sessions and reminders. Competence in BSE was tested by the ability to detect lumps in silicone breast models, and the intervention group demonstrated greater specificity and sensitivity than women in the control group (67). Although the BSE training led to much greater proficiency in BSE, this was still unable to impact upon prognosis of tumours or breast cancer mortality rates. The self-reported compliance rates from the Russian study diminished over time to just 55% after 5 years (68). This also suggests investment in population-wide BSE education programmes would not be worthwhile in the UK, a conclusion which is supported by the UK Trial of Early Detection of Breast Cancer (70). This population-based trial was not included in the Cochrane review since the assignment of different geographical districts to the BSE or non-BSE group meant the study was non-randomised, with differences in socioeconomic status. However, 16 years of follow-up saw no reduction in breast cancer mortality and a higher rate of benign biopsy results in the BSE centres (71).

A systematic review from the Canadian Task Force on Preventive Health Care examined evidence from the two large RCTs (Shanghai, Russia), the quasi-randomised trial (UK), and a large cohort study from the United States, which also showed no benefit of self-reported BSE practice (72). This led to recommendations that routine teaching of BSE should be excluded from the periodic health examination of women in all age groups

(55). This decision was reflected in the American Cancer Society's updated guidelines in 2003, when they no longer recommended monthly self-examination beginning at age 20 due to the limited scientific evidence, from reviews discussed above, for the most effective technique or optimal frequency of BSE (56).

Structured education programmes for breast self-examination would not be worthwhile in the UK, but breast awareness is still vitally important

Overall, structured BSE education programmes would not be worthwhile or cost-effective in terms of improving breast cancer mortality rates. A one year nursing-led BSE education programme has previously been estimated to cost between US$ 574-848 per competent frequent self-examiner (73), meaning any benefits would be unlikely to outweigh the resource costs. However, the value of breast awareness should not be underestimated, which combined with mammographic screening and improved treatment, is likely to have contributed to the UK's declining rates of breast cancer mortality (66)(71). Heightened awareness of an individual's own breast composition means changes could be detected and reported promptly, which is especially important in older women since breast tumours are likely to arise in the intervening three years between scheduled mammograms (56). This role of breast awareness is also reflected by Polygeia's personal communication with CoppaFeel, an education and awareness charity which promotes self-checking, aiming for all breast cancers to be diagnosed at the earliest stage. CoppaFeel believe that since mammographic screening only occurs 'during a certain period of a woman's life, and only every three years', that all women should also be 'taking note of awareness campaigns and self-checking regularly', aiming for this to become a life-long habit.

Policy Recommendations

- Organised breast self-examination education programmes would not be worthwhile or cost-effective in the UK, however general breast awareness should be enhanced, whether restricting mammographic screening to high-risk subpopulations or retaining the target population of women aged 50-70.
- Women should be encouraged to know their own normal breast composition, and to seek medical advice if noticing any changes.
- The charity CoppaFeel targets a younger, student demographic, and this valuable breast awareness should be extended to older women, since triennial mammographic screening leaves a substantial interval within which breast tumours could arise.

Nevertheless, the researchers of the Shanghai study stated that until a trial comparing mammography alone with a mammography and BSE group, there is no reason to discourage women from self-checking whilst within the age range for mammographic screening (67). Furthermore, it was highlighted that the Shanghai trial is a study of the effectiveness of BSE education, rather than of the effectiveness of practising BSE, and highly motivated women may wish to be taught BSE. Conducting BSE competently and frequently may lead to detection of breast tumours at early clinical stages, which could be beneficial since some observational studies have demonstrated an inverse correlation between tumour size at diagnosis and frequency of BSE (67). However studies such as Newcomb et al.'s (1991), which highlighted the importance of highly proficient BSE,

should not be disregarded – meaning any teaching should be of highest quality to have optimal effects (64).

Policy Recommendations

- Women should be taught breast self-examination technique if requested, however it is especially important to ensure the instruction is thorough and that the women conduct examination in a highly proficient manner.
- The decision to conduct breast self-examination should be an informed one, and similarly to the mammographic screening programme, transparent and objective information should be made available. Women should be counselled on the risks and benefits, ensuring they are aware that they would have an increased chance of breast biopsy, and that there is limited scientific evidence to suggest a reduced risk of dying from breast cancer.

Clinical breast examination is more appropriate for developing countries

Clinical breast examination (CBE), performed by a medical professional, has also been considered as an adjunct to mammographic screening. Only one large population-based trial was included in the 2003 Cochrane Review (66) – this study was based in the Philippines and studied CBE combined with BSE, with a cohort of 404,947. However no conclusions on CBE effectiveness could be drawn since poor compliance of test-positive women in following referral appointments led to discontinuation of the trial (74). Although current evidence is insufficient to assess benefits and costs of CBE, the results may not have been applicable to the UK. CBE is thought to be more suitable for countries in economic transition, whereas technically advanced countries with adequate treatment are unlikely to benefit from a screening modality with lower sensitivity and specificity compared to mammography (34)(75)(76).

Concluding remarks

In conclusion, breast-self examination is of uncertain benefit and should not replace mammographic screening, which has proven efficacy (70). BSE could be an effective adjunct to the mammographic screening programme if conducted in a highly proficient manner, however cannot be routinely recommended until ways to minimise the detection of benign lesions have been discovered (55). Nevertheless, many breast tumours are still detected by the women themselves, and campaigns and charity work should continue to maintain the UK's high levels of breast awareness, which is vitally important whether restricting the screening programme to high-risk subpopulations, or continuing to invite all women aged 50-70.

Conclusion: Future Policy

From the outset of the project our aim has been to produce concrete policy proposals. We have looked at four key areas of controversy within the treatment and research of breast cancer: overdiagnosis, evidence collation, self-examination, and the screening of women under 50. Our critical approach to the literature has uncovered results that were both expected and others that have been more surprising. Of particular importance is how all of our research has found that much of the early research has significant methodological issues, which has meant that their findings misrepresent the risks and benefits of the breast cancer-screening programme.

The failure of early research to conduct robust RCTs has led to Gøtzsche and Olsen to propose ignoring their findings (17). This is particularly significant because these results were assumed sufficient enough to justify the NHS screening programme. If we therefore have grounds to question the reliability of these earlier research papers' findings then it must also mean in turn that the screening programme itself must, at the very least, come under closer scrutiny. In particular the dispute over the findings affects the calculation of the QALYs, according to the Forrest report the screening programme is effective at £8,300 per year (4). Whilst this figure is far below the £20,000 figure used by NICE to justify intervention it has since been called into question due to uncertainties surrounding the reliability of these findings. In particular the most pertinent criticism is the underestimation of the cost of false positives; due to the lack of sensitivity of early research they likely grossly underestimated the frequency of false positives. Since false positives require unneeded invasive surgery and involve costly treatment any increase in false positives will vastly decrease the cost-effectiveness. To address this problem, we and other authors believe that screening should be restricted to the at-risk populations. Restricting treatment will lower the proportion of false positives, and ensure scarce resources are focused on the women who are most at risk.

In addition to the issue of false positives we have found that most reports used to justify the screening programme both underestimate the number of overdiagnosis and the cost of overdiagnosis. With newer evidence suggesting that for every woman saved 10 healthy women receive an overdiagnosis, it is clear this is not just a minor problem as earlier reports suggested (32). Indeed the Marmot review found that the majority of newer economic models put the cost-effectiveness of the screening programme at significantly over the £20,000-£30,000 per QALY threshold (2). These findings therefore require debate about how we can optimise the screening programme. We find that there is a strong argument and significant evidence to suggest that the programme should be restricted to only the at risk populations. The question that must be asked of future researchers is: 'How are we to define the at-risk populations?' This is not something we have attempted to answer in detail but it is clear there needs to be a move away from population based screening.

Restricting the screening programme is not the only proposal that has been suggested to restrict the costs of the screening programme. We also explored the potential for self-examination to replace some of the mammography screenings. Our findings are clear in that self-examination is not a suitable substitute for mammography screenings, indeed

there is evidence which suggests self-examination has no impact upon mortality even when compared with inaction (68). Moreover many of the costs associated with mammography screenings, such as overdiagnosis and false positives, are still present in a self-examining programme. Thus it is clear that the UK government should resist any calls to move to a self-examination centred programme.

Whilst the majority of our paper has focused on the current screening programme we also thought it was pertinent to discuss the controversy surrounding the expansion of the UK screening programme. Although the majority of the literature, as discussed earlier, focuses on radiation risks, which we found to be less of a concern than previously thought, we believe that the literature ignores the more serious issues surrounding overdiagnosis and false positives. We are not the first to call for a rethink on the expansion of screening programmes, in 2009 the US Preventive Services Task Force recommended against routine screening for women aged 40-49 (76). We agree with this position. Firstly, due to the difference in density of breast tissue of younger women, it is harder to accurately screen women in the expanded age category, which can lead to more false positives compared to older women (76). Moreover due to the lower cancer rate among younger women there is heightened sensitivity problems. Similar to how we recommended restricting the screening programme to at-risk populations we also criticise the expansion of the programme for leading to a dilution of cancer cases in the screened population. Thus we believe it vital to reconsider the expansion programme; or at the very least ensure that the current ongoing trial examines, in a more comprehensive than previous trials, the 'unseen' costs of overdiagnosis and false positives.

Policy Recommendations

Randomised Controlled Trials for Breast Cancer Screening

- Use overall lifespan as the measurement in future RCTs.
- Collect more data on overdiagnosis from screening to use for future cost-benefit analysis.

Radiation Risk: Should We Screen Women Under the Age of 50?

- Radiation risk is low-negligible meaning it cannot be used as grounds to not widen the age range of the screening programme.
- Further, more public, research is needed on the ethical assumptions that we make when valuing lives.
- Although low-risk there needs to be more investment into screening technology that uses lower doses of radiation.
- Although low-risk unnecessary screenings should be avoided.

Overdiagnosis and Overtreatment

- Advise against expanding the age range that is being trialled as this would likely further increase the issue of overdiagnosis and reduce cost-effectiveness, especially when extending to an even older population who are more susceptible to developing cancers that are often overdiagnosed. Moreover, increased time in a screening population increases cumulative radiation dose, and increases the chance of women being diagnosed with false positives.
- Consider targeting screening programme on an individual risk-basis, rather than a blanket population level. Targeting high-risk patients specifically would not only reduce the overall cost of the screening programme but also minimise the issue of overdiagnosis.
- Encourage patients to be mindful of symptoms, whilst still providing mammography on the NHS to those who request it, with the aim to still detect cancers early.

The Role of Breast Self-Examination

- Organised breast self-examination education programmes would not be worthwhile or cost-effective in the UK, however general breast awareness should be enhanced, whether restricting the mammographic screening programme to high-risk subpopulations or retaining the target population of women aged 50-70.
- Women should be encouraged to know their own normal breast composition, and to seek medical advice if noticing any changes.
- The charity CoppaFeel targets a younger, student demographic, and this valuable breast awareness should be extended to older women, since triennial mammographic screening leaves a substantial interval within which breast tumours could arise.
- Women should be taught breast self-examination technique if requested, however it is especially important to ensure the instruction is thorough and that the women conduct examination in a highly proficient manner.
- The decision to conduct breast self-examination should be an informed one, and similarly to the mammographic screening programme, transparent and objective information should be made available. Women should be counselled on the risks and benefits, ensuring they are aware that they would have an increased chance of breast biopsy, and that there is limited scientific evidence to suggest a reduced risk of dying from breast cancer.

References

1. Cancer Research UK. Breast screening statistics [Internet]. 2012 [cited 2016 Apr 29]. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/screening
2. Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. Lancet (London, England) [Internet]. Elsevier; 2012 [cited 2016 Jul 15];380(9855):1778–86. Available from: http://www.thelancet.com/article/S0140673612616110/fulltext
3. Pharoah PDP, Sewell B, Fitzsimmons D, Bennett HS, Pashayan N. Cost effectiveness of the NHS breast screening programme: life table model. BMJ [Internet]. 2013 [cited 2016 Feb 13];346(may09_1):f2618. Available from: http://www.bmj.com/content/346/bmj.f2618
4. Forrest P. Breast cancer screening. Report to the Health Ministers of England Wales Scotland and N Ireland by a working group chaired by Professor Sir Patrick Forrest [Internet]. HMSO. 1986 [cited 2016 Apr 29]. Available from: http://webarchive.nationalarchives.gov.uk/20150506150512/http://www.cancerscreening.nhs.uk/breastscreen/publications/forrest-report.html
5. Madan J, Rawdin A, Stevenson M, Tappenden P. A rapid-response economic evaluation of the UK NHS Cancer Reform Strategy breast cancer screening program extension via a plausible bounds approach. Value Health [Internet]. 2010 [cited 2016 Nov 1];13(2):215–21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19878494
6. Faculty of Clinical Radiology. Guidance on screening and symptomatic breast imaging. 2013;(Third edition).
7. NICE. Clinical guidance: Early and locally advanced breast cancer: diagnosis and treatment [Internet]. NICE; 2009 [cited 2016 Nov 10]. Available from: https://www.nice.org.uk/guidance/cg80/chapter/1-guidance#follow-up
8. McCabe C, Claxton K, Culyer AJ. The NICE cost-effectiveness threshold: what it is and what that means. Pharmacoeconomics [Internet]. 2008 [cited 2016 Nov 10];26(9):733–44. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18767894
9. NICE. Supporting investment in public health: Review of methods for assessing cost effectiveness, cost impact and return on investment. Proof of concept report. 2011;
10. NICE Guidance. Familial breast cancer: classification, care and managing breast cancer and related risks in people with a family history of breast cancer. NICE; 2013 [cited 2016 Nov 1]; Available from: https://www.nice.org.uk/guidance/cg164/resources/familial-breast-cancer-classification-care-and-managing-breast-cancer-and-related-risks-in-people-with-a-family-history-of-breast-cancer-35109691767493
11. Rafia R, Madan J, Brennan A, Collins K, Reed M, Lawrence G. Option appraisal: Modelling the effectiveness and cost-effectiveness of screening policies for Breast Cancer in elderly women in England and Wales. 2011;
12. Tabar L, Fagerberg G, Chen H-H, Duffy SW, Smart CR, Gad A, et al. Efficacy of breast cancer screening by age. New results swedish two-county trial. Cancer [Internet]. Wiley Subscription Services, Inc., A Wiley Company; 1995 May 15 [cited 2016 Oct 15];75(10):2507–17. Available from: http://doi.wiley.com/10.1002/1097-0142%2819950515%2975%3A10%3C2507%3A%3AAID-CNCR2820751017%3E3.0.CO%3B2-H
13. Banks E, Reeves G, Beral V, Bull D, Crossley B, Simmonds M, et al. Influence of personal characteristics of individual women on sensitivity and specificity of mammography in the Million Women Study: cohort study. BMJ [Internet]. 2004 Aug;329(7464):477. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15331472
14. NHS Choices. Predictive genetic tests for cancer risk genes - NHS Choices [Internet]. Department of Health; 2015 [cited 2016 Oct 15]. Available from: http://www.nhs.uk/Conditions/predictive-genetic-tests-cancer/Pages/Introduction.aspx
15. Cancer Research UK. Breast cancer incidence (invasive) statistics | Cancer Research UK [Internet]. 2012. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/incidence-invasive#heading-One

16. WebMD. Ductal Carcinoma (Invasive and in Situ) [Internet]. 2015. Available from: http://www.webmd.com/breast-cancer/ductal-carcinoma-invasive-in-situ
17. Gøtzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? Lancet [Internet]. 2000 Jan;355(9198):129–34. Available from: http://www.sciencedirect.com/science/article/pii/S0140673699060651
18. NHS Choices. Breast cancer screening - when it's offered - NHS Choices. Department of Health; 2015;
19. Yoo K-B, Kwon JA, Cho E, Kang M-H, Nam J-M, Choi K-S, et al. Is mammography for breast cancer screening cost-effective in both Western and asian countries?: results of a systematic review. Asian Pac J Cancer Prev [Internet]. 2013 [cited 2016 Oct 1];14(7):4141–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23991967
20. Robberstad B. QALYs vs DALYs vs LYs gained: What are the differences, and what difference do they make for health care priority setting? Nor Epidemiol [Internet]. 2009 Oct 9 [cited 2016 Oct 31];15(2). Available from: http://www.ntnu.no/ojs/index.php/norepid/article/view/217
21. Altobelli E, Lattanzi A. Breast cancer in European Union: an update of screening programmes as of March 2014 (review). Int J Oncol [Internet]. 2014 Nov [cited 2016 Oct 1];45(5):1785–92. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25174328
22. Sankatsing VD V, Heijnsdijk EAM, van Luijt PA, van Ravesteyn NT, Fracheboud J, de Koning HJ. Cost-effectiveness of digital mammography screening before the age of 50 in The Netherlands. Int J cancer [Internet]. 2015 Oct 15 [cited 2016 Oct 1];137(8):1990–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25895135
23. Carles M, Vilaprinyo E, Cots F, Gregori A, Pla R, Román R, et al. Cost-effectiveness of early detection of breast cancer in Catalonia (Spain). BMC Cancer [Internet]. BioMed Central; 2011 Dec 23 [cited 2016 Oct 1];11(1):192. Available from: http://bmccancer.biomedcentral.com/articles/10.1186/1471-2407-11-192
24. Gocgun Y, Banjevic D, Taghipour S, Montgomery N, Harvey BJ, Jardine AKS, et al. Cost-effectiveness of breast cancer screening policies using simulation. Breast [Internet]. 2015 Aug [cited 2016 Oct 1];24(4):440–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25866350
25. Moss SM, Wale C, Smith R, Evans A, Cuckle H, Duffy SW, et al. 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 [Internet]. Elsevier; 2015 Sep [cited 2016 Nov 3];16(9):1123–32. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26206144
26. 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 [Internet]. 2011 Feb 15 [cited 2016 Nov 3];117(4):714–22. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20882563
27. Djulbegovic B, Lyman GH. Screening mammography at 40–49 years: regret or no regret? Lancet [Internet]. 2006 Dec [cited 2016 Nov 3];368(9552):2035–7. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0140673606698164
28. de Gelder R, Draisma G, Heijnsdijk EAM, de Koning HJ. Population-based mammography screening below age 50: balancing radiation-induced vs prevented breast cancer deaths. Br J Cancer [Internet]. 2011 Mar 29 [cited 2016 Nov 3];104(7):1214–20. Available from: http://www.nature.com/doifinder/10.1038/bjc.2011.67
29. Seigneurin A, Labarère J, François O, Exbrayat C, Dupouy M, Filippi M, et al. Overdiagnosis and overtreatment associated with breast cancer mammography screening: A simulation study with calibration to population-based data. Breast [Internet]. 2016 Aug [cited 2016 Nov 1];28:60–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27240167
30. Løberg M, Lousdal ML, Bretthauer M, Kalager M, Raffle A, Gray J, et al. Benefits and harms of mammography screening. Breast Cancer Res [Internet]. BioMed Central; 2015 Dec 1 [cited 2016 Nov 1];17(1):63. Available from: http://breast-cancer-research.com/content/17/1/63
31. Barratt A. Overdiagnosis in mammography screening: a 45 year journey from shadowy idea to acknowledged reality. BMJ. 2015;350.
32. Gøtzsche PC, Jørgensen KJ. Screening for breast cancer with mammography. In: Gøtzsche PC, editor. Cochrane Database of Systematic Reviews [Internet]. Chichester, UK: John Wiley & Sons, Ltd; 2013 [cited 2016 Nov 1]. Available from: http://doi.wiley.com/10.1002/14651858.CD001877.pub5

33. Puliti D, Duffy SW, Miccinesi G, de Koning H, Lynge E, Zappa M, et al. Overdiagnosis in mammographic screening for breast cancer in Europe: a literature review. J Med Screen [Internet]. 2012 [cited 2016 Nov 1];19 Suppl 1:42–56. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22972810
34. PDQ® Screening and Prevention Editorial Board. Breast Cancer Screening (PDQ®): Health Professional Version [Internet]. PDQ Cancer Information Summaries. Bethesda, MD: National Cancer Institute; 2016 [cited 2016 Sep 2]. Available from: http://www.cancer.gov/types/breast/hp/breast-screening-pdq
35. Feig SA. Overdiagnosis of breast cancer at screening is clinically insignificant. Acad Radiol [Internet]. 2015 Aug [cited 2016 Nov 1];22(8):961–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25797300
36. Segnan N, Minozzi S, Armaroli P, Cinquini M, Bellisario C, González-Lorenzo M, et al. Epidemiologic evidence of slow growing, nonprogressive or regressive breast cancer: A systematic review. Int J cancer [Internet]. 2016 Aug 1 [cited 2016 Nov 1];139(3):554–73. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27004723
37. Cancer Australia. Overdiagnosis from mammographic screening: Position statement. 2008;
38. Esserman LJ, Thompson IM, Reid B. Overdiagnosis and overtreatment in cancer: an opportunity for improvement. JAMA [Internet]. 2013 Aug 28 [cited 2016 Nov 1];310(8):797–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23896967
39. Esserman LJ, Thompson IM, Reid B, Nelson P, Ransohoff DF, Welch HG, et al. Addressing overdiagnosis and overtreatment in cancer: a prescription for change. Lancet Oncol [Internet]. 2014 May [cited 2016 Nov 1];15(6):e234-42. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24807866
40. van Luijt PA, Heijnsdijk EAM, Fracheboud J, Overbeek LIH, Broeders MJM, Wesseling J, et al. The distribution of ductal carcinoma in situ (DCIS) grade in 4232 women and its impact on overdiagnosis in breast cancer screening. Breast Cancer Res [Internet]. 2016 May 10 [cited 2016 Nov 1];18(1):47. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27160733
41. Morris E, Feig SA, Drexler M, Lehman C. Implications of Overdiagnosis: Impact on Screening Mammography Practices. Popul Health Manag [Internet]. 2015 Sep [cited 2016 Nov 1];18 Suppl 1:S3-11. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26414384
42. Cowell CF, Weigelt B, Sakr RA, Ng CKY, Hicks J, King TA, et al. Progression from ductal carcinoma in situ to invasive breast cancer: revisited. Mol Oncol [Internet]. 2013 Oct [cited 2016 Nov 1];7(5):859–69. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23890733
43. Francis A, Thomas J, Fallowfield L, Wallis M, Bartlett JMS, Brookes C, et al. Addressing overtreatment of screen detected DCIS; the LORIS trial. Eur J Cancer [Internet]. 2015 Nov [cited 2016 Nov 1];51(16):2296–303. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26296293
44. Francis A, Fallowfield L, Rea D, Jorgensen KJ, Keen JD, Gotzsche PC, et al. The LORIS Trial: Addressing overtreatment of ductal carcinoma in situ. Clin Oncol (R Coll Radiol) [Internet]. Elsevier; 2015 Jan [cited 2016 Nov 1];27(1):6–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25445552
45. Hayse B, Hooley RJ, Killelea BK, Horowitz NR, Chagpar AB, Lannin DR. Breast cancer biology varies by method of detection and may contribute to overdiagnosis. Surgery [Internet]. 2016 Aug [cited 2016 Nov 1];160(2):454–62. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27241117
46. Onega T, Beaber EF, Sprague BL, Barlow WE, Haas JS, Tosteson ANA, et al. Breast cancer screening in an era of personalized regimens: a conceptual model and National Cancer Institute initiative for risk-based and preference-based approaches at a population level. Cancer [Internet]. 2014 Oct 1 [cited 2016 Nov 1];120(19):2955–64. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24830599
47. Nelson HD, Pappas M, Cantor A, Griffin J, Daeges M, Humphrey L. Harms of Breast Cancer Screening: Systematic Review to Update the 2009 U.S. Preventive Services Task Force Recommendation. Ann Intern Med [Internet]. 2016 Feb 16 [cited 2016 Nov 1];164(4):256–67. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26756737
48. Zahl P-H, Jørgensen KJ, Gøtzsche PC. Overestimated lead times in cancer screening has led to substantial underestimation of overdiagnosis. Br J Cancer [Internet]. 2013 Oct 1 [cited 2016 Nov 1];109(7):2014–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23963144
49. Carter JL, Coletti RJ, Harris RP. Quantifying and monitoring overdiagnosis in cancer screening: a systematic review of methods. BMJ. 2015;350.

50. Baines CJ, To T, Miller AB. Revised estimates of overdiagnosis from the Canadian National Breast Screening Study. Prev Med (Baltim) [Internet]. 2016 Sep [cited 2016 Nov 1];90:66–71. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27374944
51. Koleva-Kolarova RG, Zhan Z, Greuter MJW, Feenstra TL, De Bock GH. To screen or not to screen for breast cancer? How do modelling studies answer the question? Curr Oncol [Internet]. Multimed Inc.; 2015 Oct [cited 2016 Nov 1];22(5):e380-2. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26628880
52. Jacklyn G, Glasziou P, Macaskill P, Barratt A. Meta-analysis of breast cancer mortality benefit and overdiagnosis adjusted for adherence: improving information on the effects of attending screening mammography. Br J Cancer [Internet]. 2016 May 24 [cited 2016 Nov 1];114(11):1269–76. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27124337
53. Raftery J, Chorozoglou M. Possible net harms of breast cancer screening: updated modelling of Forrest report. BMJ. 2011;343.
54. Depypere H, Desreux J, Pérez-López FR, Ceausu I, Erel CT, Lambrinoudaki I, et al. EMAS position statement: individualized breast cancer screening versus population-based mammography screening programmes. Maturitas [Internet]. 2014 Dec [cited 2016 Nov 1];79(4):481–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25277123
55. Baxter N, Canadian Task Force on Preventive Health Care. Preventive health care, 2001 update: should women be routinely taught breast self-examination to screen for breast cancer? CMAJ [Internet]. 2001 Jun 26 [cited 2016 Sep 3];164(13):1837–46. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11450279
56. Smith RA, Cokkinides V, Eyre HJ. American Cancer Society Guidelines for the Early Detection of Cancer, 2004. CA Cancer J Clin [Internet]. John Wiley & Sons, Ltd.; 2004 Jan 1 [cited 2016 Sep 7];54(1):41–52. Available from: http://doi.wiley.com/10.3322/canjclin.54.1.41
57. Mant D, Vessey MP, Neil A, McPherson K, Jones L. Breast self examination and breast cancer stage at diagnosis. Br J Cancer [Internet]. Nature Publishing Group; 1987 Feb [cited 2016 Sep 3];55(2):207–11. Available from: http://www.ncbi.nlm.nih.gov/pubmed/3814490
58. Miller AB, Chamberlain J, Tsechkovski M, Haagensen C, Bodian C, Haagensen D, et al. Self-examination in the early detection of breast cancer. J Chronic Dis [Internet]. Elsevier; 1985 Jan [cited 2016 Sep 7];38(6):527–40. Available from: http://linkinghub.elsevier.com/retrieve/pii/0021968185900372
59. Hill D, White V, Jolley D, Mapperson K. Self examination of the breast: is it beneficial? Meta-analysis of studies investigating breast self examination and extent of disease in patients with breast cancer. BMJ [Internet]. BMJ Group; 1988 Jul 23 [cited 2016 Sep 2];297(6643):271–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/3416148
60. Foster RS, Lang SP, Costanza MC, Worden JK, Haines CR, Yates JW. Breast Self-Examination Practices and Breast-Cancer Stage. N Engl J Med [Internet]. Massachusetts Medical Society ; 1978 Aug 10 [cited 2016 Sep 3];299(6):265–70. Available from: http://www.nejm.org/doi/abs/10.1056/NEJM197808102990601
61. Foster RS, Costanza MC. Breast self-examination practices and breast cancer survival. Cancer [Internet]. Wiley Subscription Services, Inc., A Wiley Company; 1984 Feb 15 [cited 2016 Sep 11];53(4):999–1005. Available from: http://doi.wiley.com/10.1002/1097-0142%2819840215%2953%3A4%3C999%3A%3AAID-CNCR2820530429%3E3.0.CO%3B2-N
62. Greenwald P, Nasca PC, Lawrence CE, Horton J, McGarrah RP, Gabriele T, et al. Estimated Effect of Breast Self-Examination and Routine Physician Examinations on Breast-Cancer Mortality. N Engl J Med [Internet]. Massachusetts Medical Society ; 1978 Aug 10 [cited 2016 Sep 3];299(6):271–3. Available from: http://www.nejm.org/doi/abs/10.1056/NEJM197808102990602
63. Harvey BJ, Miller AB, Baines CJ, Corey PN. Effect of breast self-examination techniques on the risk of death from breast cancer. CMAJ [Internet]. 1997 Nov 1 [cited 2016 Sep 2];157(9):1205–12. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9361639
64. Newcomb PA, Weiss NS, Storer BE, Scholes D, Young BE, Voigt LF. Breast self-examination in relation to the occurrence of advanced breast cancer. J Natl Cancer Inst [Internet]. 1991 Feb 20 [cited 2016 Sep 2];83(4):260–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/1994055
65. Gastrin G, Miller AB, To T, Aronson KJ, Wall C, Hakama M, et al. Incidence and mortality from breast cancer in the Mama Program for Breast Screening in Finland, 1973-1986. Cancer [Internet]. 1994 Apr 15 [cited 2016 Sep 2];73(8):2168–74. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8156521

66. Kösters JP, Gøtzsche PC. Regular self-examination or clinical examination for early detection of breast cancer. In: Kösters JP, editor. Cochrane Database of Systematic Reviews [Internet]. Chichester, UK: John Wiley & Sons, Ltd; 2003 [cited 2016 Sep 5]. Available from: http://doi.wiley.com/10.1002/14651858.CD003373
67. Thomas DB, Gao DL, Ray RM, Wang WW, Allison CJ, Chen FL, et al. Randomized Trial of Breast Self-Examination in Shanghai: Final Results. JNCI J Natl Cancer Inst [Internet]. Oxford University Press; 2002 Oct 2 [cited 2016 Sep 5];94(19):1445–57. Available from: http://jnci.oxfordjournals.org/cgi/doi/10.1093/jnci/94.19.1445
68. Semiglazov VF, Manikhas AG, Moiseenko VM, Protsenko SA, Kharikova RS, Seleznev IK, et al. [Results of a prospective randomized investigation [Russia (St.Petersburg)/WHO] to evaluate the significance of self-examination for the early detection of breast cancer]. Vopr Onkol [Internet]. 2003 [cited 2016 Aug 29];49(4):434–41. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14569932
69. Willett L. Review: limited evidence on regular breast examination does not support its effectiveness for reducing breast cancer deaths. Evid Based Med [Internet]. BMJ Publishing Group Ltd; 2003 Nov 1 [cited 2016 Sep 3];8(6):169–169. Available from: http://ebm.bmj.com/cgi/doi/10.1136/ebm.8.6.169
70. UK Trial of Early Detection of Breast Cancer. Breast cancer mortality after 10 years in the UK Trial of Early Detection of Breast Cancer. The Breast [Internet]. Churchill Livingstone; 1993 Mar [cited 2016 Sep 3];2(1):13–20. Available from: http://linkinghub.elsevier.com/retrieve/pii/096097769390031A
71. UK Trial of Early Detection of Breast Cancer Group. 16-year mortality from breast cancer in the UK Trial of Early Detection of Breast Cancer. Lancet (London, England) [Internet]. 1999 Jun 5 [cited 2016 Sep 3];353(9168):1909–14. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10371568
72. Holmberg L, Ekbom A, Calle E, Mokdad A, Byers T. Breast cancer mortality in relation to self-reported use of breast self-examination. A cohort study of 450,000 women. Breast Cancer Res Treat [Internet]. Kluwer Academic Publishers; 1997 Apr [cited 2016 Sep 3];43(2):137–40. Available from: http://link.springer.com/10.1023/A:1005788729145
73. O'Malley MS. Cost-effectiveness of two nurse-led programs to teach breast self-examination. Am J Prev Med [Internet]. 1993 [cited 2016 Sep 3];9(3):139–45. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8347364
74. Pisani P, Parkin DM, Ngelangel C, Esteban D, Gibson L, Munson M, et al. Outcome of screening by clinical examination of the breast in a trial in the Philippines. Int J cancer [Internet]. 2006 Jan 1 [cited 2016 Sep 7];118(1):149–54. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16049976
75. Miller A, Baines CJ. The role of clinical breast examination and breast self-examination. Prev Med (Baltim). 2011;53(3):118–20.
76. US Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med [Internet]. 2009 Nov 17 [cited 2016 Sep 12];151(10):716–26, NaN-236. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19920272