Effectiveness of and overdiagnosis from mammography screening in the Netherlands: population based study

Philippe Autier,[1,2] Magali Boniol,[2] Alice Koechlin,[1,2] Cécile Pizot,[2] Mathieu Boniol[1,2]

[1]University of Strathclyde Institute of Global Public Health at iPRI, Allée Claude Debussy, 69130 Ecully, Lyon, France

[2]International Prevention Research Institute, Lyon, France

Correspondence to: P Autier philippe.autier@i-pri.org

Additional material is published online only. To view please visit the journal online.

ABSTRACT

OBJECTIVE

To analyse stage specific incidence of breast cancer in the Netherlands where women have been invited to biennial mammography screening since 1989 (ages 50-69) and 1997 (ages 70-75), and to assess changes in breast cancer mortality and quantified overdiagnosis.

DESIGN

Population based study.

SETTING

Mammography screening programme, the Netherlands.

PARTICIPANTS

Dutch women of all ages, 1989 to 2012.

MAIN OUTCOME MEASURES

Stage specific age adjusted incidence of breast cancer from 1989 to 2012. The extra numbers of in situ and stage 1 breast tumours associated with screening were estimated by comparing rates in women aged 50-74 with those in age groups not invited to screening. Overdiagnosis was estimated after subtraction of the lead time cancers. Breast cancer mortality reductions during 2010-12 and overdiagnosis during 2009-11 were computed without (scenario 1) and with (scenario 2) a cohort effect on mortality secular trends.

RESULTS

The incidence of stage 2-4 breast cancers in women aged 50 or more was 168 per 100 000 in 1989 and 166 per 100 000 in 2012. Screening would be associated with a 5% mortality reduction in scenario 1 and with no influence on mortality in scenario 2. In both scenarios, improved treatments would be associated with 28% reductions in mortality.

Overdiagnosis has steadily increased over time with the extension of screening to women aged 70-75 and with the introduction of digital mammography. After deduction of clinical lead time cancers, 32% of cancers found in women invited to screening in 2010-12 and 52% of screen detected cancers would be overdiagnosed.

CONCLUSIONS

The Dutch mammography screening programme seems to have little impact on the burden of advanced breast cancers, which suggests a marginal effect on breast cancer mortality. About half of screen detected breast cancers would represent overdiagnosis.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Since 1989 in the Netherlands women aged 50-75 have been invited to biennial mammography screening; digital mammography was introduced after 2006

A decrease in incidence of advanced cancers is the first sign of the effectiveness of screening and is independent of the effectiveness of treatments

The last evaluation of the national breast screening programme, published in 1993, suggested moderate decreases in the incidence of some categories of advanced cancer from 1989 to 1997

WHAT THIS STUDY ADDS

From 1989 to 2012, no significant decrease in the incidence of stage 2-4 breast cancers has been observed in women aged ≥50

Screening would be associated with 0 to 5% reductions in breast cancer mortality in women aged ≥50, whereas improved patient management would be associated with a 28% reduction

In 2010-12 about one third of breast cancers among women invited to screening represented overdiagnosis

Introduction

The primary goal of cancer screening is to decrease cancer mortality. Cancer screening affects cancer mortality by reducing the number of advanced cancers with poor prognosis. In populations where screening is widespread, decreases in the incidence of advanced cancer should be the first sign that screening effectively reduces cancer mortality.[1-4] This indicator has the advantage of being independent of treatment. It was recommended for the monitoring of the effectiveness of breast screening by the International Agency for Research on Cancer handbook on breast screening published in 2002 and by proponents of mammography screening.[5-8]

In 1988 a nationwide, population based organised cancer screening programme was launched in the Netherlands.[9] Deaths from breast cancer in Dutch women of all ages started to noticeably decrease after 1995.[10] In 2004 the Dutch National Evaluation Team reported reductions in the incidence of tumours of 20 mm or more diameter with positive lymph node or distant metastases, or both from 1989 to 1997 in women aged 50-69.[11] However, the incidence of smaller tumours with regional or distant metastases and tumours of 20 mm or more diameter without regional or distant metastases had remained stable.

Using methods similar to those of the Dutch National Evaluation Team, we updated the stage specific incidence of breast cancer in the Netherlands until 2012. From these updates we estimated the numbers of deaths from breast cancer that were avoided because of screening, and breast cancer overdiagnosis—that is, the number of breast cancers that would have never been detected during a woman's lifetime in the absence of mammography screening.

Methods

National breast screening programme, the Netherlands

From 1989 to 1996, women aged 50-69 were invited to attend biennial mammography screening. In 1997 women aged 70-75 started to be invited. Digital

mammography replaced film based mammography after 2006. Since the inception of the programme around 80% of invited women have attended, and screening of women aged less than 50 or more than 75 is uncommon. In 2000, only 2% of Dutch women aged 45 to 69 used oral combined oestrogen/progestogen hormone replacement therapy, decreasing to 0.5% or less after 2003,[12] a low utilisation rate that had no visible impact on the incidence of breast cancer.[13]

Data sources

We extracted data on breast cancer mortality for the Netherlands for 1987 to 2013 (international classification of diseases, 10th revision (ICD-10) code C50) from the World Health Organization mortality database.[14]

In January 2015, the Integraal Kankercentrum Nederland provided the data on the incidence of in situ and invasive breast cancer by size, stage, and five year age group from 1989 to 2012 (data for all the Netherlands before 1989 are not available). The threshold diameter size for distinguishing cancer stage 1 and 2 is 20 mm.[15 16] We therefore considered in situ and stage 1 cancers as early cancers and stage 2-4 cancers as advanced cancers. This delineation between early and advanced breast cancers is common and was used in mammography screening trials (for example[5]). Moreover, even in the era of efficient adjuvant treatments, the survival of women with stage 1 breast cancer is close to that of women without breast cancer, whereas the transition to stage 2 cancer is associated with lower survival.[17] We obtained data on age groups invited to screening by period, the type of mammography materials, and numbers of screen detected breast cancers in the reports of the Dutch National Evaluation Team.[18 19]

Trend analyses

Using the European Standard Population we age standardised the crude incidence and mortality rates.[20] We computed incidence trends of all breast cancers and advanced breast cancers for women of all ages, women aged less than 50, and women aged 50 or more. We did not split trends of stage 2-4 cancers by narrower age groups because screen detected cancers among women invited to screening could affect the rate of stage 2-4 cancers in older women not invited to screening. For estimating the influence of screening on the incidence of early cancers, we computed incidence rates for in situ and stage 1 cancers for age groups invited and not invited to screening—that is, women aged less than 50 (not invited), 50-69 (invited since 1989), 70-74 (invited since 1997), and 75 or more (not invited).

For quantifying changes in breast cancer incidence and mortality over time, we computed an overall percentage change using the annual percentage change derived from a linear regression of the log transformed age adjusted rates. Analysis of changes in temporal trends was done by joinpoint regression.[21] The joinpoint regression identifies periods when distinct linear or log linear trends are separated by inflexion points (joinpoints). Joinpoint regression was performed on data from the earliest available data until the last year of available data. A log linear regression was used with a maximum of three joinpoints. A minimum of three observations was allowed between two joinpoints and a minimum of three observations was allowed at the beginning and end of the trends.

Breast cancer deaths prevented by screening

Two complementary works have shown that the screen detection at stage 1 of three breast cancers that would have been diagnosed at stage 2-4 in the absence of screening would prevent two deaths from breast cancer.[22] Firstly, the eight randomised trials on breast screening documented that the reduction in risk of death from breast cancer is tightly correlated with reductions in the risk of having a diagnosis of an advanced cancer.[23] In these trials, for an average participation to the first screening round of 80%, the cumulative incidence of stage 2-4 cancers was on average 15% lower in the screening group than in the control group.[23] Secondly, on the basis of the SEER survival data for breast cancer, one study estimated that a 15% reduction in stage 2-4 cancers is associated with a 10% reduction (95% confidence interval 9% to 11%) in breast cancer mortality, and this relative reduction is independent of treatment efficacy.[22] We have used this 15% to 10% (i.e, 3:2 ratio) relation between the numbers of advanced cancers and breast cancer deaths for estimating mortality reductions associated with screening. So, for three fewer stage 2-4 cancers, there would be two fewer breast cancer deaths, and reciprocally. We were therefore willing to estimate the difference between the observed rates of stage 2-4 breast cancers and the rate that would have existed in the absence of screening.

The incidence of advanced breast cancer in the Netherlands before screening started is unknown because national data on cancer incidence before 1989 are not available. Because of the close relation between the risk of cancer death and the risk of advanced cancer when no efficient treatment exists, and because efficient adjuvant, chemotherapeutic, and radiotherapy treatments were uncommon before 1990,[24] we opted for an indirect approach by considering that breast cancer mortality trends in the 1970s and 1980s paralleled trends of advanced breast cancers. Because the analyses showed that mortality started to drop in 1995, we made a linear projection of mortality to 2012 using the annual percentage change prevailing before 1994 (fig 1) and considered that the difference between mortality in 1989 and mortality extrapolated to 2012 was part of mortality reductions due to screening. We then considered the difference in stage 2-4 cancer rates from 1989 to 2012 (fig 1) that we also attributed to screening. The corresponding reduction in breast cancer mortality was derived by multiplying the reduction in stage 2-4 cancers by 2/3 as explained above (fig 1).

The projection to 2012 assumed that the ascending linear trend would persist over time. For testing whether the linearity assumption was valid, we
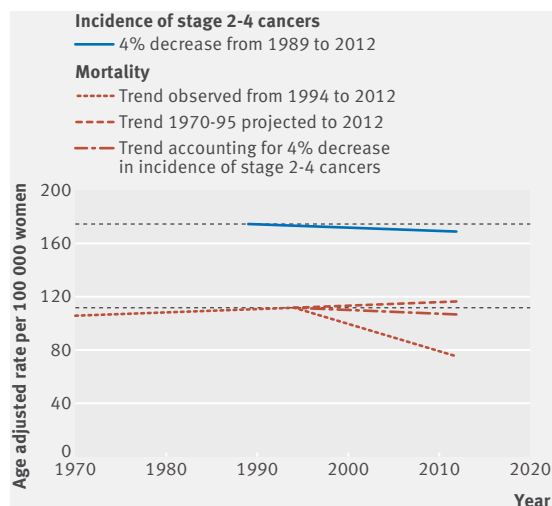
Fig 1 | Estimations of changes in breast cancer mortality in women aged 50 or more in the Netherlands due to mammography screening. Age adjusted rates are per 100 000 women aged 50 or more. The black dashed lines correspond to trends in the absence of change in rates over time

fitted age-period-cohort models to mortality data of women until the inflexion of 1995 (see supplementary materials).

Estimation of overdiagnosis

We considered all years, from 1989 to 2012, during which there was screening. We defined overdiagnosis in women aged 50 to 74 as the incidence of in situ and stage 1 breast cancer in excess to the incidence that would be expected in the absence of screening, corrected for the clinical lead time. In the absence of national incidence figures before 1989, we elected to take breast cancer incidence trends in women aged less than 50 as the trends that should be expected in the absence of screening. We also made estimations using rates in women aged 75 or more and 75 to 84.

Like the Dutch National Evaluation Team,[11] we took the age adjusted incidence rates observed in 1989 as the starting point for computing the expected incidence rates in the absence of screening. We applied the annual percentage change computed for the three age groups not invited to screening to the incidence rates of women aged 50 to 69 for 1989-97, and to the incidence rates of women aged 50 to 74 for 1998-2012.

We computed the proportions of extra cancers diagnosed among women invited to screening for three periods: 1989-2012, when women aged 50-69 were invited to screening, 1998-2012 when women aged 70-75 were also invited, and 2007-12 when digital mammography replaced film based mammography.

The extra cancers included screen detected cancers that would have been diagnosed at stage 2-4 in the absence of screening. Hence, the stage 2-4 cancers that were detected when in situ or at stage 1 (ie, the clinical lead time) need to be deduced from estimation of overdiagnosis.[25] We computed the number of advanced cancers that was detected at an earlier

stage as the mortality reduction due to screening in all women aged 50 or more multiplied by 3/2.

Other extra cancers in women aged 50-74 are the in situ and stage 1 cancers that could have been symptomatic after age 74 or more (when women are no longer invited to screening) but that were screen detected before age 75. These extra cancers should not be considered as overdiagnoses. If these cancers were detected in women invited to screening, however, their detection should be followed by a compensatory drop of in situ and stage 1 cancers in older women not invited to screening. We assessed the possibility of a compensatory drop by running a joinpoint regression on in situ and stage 1 cancers diagnosed after age 74.

Breast cancer mortality reduction in 2010-12 and overdiagnosis in 2009-11

To assess the balance between breast cancer deaths averted owing to screening and the extent of overdiagnosis, we estimated the reductions in the numbers of breast cancer deaths in 2010-12 that resulted from screening or more efficient treatments. A first scenario did not consider cohort effects on mortality and a second scenario considered cohort effects. Overdiagnosis was computed in 2009-11 because the number of screen-detected cancers in 2012 was not yet available.[19]

Results

Breast cancer mortality

The age adjusted breast cancer mortality of 39.0 per 100 000 women of all ages in 1987-89 decreased to 25.8 per 100 000 in 2011-13, hence a 38.1% mortality reduction from 1989 to 2013. The reductions over 24 years were 45.2% in women aged less than 50, 39.1% in women aged 50-69, and 33.3% in women aged 70 or more.

The trends that prevailed before the downward inflexions observed in the 1990s were not uniform across age groups (fig 2): in women aged less than 50, a moderate downward trend was already present from 1970 onwards, whereas rates in women aged 50 to 69 remained stable until 1994. In women aged 70 or more, trends increased until 1997. In women aged 50 or more, after a period of relatively stable incidence (average percentage change 0.09%, 95% confidence interval −0.02% to 0.21%) from 1967 to 1995, a decline of −2.25% per year (95% confidence interval −2.46% to −2.04%) was observed after 1995. The further decrease of −2.76% (ie, −3.38% minus −0.62%) per year observed in women aged less than 50 after 1998 is of greater magnitude than the decrease observed after 1995 in women aged 50 or more.

The age group contrasts in trends denotes gradual decreases in the lifetime risk of breast cancer death in successive generations—that is, a cohort effect— interrupted by noticeable downward inflexions occurring across all age groups in 1993 to 1996. Supplementary figures S1 to S3 in the supplementary materials display the results of the age-period-cohort modelling. A forecast of mortality trends in women
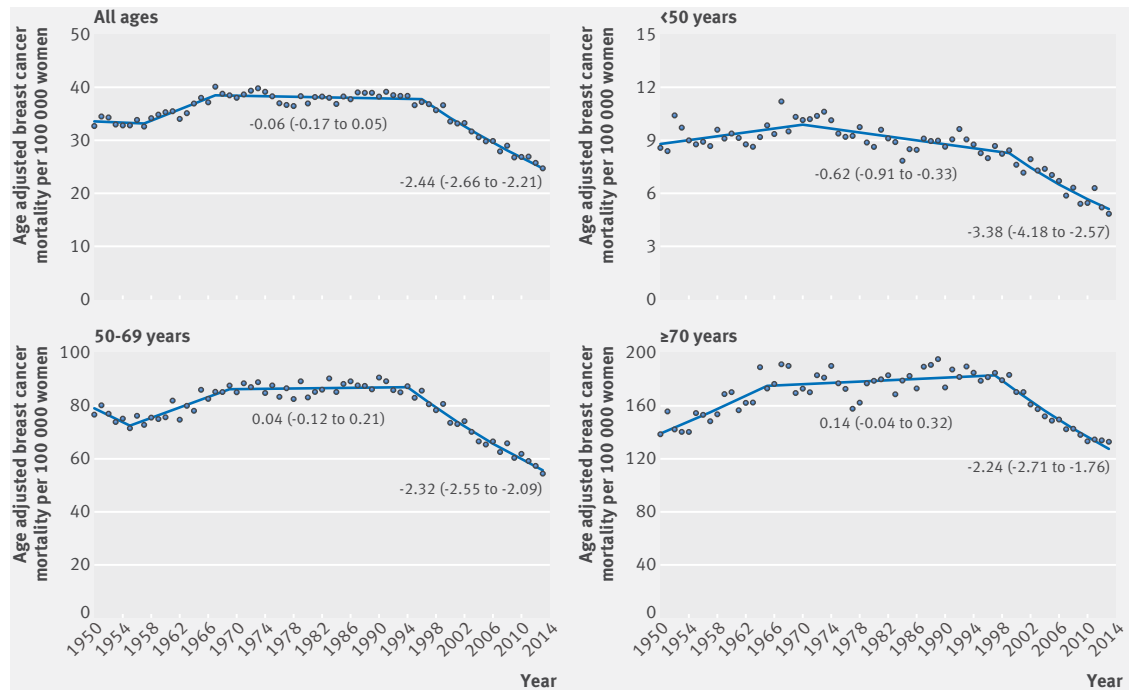
Fig 2 | Trends in age adjusted breast cancer mortality by age group in women in the Netherlands, 1950 to 2013. Figures are the annual percentage changes (95% confidence intervals) for the two last periods with stable trend according to joinpoint analysis

aged 50 or more based on the age-period-cohort model indicates a smooth, progressive tendency towards lower mortality after 1994 owing to a steadily lower risk of breast cancer death among successive generations of women born after 1923 (see supplementary figure S3). This steadily lower risk would be associated with long term changes in genetic, environmental, and lifestyle risk factors associated with breast cancer death. The reduction of mortality due to cohort effects would have happened independently of screening or of improved treatments. This reduction would have ended up as a mortality about 5% lower in 2012 than in 1995.

Stage specific breast cancer incidence

From 1989 to 2012, a total of 27 909 in situ and 267 989 invasive breast cancers were diagnosed among Dutch women of all ages. Trends from 1989 to 2012 show considerable increases in the incidence of in situ tumours and stage 1 cancers, whereas trends in stage 2-4 cancers have remained relatively stable (see supplementary figure S4). The adoption of the axillary sentinel node biopsy procedure by all medical institutions in 2002 induced a stage migration that led to the upstaging of a sizeable fraction of stage 2 cancers into stage 3 cancers. This phenomenon is not perceptible for stage 1 cancers. A joinpoint analysis of stage 1 cancer rates in women of all ages showed annual percentage changes of 11.57% (95% confidence interval 8.96% to 14.23%) in 1989-93, 1.36% (0.85% to 1.87%) in 1993-2005, and 3.48% (2.45% to 4.52%) in 2005-12. There is thus no evidence for a downward inflexion in stage 1 cancer rates in years during which the sentinel node biopsy procedure was generalised.

In women aged 50 or more, the incidence of stage 2-4 cancers remained practically unchanged over time, from 168 per 100 000 in 1989 to 166 per 100 000 in 2012 (fig 3), with a non-significant annual percentage change of −0.16% (95% confidence interval −0.36% to 0.04%, table 1). From 1995 to 2012 the screening programme was fully operational nationwide, with participation constantly around 80%. During that 18 year period the incidence of stage 2-3, stage 4, and stage 2-4 cancers remained stable (table 1). In contrast, the incidence of in situ and stage 1 cancers increased sharply. The comparison of incidence between age groups shows noticeable differences in trends over time that correlate with invitation to screening and the replacement of film based mammography with digital mammography (figs 4 and 5). In women aged
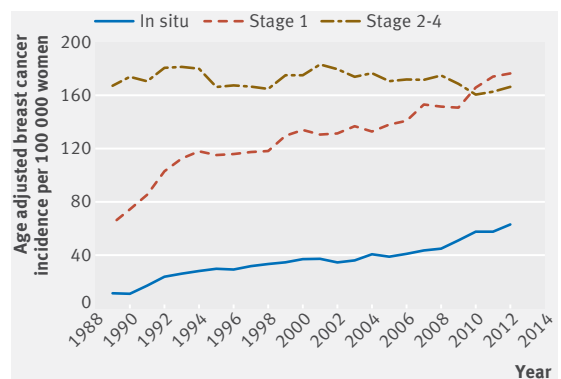

Fig 3 | Trends in age adjusted breast cancer incidence by stage in women aged 50 or more in the Netherlands, 1989 to 2012

Table 1 | Changes in age adjusted breast cancer incidence in women aged 50 or more in the Netherlands

	1989-2012			1995-2012
Cancer stage	Rate in 1989	Rate in 2012	Annual % change (95% CI)	Annual % change (95% CI)
In situ and invasive cancers	252.15	407.98	1.50 (1.22 to 1.79)	1.27 (1.07 to 1.46)
In situ cancers	11.52	63.24	5.88 (4.78 to 7.00)	4.19 (3.56 to 4.82)
Invasive cancers	240.64	344.74	1.06 (0.81 to 1.32)	0.88 (0.70 to 1.06)
Stage 1	64.02	176.25	3.24 (2.64 to 3.85)	2.44 (2.12 to 2.75)
Stage 2-3	150.08	150.75	−0.08 (−0.29 to 0.14)	−0.10 (−0.43 to 0.23)
Stage 4	17.45	15.47	−0.97 (−1.35 to −0.59)	−0.39 (−0.95 to 0.16)
Stage 2-4	167.53	166.22	−0.16 (−0.36 to 0.04)	−0.13 (−0.44 to (0.19)

Rates are per 100 000 women.

50-69, the steep increases in incidence in 1990-96 were followed by further increases until 2007, after which other steep increases occurred when digital mammography was introduced.

After a nine year period of stable incidence in women aged 70-74 similar to that in women aged 75 or more, invitation to screening in 1998 led to a sudden increase in rates followed by further increases until the advent of digital mammography, which prompted new noticeable rate increases. Overall, during the 23 year period, a sixfold increase in the incidence of in situ cancers occurred among women aged 50-74. The incidence of stage 1 cancers increased by threefold in women aged 50-69 and by 3.5-fold in women aged 70-74. During the same period, rates in women aged less than 50 or more than 75 increased by 2.6-fold for in situ cancer and by 1.3-fold for stage 1 cancers. Of note, among women aged 75 or more, joinpoint analyses for the incidence of in situ and stage 1 cancers combined found two successive periods of increasing annual percentage changes, 6.4% per year (95% confidence interval 1.3% to 11.8% per year) from 1989 to 1993, and 0.7% per year (95% confidence interval 0.2% to 1.1% per year) from 1993 to 2012. Hence, the extra in situ and stage 1 cancers diagnosed in women invited to screening were not followed by a downward inflexion in trends among older women not invited to screening, which could suggest the presence of a compensatory decrease at any moment between 1989 and 2012.

Among women aged less than 50, the incidence in stage 2-4 cancers has remained stable, except between 1998 and 2002 when an increase in stage 2 cancers took place, which was interrupted by a sudden decrease corresponding to the adoption of the sentinel node biopsy procedure (see supplementary figure S5).

Causes of mortality reductions

In a first scenario, we assumed that in the absence of efficient treatment the ascending trend of 0.09% per year observed in 1967-95 among women aged 50 or more would have persisted until 2012 (fig 1). Hence the rates would have changed from 110 per 100 000 women in 1995 to 112 per 100 000 women in 2012—that is, by about 2%. We considered that screening did prevent this 2% increase. We then estimated that screening was responsible for the 0.16% per year decline in the incidence of stage 2-4 cancers after 1989, which resulted in a 4% reduction in incidence over the 23 year period (fig 1). This reduction equated to a mortality decrease of about 3% (ie, 4%×2/3, fig 1).

These computations are recapitulated in table 2 for 2010-12, with their translation in numbers of breast cancer deaths. In the first scenario without a cohort effect on mortality, screening would be associated with a 5% reduction in breast cancer mortality, whereas improved treatments would be associated with a 28% reduction in breast cancer mortality. Hence screening would have contributed to 15% of reductions in breast cancer mortality, whereas improved treatments and patient management would have contributed to 85% of reductions.
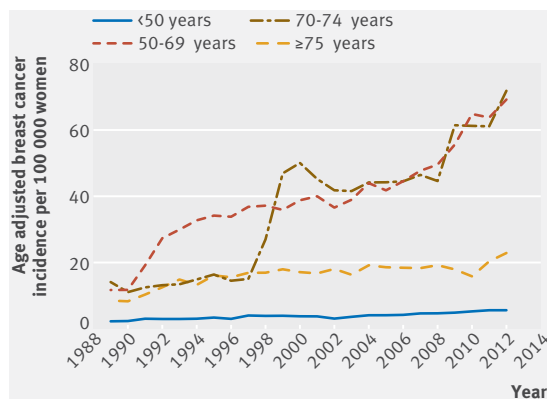

Fig 4 | Trends in age adjusted in situ breast cancer incidence in women by age group in the Netherlands, 1989 to 2012
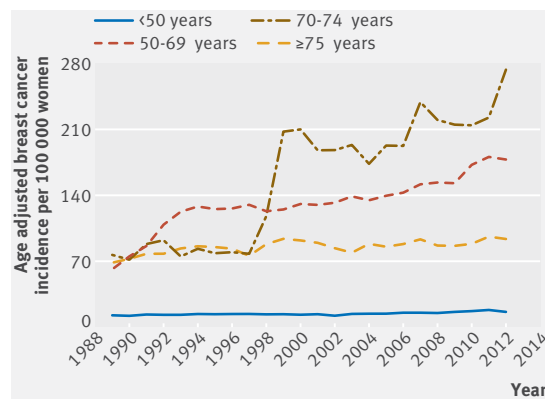

Fig 5 | Trends in age adjusted stage 1 breast cancer incidence in women by age group in the Netherlands, 1989 to 2012

Table 2 | Change in breast cancer mortality in women aged 50 or more living in the Netherlands in 2010-12, assuming no cohort effect on mortality trends

Parameters	Estimate	Row	Comments
No of breast cancer deaths, 2010-12*	8574	(1)	
Decrease in breast cancer mortality from 1995 to 2010-12 (%)	31	(2)	
Additional breast cancer deaths if trends before 1995 had persisted until 2010-12:			
Increase in breast cancer mortality (%)	2	(3)	See fig 1
Reduction in breast cancer mortality (%)	33	(4)	(2)+(3)
No of fewer breath cancer deaths	4223	(5)	[(1)/(1−(4))]−(1)
Mortality changes due to changes in incidence of advanced breast cancers, 1989 to 2012:			
Decrease in incidence of stage 2-4 breast cancers (%)	4	(6)	See fig 1
Decrease in breast cancer mortality due to decreases in stage 2-4 cancers (%)	3	(7)	(6)×2/3 (see methods and fig 1)
Reduction in breast cancer mortality as a result of screening (%)	5	(8)	(3)+(7)
Reduction in breast cancer mortality owing to factors other than screening (%)	28	(9)	(4)−(8)=(2)−(7)
Breast cancer deaths because of screening and other factors:			
No of breast cancer deaths avoided because of screening	640	(10)	(5)×(8)/(4)
No of breast cancer deaths avoided because of factors other than screening	3583	(11)	(5)×(9)/(4)
% of reduction in breast cancer mortality attributable to screening	15	(12)	(10)/(5)
% of reduction in breast cancer mortality attributable to improved treatments and patient management	85	(13)	(11)/(5)

*Source: WHO mortality database.

In a second scenario we considered that mortality trends from 1995 to 2012 declined by about 5% (fig 1) because of a cohort effect (see supplementary materials). Hence the cohort effect would replace the putative effect of screening in the first scenario, whereas factors other than screening would still be associated with a 28% reduction in breast cancer mortality.

Estimation of overdiagnosis

If incidence in women aged less than 50 is taken as the comparison group, during 1989 to 2012 around 50% extra in situ and stage 1 cancers were diagnosed among women in age groups invited to screening (see supplementary tables S1 and S2). Extra cancers amounted to 56% of in situ and stage 1 cancers in 2007-12, when digital mammography totally replaced film based mammography. Using women of older age groups not invited to screening as reference led to slightly higher proportions of extra cancers.

In 2009-11, 27 682 in situ and invasive breast cancers were diagnosed among women aged 50 to 74 in the Netherlands, 16 965 (61.3%) of which were screen detected (table 3). Extra in situ and stage 1 cancers amounted to 9819 cancers, a portion of which represented clinical advances in diagnosis, which we estimated as the 640 screen prevented deaths from breast cancer (table 2) multiplied by 3/2, for obtaining an estimate of 960 cancers (table 3). In the first scenario, after deduction of the 960 cancers that were detected at an earlier stage we estimated that overdiagnosed cancers represented 32% of all breast cancers diagnosed among women in age groups invited to screening and 52% of screen detected breast cancers. Hence, in the era of digital mammography and in the best case scenario, there would be 8859 overdiagnosed cancers for 640 averted breast cancer deaths as a result of screening. Thus for one woman who would not die from breast cancer because of screening, about 14 women would be overdiagnosed with an in situ or a stage 1 cancer. In the second scenario, owing to cohort effects there would be overdiagnosis without earlier detection of cancers that would be stage 2-4 in the absence of screening.

Discussion

Our study is based on the analysis of incidence trends of advanced cancer, a method previously utilised by the Dutch National Evaluation Team for providing evidence that screening was involved in reductions in breast cancer mortality.[11] Our study shows that after 23 years of screening, the reduction in breast cancer mortality in women aged 50 or more associated with screening ranges from 0 to 5%. In contrast, factors other than screening, presumably improvements in the treatment of patients with breast cancer, would have led to a 28% reduction in breast cancer mortality. The likelihood of the predominant influence of treatment effectiveness on mortality reductions is supported by the observation that the timing and magnitude of decreases in mortality were similar among women aged less than 50 and among women aged 50-69. Moreover, the noticeable reduction in breast cancer mortality in women aged less than 50 was independent of the incidence of stage 2-4 cancers and thus entirely resulted from improved treatments.

The computations in table 2 suggesting a 5% reduction in mortality associated with screening are based on two suppositions in favour of screening: firstly, we considered that screening had somehow affected trends of stage 2-4 cancers towards lower rates despite the absence of a statistically significant decrease in incidence of these cancers. Secondly, we posited that the difference between the mortality rate linearly extrapolated to 2012 and the mortality rate observed in 1995 (fig 1) was as a result of screening. Objections against the latter supposition could be raised because the difference in rates between 1995 and 2012 could also result from improved treatments. However, the age-period-cohort model (see supplementary materials) indicates that the linear extrapolation of trends prevailing before 1995 is probably not the most

Table 3 | Estimation of overdiagnosis in women aged 50-74 years in the Netherlands, 2009-11

Parameters	Estimate	Row	Comments
No of breast cancers diagnosed in women aged 50-74 in 2009-11*:			
In situ	4422	(1)	
Stage 1	12 804	(2)	
Stage 2+	10 270	(3)	
Not staged	186	(4)	
Total	27 682	(5)	
No of screen detected breast cancers†	16 965	(6)	
% of screen detected breast cancers	61	(7)	(6)/(5)
Estimation of overdiagnosed in situ and stage 1 breast cancers:			
No of extra cancers, considering 57% extra cancers	9819	(8)	((1)+(2))×0.59 (extra cancer proportion of 59% from supplementary table S2)
No of screen detected breast cancers representing clinical lead time	960	(9)	612 breast cancer deaths avoided because of screening (from row 10 of table 2 multiplied by 3/2)
No of overdiagnosed breast cancers	8859	(10)	(8)−(9)
Breast cancers overdiagnosed among women invited to screening (%)	32	(11)	(10)/(5)
Breast cancers overdiagnosed among women with screen detected breast cancer (%)	52	(12)	(10)/(6)
Overdiagnosed in situ and stage 1 cancer among women invited to screening (%)	51	(13)	(10)/((1)+(2))

*Source: Integraal Kankercentrum Nederland.
†Source: Dutch National Evaluation Team, 2014.

likely scenario, and that the long lasting cohort effect is likely to contribute to declines in mortality.

The Dutch National Evaluation Team acknowledged that after 1997 the incidence of advanced breast cancer no longer decreased but tended to return to prescreening levels.[26] No explanation was, however, proposed that would not challenge the effectiveness of screening. Other studies in the Netherlands have documented the poor influence of screening on the incidence of advanced breast cancer, one study in women participating in screening in a southern province and another in all women aged 70 to 75.[27 28]

Stage migration as a result of the sentinel node biopsy procedure has led to an upstaging in only a fraction of patients, mainly from stage 2 to stage 3 cancers.[29-31] Hence stage migration is unlikely to explain the quasi-stability of the incidence of stage 2-4 cancers.[32] We found no explanation for the transient increase in stage 2 cancers observed among women aged less than 50. Probably this phenomenon was due to changes in the histological assessment of surgically removed materials. No data on stage specific incidence of breast cancer are available before 1989, and thus it is impossible to assess whether the continuous decrease in mortality that started in 1970 in women aged less than 50 was associated with changes in the incidence of advanced cancers among young Dutch women.

These observations on the limited ability of mammography screening to reduce the incidence of advanced cancers are supported by studies on breast cancer mortality. Twenty years after the introduction of mammography screening in the Dutch city of Nijmegen, deaths from breast cancer did not decline any earlier than in the Dutch city of Arnhem or in the Netherlands generally.[33] The healthcare systems and the socioeconomic environments in the Netherlands and Belgium are comparable. Breast screening in

Belgium was implemented about 15 years later than in the Netherlands, and in 2000-05 participation to screening was still around 50% in Belgium, whereas in the Netherlands it has remained around 80% since 1989.[34] The age adjusted breast cancer mortality of 37.5 per 100 000 in Belgian women of all ages in 1987-89 decreased to 26.0 per 100 000 in 2011-13, hence a 24 year mortality reduction of 34%, which is comparable to the 38% reduction observed in the Netherlands.

The steep increases of in situ tumours and stage 1 cancers in women aged 50 to 74 after invitation to screening are typical of period effects affecting age groups rapidly exposed to a new detection method. Because changes in the incidence of advanced cancers were marginal, and because no compensatory decrease in the incidence of early stage breast cancer was observed in women aged 75 or more, most of the increases in in situ tumours and stage 1 cancers in women aged 50-74 would represent overdiagnosis.

Overdiagnosis has increased over time with the extension of screening to women aged 70-75, and with the replacement of film based mammography by digital mammography. In 2009-11, 51% of in situ tumours and stage 1 cancers found in women aged 50-74 and 52% of screen detected cancers would represent overdiagnosis.

A recent study of the Dutch National Evaluation Team estimated that 22% of cancers found in women invited to screening from 1989 to 2009 represented overdiagnosis.[35] We found that 22% to 24% of breast cancers found among women invited to screening during 1989-2009 represented overdiagnosis (see supplementary table S3).

The overdiagnosis of stage 1 cancers was about two times higher in women aged 70-74 than in women aged 50-69 (fig 5). This age contrast accords with the

50% overdiagnosis rate found in Dutch women aged 70-75 by another study.[27]

Observations in the Netherlands are in line with other data showing that the replacement of film based mammography by digital mammography has substantially increased the burden of in situ breast tumours without decreasing rates of interval cancers.[36-39] Hence the advent of digital technologies has probably worsened the overdiagnosis problem without clear evidence for improvements in the ability of screening to curb the risk of breast cancer death. The gradual increase in overdiagnosis over time is the major factor involved in the increasing lifetime risk of in situ or invasive breast cancer that shifted from 1 in 10 Dutch women in 1990 to 1 in 6.6 in 2010.[40]

Strengths and limitations of this study

Our study has several strengths. Cancer registration in the Netherlands is of high quality, and a study estimated that the reliability of cause of death statistics in the Netherlands was high (>90%) for major causes of death such as cancers.[41] The 3:2 relation between reductions in stage 2-4 breast cancers and death from breast cancer we used is derived from results of randomised trials, cancer registries, and cause of death registries.

Our study has several limitations. We assumed that incidence trends would follow trends in age groups not invited to screening and that mortality trends would follow trends prevailing before the shifts observed in the 1990s. Unlike other studies[42] we had no contemporary comparison population with or without low participation to screening. Moreover, nationwide data on breast cancer incidence did not exist before 1989. The choice to use trends in women aged less than 50 and older than 74 was probably the most acceptable surrogate for trends that would have existed in the absence of screening. Our choice is supported by several observations: trends in incidence in women aged less than 50, 70-74, and 75 or more were similar until 1997. From 1975 to 1998 in the Eindhoven area (southwest of the Netherlands), the incidence of breast cancer among women aged less than 50, 50-69, and 70 or more increased at about the same pace.[43]

Our assumption that in the absence of screening and effective patient treatment the incidence of stage 2-4 cancers would have paralleled that of breast cancer mortality is justified by the close relation between the risk of death from breast cancer and the risk of having a diagnosis of an advanced breast cancer observed in randomised trials.[23]

Conclusion

The data on advanced breast cancer in the Netherlands indicate that the Dutch national mammography screening programme would have had little influence on the decreases in breast cancer mortality observed over the past 24 years. This conclusion accords with the mounting evidence that randomised trials have overestimated the ability of mammography screening to reduce the risk of deaths from breast cancer in the entire life period after first exposure to mammography screening.[44-46] In contrast, the extent of overdiagnosis is continuously increasing with the invitation of older women to screening and the adoption of imaging technologies able to detect increasingly smaller breast tumours, most of which are of unknown clinical importance.

1   Morrison AS. Intermediate determinants of mortality in the evaluation of screening. *Int J Epidemiol* 1991;20:642-50. doi:10.1093/ije/20.3.642
2   World Health Organization. *Cancer IAfRo. Cancer: Causes, Occurrence and Control.* International Agency for Research on Cancer, 1990.
3   Cole P, Morrison AS. Basic issues in population screening for cancer. *J Natl Cancer Inst* 1980;64:1263-72.
4   Greenwald P, Sondik EJ, Young JLJr. Emerging roles for cancer registries in cancer control. *Yale J Biol Med* 1986;59:561-6.
5   Tabár 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. doi:10.1016/S0140-6736(85)92204-4
6   Smith RA, Duffy SW, Gabe R, Tabar L, Yen AM, Chen TH. The randomized trials of breast cancer screening: what have we learned? *Radiol Clin North Am* 2004;42:793-806, v.
7   Day NE, Williams DR, Khaw KT. Breast cancer screening programmes: the development of a monitoring and evaluation system. *Br J Cancer* 1989;59:954-8. doi:10.1038/bjc.1989.203
8   IARC. Breast Cancer Screening: IARC Press 2002.
9   Fracheboud J, de Gelder R, Otto SJ, et al. National evaluation of breast cancer screening in the Netherlands 1990-2007. The Netherlands: National Evaluation Team for Breast Cancer Screening (NETB), 2009.
10 Autier P, Boniol M, La Vecchia C, et al. Disparities in breast cancer mortality trends between 30 European countries: retrospective trend analysis of WHO mortality database. *BMJ* 2010;341:c3620. doi:10.1136/bmj.c3620.
11 Fracheboud J, Otto SJ, van Dijck JA, Broeders MJ, Verbeek AL, de Koning HJ. National Evaluation Team for Breast cancer screening (NETB). Decreased rates of advanced breast cancer due to mammography screening in The Netherlands. *Br J Cancer* 2004;91:861-7. doi:10.1038/sj.bjc.6602075.
12 Faber A, Bouvy ML, Loskamp L, van de Berg PB, Egberts TC, de Jong-van den Berg LT. Dramatic change in prescribing of hormone replacement therapy in The Netherlands after publication of the Million Women Study: a follow-up study. *Br J Clin Pharmacol* 2005;60:641-7. doi:10.1111/j.1365-2125.2005.02502.x.

13 Soerjomataram I, Coebergh JW, Louwman MW, Visser O, van Leeuwen FE. Does the decrease in hormone replacement therapy also affect breast cancer risk in the Netherlands? *J Clin Oncol* 2007;25:5038-9, author reply 5039-40. doi:10.1200/JCO.2007.13.7281.

14 World Health Organization. *Mortality database.* WHO, 2013.

15 American Joint Committee on Cancer. *AJCC Cancer Staging Manual*, 6th ed. Springer, New York, 2002:171-80.

16 International Union Against Cancer. *TNM Classification of Malignant Tumours.* 6th ed. Wiley-Liss, 2002.

17 Walters S, Maringe C, Butler J, et al. ICBP Module 1 Working Group. Breast cancer survival and stage at diagnosis in Australia, Canada, Denmark, Norway, Sweden and the UK, 2000-2007: a population-based study. *Br J Cancer* 2013;108:1195-208. doi:10.1038/bjc.2013.6.

18 (NETB) NETfBcsitN. *NETB interim report 2007: Main results breast cancer screening programme in the Netherlands.* Erasmus MC, Rotterdam, and UMC Radboud, 2007.

19 NETB. NETfBcs. National evaluation of breast screening in the Netherlands, 1990-2011/12. Rotterdam, Nijmegen, 2014.

20 Waterhouse J, Muir C, Correa P, et al. *Cancer incidence in five continents.* Vol III. World Health Organization, 1976.

21 Jointpoint regression program [program]. Version 3.4.2. version, 2009.

22 Birnbaum J, Gadi VK, Markowitz E, Etzioni R. The Effect of Treatment Advances on the Mortality Results of Breast Cancer Screening Trials: A Microsimulation Model. *Ann Intern Med* 2016;164:236-43. doi:10.7326/M15-0754.

23 Autier P, Héry C, Haukka J, Boniol M, Byrnes G. Advanced breast cancer and breast cancer mortality in randomized controlled trials on mammography screening. *J Clin Oncol* 2009;27:5919-23. doi:10.1200/JCO.2009.22.7041.

24 Vervoort MM, Draisma G, Fracheboud J, van de Poll-Franse LV, de Koning HJ. Trends in the usage of adjuvant systemic therapy for breast cancer in the Netherlands and its effect on mortality. *Br J Cancer* 2004;91:242-7. doi:10.1038/sj.bjc.6601969.

25 Zahl PH, Jørgensen KJ, Gøtzsche PC. Overestimated lead times in cancer screening has led to substantial underestimation of overdiagnosis. *Br J Cancer* 2013;109:2014-9. doi:10.1038/bjc.2013.427.

26 de Gelder R, van As E, Tilanus-Linthorst MM, et al. Breast cancer screening: evidence for false reassurance? *Int J Cancer* 2008;123:680-6.

27 de Glas NA, de Craen AJ, Bastiaannet E, et al. Effect of implementation of the mass breast cancer screening programme in older women in the Netherlands: population based study. *BMJ* 2014;349:g5410. doi:10.1136/bmj.g5410.

28 Nederend J, Duijm LE, Voogd AC, Groenewoud JH, Jansen FH, Louwman MW. Trends in incidence and detection of advanced breast cancer at biennial screening mammography in The Netherlands: a population based study. *Breast Cancer Res* 2012;14:R10. doi:10.1186/bcr3091.

29 Maaskant AJ, van de Poll-Franse LV, Voogd AC, Coebergh JW, Tutein Nolthenius-Puylaert MC, Nieuwenhuijzen GA. Stage migration due to introduction of the sentinel node procedure: a population-based study. *Breast Cancer Res Treat* 2009;113:173-9. doi:10.1007/s10549-008-9913-8.

30 Tvedskov TF. Staging of women with breast cancer after introduction of sentinel node guided axillary dissection. *Dan Med J* 2012;59:B4475.

31 Vanderveen KA, Schneider PD, Khatri VP, Goodnight JE, Bold RJ. Upstaging and improved survival of early breast cancer patients after implementation of sentinel node biopsy for axillary staging. *Ann Surg Oncol* 2006;13:1450-6. doi:10.1245/s10434-006-9109-6.

32 Narod SA, Iqbal J, Miller AB. Why have breast cancer mortality rates declined? *J Cancer Policy* 2015;5:8-17. doi:10.1016/j.jcpo.2015.03.002.

33 Broeders MJ, Peer PG, Straatman H, et al. Diverging breast cancer mortality rates in relation to screening? A comparison of Nijmegen to Arnhem and the Netherlands, 1969-1997. *Int J Cancer* 2001;92:303-8. doi:10.1002/1097-0215(200102)9999:9999<::AID-IJC1186>3.0.CO;2-B

34 Autier P, Boniol M, Gavin A, Vatten LJ. Breast cancer mortality in neighbouring European countries with different levels of screening but similar access to treatment: trend analysis of WHO mortality database. *BMJ* 2011;343:d4411. doi:10.1136/bmj.d4411.

35 Ripping TM, Verbeek AL, Fracheboud J, de Koning HJ, van Ravesteyn NT, Broeders MJ. Overdiagnosis by mammographic screening for breast cancer studied in birth cohorts in The Netherlands. *Int J Cancer* 2015;137:921-9. doi:10.1002/ijc.29452.

36 Nederend J, Duijm LE, Louwman MW, et al. Impact of the transition from screen-film to digital screening mammography on interval cancer characteristics and treatment - a population based study from the Netherlands. *Eur J Cancer* 2014;50:31-9. doi:10.1016/j.ejca.2013.09.018.

37 Hoff SR, Abrahamsen AL, Samset JH, Vigeland E, Klepp O, Hofvind S. Breast cancer: missed interval and screening-detected cancer at full-field digital mammography and screen-film mammography--results from a retrospective review. *Radiology* 2012;264:378-86. doi:10.1148/radiol.12112074.

38 Sala M, Domingo L, Macià F, Comas M, Burón A, Castells X. Does digital mammography suppose an advance in early diagnosis? Trends in performance indicators 6 years after digitalization. *Eur Radiol* 2015;25:850-9. doi:10.1007/s00330-014-3431-3.

39 Hofvind S, Skaane P, Elmore JG, Sebuødegård S, Hoff SR, Lee CI. Mammographic performance in a population-based screening program: before, during, and after the transition from screen-film to full-field digital mammography. *Radiology* 2014;272:52-62. doi:10.1148/radiol.14131502.

40 van der Waal D, Verbeek AL, den Heeten GJ, Ripping TM, Tjan-Heijnen VC, Broeders MJ. Breast cancer diagnosis and death in the Netherlands: a changing burden. *Eur J Public Health* 2015;25:320-4. doi:10.1093/eurpub/cku088.

41 Harteloh P, de Bruin K, Kardaun J. The reliability of cause-of-death coding in The Netherlands. *Eur J Epidemiol* 2010;25:531-8. doi:10.1007/s10654-010-9445-5.

42 Jørgensen KJ, Gøtzsche PC, Kalager M, Zahl PH. Breast cancer screening in denmark: A cohort study of tumor size and overdiagnosis. *Ann Intern Med* 2017;166:313-23. doi:10.7326/M16-0270.

43 Louwman WJ, Voogd AC, van Dijck JA, et al. On the rising trends of incidence and prognosis for breast cancer patients diagnosed 1975-2004: a long-term population-based study in southeastern Netherlands. *Cancer Causes Control* 2008;19:97-106. doi:10.1007/s10552-007-9075-8.

44 Gøtzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet* 2000;355:129-34. doi:10.1016/S0140-6736(99)06065-1.

45 Autier P, Boniol M, Smans M, Sullivan R, Boyle P. Statistical analyses in Swedish randomised trials on mammography screening and in other randomised trials on cancer screening: a systematic review. *J R Soc Med* 2015;108:440-50. doi:10.1177/0141076815593403.

46 Autier P, Boniol M, Smans M, Sullivan R, Boyle P. Observed and Predicted Risk of Breast Cancer Death in Randomized Trials on Breast Cancer Screening. *PLoS One* 2016;11:e0154113. doi:10.1371/journal.pone.0154113.

Supplementary information: additional tables and figures