Cancer Epidemiology 42 (2016) 147–153

Contents lists available at ScienceDirect

Cancer Epidemiology

The International Journal of Cancer Epidemiology, Detection, and Prevention

journal homepage: www.cancerepidemiology.net

Extrapolation of pre-screening trends: Impact of assumptions on overdiagnosis estimates by mammographic screening

T.M. Ripping[a], A.L.M. Verbeek[a], K. ten Haaf[b], N.T. van Ravesteyn[b], M.J.M. Broeders[a,c,*]

[a] Radboud Institute for Health Sciences, Radboud University Medical Center, PO Box 9101 (route 133), 6500HB Nijmegen, The Netherlands
[b] Department of Public Health, Erasmus MC, University Medical Centre, PO Box 2040, 3000CA Rotterdam, The Netherlands
[c] Dutch Reference Centre for Screening, PO Box 6873, 6503 GJ Nijmegen, The Netherlands, The Netherlands

ARTICLE INFO

Article history:
Received 17 December 2015
Received in revised form 22 March 2016
Accepted 25 April 2016
Available online 3 May 2016

Keywords:
Breast cancer
Excess incidence
Mammographic screening
Netherlands
Overdiagnosis
Pre-screening trend
Screening

ABSTRACT

Background: Overdiagnosis by mammographic screening is defined as the excess in breast cancer incidence in the presence of screening compared to the incidence in the absence of screening. The latter is often estimated by extrapolating the pre-screening incidence trend. The aim of this theoretical study is to investigate the impact of assumptions in extrapolating the pre-screening incidence trend of invasive breast cancer on the estimated percentage of overdiagnosis.
Methods: We extracted data on invasive breast cancer incidence and person-years by calendar year (1975–2009) and 5-year age groups (0–85 years) from Dutch databases. Different combinations of assumptions for extrapolating the pre-screening period were investigated, such as variations in the type of regression model, end of the pre-screening period, screened age range, post-screening age range and adjustment for a trend in women <45. This resulted in 69,120 estimates of the percentage of overdiagnosis, i.e. excess cancer incidence in the presence of screening as a proportion of the number of screen-detected and interval cancers.
Results: Most overdiagnosis percentages are overestimated because of inadequate adjustment for lead time. The overdiagnosis estimates range between −7.1% and 65.1%, with a median of 33.6%. The choice of pre-screening period has the largest influence on the estimated percentage of overdiagnosis: the median estimate is 17.1% for extrapolations using 1975–1986 as the pre-screening period and 44.7% for extrapolations using 1975–1988 as the pre-screening period.
Conclusion: The results of this theoretical study most likely cover the true overdiagnosis estimate, which is unknown, and may not necessarily represent the median overdiagnosis estimate. This study shows that overdiagnosis estimates heavily depend on the assumptions made in extrapolating the incidence in the pre-screening period, especially on the choice of the pre-screening period. These limitations should be acknowledged when adopting this approach to estimate overdiagnosis.
© 2017 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Mammographic screening benefits some women and harms a number of others [1,2]. The major benefit of mammographic screening is the prevention of breast cancer deaths [1] by detecting breast cancers at an early stage with better treatment outcomes [3]. However, a major drawback of mammographic screening is the detection of cancers that would not be clinically detected during a woman's lifetime if screening had not occurred, i.e. overdiagnosed cancers.

There is much debate on the extent of overdiagnosis in mammographic screening, with estimates ranging from 0 to 57% [4,5]. According to Carter et al. [6], ecological and cohort studies are the most suitable method for estimating overdiagnosis. There is, however, a wide variability in the design of these studies, which are related to the methods used to adjust for lead time and the choice of the unscreened reference population [6]. The unscreened reference population is often obtained through extrapolating the incidence in the pre-screening period [5,7–10]. However, studies utilize different assumptions in order to estimate the pre-screening incidence trend [11]. Some studies apply linear regression to incidence rates, while others apply poisson regression to absolute numbers. Furthermore, studies differ in the age groups modeled, choice of pre-screening period and whether to adjust for a trend in non-screened ages.

* Corresponding author at: Radboud Institute of Health Sciences, Radboud university medical center, PO Box 9101 (route 133), 6500HB Nijmegen, The Netherlands.
E-mail address: Mireille.broeders@radboudumc.nl (M.J.M. Broeders).

http://dx.doi.org/10.1016/j.canep.2016.04.015
1877-7821/© 2017 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

The Independent UK Panel on Breast Cancer Screening [11] estimated the absolute number of overdiagnosed cases in the UK National Health Service breast cancer screening program using several different assumptions. They showed that the estimated number of overdiagnosed cases depends on the specification of the model used for the estimation. Although this indicates that the choice of the model influences the estimated percentage of overdiagnosis, the panel only discussed the effects of a limited number of model assumptions. Furthermore, the percentage of overdiagnosis does not only depend on the estimated number of overdiagnosed cases, but also on the number of cancers in the denominator [12]. Therefore, this theoretical study investigates the influence of a large number of assumptions in extrapolating pre-screening incidence trends on the estimated percentage of overdiagnosis by mammographic screening in the Netherlands.

2. Methods

2.1. Setting

In 1989, a biennial mammographic screening program was gradually implemented in the Netherlands, inviting women aged 50–69 years. Nationwide full coverage was reached in 1997 and the upper age limit was gradually extended to age 75 in the period 1998–2001. Because women are invited per region and receive their first invitation in the year they turn 50, 51 or 52, women aged 49 can be screened. The attendance rates in the Dutch program have always been high, ranging from 72% in 1990 to about 80% from 1997 onwards [13]. Until 2014, initial screens consisted of two view mammography and subsequent screens of one view, an oblique view, unless a second cranio-caudal view was required. From 2014 onwards, two view mammography became the standard for subsequent screening. Mammograms are independently read by two radiologists who decide in consensus on recall. Digital mammography was introduced in 2004 and reached full coverage in 2010 [14].

Because the incidence of *ductal carcinoma in situ* (DCIS) was not registered before 1989, we limited our estimates of overdiagnosis to invasive breast cancer. Data on the number of invasive breast cancers were obtained from *Stichting Medische Registratie* for the period 1975–1988 (ages 0–85 years) and the website of the National Cancer Registry [15] in the Netherlands for the period 1989–2013 (ages 0–99 years). The number of screen-detected breast cancers and interval cancers were collected centrally from the screening organizations (1975–2009) [14] and the information on the number of women living in the Netherlands were obtained from Statistics Netherland (1975–2013) [16]. All data was provided by calendar year and 5-year age groups (0–85 years) (see supplement A). Fig. 1 presents the invasive breast cancer incidence rate per 100,000 women-years by calendar period and age group.

2.2. Percentage of overdiagnosis

The percentage of overdiagnosis was defined as '*the percentage of cancers detected during the screening period that would not present symptomatically during one's lifetime in the absence of screening*', in line with previous work [17]. The nominator is the absolute number of overdiagnosed cases estimated by subtracting the cumulative incidence in the absence of screening from the cumulative incidence in the presence of screening. In this study, the cumulative incidence in the presence of screening is the observed breast cancer incidence in the screened age group during the screening period. The cumulative incidence in the absence of screening could not be observed and was estimated by extrapolation of pre-screening incidence trends. This approach is called the cumulative incidence method or excess-incidence method [18,19]. The cumulative incidence approach needs to fulfill two conditions to adequately adjust for lead time: 1) the follow-up after screening cessation should include the maximum length of lead time, and 2) the excess incidence during screening and the compensatory drop after screening cessation should be estimated from women who had the same screening participation rates and experienced the

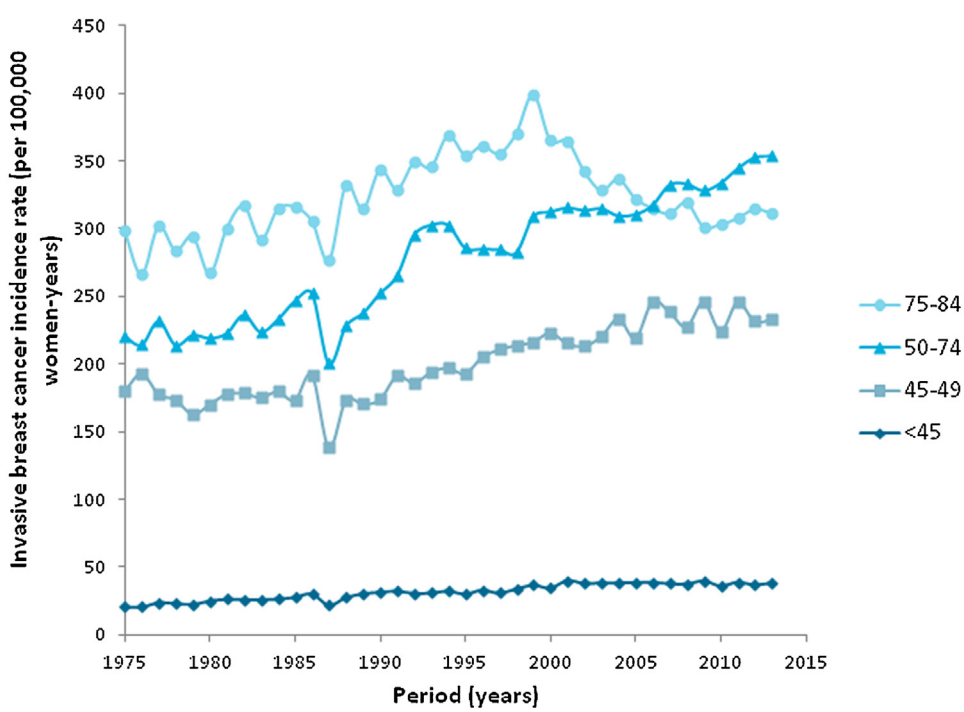

Fig. 1. Invasive breast cancer incidence rate per 100,000 women-years by calendar period.

Table 1
Assumptions used to estimate the percentage of overdiagnosis.

Category	Options
Model	
Type of regression model	poisson regression, linear regression
End of pre-screening period	1984–1989 *(the start of the pre-screening era is 1975 for all analyses)*
Age groups to estimate period trend	20–45, 45–49, 45–74, 50–69, 50–74, 70–74, 75–79, 75–84, 80–84
Calculation	
Screened age range	45–74, 50–74
Post-screening age range	75–79, 75–84
Starting year screening period	1990–2001
Stopping year screening period	2005–2009
Numbers	absolute, rates
Risk adjustment	no, trend in women aged 20–45

same screening practice [20,21]. In this study, we could not fulfill the last condition – and probably also not the first condition – because we estimated overdiagnosis in periods with changing participation rates and screening practice rather than in birth cohorts [20]. Knowing that the participation and detection rates increased over time, the overdiagnosis estimates are likely to be overestimated [18,19]. Furthermore, the cumulative incidence approach can result in negative overdiagnosis estimates when the observed incidence in the presence of screening is lower than the expected incidence in the absence of screening. Overdiagnosis can, however, never be negative; therefore, values below zero are to be interpreted as overdiagnosis being non-existent (no overdiagnosis). For the denominator we used the number of cancers detected in women participating in the screening program, i.e. screen-detected breast cancers and interval cancers (i.e. breast cancers diagnosed in screened women during the interval between two screening rounds) [17].

2.3. Models

We estimated overdiagnosis through extrapolating the pre-screening incidence trend, using 69,120 different combinations. These combinations varied with regards to the: type of regression model used, end of the pre-screening period, and the age-groups used to estimate the period trend (see Table 1). The percentage of overdiagnosis was estimated for different starting and stopping years of the screening period, screened age ranges, post-screening age groups, and with/without adjustment for a trend in women younger than 45. We adjusted for a trend in women younger than 45 years by dividing the expected incidence in the screened age range by the relative excess, i.e. the ratio between the expected and observed incidence, in women younger than 45 [8]. The percentage of overdiagnosis was calculated from absolute numbers and rates, regardless of the type of regression model used to estimate the pre-screening trend.

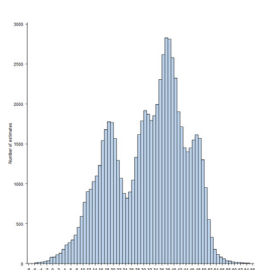

Fig. 2. Overdiagnosis estimates obtained from the extrapolation of pre-screening trends using different model specifications.

3. Results

The percentage of overdiagnosis obtained by the different assumptions to estimate the pre-screening trend ranges from −7.1% to 65.1% (Fig. 2). Fig. 2 shows that the range of estimates form a multimodal distribution with peaks at 18%, 31%, 38% and 48%. These peaks represent a group of small normal distributions, which are obtained when stratifying for pre-screening period, screened age range and adjustment for a trend in women below 45. Overall, the estimates are not normally distributed and the median estimated percentage of overdiagnosis is 33.6%

Table 2 presents the median percentages of overdiagnosis that were derived from 34,560 model combinations without adjustment for a trend in women younger than 45 for different assumptions. The table shows that the end of the pre-screening era has the largest influence on the percentage of overdiagnosis. For the screened age range 50–74, the median percentage of overdiagnosis is 14.3% for the estimates using 1975–1986 as the pre-screening period and 39.6% for estimates using 1975–1988 as the pre-screening period. Fig. 1 gives an explanation for this difference: it shows a peak in the breast cancer incidence rate in 1986, which causes a higher trend and thus low overdiagnosis estimates, and a drop in the breast cancer incidence rate in 1987, which has an opposite effect on the trend and overdiagnosis estimates. Table 2 also demonstrates that the overdiagnosis estimates are higher for the screened age range 45–74 than for the screened age range 50–74, which is caused by an increasing breast cancer incidence in women aged 45–49 (see Fig. 1).

Other model specifications, i.e. type of regression model, calculation method, age groups used to estimate period trend, selected post-screening age groups, and starting and stopping year of the screening period, also influence the percentage of overdiagnosis, but to a smaller extent. For example, overdiagnosis estimates decline with later starting and stopping years of the screening era. The effect of age groups in the model to estimate the period trend seems to depend on two factors, namely the number of groups and the specific age ranges in each group. If the number of age groups in the model is smaller and if the screened age range is grouped together (i.e. either 45–74 or 50–74), the overdiagnosis estimates are lower. The estimates of overdiagnosis are also slightly lower when linear regression is used instead of poisson regression, overdiagnosis is calculated based on rate ratios rather than absolute numbers and the post-screening age group was 75–84 years (10-year follow-up) rather than 75–79 years (5-year follow-up).

Table 3 presents the median overdiagnosis percentage of 34,560 model combinations adjusted for a trend in women below 45. The adjusted overdiagnosis estimates are generally lower than the unadjusted overdiagnosis estimates (30.2% versus 37.0%). When comparing the results from Table 2 and 3, it becomes apparent that the effect of each model specification is similar for

Table 2
Median (quartile 1–quartile 3) percentage of overdiagnosis unadjusted for a trend in women <45.

	Screened age range		Total
	45–75	50–75	
Overall	43.0 (29.2–47.8)	35.0 (22.1–38.8)	37.0 (26.5–44.0)
Type of regression			
Linear regression	42.6 (29.0–46.8)	34.1 (22.1–37.9)	36.5 (25.9–43.2)
Poisson regression	44.2 (29.4–48.6)	36.1 (22.0–39.7)	37.4 (27.2–45.1)
Calculation overdiagnosis			
Based on rates	42.1 (28.2–46.8)	34.1 (20.9–37.7)	35.9 (25.6–43.0)
Based on absolute numbers	44.2 (30.3–48.8)	36.2 (22.9–39.8)	38.0 (27.8–45.1)
End of pre-screening period			
1984	38.7 (36.7–40.9)	33.1 (31.0–35.4)	36.1 (33.0–39.0)
1985	29.2 (27.5–30.9)	22.0 (19.8–24.9)	26.6 (22.0–29.4)
1986	18.7 (17.1–20.3)	14.3 (12.2–16.9)	17.1 (14.2–19.2)
1987	47.9 (46.3–49.4)	38.2 (36.3–40.6)	43.8 (38.2–47.9)
1988	48.7 (47.1–50.2)	39.6 (37.7–41.7)	44.7 (39.6–48.7)
1989	46.8 (45.3–48.3)	38.2 (36.3–40.2)	43.1 (38.2–46.8)
Age groups to estimate period trend			
<45, 45–74, 75–79, 80–84	43.3 (28.3–47.2)	38.5 (26.5–41.8)	39.9 (37.4–44.6)
<45, 45–74, 75–84	42.2 (27.5–46.4)	37.5 (25.8–40.9)	38.9 (26.6–43.7)
<45, 45–49, 50–69, 70–74, 75–79, 80–84	44.9 (30.8–48.9)	35.0 (21.5–38.2)	36.8 (25.7–44.9)
<45, 45–49, 50–69, 70–74, 75–84	43.8 (30.0–48.0)	34.0 (20.7–37.3)	35.8 (25.3–43.8)
<45, 45–49, 50–74, 75–79, 80–84	43.8 (29.7–48.3)	34.2 (30.3–37.7)	36.3 (25.5–44.1)
<45, 45–49, 50–74, 75–84	42.5 (28.9–47.4)	33.1 (19.7–36.7)	35.3 (24.6–42.8)
Post-screening age group			
75–79	43.1 (29.5–47.8)	35.1 (22.3–38.8)	37.0 (26.8–44.0)
75–84	43.0 (29.0–47.7)	34.9 (21.8–38.8)	36.9 (26.3–44.0)
Starting year of screening era			
1990	45.2 (31.2–49.8)	37.5 (23.8–40.8)	39.0 (28.8–46.0)
1991	44.9 (31.3–49.5)	37.2 (23.8–40.5)	38.8 (28.7–45.7)
1992	44.6 (31.0–49.2)	36.8 (23.6–40.1)	38.5 (28.5–45.3)
1993	43.9 (30.3–48.5)	36.0 (22.7–39.3)	37.7 (27.7–44.60
1994	43.5 (29.9–48.1)	35.5 (22.3–39.3)	37.4 (27.1–44.7)
1995	42.7 (28.9–47.3)	34.6 (21.6–38.1)	36.5 (26.4–43.4)
1996	42.1 (28.5–46.9)	34.2 (20.7–37.5)	36.0 (25.9–42.8)
1997	42.1 (28.5–46.9)	33.9 (20.2–37.6)	36.0 (25.6–42.9)
1998	42.1 (28.4–46.9)	34.2 (20.5–37.5)	36.0 (25.8–42.9)
1999	42.3 (28.5–47.2)	34.4 (20.7–37.8)	36.2 (25.9–43.1)
2000	42.3 (28.3–47.1)	34.1 (20.7–37.7)	36.1 (25.6–43.0)
2001	41.5 (27.2–46.5)	33.9 (20.7–38.5)	36.0 (24.9–43.1)
Stopping year of screening era			
2005	44.7 (30.9–49.3)	36.9 (23.7–40.5)	38.7 (28.4–45.7)
2006	43.6 (29.8–48.2)	35.5 (22.3–39.2)	37.4 (27.2–44.4)
2007	43.1 (29.2–47.7)	35.0 (21.8–38.6)	36.9 (26.7–43.8)
2008	42.5 (28.7–47.2)	34.6 (21.3–38.2)	36.5 (26.2–43.3)
2009	41.8 (27.9–46.5)	33.7 (20.5–37.4)	35.8 (25.3–42.6)

the unadjusted and adjusted estimates except for differences in the end of the specified pre-screening period and the calculation of overdiagnosis. The adjusted overdiagnosis estimates from Table 3 are higher than the unadjusted estimates from Table 2 when the pre-screening period ends between 1984 and 1986 and vice versa for the other pre-screening periods. Furthermore, in Table 2 overdiagnosis estimates are higher when based on rate ratios rather than absolute numbers, but the opposite is true for the adjusted estimates in Table 3.

4. Discussion

The percentage of overdiagnosis estimated through the extrapolation of pre-screening trends varies from −7.1% to 65.1% in this study. The period used to estimate the pre-screening trend has the largest influence on the percentage of overdiagnosis, but the influence of other factors should not be neglected.

4.1. Comparision with other studies

Several studies estimated the percentage of overdiagnosis using the extrapolation of pre-screening trends. The estimates of these studies vary from less than 0% [9] to 57% [5]. Because each study uses a different denominator to estimate the percentage of overdiagnosis, comparison of our estimates with previously reported estimates is not straightforward. Our study, however, shows that a wide variation in overdiagnosis estimates can even solely be the result of using different assumptions for the extrapolation of pre-screening trends, regardless of the denominator used to estimate overdiagnosis.

In addition, we showed that variation in overdiagnosis estimates based on extrapolation of pre-screening trends depends on assumptions such as the pre-screening period, the age groups used to model the pre-screening trend, the screened age group, the starting and stopping year of the screening era, adjustment for a trend in women below 45, and the method to calculate overdiagnosis (absolute numbers vs rates). The Independent UK panel estimated the absolute number of overdiagnosed cases using different models [11]. They found, similar to this study, that the number of overdiagnosed cases decreased with a smaller screened age range (50–64 vs 45–64), when the estimates were adjusted for a trend in women below 45, and when linear regression was used rather than poisson regression [11]. Furthermore, the UK panel found an inverse relationship between length of pre-screening period and number of overdiagnosed cases, which was not observed in our study. Because this relation depends on the

Table 3
Median (quartile 1–quartile 3) percentage of overdiagnosis adjusted for a trend in women <45.

	Screened age range		Total
	45–75	50–75	
Overall	35.5 (20.6–40.6)	28.0 (16.2–33.4)	30.2 (18.4–37.9)
Type of regression			
Linear regression	33.1 (30.4–39.6)	26.4 (15.7–31.9)	28.7 (18.0–36.3)
Poisson regression	37.1 (30.7–41.8)	39.8 (16.6–34.7)	31.9 (18.9–39.2)
Calculation overdiagnosis			
Based on rates	37.9 (29.2–40.7)	30.4 (22.1–33.6)	32.7 (26.3–38.3)
Based on absolute numbers	26.5 (16.0–40.2)	20.1 (11.4–32.5)	23.6 (13.7–36.5)
End of pre-screening period			
1984	43.7 (39.7–50.1)	37.2 (33.4–42.7)	40.8 (36.4–46.7)
1985	31.6 (28.6–36.6)	24.8 (21.0–29.2)	28.8 (24.4–33.3)
1986	23.9 (19.1–36.4)	20.6 (14.5–29.2)	22.2 (17.4–32.7)
1987	32.7 (19.0–40.3)	26.4 (13.6–32.2)	37.7 (16.2–36.8)
1988	29.2 (12.6–39.7)	24.5 (8.6–32.2)	25.4 (10.5–36.5)
1989	30.5 (16.4–39.2)	25.5 (11.9–31.8)	26.8 (14.0–36.0)
Age groups to estimate period trend			
<45, 45–74, 75–79, 80–84	35.2 (19.6–40.6)	31.6 (19.5–36.3)	32.4 (19.5–38.6)
<45, 45–74, 75–84	34.1 (18.9–39.9)	30.6 (18.8–35.5)	31.5 (18.9–37.7)
<45, 45–49, 50–69, 70–74, 75–79, 80–84	36.8 (21.8–41.4)	28.0 (14.9–32.6)	30.5 (18.9–38.3)
<45, 45–49, 50–69, 70–74, 75–84	36.0 (21.1–40.6)	27.1 (14.3–31.8)	29.6 (18.1–37.4)
<45, 45–49, 50–74, 75–79, 80–84	36.4 (20.9–40.9)	27.1 (14.2–32.1)	29.6 (18.0–37.8)
<45, 45–49, 50–74, 75–84	35.4 (20.3–40.0)	26.1 (13.5–31.3)	28.6 (17.2–36.9)
Post-screening age group			
75–79	35.5 (20.9–40.6)	28.1 (16.6–33.4)	30.3 (18.9–37.9)
75–84	35.5 (20.1–40.6)	27.9 (15.6–33.3)	30.1 (18.0–37.8)
Starting year of screening era			
1990	39.9 (23.8–44.3)	32.5 (19.7–36.8)	34.1 (22.0–41.9)
1991	39.6 (23.9–43.8)	32.2 (19.8–36.4)	33.9 (22.1–41.5)
1992	39.3 (23.7–43.5)	31.9 (19.6–36.0)	33.5 (21.9–41.2)
1993	38.4 (22.5–42.5)	30.8 (18.2–34.9)	32.5 (20.8–40.3)
1994	37.2 (20.8–41.4)	29.8 (16.7–33.9)	31.5 (19.2–39.1)
1995	36.3 (19.6–40.4)	28.8 (15.1–32.7)	30.4 (18.1–38.1)
1996	35.4 (18.6–39.4)	27.8 (13.8–32.0)	29.6 (17.2–37.3)
1997	34.9 (18.1–38.9)	27.3 (13.2–31.5)	29.1 (16.8–36.8)
1998	33.9 (17.0–38.2)	26.3 (12.0–31.3)	28.4 (15.5–36.3)
1999	33.5 (16.7–38.4)	26.0 (12.3–31.5)	28.5 (14.9–36.0)
2000	32.9 (15.9–37.6)	25.4 (11.1–30.8)	27.7 (14.0–35.5)
2001	30.9 (13.9–37.1)	23.3 (9.1–29.9)	26.4 (10.9–33.9)
Stopping year of screening era			
2005	37.5 (23.0–42.4)	30.1 (18.6–35.3)	31.9 (20.7–39.8)
2006	36.2 (21.2–41.0)	28.5 (16.8–33.8)	30.5 (18.9–38.3)
2007	35.8 (20.3–40.5)	28.1 (15.7–33.3)	30.1 (18.1–37.8)
2008	35.3 (19.5–39.9)	27.7 (15.1–32.8)	29.7 (17.6–37.3)
2009	34.1 (18.6–39.0)	26.5 (14.1–31.7)	28.7 (16.6–36.3)

increase in the breast cancer incidence rates before the introduction of screening, it is likely to vary between countries. For example, in the Netherlands the breast cancer incidence peaks at 1986 and drops in 1987 causing the lowest and highest overdiagnosis estimates, when using these years as the end of the pre-screening period. Unfortunately, it is unknown whether this sudden fluctuation is real or due to data errors (i.e. registration issues). Our study also showed that overdiagnosis is influenced by the age groups included in the model and shows the impact of each assumption in perspective. Unlike estimates for absolute numbers of overdiagnosed cases, estimates for percentage of overdiagnosis are not dependent on the number of women screened and illustrate the impact of each assumption in a standardized way.

4.2. Limitations of utilizing pre-screening trends

The overdiagnosis estimates in this theoretical study range from −7.1% to 65.1% and will therefore most likely cover the true overdiagnosis estimate. The true overdiagnosis estimate however remains unknown and is not necessarily represented by the median overdiagnosis estimate of this study. The use of extrapolation of pre-screening trends has two important assumptions when it comes to estimating overdiagnosis. Firstly, this method assumes a constant trend in the (breast) cancer incidence rate after the pre-screening period ends [11]. The validity of this assumption can be questioned because it requires that all risk factors for (breast) cancer, both birth cohort and period related, cause a constant increase in the (breast) cancer incidence rate by calendar year. Besides the validity of this assumption, this study showed that the magnitude of the constant increase is not a set value and is highly dependent on the choice of the pre-screening period. Secondly, the use of extrapolation of pre-screening trends assumes that the quality of case ascertainment remains the same [11]. This is unlikely given the improvements in diagnostic procedures over the last decades.

In addition to the assumptions required to extrapolate pre-screening trends, an adjustment for a trend in women younger than 45 assumes that the breast cancer incidence rate increases at the same pace in women in the screened age range and below 45. In other words, the risk factors for pre- and postmenopausal breast cancers are assumed to be similar or at least cause a similar proportional increase in breast cancer incidence in each age group. This assumption can be questioned, because some risk factors, i.e. hormone-replacement therapy, only have an effect on post-

menopausal breast cancers. Adjustment for a trend in women younger than 45 may however be useful when the unadjusted and adjusted trends are compared. Similar overdiagnosis estimates for unadjusted and adjusted trends may indicate the more reliable trends and therefore more reliable overdiagnosis estimates. Another way to obtain more reliable and stable trends is the utilization of 5-year smoothed averages.

Another limitation of using of pre-screening trends is that it can often only be used to estimate overdiagnosis from invasive cancer. DCIS is more likely to be overdiagnosed than invasive breast cancer. However, in most countries DCIS was not registered before the introduction of mammographic screening and can therefore not be included in the overdiagnosis estimate. This leads to an underestimation of overdiagnosis.

4.3. Limitations of the current study

Besides the limitations of the assumptions needed to extrapolate pre-screening trends, other factors may cause an overestimation of the range of overdiagnosis estimates provided in this study. We estimated overdiagnosis in a population that is still screened, which leads to an inadequate adjustment for lead time. In order to adequately adjust for lead time and obtain a reliable estimate of overdiagnosis using a cumulative incidence approach, overdiagnosis should be estimated in a population in which screening has ceased and is followed up until the maximum length of lead time after screening has ceased. Such a population can be obtained either by measuring overdiagnosis in birth cohorts that stop screening at a certain age [20] or by measuring overdiagnosis in countries that once screened for (breast) cancer but have ceased screening. Because we could not adequately adjust for lead time in this study, the overdiagnosis estimates are most likely overestimated. This is also reflected by the decreasing percentage of overdiagnosis with later starting and stopping years of the screening period: overdiagnosis is expected to increase within this period based on advances in technology (i.e. digital mammography) [22]. However, this does not occur because the compensatory drop starts to compensate the excess cases from about 2002 onwards in the age group 75–79 and even later in the age group 80–84. Other overdiagnosis estimates from the Netherlands with a more adequate adjustment for lead time have also reported estimates towards the lower end of the range of estimates presented here and also included DCIS [12,23].

Another limitation of this study is that we analyzed the data for the whole population of the Netherlands rather than per region, even though screening was implemented in different years in different regions/municipalities. Incidence trends per region will prevent dilution and show a more pronounced shift in the breast cancer incidence after introduction of mammographic screening [24]. The effect of such analyses on the overdiagnosis estimate is not straightforward, because the compensatory drop may also be more pronounced and the denominator depends – just as the number of overdiagnosed cases – on the implementation. Analyses per region are, however, likely to result in more accurate overdiagnosis estimates.

Furthermore, we would like to point out that we used the cumulative incidence method [18], also called excess incidence method [19], to estimate the percentage of overdiagnosis. Extrapolation of pre-screening trends is also used to create a reference population for estimating overdiagnosis [25,26] and to estimate the percentage of overdiagnosis using the lead time approach [27,28]. It can be expected that the influence of the pre-screening trends on the percentage of overdiagnosis in such studies is smaller than the range of estimates presented here [25], because the current estimates are also affected by an inadequately measured drop after leaving screening.

4.4. Conclusion

To conclude, extrapolation of pre-screening trends are commonly used to estimate the percentage of overdiagnosis. This study shows that overdiagnosis estimates are heavily dependent on the assumptions made, especially those for the pre-screening period. Researchers should acknowledge the limitations of extrapolation of pre-screening trends and adjust adequately for lead time.

Author contribution

T.M. Ripping conceived the idea for this study and designed the study. She obtained data and carried out the analyses. She drafted the manuscript and approved the final versions of the manuscript.

A.L.M. Verbeek contributed to the interpretation of data and revised the manuscript critically for important intellectual content. He approved the final version of the manuscript.

K. the Haaf contributed to the interpretation of data and revised the manuscript critically for important intellectual content. He approved the final version of the manuscript.

N.T. van Ravesteyn contributed to the design of the study, interpretation of data and revised the manuscript critically for important intellectual content. She approved the final version of the manuscript.

M.J.M. Broeders contributed to the design of the study, interpretation of data and revised the manuscript critically for important intellectual content. She approved the final version of the manuscript.

Conflicts of interest

The authors declare that they have no conflicts of interest

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.canep.2016.04.015.

References

[1] Independent UK panel on Breast Cancer Screening, The benefits and harms of breast cancer screening: an independent review, Lancet 380 (2012) 1778–1786.
[2] E. Paci, M. Broeders, S. Hofvind, D. Puliti, S.W. Duffy, E.W. Group, European breast cancer service screening outcomes: a first balance sheet of the benefits and harms, Cancer Epidemiol. Biomarkers Prev. 23 (2014) 1159–1163.
[3] N. Perry, M. Broeders, C. de Wolf, S. Törnberg, R. Holland, L. von Karsa, European guidelines for quality assurance in breast cancer screening and diagnosis: fourth edition – summary document, Ann. Oncol. 19 (2008) 614–622.
[4] D. Puliti, S.W. Duffy, G. Miccinesi, H. de Koning, E. Lynge, M. Zappa, et al., Overdiagnosis in mammographic screening for breast cancer in Europe: a literature review, J. Med. Screen. 19 (2012) 42–56.
[5] K.J. Jørgensen, P.C. Gøtzsche, Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends, BMJ (Online) 339 (2009) 206–209.
[6] J.L. Carter, R.J. Coletti, R.P. Harris, Quantifying and monitoring overdiagnosis in cancer screening: a systematic review of methods, BMJ 350 (2015) g7773.
[7] D. Puliti, M. Zappa, G. Miccinesi, P. Falini, E. Crocetti, E. Paci, An estimate of overdiagnosis 15 years after the start of mammographic screening in Florence, Eur. J. Cancer 45 (2009) 3166–3171.
[8] S.W.T.L. Duffy, A.H. Olsen, B. Vitak, P.C. Allgood, T.H.H. Chen, A.M.F. Yen, R.A. Smith, Absolute numbers of lives saved and overdiagnosis in breast cancer screening, from a randomized trial and from the Breast Screening Programme in England, J. Med. Screen. 17 (2010) 25–30.
[9] A. Coldman, N. Phillips, Incidence of breast cancer and estimates of overdiagnosis after the initiation of a population-based mammography screening program, CMAJ 185 (2013) E492–8.
[10] P.H. Zahl, J. Maehlen, Overdiagnosis of breast cancer after 14 years of mammography screening, Tidsskrift for den Norske laegeforening: tidsskrift for praktisk medicin, ny raekke 132 (2012) 414–417.

[11] M.G. Marmot, D.G. Altman, D.A. Cameron, J.A. Dewar, S.G. Thompson, M. Wilcox, The benefits and harms of breast cancer screening: an independent review, Br. J. Cancer (2013).
[12] R. De Gelder, E.A.M. Heijnsdijk, N.T. Van Ravesteyn, J. Fracheboud, G. Draisma, H.J. De Koning, Interpreting overdiagnosis estimates in population-based mammography screening, Epidemiol. Rev. 33 (2011) 111–121.
[13] National Evaluation Team for Breast cancer screening, National evaluation of breast cancer screening in the Netherlands. 1990–2007 (XII), Twelfth Evaluation Report, Erasmus MC, Rotterdam, 2009.
[14] National Evaluation Team for Breast cancer screening, National Evaluation of Breast Cancer Screening in the Netherlands 1990–2011/2012, Erasmus MC, Rotterdam, 2014.
[15] Integraal kankercentrum Nederland, http://www.iknl.nl (accessed 18.06.2015).
[16] Centraal Bureau voor de Statistiek, http://statline.cbs.nl/statweb/ (accessed 25.04.2015).
[17] T.M. Ripping, A.L. Verbeek, M.J. Broeders, Overdiagnosis in cancer screening: the need for a standardized denominator, J. Med. Screen. (2015), doi:http://dx.doi.org/10.1177/0969141315604864).
[18] C. Biesheuvel, A. Barratt, K. Howard, N. Houssami, L. Irwig, Effects of study methods and biases on estimates of invasive breast cancer overdetection with mammography screening: a systematic review, Lancet Oncol. 8 (2007) 1129–1138.
[19] R. Etzioni, R. Gulati, L. Mallinger, J. Mandelblatt, Influence of study features and methods on overdiagnosis estimates in Breast and prostate cancer screening, Ann. Int. Med. 158 (2013) 831–838.
[20] T.M. Ripping, A.L. Verbeek, D. van der Waal, J.D. Otten, G.J. den Heeten, J. Fracheboud, et al., Immediate and delayed effects of mammographic screening on breast cancer mortality and incidence in birth cohorts, Br. J. Cancer 109 (2013) 2467–2471.
[21] A. Seigneurin, J. Labarere, S.W. Duffy, M. Colonna, Overdiagnosis associated with breast cancer screening: a simulation study to compare lead-time adjustment methods, Cancer Epidemiol. 39 (2015) 1128–1135.
[22] R. De Gelder, J. Fracheboud, E.A.M. Heijnsdijk, G. den Heeten, A.L.M. Verbeek, M.J.M. Broeders, et al., Digital mammography screening: weighing reduced mortality against increased overdiagnosis, Prev. Med. 53 (2011) 134–140.
[23] T.M. Ripping, A.L. Verbeek, J. Fracheboud, H.J. de Koning, N.T. van Ravesteyn, M. J. Broeders, Overdiagnosis by mammographic screening for breast cancer studied in birth cohorts in The Netherlands, Int. J. Cancer 137 (2015) 921–929.
[24] S.J. Otto, J. Fracheboud, C.W.N. Looman, M.J.M. Broeders, R. Boer, J.H.C.L. Hendriks, et al., Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality: a systematic review, Lancet 361 (2003) 1411–1417.
[25] K.R. Beckmann, J.W. Lynch, J.E. Hiller, G. Farshid, N. Houssami, S.W. Duffy, et al., A novel case-control design to estimate the extent of over-diagnosis of breast cancer due to organised population-based mammography screening, Int. J. Cancer 136 (2015) 1411–1421.
[26] R.S. Falk, S. Hofvind, P. Skaane, T. Haldorsen, Overdiagnosis among women attending a population-based mammography screening program, Int. J. Cancer 133 (2013) 705–712.
[27] A.H. Olsen, O.F. Agbaje, J.P. Myles, E. Lynge, S.W. Duffy, Overdiagnosis, sojourn time, and sensitivity in the Copenhagen mammography screening program, Breast J. 12 (2006) 338–342.
[28] E. Paci, G. Miccinesi, D. Puliti, P. Baldazzi, V. De Lisi, F. Falcini, et al., Estimate of overdiagnosis of breast cancer due to mammography after adjustment for lead time. A service screening study in Italy, Breast Cancer Res. 8 (2006) R68.