Breast cancer screening

Philippe Autier

International Prevention Research Institute (iPRI), Lyon, France

Preamble: Breast screening for the prevention of late stage cancer

Clinical data show a strong correlation between the size of breast cancer and the likelihood of metastases in axillary lymph nodes or in distant organs. Breast cancer screening is based on the principle that detection of a cancer when still small and not symptomatic would prevent progression to clinically evident, advanced disease associated with positive lymph nodes or distant metastases. The consequence of early detection would be a reduction of the risk to die from breast cancer.

In this educational paper, we first describe the randomised trials and controversies that influenced the adoption of mammography screening by high-income countries. We then briefly examine practical modalities of breast screening in general populations including issues like "overdiagnosis". We finally outline studies that assessed the efficiency of breast screening, i.e., how much of the breast cancer mortality decline observed in many high-income countries is due to screening. We focus on screening of women at average risk of breast cancer and we do not consider breast surveillance of women with a hereditary risk of breast cancer.

The testing of breast screening

In the 1960 and 1970s, three main screening modalities were considered: X-ray based mammography, breast self examination (BSE) and breast clinical examination (BCE) by a health professional.

BSE and BCE

Studies in the 1980s and 1990s on BSE and BCE were usually limited in scope and in size. A large randomised trial in Shanghai, China, concluded that BSE did not reduce breast cancer mortality while it increased the risk of unnecessary breast biopsy [1]. A review performed by an international expert group convened by the IARC concluded there was inadequate evidence that BSE or BCE can reduce mortality from breast cancer [2].

Randomised trials on mammography screening

From 1962 to 2006, ten randomised trials, one in the USA, five in Sweden, two in Canada and two in the UK, tested the ability of breast screening to prevent breast cancer death. One trial in Edinburgh (UK) is usually considered as uninformative because of socio-economic imbalances in cluster randomisation that resulted in non-comparable intervention and control groups [2]. Altogether, the nine trials in Table 1 included 298,765 women 40–74 years of age in intervention groups, and 317,418 women of the same age in the control groups. The trial designs were quite heterogeneous with respect to age groups included, constitution of control groups (seven trials had "usual care" control women without informed consent) screening modality (e.g., single view or 2-view mammography, mammography alone or combined with BCE), time interval between screens (from 1 to 3 years), and number of screening rounds (2 to 9). In addition, the Swedish trials offered screening to control women after women in the intervention group had received the scheduled screening intervention.

The HIP trial that started in 1962 tested BCE and mammography and each exam was done in total independence from the other. Results were reported for each screening modality, showing that 55% of screen-detected cancers were found by BCE and 77% were found by mammography [3]. Indeed, the mammography technology used in this trial did not match the performance of machines used in trials that started ten years later.

The five Swedish trials used mammography only. Four trials used various types of individual randomisation procedures, and the Two-County Trial randomised women by geographical cluster, each cluster comprising about 2,700 women in Kopparberg (currently named Dalarna) county and about 3,200 women in Östergötland county [4].

Table 1
Summary of design and results of randomised trials on breast screening that used mammography with or without BSE or BCE[a]

Trial, country	Age at entry	Popula-tion-based	Method of randomisation	No of women in intervention group	No. of women in the control group	Screening in intervention group	Type of MM	Screening in control group	Median duration (years) of trial [b]	Attendance rate, first round	Relative risk of breast cancer death, intervention versus control group	95% CI
HIP, USA	40–64	Yes	Individual	30239	30256	MM+BCE, 4 rounds every 12 months	2 views	None, usual care	5	67%	0.83	0.70–0.98
TCT, Sweden [d]	40–74	Yes	Cluster	77080	55985	MM, 2–4 rounds every 24 months for 40–49, and every 33 months for 50–74	1 view	None, usual care	8	85%	0.68	0.59–0.81
MMST I, Sweden	45–70	Yes	Individual	21088	21195	MM, 9 rounds every 18–24 months	2 views	None, usual care	14	74%	0.82	0.67–1.00
MMST II, Sweden	43–49	Yes	Individual	9581	8212	MM, 1 to 7 rounds every 18–24 months	2 views	None, usual care	9.1	78%	0.65	0.39–1.08
Stockholm, Sweden	39–65	Yes	Individual randomisation by day of birth	40318	19943	MM, 2 rounds every 24–28 months	1 view	None, usual care	5	82%	0.91	0.65–1.27
Göteborg, Sweden	39–59	Yes	Individual randomisation by day of birth	21650	29961	MM, 5 rounds every 18 months	2 views at initial screen, then 1 or 2 views	None, usual care	7	84%	0.76	0.56–1.04
Age Trial, UK	39–41	Yes	Individual	53884	106956	MM, 6 rounds every 12 months	2 views at first screen, then 1 view	None, usual care	7	68% [c]	0.83	0.66–1.04
NBSS I, Canada	40–49	No	Individual	25214	25216	[MM+BCE] 4–5 rounds every 12 months + BSE	2 views	BCE at first round + BSE	5	86%	0.97	0.74–1.27
NBSS II, Canada	50–59	No	Individual	19711	19694	[MM+BCE] 4–5 rounds every 12 months + BSE	2 views	BCE 4–5 rounds every 12 months + BSE	5	87%	1.02	0.78–1.33

HIP, Greater New-York Health Insurance Plan; TCT, Two-County Trial; MMST, Malmö Mammographic Screening Trial; NBBS, National Breast Screening Study.
MM = mammography; BCE: breast clinical examination; BSE: breast self-examination.
[a] Exhaustive references in references [2,5,6] plus reference [7] for the Age trial. The Edinburgh trial is not summarised (see text for justification).
[b] Period during which screening was offered to women in intervention group and not to women in control group, follow-up may have lasted for several more years.
[c] Overall, 81% of women attended at least one MM round.
[d] The two counties were Kopparberg (currently named Dalarna) and Östergötland.

Table 1 shows relative risk of breast cancer death associated with screening detailed in a recent review article [5]. The most important breast cancer mortality reduction was observed in the county of Kopparberg (Dalarna) that was part of the Two-County Trial, where a 37% mortality reduction was reported [8].

The first overview of Swedish trials, published in 1993 [9], indicated that mammography screening of women 40–74 years of age was associated with a relative risk of breast cancer death of 0.77 (95% confidence interval [95% CI]: 0.67–0.88), which can also be read as a 23% (95% CI: 12–33%) reduction in the risk of breast cancer death.

The two Canadian trials NBSS I and NBSS II compared mammography screening to BSE (women 40–49) or BCE (women 50–59). Women were volunteers, and randomisation was individual after women had signed a consent form. The two trials found no decrease in breast cancer mortality associated with mammography screening.

Hence, the Swedish trials provided the bulk of experimental evidence supporting the use of mammography for preventing breast cancer death. In contrast, the HIP trial did not really contribute to this evidence as most cancers were found by BCE, and the two Canadian trials did not support a role for mammography screening.

The two Canadian trials were much criticised for sub-optimal mammography quality and selection of women included in the trials. However, independent investigators who had access to the original data at individual level showed that these criticisms were not grounded. For instance, the size of breast cancers in the Canadian trials was on average smaller than in the Two-County Trial [10]. Such difference would probably not exist if the quality of mammography in the Canadian trials had been sub-optimal.

Screening of women 40 to 49 years old

The efficacy of mammography screening has usually been considered as low in pre-menopausal women, for a number of reasons. First, before menopause, the radiological density of breasts is higher, which decreases the sensitivity of mammography. Second, throughout the menstrual cycle, sex hormones influence breast glands and breast lesions, leading to high recall rates. Third, women included in randomised trials in their forties and diagnosed with breast cancer often died from breast cancer above the age of 50, suggesting that what matters is being screened after 49. Fourth, breast cancer incidence is lower in women 40–49 years of age than in older women; as a consequence, for a similar specificity as in older women, a higher proportion of women under 50 with suspicious mammograms are in fact false positives (Bayes theorem). Lastly, because of the low breast cancer incidence, even if reductions in the relative risk of death would be similar as in women 50 to 69, the number of breast cancer deaths prevented would be lower, and therefore, the cost per year of life saved would be much higher than in older women.

Two Swedish trials found evidence for efficacy of mammography screening starting before 50 with risk reduction of 44% (95% CI: 1–69%) in Gothenburg and 36% (95% CI: 11–55%), in Malmö [6,11]. However, these trials warned against the potential harmful effects of screening pre-menopausal women, such as the need to call back 63 cancer-free women for further examination for each breast cancer death prevented [11].

For disentangling the issue of age at screening and age at dying from breast cancer, the Age Trial, organised in the UK [7] (Table 1), randomised women 39 to 41 years of age, which allowed monitoring breast cancer deaths of women in their forties. After a mean follow-up of 10.7 years, a statistically non-significant reduction in breast-cancer mortality of 17% (95% CI: –4–34%, $P = 0.11$) was observed.

Relative and absolute risk reduction

Relative risks provide an estimate of the risk reduction to be expected from screening, but they tell little about the public health efforts needed for preventing one breast cancer death. This is better achieved using the absolute risk reduction, that is, the number of women that need to be screened in order to save one life.

The US Preventive Services Task Force (USPSTF) [12] calculated the number needed to screen to prevent one death from breast cancer after approximately 14 years of observation. These numbers were 1,224 for women 40–74, 1,792 for women 40–49, and 838 for women 50 years old at screening start. Using results from the Age Trial [7], the number of women 39 to 41 at screening start needed to be screened and followed during 13 years was 1,785, a figure quite close to the USPSTF estimates.

The controversy that started in 2000

The Cochrane review of randomised trials

A systematic review of breast screening randomised trials by a Cochrane Group was highly critical of trial methodologies (Gøtzsche and Olsen, 2000 [13]).

Main critiques concerned the randomisation methods used and the cause-of-death assessment. The review authors considered that the randomisation procedure adopted by all trials but three (NBSS I, NBSS II and MSST I) was flawed. Cause-of-death assessment by local endpoint committees may have been biased if assessors were aware of whether dead women had been allocated to the intervention or to the control group (i.e., lack of concealment of the randomisation group). Gøtszche and Olsen [13] performed a meta-analysis restricted to the three trials with adequate randomisation procedure, which resulted in a relative risk of 1.04 (95% CI: 0.94–1.05). Hence, according to this review, mammography screening was not justifiable.

The re-evaluations by the IARC and the US Preventive Service Task Force

The critiques of Gøtzsche and Olsen prompted a number of works intended to re-evaluate the mammography trials. The two main published re-evaluation works were done by a group of experts convened by the International Agency for Research on Cancer [2] and by the US Preventive Services Task Force [12]. Both re-evaluations acknowledged that several aspects of trials were not optimal and complicated the interpretation of their findings. But overall, except for the Edinburgh trial, results on efficacy of mammography screening were not the consequence of methodological flaws. The main critiques by Gøtzsche and Olsen concerning the trials and the answers provided by trial investigators have been summarised by Elmore and Fletcher [14].

Meta-analyses of randomised trials

A meta-analysis is done using published relative risks and confidence intervals, while a pooled analysis uses data of individual women included in trials. After 2000, one pooled and about 20 meta-analyses of breast screening randomised trials were performed and published in English-language medical journals.

The second overview of Swedish trials with an update of results as of year 2000 was a pooled analysis using original data files which served as the basis for the randomisation [15]. Causes of death were taken directly from the Swedish Cause of Death Registry. Of note, data of the Kopparberg (Dalarna) trial (that was part of the Two-County Trial) were not made available to the Swedish review. Also, mortality reductions in the Östergötland trial (the second county of the Two-County Trial) were lower than reported by investigators, i.e., from 0.76 (95% CI: 0.62–0.94) in Tabar and colleagues [16] to 0.90 (95% CI: 0.73–1.11) in the 2002 overview. Overall, the pooled analysis showed a summary relative risk of breast cancer death of 0.80 (95% CI: 0.71–0.90) associated with invitation to screening of women 40 to 74 years of age.

Meta-analyses yielded quite similar results. Thus, the USTFPS and Cochrane reviews concluded that after 13 or 14 years of follow-up mammography screening was associated with a summary relative reduction in breast cancer death of 20–22% in women 40–74 years of age at screening start [12,17]. The relative reduction was 22–24% for women 50–74, and 15% for women 40–49 at screening start.

From trials to population breast screening

Organised and non-organised screening

The HIP trial and the Swedish trials gave the impetus for screening programmes based on invitation of women (Sweden, Finland, UK, the Netherlands) or for active promotion of screening (e.g., the USA). However, in many countries, the usefulness of breast screening was still a matter of debate (e.g., Germany, Switzerland) or health insurances or social security offices were reluctant to cover costs (e.g., France, Belgium).

Around 2000, mammography screening was widespread in many high-income countries, but the way it was implemented was highly variable within and between countries. In 2011, the two main differences in the way breast screening works are, first, whether screening is centrally organised with invitation of women in target age groups (Centrally organised screening: COS), or whether screening is left to women and doctor's initiative (non-organised or "opportunistic" screening [NOS]). Second, the age at which screening should start or end has always been a subject of contention, with considerable literature on pros and cons of screening before age 50 or after 70.

Typically, in most countries were COS has been established (e.g., the Netherlands, UK, Norway, Finland, Denmark, Canada) screening of women under 50 is rare, and a screening schedule every 2 years (3 years in the UK) is the rule.

In countries where NOS prevails (e.g., USA, Germany, Austria), screening of women under 50 is common and the screening schedule may be annual (mainly in women under 50). Furthermore, exams other than mammography are often done as part of the screening examination, such as BCE in France. Finally, COS and NOS coexist in many countries, e.g., in France, Belgium, Italy, Spain and Ireland.

Guidelines and quality assurance

Quality assurance for installation and operation of mammography units and X-ray film developers has received much attention in the 1990s when screening expanded rapidly. In the USA, the enforcement of the Mammography Quality Standard Act issued in 1992 probably led to the creation of mammography facilities that could better apply quality assurance requirements [18,19].

In Europe, the European Guidelines for Quality Assurance in Breast Cancer Screening and Diagnosis exist since 1993 [20]. These guidelines constitute essential benchmarks guiding health professionals and decision makers. The European guidelines include a number of recommendations and standards for the screening process that do not exist in the USA.

Guidelines in Europe and in the USA did not prevent considerable differences in screening practices between and within countries, and this variability is correlated to the supply of screening services (see below). In the USA, seven bodies have issued different recommendations on age and frequency of screening [21]. In Sweden, age groups invited to screening vary from county to county. Screening before age 50 is rare in the UK, the Netherlands, Norway, Finland and Denmark. In contrast, in Germany, 18% of first mammographies were in women below 30 years and 31% were in women aged 30–39 years [22]. In the USA and in France, 47% and 45% respectively of first mammographies were in women under 40 [23,24].

The variability also concerns use of breast clinical examination, ultrasound examination, and other techniques. Women may be offered so-called "personalised screening" that is costly and time consuming as it systematically includes mammography (often with more than two views), clinical examination, ultrasonography and other methods.

Little evidence supports the efficacy and cost-effectiveness of these screening schemes, and the harm (overdiagnosis, overtreatment) they may cause is unknown.

The (over)supply of mammography machines

A study in 31 countries has estimated that around 2005, there were from 14 to 100 mammography machines per million women [25]. This considerable variability is tightly correlated with the number of radiologists working in a country. For biennial mammography of women 50 to 69 years old, the Dutch National Breast Screening programmes is equipped with 21 mammography units per million women. Hence, a situation of oversupply of mammography units prevails in many countries, which may have undesirable consequences like the broadening of age ranges to whom screening is proposed (mainly women under the age of 40), an increasing frequency of mammography, and increased costs of screening for the community.

The performance of mammography screening

Sensitivity of mammography

Sensitivity of screening is the ability to detect a cancer that is actually in the breast. Cancer missed by screening or developing in years after screening are the interval cancers that are diagnosed in the time period between the last and the following screening rounds.

Sensitivity is usually estimated as the proportion of cancers that are screen-detected among all cancers diagnosed in women attending screening. Using this method, the sensitivity of mammography screening ranged from 61% to 81% in Swedish screening trials [12]. Most screening programmes report sensitivities ranging from 60% to 80% in post-menopausal women. High radiological breast density reduces sensitivity; it is usually lower than 70% in women under 50. Sensitivity is also influenced by radiologist's experience and technical quality of mammograms.

The breast cancer incidence generally increases with screening coverage and participation, first because some cancers are detected earlier (lead time), and second, because a proportion of screen-detected cancers are slow growing or indolent small tumours that would probably have never become clinically apparent (length time – see below). It is therefore more exact to estimate sensitivity as the ratio between incidence rate of interval cancer and the incidence rate of cancer before screening start. According to European guidelines, this ratio should be less than 0.3 and 0.5 in the first and second years following screening, respectively [20].

Mammography density

The amount of radiological dense breast tissue appearing on a mammogram is called the mammography density. The density increases with increasing content of collagen and epithelium, and decreases with increasing content of fat [26]. Mammography density is a strongly genetically determined phenotype. A high density is associated with a higher risk of breast cancer, and independently from its contribution to breast cancer risk, it decreases the ability of mammography to detect cancers [27]. Mammography density is usually greater among pre-menopausal women, which partly explains the lower sensitivity of mammography screening in young women.

False positive test results

A false positive screening test occurs when a woman who actually has no breast cancer is recalled for further exams because her screening test was positive. False positive screening results generate a lot of anxiety and lead to additional imaging exams and invasive procedures (e.g., biopsy). False positive rates vary widely, from a low 2% in women 50–69 years of age in the Netherlands, 4% in the UK [28], up to 11% in some areas in the USA [29]. False positive screenings are more frequent in women under 50, when radiologists are less experienced, and when failure to detect cancer may have serious legal implications, like in the USA. It needs to be underlined that higher recall rates do not increase cancer detection rates [30,31].

False positive rates and higher screening frequency multiply their effect on the likelihood to experience a false positive result over the entire screening period. For instance, it was estimated that after 10 years of annual screening in the United States, half of screened women will have at least 1 false-positive mammogram result [14,29–31]. With triennial screening, 1 of 8 screened English women will have a false positive result after 10 years [28,30,31].

Type of breast cancers detected by mammography screening

Breast carcinomas arise from ductal or from lobular tissues. Lobular carcinomas represent less than 10% of breast carcinomas but they have worse prognosis than ductal carcinomas. Mammography has a greater ability to detected ductal than lobular carcinomas, one reason being that microcalcifications are more frequent in the former type of cancer.

Mammography discovers breast cancers quite different from those showing up clinically as interval cancer. Interval cancers are bigger, more undifferentiated (i.e., of higher grade), sex hormone receptor negative and more frequently associated with distant metastases [32–34]. In addition, when all known outcome predictors are accounted for, the survival of women with screen-detected cancer is better than that of women with interval cancer, meaning that screen-detection is by itself an independent good prognosis factor, whose prognosis significance goes beyond stage migration [34]. The screening interval seems not to influence sensitivity for cancers missed by mammography [35]. There is thus accumulating data suggesting that mammography screening often detects cancer of rather low malignant potential and this raises questions on the real capacity of this technique to achieve early detection of the most life-threatening breast cancers.

In situ breast cancer

Before the implementation of mammography screening, in situ cancers represented 2–5% of all breast cancers. In the years following generalisation of screening, the proportion of in situ breast cancer rose to 10–20% of all breast cancers. This increase has mainly concerned ductal cancer in situ (DCIS) because calcium deposits are frequent in these lesions. The increase in the incidence of DCIS with mammography screening has been particularly spectacular in the USA, where its incidence rose from 1.9 per 100,000 women in 1973–75 (pre-screening epoch) to 32.5 in 2004 [36]. In Europe, changes were of lower magnitude, e.g., from 4 cases per 100,000 Norwegian women in 1993 to 11 in 2007 [37].

Two-thirds of in situ breast cancers are detected during the initial (prevalent) screening round, and one third during subsequent rounds. This illustrates the slow rate of development of this type of cancerous lesion.

Women with (ductal or lobular) in situ breast cancer have a 4-fold increased risk to develop an invasive breast cancer [38]. In situ breast cancer may also be considered as a marker of the propensity to develop an invasive breast cancer because about 40% of invasive cancers diagnosed in women with a history of in situ breast cancer are found in the contralateral breast [38]. However, most in situ breast cancers will not become invasive and the clinical management of these lesions remains controversial.

Overdiagnosis

Overdiagnosis is the diagnosis of a condition that would not have become clinically significant had it not been detected by screening [39]. Autopsy studies have shown that post-menopausal women often harbour undiagnosed breast cancers. These observations indicate the existence of reservoirs of indolent invasive breast cancers, i.e., usually small tumours (less than 20 mm size) having the histological characteristics of invasive cancer but that did not grow and probably had no metastatic potential. These cancers could remain clinically silent during lifetime, they would be seldom life-threatening, and they could even regress and disappear [40]. Detection of indolent cancers does not contribute to the prevention of breast cancer death but it may lead to unnecessary treatment, poorer quality of life, and higher health expenditures.

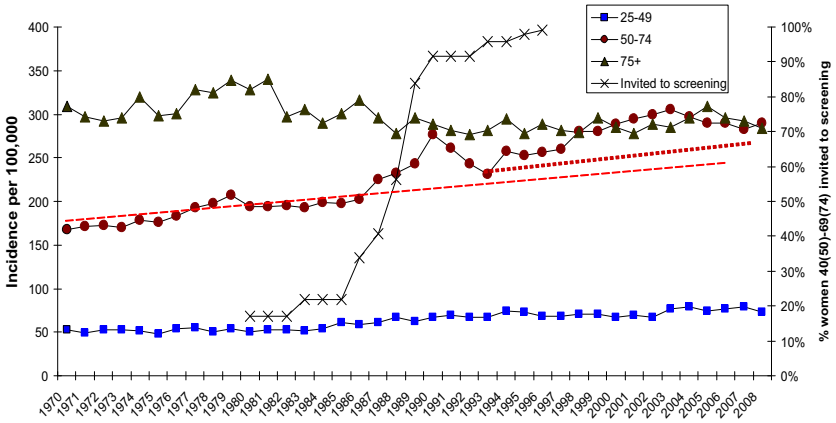

Fig. 1. Breast cancer incidence trends in Swedish women and percentage of women 40–74 years of age invited to screening for the first time (Swedish cancer registry data, NordCan database [41], www.dep-iarc.fr; invitation percentages after Olsson and colleagues, 2000 [42]). The dashed line is a linear trend calculated from incidence rates in women 50–74 years of age from 1970 to 1985, and the dotted line is the expected incidence with screening, after the initial (prevalent) screening round.

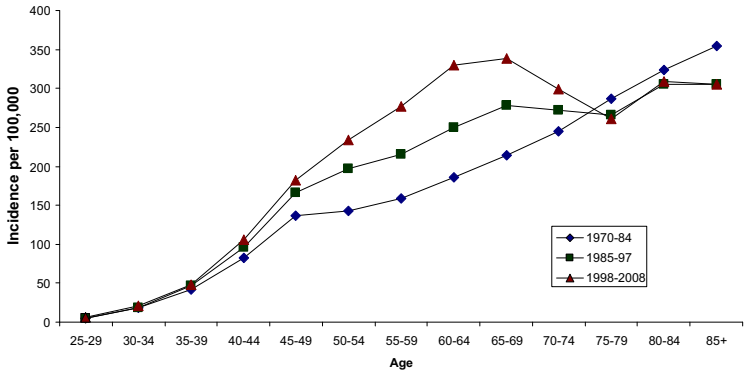

Fig. 2. Age-specific breast cancer incidence in Swedish women, 1970–2008 (Swedish cancer registry data, NordCan database [41], www.dep-iarc.fr).

Early ideas about breast cancer incidence in the screening era

In randomised trials, the numbers of breast cancers diagnosed in the intervention and the control groups were equivalent in the HIP and the Stockholm trials. The number in the intervention group was 30% higher in the Two-County Trial and in the Malmö I trial. It was 11% lower in the Göteborg trial. The higher number of cancers in the intervention groups was deemed to be due to the high screening sensitivity enabling substantial advances in the diagnosis of cancer (lead time). So, in the 1990s, it was thought that rapid nationwide implementation of mammography screening programmes would first result in a hike in cancer incidence due to initial screen that would detect many subclinical cancers already present years before, the so-called prevalent cancers. Further screening rounds would detect cancers developing between screening rounds, the so-called incident cancers. After rolling-out of the screening programme, the screening population would be a mix including a minority of women attending screening for the first time (prevalent screening) and a majority of women participating in further rounds (incident screening). It was thus thought that after the incidence hike due to rolling-out of initial screening rounds, the incidence would return to rates slightly higher (around 5%) than in the pre-screening period.

Small non life-threatening breast cancer

Swedish cancer registry data (Figs 1 and 2) illustrate expected and observed breast cancer incidence trends [41]. Nationwide mammography screening that started after 1985 was followed in 1988–92 by a prevalent incidence peak among women 50–74 invited for the first time to screening. The expected incidence rate after the initial screening round is represented by the dotted line that is slightly above the projected

incidence trend in the absence of screening. After 1992, the incidence in women 50–74 never came back to pre-screening levels but continued to rise, and after 1997, it catched up and even surpassed the incidence in women over the age of 74. After 1985, the incidence remained stable in women 74 years of age and over, an age group that was not invited to screening. Figure 2 shows how incidence rates have steadily increased in Swedish women groups invited to screening. If the sustained raised incidence in women 50–74 was due to earlier detection of cancer that would show up clinically later in life (lead time), then the incidence in women older than 74 should have decreased over time, which was not the case.

The difference between the dotted linear trend and observed rates in Fig. 1 reflects the amount of breast cancer detected by screening that would have remained undetected in the absence of screening. Such situation is often observed in countries where screening is widespread, that is, the incidence of cancer in women 50–69 years of age is now higher than in older, less screened women [43].

What is the amount of screen-induced overdiagnosis?

Some appraisals of incidence trends after screening start concluded that in Swedish randomised trials, overdiagnosis accounted for 1% of screen-detected cancers [44], and that in breast cancer screening programmes, overdiagnosis was probably a minor phenomenon [45]. These estimations corrected for lead time incidence rates observed after screening start, assuming that increased incidence was mainly due to advancement in time of cancer that would have been diagnosed at late stage in the absence of screening. If lead time was the true underlying reason, the total number of breast cancers diagnosed in a population over, say, a 10- or 20-year period after screening start should remain constant. In this respect, the screening-induced increased incidence should be transient and after several years, incidence should come back to levels close to pre-screening levels. Such return to pre-screening levels has yet not been observed in any area where mammography screening was introduced. Hence, the driving force of increased incidence is not lead time, but length time, implying that additional cancers found after screening start are rather slow growing tumours with low malignant potential.

Many works have estimated the amounts of overdiagnosis, but the method to do this are not straightforward [46]. For instance, hormone replacement therapy use by women in screening ages increased during in the 1990s, which also contributed to increase incidence rates. Notwithstanding, most studies that evaluated overdiagnosis in areas with high participation to screening without correction for lead time estimated that 30–50% of invasive breast cancers found by screening in women 50–60 could be considered as overdiagnosis [46–50].

Consequences of overdiagnosis: overtreatment, loss of quality of life and costs

Overdiagnosis usually leads to overtreatment when women detected with small good-prognosis cancer are treated with lumpectomy, hormones or radiotherapy when there is in fact no evidence that treating these small cancers would actually prevent death from breast cancer.

Overtreatment is also a critical issue for in situ cancer. In spite of uncertainties on the natural course of these cancers, surgery for in situ cancer is often more extensive than for small invasive cancers, and many clinicians treat them like invasive cancers, including radiotherapy. Reasons often advanced for advocacy of similar clinical management of in situ and invasive cancers are first, the similarity of risk factors for both types of breast cancer, and second, the known possibility that histological examination of the removed in situ cancer would have missed an area of invasive cancer [36].

For both the small invasive and the in situ breast cancer, a tool that could help predicting the long-term evolution of these lesions would for sure prevent a large proportion of unnecessary treatment.

Breast screening methods

The mammography technology has steadily improved over time. Improvements have mainly concerned doses of X-rays delivered and the quality of mammograms. Despite these improvements, numerous factors may influence the performance of screening mammography, including material quality, radiologist's experience and breast radiological density. In this section, we just summarise some aspects more essential to clinicians.

Techniques for improving mammography screening performance

A constant trend in mammography screening has been the optimisation of cancer detection. Three main techniques have been used for this purpose. First, double reading of mammograms by two independent radiologists increases cancer detection by 5–15% [2].

Second, Computer Aided Diagnostic (CAD) tools aim at assisting radiologists by indicating details of the mammogram that need attention. Large-scale randomised and non-randomised studies did not provide solid evidence for the usefulness of CAD technologies [51]. However, combination of a single radiologist plus CAD seems to have at least similar sensitivity as that obtained by double reading of mammograms [52].

Third, digital mammography has been viewed as a technological breakthrough because, among other things, it was deemed to increase sensitivity in dense breast and it allowed delivering lower X-ray doses. A large North American study showed that digital mammography was more sensitive in women under 50, in women with radiographycally dense breasts, and in pre- or peri-menopausal women [53]. However, the ten studies that compared digital to film-based mammography have obtained conflicting results and it is still uncertain whether digital mammography would not lead to more false positive results and overdiagnosis [54]. Nevertheless, there is a rapid conversion to digital mammography in many high-income countries.

Echography

In many settings, echography has become part of the screening arsenal, mainly as an adjunct to mammography in radiographically dense breasts and for assessing cystic images in young women. This technique is often supplemented with "elastrography", a technique based on the knowledge that the stiffness of benign cysts is different from that of malignant lesions. Also, many breast biopsies are done under ultrasonography guidance. Ultrasonography is however highly operator dependent, and there is a lack of well-conducted studies evaluating its real contribution to screening.

Magnetic resonance imaging (MRI)

MRI is mainly used for assessment of complex lesions found by screening, and for the follow-up of women with high breast cancer risk (e.g., women with BRCA mutation). There is no good evaluation of the benefit to be expected from this costly technique in terms of breast cancer mortality reduction while there is evidence that MRI leads to false positive results 3 to 5 times higher than mammography. Moreover, MRI seems particularly prone to overdiagnosis as it frequently discovers satellite cancerous lesions of unknown clinical significance.

Effectiveness of the screening programme

Effectiveness is the ability of screening to actually reduce breast cancer mortality in the general population. After 1985, breast cancer mortality has started to decrease in most high-resource countries [55,56]. Generalisation of efficient breast cancer treatments took place after 1985, at the same time implementation of breast screening took place in many countries. The simultaneous implementation of screening and of modern treatments has made it difficult to appraise their respective role in mortality reductions. Therefore various approaches have been adopted for assessing how much screening is involved in the reduction of breast cancer mortality. We summarise the most common ones here.

A frequent mistake

A lower proportion of advanced breast cancer or a lower average size of invasive cancers after the introduction of screening is frequently taken as evidence for screening efficiency. This reasoning is erroneous because the increase of the number of slow growing or indolent screen-detected cancers will spuriously lead to reductions in the proportion of advanced (or of big) cancers, even if actual numbers of advanced or of big cancer do not decrease [2,57].

Model-based estimations of impact on breast cancer mortality

The Cancer Intervention and Surveillance modeling NETwork (CISNET) is a consortium of investigators sponsored by the US National Cancer Institute whose purpose is to measure the effect of cancer-control interventions on the incidence of, and risk of death from, cancer in the general population [58,59]. Seven independent statistical models of breast-cancer incidence and mortality concluded that the proportion of the total reduction in the rate of death from breast cancer attributed to screening varied from 28% to 65%, with adjuvant treatment contributing the rest [58].

In the Netherlands, the simulation models known as MISCAN (for microsimulation screening analysis) and its extension MISCAN-Fadia (for microsimulation screening analysis – fatal diameter) [60,61], has been extensively used for estimating likely impact of mammography screening on breast cancer mortality. It found relative risk estimates close to those found by the Two-County Trial, and estimated that in the US, screening led to a 15%, and adjuvant treatment to a 21%, mortality reduction in the year 2000.

Comparison of breast cancer mortality before and after implementation of screening with use of adjustment methods

This method compares breast cancer mortality of women diagnosed with breast cancer who died from it during a period after screening introduction, with women from the same areas who were diagnosed with breast cancer and died from it during a period of same duration before screening introduction. This method is referred to as "incidence-based" or "refined mortality". Most studies of this type were done in Sweden in the years following generalisation of mammography screening; they found 30–45% reductions in breast cancer mortality associated with screening [62–65]. These studies had recourse to variable adjustments for calculation of the risk of breast cancer death during the screening epoch relative to the pre-screening epoch. For instance, adjustments for changes in breast cancer incidence or for lead time were commonly used.

Case–control studies

Case–control studies consist in comparing the screening history of women who died from breast cancer and of women who did not die from breast cancer. With few exceptions, case–control studies have always suggested strong decreases (in the range of 30–50%) in breast cancer mortality associated with screening.

Comparison of trends in breast cancer mortality in areas where screening was and was not implemented

A study in Norway compared the breast cancer mortality rates in areas where screening was implemented to areas without screening and also with rates prevailing before screening start [66]. This study found that screening and patient management had reduced breast cancer mortality by 10% and 20%, respectively. A similar study in Denmark found no difference in breast cancer mortality trends between areas where screening was widespread and areas without screening [67].

A recent study compared trends in breast cancer mortality within three pairs of neighbouring European countries: Northern Ireland (United Kingdom) versus the Republic of Ireland, The Netherlands versus Belgium and Flanders (the Belgian region just south of The Netherlands), and Sweden versus Norway [68]. Countries of each pair were similar with respect to health care services and prevalence of risk factors for breast cancer death, but different in that mammography screening was implemented about ten to fifteen years later in the second country. From 1989 to 2006, no difference in breast cancer mortality decrease was noticeable within each country pair. The contrast between the timing of implementing breast cancer screening and the similarity in mortality reduction between the country pairs suggested that mammography screening did not play a direct role in the mortality reductions of breast cancer.

Changes in incidence rates of advanced breast cancer in areas where breast screening is widespread

Breast screening aims at reducing breast cancer mortality through early detection of cancer, before it has grown into a more life-threatening advanced stage. The randomised trials have shown that reductions in the risk of breast cancer death associated with mammography screening were directly proportional to reductions in the risk of being diagnosed with an advanced breast cancer [5]. Hence, logically, if screening actually reduces breast cancer mortality, it should first reduce the incidence of advanced breast cancer.

Systematic examination of cancer registry data in areas in the USA, Europe and Australia where at least 70% of women in the eligible age group participate in screening since at least 7 years show no or modest decreases in incidence trends of advanced breast cancer [57,69].

Commentary on methods for efficiency evaluation

The model approaches require assumptions about key variables. For instance, breast cancer is a complex, heterogeneous disease and it is not well appreciated up to which point complexity and heterogeneity are appraised by parameters composing these models. Many models estimate reduction in breast cancer deaths form gains in survival due to detection of cancer at an earlier stage (lead time effect). However, mammography screening detects many non life-threatening cancers (length time bias). This phenomenon will increase the overall prognosis of breast cancers and make believe that lives have been saved, when in reality, mortality did not change [70]. The same remark holds for assumptions on screening sensitivity.

Some adjustments of the risks of breast cancer death used in the comparison of mortality before and after screening implementation are problematic. For instance, adjusting breast cancer mortality for changes in breast cancer incidence implies that in the absence of screening, trends in mortality would have paralleled trends in incidence. The validity of this assumption

is questionable since the increases in incidence after 1990 were largely due to mammography screening itself. In addition, incidence-based mortality cannot accurately take into account changes in mortality brought by changes in patient management over time.

Case–control studies for the evaluation of screening efficiency have complex designs, requiring multiple steps and adjustment procedures. For instance, women participating in screening are different from women not participating in screening (e.g., health consciousness, lifestyle habits, socio-economic status), which constitutes a "confounding by indication", that is, in the absence of screening, women participating in screening would be less likely to die from breast cancer than women not participating in screening. Confounding by indication is intractable as no adjustment method fully resolves this source of error in observational studies [71]. Hence, many scientists are sceptical about the interpretation of results from these studies [72].

The method used by Kalager and colleagues [66], Jørgensen and colleagues [67], and Autier and colleagues [68] makes use of historical and contemporary control populations where no screening exist, which allows distinguishing the effects of screening and of patient management. These studies did not use any adjustment of mortality risks on incidence or on lead time.

The method based on incidence trends of advanced breast cancer is independent from treatments [69]. Furthermore, a wide consensus exists since long for considering that in populations where screening is widespread, a decrease in the incidence of advanced breast cancer is the best indicator of the contribution of screening to decreases in breast cancer mortality (see references in Autier and colleagues [69]). The validity of this reasoning is reinforced by the knowledge that decreases in cervical cancer mortality in Nordic countries were preceded by decreases in the rates of advanced cervical cancer [73]. Likewise, fast decreases in the rates of advanced colorectal cancer observed in some countries like the USA are correlated with similar decreases in the incidence of advanced colorectal cancer, most of which would be due to screening [74].

Hence, there is a clear contrast between studies based on sophisticated models and multiple adjustments and studies based on more classic epidemiological analysis of incidence and mortality data. It remains to be settled which of these method is telling the truth. If the simpler methods proved to be right, then the question would be why is there such a discrepancy between results of randomised trials on efficacy and general population studies on efficiency.

Efficiency of breast screening

Efficiency is the balance between health benefits on the one hand and the side effects, the costs of implementation of the screening test or programme, and the changes in the health services induced by screening activities on the other. Efficiency is based on the comparison of health gains and of side effects associated with screening that is summarised in Table 2, taking the Dutch screening programme as an example [75].

Table 2
Advantages and disadvantages of mammography screening of 10,000 women of 50 years and older; the figures are calculated for the Dutch situation in the assumption that the attendance to breast cancer screening is 80% (after ref. [75]).

Advantages (per 10,000 mammograms[a])	Disadvantages (per 10,000 mammograms[a])
55 women have a breast cancer found by screening, this is 5.5 per 1,000 screened women	
4 women with earlier detection of breast cancer will not die from it and will live on average 18 years longer	51 women get the diagnostic 'breast cancer' without changes to the prognosis; they know they have breast cancer for more years than if they had not been screened
	125 women with positive mammogram have no breast cancer and are unnecessarily referred
	16–20 of the screen-detected cancers are small, indolent cancers that are treated unnecessarily (overdiagnosis and overtreatment)
	21 women get interval cancer
	15 women die from breast cancer despite participating in screening[b]

[a] This may also be read 'per 1,000 lifelong participating women', assuming that every participant during 25 years undergoes 10 mammograms.
[b] Including death from interval cancer.

In addition to harms directly related to the screening process displayed in Table 2, other longer-term harmful effects may be due to treatment of many in situ or small invasive cancers that would never have been life-threatening. For instance, the higher risk of long-term cardiac side effects of breast radiotherapy are well documented [76], and greater recourse to radiotherapy for in situ breast cancer may increase the likelihood to die from cardiac disease. Hence, in the long run, the putative modest decrease in breast cancer mortality expected from aggressive treatment of in situ and small invasive cancers could be offset by non breast-cancer mortality associated with the consequences of treatments.

A main difficulty with screening in general is to communicate on efficiency issues, with the challenge of explaining to women the pros and cons of screening. This topic is nowadays the centre of considerable debate, and new data on screening efficiency and side effects to be published in coming years will influence the way virtues and drawbacks of screening will be communicated to decision makers and the public.

Conflict of interest statement

The author has no conflict of interest to declare.

References

1 Thomas DB, Gao DL, Ray RM, et al. Randomized trial of breast-self-examination in Shanghai: Final results. *J Natl Cancer Inst* 2002;**94**:1445–57.

2 IARC. Vainio H, Bianchini F, eds. *IARC Handbooks of Cancer Prevention. Volume 7. Breast Cancer Screening*. Lyon: IARC Press; 2002.

3 Shapiro S. Evidence on screening for breast cancer from a randomized trial. *Cancer* 1977;**39**:2772–82.

4 Tabar L, Fagerberg G, Duffy SW, et al. Update of the Swedish two-county program of mammographic screening for breast cancer. *Radiol Clin North Am* 1992;**30**:187–210.

5 Autier P, Héry C, Haukka J, et al. Advanced breast cancer and breast cancer mortality in randomized controlled trials on mammography screening. *J Clin Oncol* 2009;**27**:5919–23.

6 Bjurstam N, Bjorneld L, Duffy SW, et al. The Gothenburg breast screening trial: First results on mortality, incidence, and mode of detection for women 39–49 years at randomization. *Cancer* 1997;**80**:2091–9.

7 Moss SM, Cuckle H, Evans L, et al. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. *Lancet* 2006;**368**:2053–60.

8 Tabar L, Fagerberg CJ, Gad A, et al. Reduction in mortality from breast cancer after mass screening with mammography. Randomised trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health and Welfare. *Lancet* 1985;**1**:829–32.

9 Nyström L, Rutqvist LE, Wall S, et al. Breast cancer screening with mammography: overview of Swedish randomised trials. *Lancet* 1993;**341**:973–8.

10 Narod SA. On being the right size: a reappraisal of mammography trials in Canada and Sweden. *Lancet* 1997;**349**: 1846.

11 Andersson I, Janzon L. Reduced breast cancer mortality in women under age 50: updated results from the Malmö Mammographic Screening Program. *J Natl Cancer Inst Monogr* 1997;63–7.

12 Humphrey LL; US Preventive Services Task Force. Screening for breast cancer: recommendations and rationale. *Ann Intern Med* 2002;**137**:344–6.

13 Gøtzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet* 2000;**355**:129–34.

14 Fletcher SW, Elmore JG. Mammographic screening for breast cancer. *N Engl J Med* 2003;**348**:1672–80.

15 Nyström L, Andersson I, Bjurstam N, et al. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet* 2002;**359**:909–19.

16 Tabar L, Vitak B, Chen HH, et al. The Swedish Two-County Trial twenty years later. *Radiol Clin North Am* 2000;**38**:625–51.

17 Gøtzsche PC, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2006;**4**: CD001877.

18 Fischer R, Houn F, Van De Griek A, et al. The impact of the mammography quality standards act on the availability of mammography facilities. *Prev Med* 1998;**27**:697–701.

19 Destouet JM, Bassett LW, Yaffe MJ, Butler PF, Wilcox PA. The ACR's Mammography Accreditation Program: ten years of experience since MQSA. *J Am Coll Radiol* 2005;**7**:585–94.

20 Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis. *Ann Oncol* 2008;**19**:614–22.

21 U.S. Preventive Services Task Force Screening for Breast Cancer. U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med* 2009;**151**:716–26.

22 Klug SJ, Hetzer M, Blettner M. Screening for breast and cervical cancer in a large German city: participation, motivation and knowledge of risk factors. *Eur J Public Health* 2005;**15**:70–7.

23 Spyckerelle Y, Kuntz C, Giordanella JP, Ancelle-Park R. Pratiques de la mammographie chez les femmes de 35 à 75 ans: étude descriptive dans la population consultant les centres d'examens de santé. *Bull Cancer* (Paris) 2002;**89**:957–62. Article in French.

24 Colbert JA, Kaine EM, Bigby JA. The age at which women begin mammographic screening. *Cancer* 2004;**101**:1850–9.

25 Autier P, Ait Ouakrim DA. Determinants of the number of mammography units in 31 countries with significant mammography screening. *Br J Cancer* 2008;**99**:1185–90.

26 Boyd NF, Rommens JM, Vogt K, et al. Mammographic breast density as an intermediate phenotype for breast cancer. *Lancet Oncol* 2005;**6**:798–808.

27 Mandelson MT, Nina O, Porter PL, et al. Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. *J Natl Cancer Inst* 2000;**92**:1081–7.

28 NHSBSP. Advisory Committee on Screening for Breast Cancer in England: past and future. *NHSBSP Publication* 2006; No 61.

29 Elmore JG, Armstrong K, Lehman CD, et al. Screening for breast cancer. *JAMA* 2005;**293**:1245–56.

30 Smith-Bindman R, Chu PW, Miglioretti DL. Comparison of screening mammography in the United States and the United Kingdom. *JAMA* 2003;**290**:2129–37.

31 Smith-Bindman R, Ballard-Barbash R, Miglioretti DL, et al. Comparing the performance of mammography screening in the USA and the UK. *J Med Screen* 2005;**12**:50–4.

32 Joensuu H, Lehtima T, Holli K. Risk for distant recurrence of breast cancer detected by mammography screening or other methods. *JAMA* 2004;**292**:1064–73.

33 Gilliland FD, Joste N, Stauber PM, et al. Biologic characteristics of interval and screen-detected breast cancers. *J Natl Cancer Inst* 2000;**92**:743–9.

34 Mook S, Van 't Veer LJ, Rutgers EJ, et al. Independent prognostic value of screen detection in invasive breast cancer. *J Natl Cancer Inst* 2011;**103**:585–97.

35 Narod SA, Dube MP. Re: Biologic characteristics of interval and screen-detected breast cancers. *J Natl Cancer Inst* 2001;**93**: 151.

36 Virnig BA, Tuttle TM, Shamliyan T, et al. Ductal carcinoma in situ of the breast: A systematic review of incidence, treatment and outcomes. *J Natl Cancer Int* 2010;**102**:170–8.

37 Sørum R, Hofvind S, Skaane P, et al. Trends in incidence of ductal carcinoma in situ: the effect of a population-based screening programme. *Breast* 2010;**19**:499–505.

38 Wärnberg F, Yuen J, Holmberg L. Risk of subsequent invasive breast cancer after breast carcinoma in situ. *Lancet* 2000;**355**: 724–5.

39 Black WC. Overdiagnosis: An underrecognized cause of confusion and harm in cancer screening. *J Natl Cancer Inst* 2000;**92**:1280–82.

40 Zahl P, Maehlen J, Welch HG. The natural history of invasive breast cancers detected by screening mammography. *Arch Intern Med* 2008;**168**:2311–16.

41 Engholm G, Ferlay J, Christensen N, et al. *NORDCAN: Cancer incidence, mortality and prevalence in the nordic countries*. Danish Cancer Society; 2008 (www.dep-iarc.fr/nordcan.htm).

42 Olsson S, Andersson I, Karlberg I, Bjurstam N, Frodis E, Hakansson S. Implementation of service screening with mammography in Sweden: from pilot study to nationwide programme. *J Med Screen* 2000;**7**:14–8.

43 Hery C, Ferlay J, Boniol M, Autier P. Quantification of changes in breast cancer incidence and mortality since 1990 in 35 countries with Caucasian-majority populations. *Ann Oncol* 2008;**19**:1009–18.

44 Duffy SW, Agbaje O, Tabar L, et al. Estimates of overdiagnosis from two trials of mammographic screening for breast cancer. *Breast Cancer Res* 2005,**7**:258–65.

45 Paci E, Duffy S. Overdiagnosis and overtreatment in service screening. *Breast Cancer Res* 2005,**7**:266–70.

46 Biesheuvel C, Barratt A, Howard K, et al. Effects of study methods and biases on estimates of invasive breast cancer overdetection with mammography screening: a systematic review. *Lancet Oncol* 2007;**8**:1129–38.

47 Morrell S, Barratt A, Irwig L, et al. Estimates of overdiagnosis of invasive breast cancer associated with screening mammography. *Cancer Causes Control* 2010;**21**:275–82.

48 Duffy SW, Lynge E, Jonsson H, et al. Complexities in the estimation of overdiagnosis in breast cancer screening. *Br J Cancer* 2008;**99**:1176–8.

49 Jørgensen KJ, Gøtzsche PC. Over diagnosis in publicly organized mammography screening programmes: systematic review of incidence trends. *BMJ* 2009;**339**:b2587.

50 Zahl PH, Strand BH, Mæhlen J. Incidence of breast cancer in Norway and Sweden during introduction of nationwide screening: prospective cohort study. *BMJ* 2004;**328**;921–4.

51 Fenton JJ, Taplin SH, Carney PA, et al. Influence of computer-aided detection on performance of screening mammography. *N Engl J Med* 2007;**356**:1399–409.

52 Gilbert FJ, Astley SM, Gillan MGC, et al. Single reading with computer-aide detection for screening mammography. *N Engl J Med* 2008;**359**:1675–84.

53 Pisano ED, Gatsonis C, Hendrick E, et al. Diagnostic performance of digital versus film mammography for breast cancer screening. *N Engl J Med* 2005;**353**:1773–83.

54 Skaane P. Studies comparing screen-film mammography and full-field digital mammography in breast cancer screening: updated review. *Acta Radiol* 2009;**50**:3–14.

55 Héry C, Ferlay J, Boniol M, Autier P. Quantification of changes in breast cancer incidence and mortality since 1990 in 35 countries with Caucasian-majority populations. *Ann Oncol* 2008;**19**:1187–94.

56 Autier P, Boniol M, LaVecchia C, et al. Disparities in breast cancer mortality trends between thirty European countries: retrospective trend analysis of WHO mortality database. *BMJ* 2010;**341**:c3620.

57 Esserman L, Shieh Y, Thompson I. Rethinking screening for breast cancer and prostate cancer. *JAMA* 2009;**15**:1685–92.

58 Berry DA, Cronin KA, Plevritis SK, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. *N Engl J Med* 2005;**353**:1784–92.

59 Cronin KA, Feuer EJ, Clarke LD, et al. Impact of adjuvant therapy and mammography on U.S. mortality from 1975 to 2000: comparison of mortality results from the CISNET breast cancer base case analysis. *J Natl Cancer Inst Monogr* 2006;**36**: 112–21.

60 de Koning HJ, Boer R, Warmerdam PG, et al. Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials. *J Natl Cancer Inst* 1995;**87**:1217–23.

61 Tan SY, van Oortmarssen GJ, de Koning HJ, et al. The MISCAN-Fadia continuous tumor growth model for breast cancer. *J Natl Cancer Inst Monogr* 2006;**36**:56–65.

62 Tabar L, Yen MF, Vitak B, et al. Mammography service screening and mortality in breast cancer patients: 20-year follow-up before and after introduction of screening. *Lancet* 2003;**361**:1405–10.

63 SOSSEG. Reduction in breast cancer mortality from the organised service screening with mammography: 2. Validation with alternative analytic methods. *Cancer Epidemiol Biomarkers Prev* 2006;**15**:52–6.

64 SOSSEG. Reduction in breast cancer mortality from organized service screening with mammography: 1. Further confirmation with extended data. *Cancer Epidemiol Biomarkers Prev* 2006;**15**: 45–51.

65 Hellquist BN, Duffy SW, Abdsaleh S, et al. Effectiveness of population-based service screening with mammography for women ages 40 to 49 years: evaluation of the Swedish Mammography Screening in Young Women (SCRY) cohort. *Cancer* 2011;**117**:714–22.

66 Kalager M, Zelen M, Langmark F, Adami HO. Effect of screening mammography on breast-cancer mortality in Norway. *N Engl J Med* 2010;**363**:1203–10.

67 Jørgensen KJ, Zahl PH, Gøtzsche PC. Breast cancer mortality in organised mammography screening in Denmark: comparative study. *BMJ* 2010;**340**:c1241.

68 Autier P, Boniol M, Gavin A, Vatten L. Breast cancer mortality in neighbouring European countries with different level of breast screening but similar access to treatment: An analysis of WHO mortality database. *BMJ* 2011, in press.
69 Autier P, Boniol M, Middleton R, et al. Advanced breast cancer incidence following population-based mammographic screening. *Ann Oncol* 2011;**22**:1726–35.
70 Autier P, Boniol M, Héry C, et al. Cancer survival statistics should be viewed with caution. *Lancet Oncol* 2007;**8**:1050–2.
71 Bosco JLF, Silliman RA, Thwin SS, et al. A most stubborn bias: no adjustment method fully resolves confounding by indication in observational studies. *J Clin Epidemiol* 2010;**63**:64–74.
72 Cronin KA, Weed DL, Connor RJ, et al. Case–control studies of cancer screening: theory and practice. *J Natl Cancer Inst* 1998;**90**:498–504.
73 Sigurdsson K, Sigvaldason H. Effectiveness of cervical cancer screening in Iceland, 1964–2002: a study on trends in incidence and mortality and the effect of risk factors. *Acta Obstet Gynecol Scand* 2006;**85**:343–9.
74 Edwards BK, Ward E, Kohler BA, et al. Annual report to the nation on the status of cancer, 1975–2006, featuring colorectal cancer trends and impact of interventions (risk factors, screening, and treatment) to reduce future rates. *Cancer* 2010;**116**:544–73.
75 Bonneux L. Advantages and disadvantages of breast cancer screening: Time for evidence based information. *Ned Tijdschr Geneeskd* 2009;**153**:A887.
76 Bouillon K, Haddy N, Delaloge S, et al. Long-term cardiovascular mortality after radiotherapy for breast cancer. *J Am Coll Cardiol* 2011;**57**:445–52.