# Overdiagnosis of invasive breast cancer in population-based breast cancer screening

Ding, Lilu; Poelhekken, Keris; Greuter, Marcel J W; Truyen, Inge; De Schutter, Harlinde; Goossens, Mathijs; Houssami, Nehmat; Van Hal, Guido; de Bock, Geertruida H

*Published in:*
European Journal of Cancer

*DOI:*
10.1016/j.ejca.2022.06.027



*Publication date:*
2022

European Journal of Cancer 173 (2022) 1–9

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejcancer.com

Original Research

# Overdiagnosis of invasive breast cancer in population-based breast cancer screening: A short- and long-term perspective

Lilu Ding [a,b], Keris Poelhekken [a,c], Marcel J.W. Greuter [c,d], Inge Truyen [e], Harlinde De Schutter [e], Mathijs Goossens [f,g], Nehmat Houssami [h,i], Guido Van Hal [b,f], Geertruida H. de Bock [a,*]

[a] *Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands*
[b] *Department of Social Epidemiology and Health Policy, University of Antwerp, Antwerp, Belgium*
[c] *Department of Radiology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands*
[d] *Department of Robotics and Mechatronics, University of Twente, Enschede, the Netherlands*
[e] *Belgian Cancer Registry, Brussels, Belgium*
[f] *Center for Cancer Detection (CvKO) in Flanders, Belgium*
[g] *Vrije Universiteit Brussel, Brussels, Belgium*
[h] *Sydney School of Public Health, Faculty of Medicine and Health, The University of Sydney, New South Wales, Australia*
[i] *The Daffodil Centre, The University of Sydney, a Joint Venture with Cancer Council NSW, Australia*

Received 10 March 2022; received in revised form 30 May 2022; accepted 12 June 2022

**KEYWORDS**
Breast neoplasms;
Mammography;
Mass screening;
Overdiagnosis;
Modelling studies;
Invasive breast cancer

**Abstract** ***Background:*** Overdiagnosis of invasive breast cancer (BC) is a contentious issue.
***Objective:*** The aim of this paper is to estimate the overdiagnosis rate of invasive BC in an organised BC screening program and to evaluate the impact of age and follow-up time.
***Methods:*** The micro-simulation model SiMRiSc was calibrated and validated for BC screening in Flanders, where women are screened biennially from age 50 to 69. Overdiagnosis rate was defined as the number of invasive BC that would not have been diagnosed in the absence of screening per 100,000 screened women during the screening period plus follow-up time (which was set at 5 years and varied from 2 to 15 years). Overdiagnosis rate was calculated overall and stratified by age.
***Results:*** The overall overdiagnosis rate for women screened biennially from 50 to 69 was 20.1 (95%CI: 16.9–23.2) per 100,000 women screened at 5-year follow-up from stopping screening. Overdiagnosis at 5-year follow-up time was 12.9 (95%CI: 4.6–21.1) and 74.2 (95%CI: 50.9

* *Corresponding author*:
*E-mail addresses:* l.ding@umcg.nl (L. Ding), k.poelhekken@umcg.nl (K. Poelhekken), m.j.w.greuter@umcg.nl (M.J.W. Greuter), inge.truyen@kankerregister.org (I. Truyen), Harlinde.DeSchutter@kankerregister.org (H. De Schutter), mathieu.goossens@bevolkingsonderzoek.be (M. Goossens), nehmat.houssami@sydney.edu.au (N. Houssami), guido.vanhal@uantwerpen.be (G. Van Hal), g.h.de.bock@umcg.nl (G.H. de Bock).

https://doi.org/10.1016/j.ejca.2022.06.027

–97.5) per 100,000 women screened for women who started screening at age 50 and 68, respectively. At 2- and 15-year follow-up time, overdiagnosis rate was 98.5 (95%CI: 75.8–121.3) and 13.4 (95%CI: 4.9–21.9), respectively, for women starting at age 50, and 297.0 (95%CI: 264.5–329.4) and 34.2 (95%CI: 17.5–50.8), respectively, for those starting at age 68.

***Conclusions:*** Sufficient follow-up time (≥10 years) after screening stops is key to obtaining unbiased estimates of overdiagnosis. Overdiagnosis of invasive BC is a larger problem in older compared to younger women.



## 1. Introduction

Population-based mammographic screening has been implemented in most high-income countries to reduce breast cancer-specific mortality [1]. While population-based mammographic screening has the potential to achieve up to 40% breast cancer-specific mortality reduction, one of the most concerning harms in breast cancer screening is overdiagnosis [1]. Overdiagnosis of breast cancer refers to the detection of breast cancers that would not have been diagnosed in the absence of screening [2].

Overdiagnosis has gained increasing attention in the past decade, along with the advocacy for informed decision-making based on both the benefit and harms of breast cancer screening [3,4]. Studies on overdiagnosis in general combine invasive breast cancer and ductal carcinoma in situ (DCIS) [5–7]. Although there is agreement that screen-detected DCIS contributes to overdiagnosis [8], the estimation of overdiagnosis of invasive breast cancer remains a contentious issue. The published data on the proportion of overdiagnosed invasive breast cancers range from −0.2% to 54% [2,9,10]. Among the many explanations for this wide range of estimates are as follows: the applied denominator of the proportion of overdiagnosed cancers [11], different tumour growth velocities [12], varying breast cancer background incidence [13] and the follow-up time used in the analysis [11,14]. When screening is initiated, short-term breast cancer incidence will rise since many cancers are found earlier than they would have been without screening [12]. Therefore, to properly estimate overdiagnosis, it is necessary to have a sufficiently long period of follow-up time to compensate for this lead time effect [11,14]. In general, overdiagnosis is reported as one estimate for a whole population [5,6,9,15–20]. This is also the case in studies where overdiagnosis was informed as one estimate to all individual women to help them make informed decisions [3,4]. However, overdiagnosis is likely to be different for women in different age groups given the different breast cancer incidence rate and tumour growth rate [21,22].

Therefore, the aim of this study was to quantify the overdiagnosis of invasive breast cancer for women who were screened biennially from age 50 to 69 in the population-based breast cancer screening program in Flanders and to quantify the influence of age and follow-up time on overdiagnosis rate. In this study, we focused on the estimation of overdiagnosis of invasive breast cancer in a screened population, overall, as well as for women at different ages.

## 2. Methods

### 2.1. Breast cancer screening in Flanders

In Flanders, a population-based breast cancer screening program has been implemented since 2001 [23]. Every 2 years, women aged 50–69 with no history of breast cancer in the last 10 years are invited to screen by mammography unless they actively opt-out [24].

### 2.2. The SiMRiSc model: description, input, outcomes and validation

The micro-simulation model SiMRiSc [25] was applied to the population-based breast cancer screening program in Flanders and has been previously described in detail [25–28]. A virtual cohort of women was created and followed from birth. Every woman was assigned an age of death based on data of the life expectancy of women in Flanders [29,30]. For each woman, the probability of developing invasive breast cancer and the age at clinical diagnosis were derived from the invasive breast cancer incidence rate of women in Flanders [29]. A normal distributed breast cancer incidence risk and a log normal distributed clinical (self-) detection risk were assumed characterised by the distribution (geometric) mean and standard deviation. The tumour growth parameters and the tumour size at clinical self-detection were determined from literature [21,31]. For each woman who had incident breast cancer, an age-dependent tumour volume doubling time was sampled from a log-normal distribution and applied in an exponential tumour growth model to determine the individual tumour growth history [26]. During the screening period, developing invasive breast cancers

could be detected by mammography. The model used a screening sensitivity based on tumour size and breast density [32,33], where breast density was modelled as a function of age [34–36]. The risk of tumour induction due to ionising radiation from mammography was also considered based on the relative risk model described in the BEIR7 report [37]. Tumours that became clinically evident by self-detection between screening rounds were assigned as interval cancers. The preclinical period of a tumour was defined as the time the tumour developed from a minimal screen-detectable tumour diameter of 5 mm until it became clinically evident by self-detection [21]. Women with a screen-detected or self-detected invasive breast cancer were removed from the screening and assigned a breast cancer-specific death probability based on tumour diameter at diagnosis, which was based on the relative survival of breast cancer patients in the Belgian Cancer Registry (BCR) [29]. All other women stayed in the screening until the end of the screening or death.

All input parameters were derived from literature and data from the BCR and Statistics Belgium (Table S1), who also have access to reimbursed-based screening data. The incidence and relative survival rate of breast cancer for women of all ages from 2000 to 2017 were provided by the BCR. The all-cause mortality for women of all ages from 2000 to 2017 was provided by Statistics Belgium.

The model was validated by comparison of modelled outcomes to the empirical 95%CI of the observed data from the BCR and the Center for Cancer Screening (CvKO) in the first and second screening round. The outcomes used for comparison include the number of screen-detected and interval breast cancers and the size distribution of the screen-detected breast cancers in the first and second screening round from the year 2015. For model validation, the screening was simulated for women who started biennially screening between 50 and 69 years of age until age 69. The outcomes of the model were calculated for each birth cohort and subsequently calculated for the four age groups 50–54, 55–59, 60–64 and 65–69 years of age and standardised per 1000 screened women.

### 2.3. Quantification of overdiagnosis

Based on the above-noted formulation of the SiMRiSc model, a cohort was modelled from birth to death. To get an overall estimate of overdiagnosis, the simulation was performed for the biennial screening in a cohort for women who were screened biennially from age 50 to 69. As a control, the same cohort was created without screening and followed during lifetime for the incidence of breast cancer or death. In this control cohort, all observed cancers were clinically diagnosed.

Overdiagnosed cancers were defined as invasive cancers detected by screening that would not have presented clinically during the screening period and the follow-up time after screening stops. We calculated the number of overdiagnosed invasive breast cancers by comparing the number of diagnosed invasive cancers in a screened cohort to the number of diagnosed invasive cancers in a control cohort from the start of screening to the end of follow-up. The proportion of overdiagnosed cancers was defined as the number of overdiagnosed breast cancers divided by the number of screen-detected and interval cancers in the screened cohort, in which the interval cancers were defined as the breast cancers diagnosed between the current screening age and the next scheduled screening age in 2 years. Furthermore, as shown by many papers, the proportion of overdiagnosed cancers is strongly influenced by what is used as a denominator [11,38]. Therefore the absolute number of overdiagnosed cancers per 100,000 screened women was also calculated. The follow-up time was included because, in the screened cohort compared to the control cohort, more breast cancers will be diagnosed during screening and less breast cancers will be diagnosed after screening stops.

To evaluate the effect of age on overdiagnosis, simulation was also performed in single birth cohorts in which the screening start age varied by every 2 years from 50 to 69 and screened only once for each birth cohort. To evaluate the effect of prevalence mammography at age 50, the probability of overdiagnosis conditional to having had one, two and three negative mammograms was also simulated.

For all simulations, the follow-up time was varied from 2 to 15 years after the screening. For each cohort, a 100% screening uptake was applied, and 10 iterations were performed. The mean and 95% confidence interval (CI) of the overdiagnosis rate were calculated from the 10 iterations.

### 2.4. Sensitivity analysis

The robustness of the outcomes of the model was evaluated in a univariate sensitivity analysis, where the upper and lower limit of the 95%CI of the input parameters was applied. The impact of low participation was also tested in the sensitivity analysis. Screening participation of 50% was evaluated where the probability of being screened was 50% for a woman in the last 2 years. As the baseline estimate for the sensitivity analysis, the overdiagnosis rate for women screened once from age 50 with 5-year follow-up time was used. For simplicity, one iteration was performed in the sensitivity analysis. The sensitivity analysis results were summarised in a tornado plot.

## 3. Results

### 3.1. Validation of the model

The simulated number of screen-detected breast cancers corresponded well with the observed data, albeit with a slight overestimation in the younger age groups of 50–54 and 55–59 (Table 1). The simulated number of interval breast cancers in the second screening round was also slightly overestimated, whereas the simulated number of interval breast cancers in the first screening round corresponded well with the observed data. The simulated size distribution of the diagnosed breast cancers corresponded well with the observed values in the first and second screening round for all age-groups with a slight underestimation of large-sized cancers (>2 cm) in the second screening round.

### 3.2. Estimation of overdiagnosis

The overall overdiagnosis rate of invasive breast cancer for women screened biennially from 50 to 69 was 20.1 (95%CI: 16.9–23.2) per 100,000 women screened at 5 years of follow-up time, whereas overdiagnosis was 40.5 (95%CI: 36.0–45.0) and 17.8 (95%CI: 15.2–20.4) per 100,000 women screened for 2 and 15 years of follow-up time, respectively (Table 2). The overall proportion of overdiagnosed cancers decreased from 5.4% (95%CI: 4.8%–6.1%) to 2.4% (95%CI: 2.0%–2.8%) for 2 and 15 years of follow-up time, respectively (Table 2).

Overdiagnosis rate at 5 years of follow-up time was 12.9 (95%CI: 4.6–21.1) and 74.2 (95%CI: 50.9–97.5) per 100,000 women screened for women who started screening at age 50 and 68, respectively (Table 3).

For women screened at age 50, overdiagnosis rate at 2 years of follow-up time was 98.5 (95%CI: 75.8–121.3) and decreased to 13.4 (95%CI: 4.9–21.9) per 100,000 women screened at 15 years of follow-up time (Table 3). Similarly, overdiagnosis rates for women at older screening start ages (52, 54, …, 68) all decreased with longer follow-up time. Overdiagnosis rate was higher for older women than younger women (Table 3). Overdiagnosis rate for women diagnosed at age 50 was 98.5 (75.8–121.3) and 13.4 (4.9–21.9) at 2 and 15 years of follow-up, respectively. For women diagnosed at age 56 who had three negative mammograms, overdiagnosis rate decreased to 30.6 (23.4–37.8) and 7.9 (3.5–12.4) at 2 and 15 years of follow-up, respectively (Table 4).

Table 1
Results of the model validation. Comparison between the simulated and observed data.

| | Observed 95%CI[a] (Flanders data) | Simulated |
|---|---|---|
| **Number of screen-detected tumours (per 1000 screenings)** | | |
| The first screening round | | |
| 50–54 | 4.7 (4.0–5.5) | 4.4 |
| 55–59 | 7.3 (4.6–10.1) | 5.6 |
| 60–64 | 9.8 (5.5–14.1) | 6.2 |
| 65–69 | 12.1 (6.5–17.7) | 7.8 |
| The second screening round | | |
| 50–54 | 2.9 (2.3–3.5) | 4.0 |
| 55–59 | 3.8 (3.3–4.3) | 4.5 |
| 60–64 | 5.4 (4.7–6.0) | 5.1 |
| 65–69 | 6.4 (5.7–7.1) | 6.1 |
| **The tumour size distribution of screen-detected breast cancer** | | |
| First screening round | | |
| <1 cm | 24.2% (17.3%–31.0%) | 21.9% |
| 1–2 cm | 43.6% (35.7%–51.6%) | 46.9% |
| >2 cm | 32.2% (24.7%–39.7%) | 31.3% |
| The second screening round | | |
| <1 cm | 28.5% (25.1%–31.8%) | 25.9% |
| 1–2 cm | 48.9% (45.2%–52.6%) | 55.9% |
| >2 cm | 22.7% (19.6%–25.8%) | 18.5% |
| **Number of interval cancers (per 1000 screenings)** | | |
| After the first screening round | 3.5 (2.9–4.1) | 3.1 |
| After the second screening round | 2.6 (2.3–2.8) | 2.9 |

[a] Data source: The Belgian Cancer Registry [29]. Index year: 2015.

### 3.3. Sensitivity analysis

The sensitivity analysis showed that the overdiagnosis rate was most sensitive to the mean tumour volume doubling time and varied between 5.6 and 33.3 per 100,000 screened women when set at the lower and upper 95%CI limit, respectively (Fig. 1,Table S2). However, the tumour volume doubling time of women older than 70 only had a minor effect on overdiagnosis rate. A smaller mean size of self-detected tumours, a lower lifetime risk of developing breast cancer, and 50% uptake were associated with a decreased overdiagnosis rate.

## 4. Discussion

In this study, after the simulation model was successfully calibrated and validated to population-screening in Flanders, we found an overdiagnosis rate of 17.8 invasive breast cancers per 100,000 women screened biennially from age 50 to 68, at a follow-up of 15 years after screening stops. Overdiagnosis was overestimated at 40.5 per 100,000 women screened using insufficient follow-up time of 2 years. Overdiagnosis rate for women who started screening at age 68 was nearly three times higher than for women who started screening at age 50. In addition, for women of different ages at commencement of screening, overdiagnosis decreased with longer follow-up and stabilised at 10-year follow-up. The estimated overdiagnosis rate was most sensitive to changes in tumour volume doubling time, the size of self-detected tumours and lifetime risk of developing breast cancer.

Our model showed lower overdiagnosis rates for invasive breast cancer with longer follow-up time for all

Table 2
Mean and 95%CI of the number of screen-detected and interval cancer, the number of overdiagnosed breast cancer and the overdiagnosis rate of women who were screened biennially from age 50 to 69 with follow-up time after screening stops varied from 2 to 15 years.

| Follow-up time after screening stops at age 69 | Number of overdiagnosed breast cancers | Number of screen-detected breast cancers | Number of interval breast cancers | Number of mammograms performed | Overdiagnosis rate (per 100,000 women biennially screened) | Proportion of overdiagnosed cancers[a] |
|---|---|---|---|---|---|---|
| 2 years | 326 (291–361) | 3657 (3556–3758) | 2323 (2213–2433) | 804,033 (802,351–805,715) | 40.5 (36.0–45.0) | 5.4% (4.8%–6.1%) |
| 3 years | 232 (202–261) | | | | 28.8 (24.9–32.7) | 3.9% (3.3%–4.4%) |
| 4 years | 186 (161–210) | | | | 23.1 (19.9–26.3) | 3.1% (2.7%–3.6%) |
| 5 years | 162 (137–186) | | | | 20.1 (16.9–23.2) | 2.7% (2.2%–3.1%) |
| 10 years | 144 (129–158) | | | | 17.8 (16.0–19.7) | 2.4% (2.1%–2.7%) |
| 15 years | 143 (123–163) | | | | 17.8 (15.2–20.4) | 2.4% (2.0%–2.8%) |

[a] Denominator of the proportion: the number of screen-detected plus the number of interval breast cancer.

birth cohorts. In previously published studies, for a follow-up time of 10 years or more, the published proportion of overdiagnosed invasive breast cancers are generally low, in the range of 1.0%–3.0% in observed data [17,39] and 0.4–4.6% in modelling studies [15,38], which is comparable to our model estimated proportion of overdiagnosed invasive breast cancers of 2.4% at 15 years of follow-up time. For a follow-up time of 5 years, the published proportion of overdiagnosed invasive breast cancers vary between 14.7% and 56% in observed data [16,19], which is higher than our model's estimated proportion of overdiagnosed cancers at 5 years of 2.7%. The large variation of published data at short follow-up time epitomises the impact of the population characteristics and the definition of overdiagnosis; however, the decreased overdiagnosis at longer follow-up time shows the dominant role of the length of follow-up time on the estimation of overdiagnosis in breast cancer screening program.

We found that overdiagnosis was nearly three times higher in women who started screening at age 68 compared to women who started screening at age 50. This observation is new, as overdiagnosis is commonly reported as a point estimate for a whole screened population, independent of age [5,6,9,15–20]. There are some potential explanations for the more pronounced effects of overdiagnosis of invasive breast cancer with a shorter follow-up time in older women. Older women have a higher breast cancer risk and a lower average breast cancer growth rate compared to younger women [21,22]. Because of this higher risk, the incidence is higher among older women and more screen-detected tumours will be found, and because of the lower growth rate, these tumours are less likely to become symptomatic without screening and are more likely to be overdiagnosed. Moreover, older women are more likely to die of competing causes of death such as heart disease, other cancers and external causes [40,41] than younger women. Therefore, compared to younger women, older women are less likely to be diagnosed with breast cancer during follow-up time after screening stops due to higher risk of competing causes of death and are thus more likely to become overdiagnosed with breast cancer. Change in life expectancy was accounted for in the simulations since life expectancy is incorporated in our model.

We also found that overdiagnosis rate decreased with more previously negative mammograms, and was most evident for prevalence mammography. This is in line with a previous publication from the UK [42]. A strength of this study is that we applied and validated an already existing and validated model with input parameters that were independently derived from published sources. The model enabled the estimation of overdiagnosis rate via a per woman comparison of women in a screened and unscreened situation. Our

Table 3
Estimates of the mean and 95%CI of the number of overdiagnosed breast cancers per 100,000 women screened once at different follow-up times after screening stops for women with different screening start age.

| Screen start age | 2 years | 3 years | 4 years | 5 years | 10 years | 15 years |
|---|---|---|---|---|---|---|
| 50 | 98.5 (75.8–121.3) | 44.6 (34.1–55.1) | 21.6 (10.8–32.4) | 12.9 (4.6–21.1) | 12.3 (5.3–23.0) | 13.4 (4.9–21.9) |
| 52 | 107.9 (94.1–121.6) | 47.6 (36.3–59.0) | 20.3 (12.5–28.1) | 12.1 (6.5–17.8) | 12.1 (5.3–19.0) | 12.4 (5.5–19.2) |
| 54 | 121.1 (104.4–137.8) | 53.6 (39.8–67.3) | 24.9 (16.5–33.5) | 15.5 (7.4–23.7) | 12.3 (7.5–17.1) | 13.4 (6.6–20.3) |
| 56 | 138.2 (113.4–162.9) | 60.8 (49.1–72.5) | 28.4 (16.3–40.5) | 16.8 (5.2–28.3) | 13.2 (5.5–20.9) | 16.4 (3.2–29.6) |
| 58 | 139.0 (112.8–165.1) | 58.8 (42.2–75.3) | 27.6 (15.2–39.9) | 16.1 (6.6–25.6) | 12.3 (5.3–19.4) | 14.0 (6.1–22.0s) |
| 60 | 148.8 (131.3–166.4) | 70.1 (54.1–86.1) | 37.4 (23.3–51.6) | 24.2 (13.1–35.2) | 17.1 (9.3–25.0) | 16.6 (9.8–23.5) |
| 62 | 167.6 (144.0–191.3) | 82.3 (64.2–100.5) | 47.4 (34.8–59.9) | 32.9 (18.8–47.0) | 21.8 (12.4–31.2) | 22.5 (14.5–30.5) |
| 64 | 186.3 (157.0–215.7) | 95.4 (73.0–117.8) | 62.9 (48.4–77.3) | 51.0 (36.4–65.6) | 22.8 (9.1–36.4) | 21.8 (11.8–31.8) |
| 66 | 239.1 (196.0–282.2) | 155.9 (126.2–185.6) | 119.2 (98.7–139.8) | 76.0 (56.2–95.9) | 31.9 (16.8–47.0) | 29.0 (19.1–38.9) |
| 68 | 297.0 (264.5–329.4) | 176.7 (144.0–209.3) | 110.2 (80.9–139.6) | 74.2 (50.9–97.5) | 36.7 (20.6–52.9) | 34.2 (17.5–50.8) |

Table 4
Overdiagnosis conditional on having had one, two and three negative mammograms. Numbers are per 100,000 women biennially screened with follow-up time after screening stops varied from 2 to 15 years.

| Follow-up time after screening stops | Overdiagnosed breast cancers per 100,000 women biennially screened conditional to number of prior negative screens | | | |
|---|---|---|---|---|
| | Age 50 | Age 52 after one negative mammogram | Age 54 after two negative mammograms | Age 56 after three negative mammograms |
| 2 years | 98.5 (75.8–121.3) | 50.9 (40.9–61.0) | 37.2 (31.3–43.0) | 30.6 (23.4–37.8) |
| 3 years | 44.6 (34.1–55.1) | 23.6 (17.7–29.4) | 18.4 (14.5–22.3) | 15.8 (12.4–19.2) |
| 4 years | 21.6 (10.8–32.4) | 10.9 (7.1–14.6) | 9.0 (6.8–11.1) | 8.8 (5.1–12.7) |
| 5 years | 12.9 (4.6–21.1) | 10.7 (6.1–15.3) | 11.5 (8.4–14.5) | 10.2 (5.9–14.5) |
| 10 years | 12.3 (5.3–23.0) | 7.8 (4.1–11.6) | 8.8 (5.3–12.2) | 9.1 (6.4–11.8) |
| 15 years | 13.4 (4.9–21.9) | 7.6 (4.2–10.9) | 7.4 (3.5–11.3) | 7.9 (3.5–12.4) |

estimates can give quantified evidence of overdiagnosis related to the detection of invasive breast cancers by screening and can also quantify the impact of the length of follow-up time on the magnitude of overdiagnosis related to invasive breast cancers. Moreover, our estimated overdiagnosis rates were robust for most model input parameters.

This study has some limitations. First, the number of screen-detected breast cancers was underestimated in older age groups in the first screening round. This can be caused by an underestimation of the tumour doubling time in older women. The validation results also showed that the simulated number of screen-detected breast cancers was overestimated for younger age groups in the second screening round. This overestimation can be caused by overestimation of tumour doubling time in young women or an overestimation of tumour size at symptoms in this young age-group, or both. The slight underestimation of large-sized cancers in the second screening round could be related to our age dependent tumour volume doubling time model. An underestimation of the variance of cancer growth might cause an underestimation of large-sized cancers. Second, our estimate of overdiagnosis was most sensitive to a change in tumour volume doubling time in women aged 50–70 years old. In our model the tumour growth was modelled as an exponential growth with a log-normal distribution around a mean growth rate per age group. Although the growth characteristics resemble the observed age dependent growth rate of breast cancer [21], only one mean growth rate was used per age group. Extension of our growth model with distributions around slow, medium and fast growing tumours might therefore yield a more accurate estimation of overdiagnosis in breast cancer screening. Third, we assumed a 100% uptake of screening for our estimation, which is less likely to happen in population screening programs. This is because we aim to quantify overdiagnosis of invasive breast cancer from the perspective of women who will participate in the screening. The overdiagnosis rate with lower screening uptake is expected to be lower than our estimates, which has been verified in the sensitivity analysis. Due to these limitations, point estimates of overdiagnosis rate should be interpreted as approximations because of the inherent uncertainties of the microsimulation approach. Our results are particularly useful to test multiple scenarios that otherwise would be impossible to test in an observational study.

Overdiagnosis is recognised as the most serious harmful effect of screening [1,12]. It leads to unnecessary physical and mental burden and potential overtreatment of women who would not have been diagnosed with breast cancer in the absence of screening [1,12]. Since overdiagnosis is caused by the detection of cancer that would not have been diagnosed if not screened, the surge of detected cancers in a short term will be largely compensated by long term follow-up time. Therefore, a long follow-up time is needed to accurately estimate overdiagnosis [11,14]. As pointed out in some studies, the extent of overdiagnosis is overestimated if follow-up time is shorter than the maximum lead time [11,43]. Our results verify that the overdiagnosis rate decreases in the first 5 years of follow-up. Furthermore, we found that for women who started screening from all different ages, overdiagnosis rates are overestimated with insufficient follow-up time. In addition, the overdiagnosis rate in all ages stabilised at follow-up longer than 10 years. Therefore, overdiagnosis of invasive breast cancer should be estimated with at least 5 years of follow-up and estimates with 10 years or longer follow-up will be optimal. For women who started screening at older age (60+), a sufficient follow-up is even more important than for younger women because overdiagnosis of invasive breast cancer is a higher problem for older women compared to younger ones.

Future efforts are needed to estimate overdiagnosis caused by DCIS. Although there is general consensus that DCIS is an important cause of overdiagnosis, accurate estimation of overdiagnosis from DCIS is difficult, mainly because of the lack of definitive evidence about the probability of progression to invasive breast cancer - this is likely less than 40% [44]. In addition, the longer lead time of DCIS compared to invasive cancer entails different estimation approaches of the two entities [18]. These considerations led to our focus on invasive breast cancer in this study.

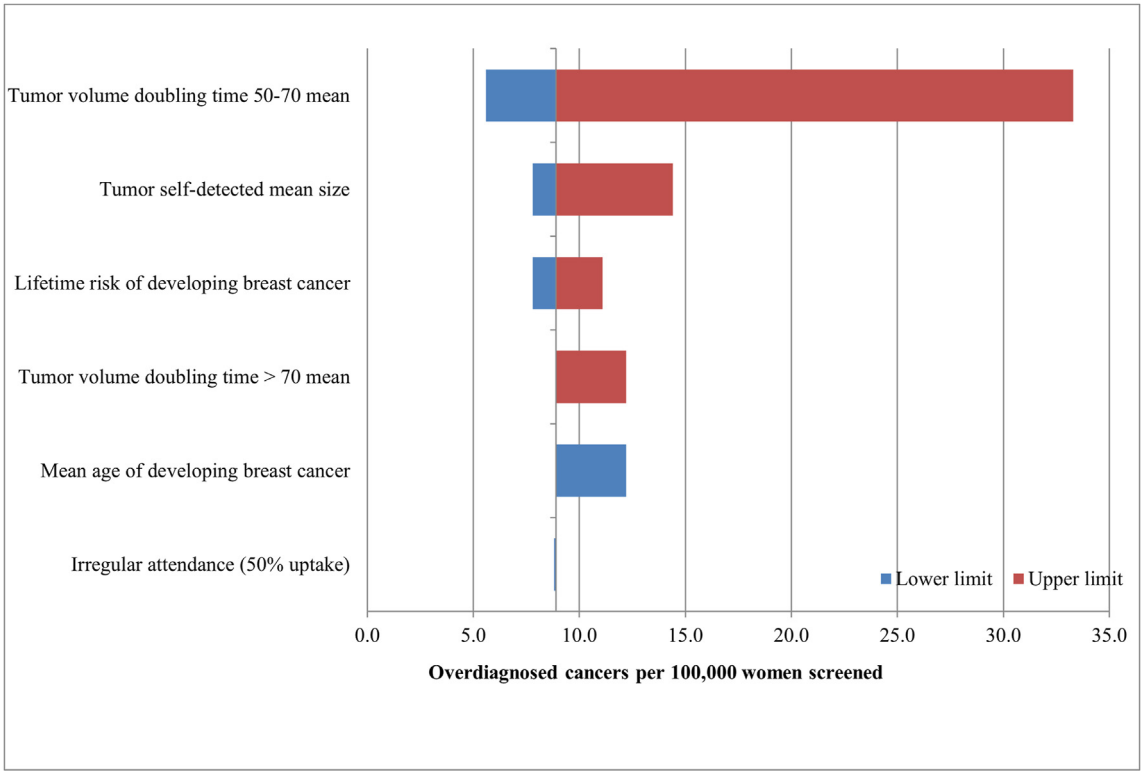


Fig. 1. Sensitivity analysis of overdiagnosis parameters set at lower and upper 95%CI limit.

## 5. Conclusion

Overdiagnosis rates from breast cancer screening are accurately estimated if a sufficient follow-up duration (10 years or longer) is used after screening stops. The risk of an overdiagnosed invasive breast cancer is < 1 in 1000 biennially screened women aged 50–69 with a 10-year follow-up time after screening stops. Overdiagnosis of invasive breast cancer is a larger problem for older women compared to younger women. Overdiagnosis decreased with more previously negative mammograms, suggesting that regular biennial screening optimises trade-off between benefit and harms (specifically overdiagnosis), no screening avoids overdiagnosis but removes benefit, whereas screening irregularly maintains the harms but reduces potential benefit.

## Author contributions

Conceptualisation: L. Ding, M.J.W. Greuter, G.H. de Bock; Data curation: I. Truyen, H. De Schutter; Formal analysis: L. Ding, I. Truyen, M.J.W. Greuter; Funding acquisition: N. Houssami; Investigation: I. Truyen, H. De Schutter, M. Goossens, K. Poelhekken; Methodology: L. Ding, M.J.W. Greuter, G.H. de Bock, G. Van Hal; Project administration: H. De Schutter; Resources: I. Truyen, H. De Schutter, M. Goossens; Software: L. Ding, M.J.W. Greuter, I. Truyen; Supervision: G.H. de Bock, G. Van Hal; Validation: I. Truyen, H. De Schutter, M. Goossens, K. Poelhekken, N. Houssami; Visualisation: L. Ding, I. Truyen; Roles/Writing – original draft: L. Ding, M.J.W. Greuter, G.H. de Bock, G. Van Hal; Writing – review & editing: all authors.

## Funding

NH was supported by the National Breast Cancer Foundation (NBCF) Chair in Breast Cancer Prevention grant (#EC-21-001).

## Conflict of interest statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

L. Ding is supported by the China Scholarship Council (Scholarship number: 201808320439) for his PhD study. The Council is not involved in the design, implementation and publication of the study.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejca.2022.06.027.

## References


[1] Lauby-Secretan B, Scoccianti C, Loomis D, et al. Breast-cancer screening — viewpoint of the IARC working group. N Engl J Med 2015;372(24):2353–8. https://doi.org/10.1056/nejmsr1504363.

[2] Puliti D, Duffy SW, Miccinesi G, et al. Overdiagnosis in mammographic screening for breast cancer in Europe: a literature review. J Med Screen 2012;19(SUPPL. 1):42–56. https://doi.org/10.1258/jms.2012.012082.

[3] Hersch J, Barratt A, Jansen J, et al. Use of a decision aid including information on overdetection to support informed choice about breast cancer screening: a randomised controlled trial. Lancet (London, England) 2015;385(9978):1642–52. https://doi.org/10.1016/S0140-6736(15)60123-4.

[4] Roberto A, Colombo C, Candiani G, et al. A dynamic web-based decision aid to improve informed choice in organised breast cancer screening. A pragmatic randomised trial in Italy. Br J Cancer 2020;123(5):714–21. https://doi.org/10.1038/s41416-020-0935-2. 2020 1235.

[5] Falk RS, Hofvind S, Skaane P, Haldorsen T. Overdiagnosis among women attending a population-based mammography screening program. Int J Cancer 2013;133(3):705–12. https://doi.org/10.1002/ijc.28052.

[6] Lynge E, Beau AB, von Euler-Chelpin M, et al. Breast cancer mortality and overdiagnosis after implementation of population-based screening in Denmark. Breast Cancer Res Treat 2020;184(3):891–9. https://doi.org/10.1007/s10549-020-05896-9.

[7] Njor SH, Olsen AH, Blichert-Toft M, Schwartz W, Vejborg I, Lynge E. Overdiagnosis in screening mammography in Denmark: population based cohort study. BMJ 2013;346(7901). https://doi.org/10.1136/bmj.f1064.

[8] van Seijen M, Lips EH, Thompson AM, et al. Ductal carcinoma in situ: to treat or not to treat, that is the question. Br J Cancer 2019;121(4):285–92. https://doi.org/10.1038/s41416-019-0478-6.

[9] Morrell S, Barratt A, Irwig L, Howard K, Biesheuvel C, Armstrong B. Estimates of overdiagnosis of invasive breast cancer associated with screening mammography. Cancer Causes Control 2010;21(2):275–82. https://doi.org/10.1007/s10552-009-9459-z.

[10] Lund E, Nakamura A, Thalabard JC. No overdiagnosis in the Norwegian Breast Cancer Screening Program estimated by combining record linkage and questionnaire information in the Norwegian Women and Cancer study. Eur J Cancer 2018;89:102–12. https://doi.org/10.1016/j.ejca.2017.11.003.

[11] Marmot M, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M. The benefits and harms of breast cancer screening: an independent review. Lancet 2012;380(9855):1778–86. https://doi.org/10.1016/S0140-6736(12)61611-0.

[12] Srivastava S, Koay EJ, Borowsky AD, et al. Cancer overdiagnosis: a biological challenge and clinical dilemma. Nat Rev Cancer 2019;19(6):349–58. https://doi.org/10.1038/s41568-019-0142-8.

[13] Jørgensen KJ, Gøtzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. BMJ 2009;339(7714):206–9. https://doi.org/10.1136/bmj.b2587.

[14] Elmore JG, Fletcher SW. Overdiagnosis in breast cancer screening: time to tackle an underappreciated harm. Ann Intern Med 2012;156(7):536–7. https://doi.org/10.7326/0003-4819-156-7-201204030-00012.

[15] Wu WYY, Törnberg S, Elfström KM, Liu X, Nyström L, Jonsson H. Overdiagnosis in the population-based organized breast cancer screening program estimated by a non-homogeneous multi-state model: a cohort study using individual data with long-term follow-up 11 Medical and Health Sciences 1117 Public Health and Health Se. Breast Cancer Res 2018;20(1). https://doi.org/10.1186/s13058-018-1082-z.

[16] Zahl PH, Strand BH, Mæshlen J. Incidence of breast cancer in Norway and Sweden during introduction of nationwide screening: prospective cohort study. Br Med J 2004;328(7445):921–4. https://doi.org/10.1136/bmj.38044.666157.63.

[17] Michalopoulos D, Duffy SW. Estimation of overdiagnosis using short-term trends and lead time estimates uncontaminated by overdiagnosed cases: results from the Norwegian breast screening programme. J Med Screen 2016;23(4):192–202. https://doi.org/10.1177/0969141315623980.

[18] Kalager M, Adami H-O, Bretthauer M, Tamimi RM. Overdiagnosis of invasive breast cancer due to mammography screening: results from the Norwegian screening program. Ann Intern Med 2012;156(7):491. https://doi.org/10.7326/0003-4819-156-7-201204030-00005.

[19] 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(5):313–23. https://doi.org/10.7326/M16-0270.

[20] Beckmann KR, Lynch JW, Hiller JE, 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 2015;136(6):1411–21. https://doi.org/10.1002/ijc.29124.

[21] Peer PGM, Van Dijck JAAM, Verbeek ALM, Hendriks JHCL, Holland R. Age-dependent growth rate of primary breast cancer. Cancer 1993;71(11):3547–51. https://doi.org/10.1002/1097-0142(19930601)71:11<3547::AID-CNCR2820711114>3.0.CO;2-C.

[22] Muss HB, Busby-Whitehead J. Older women with breast cancer: slow progress, great opportunity, now is the time. J Clin Oncol 2011;29(35):4608–10. https://doi.org/10.1200/JCO.2011.38.6888.

[23] Goossens M, Van Hal G, Van der Burg M, et al. Quantifying independent risk factors for failing to rescreen in a breast cancer screening program in Flanders, Belgium. Prev Med 2014;69:280–6. https://doi.org/10.1016/j.ypmed.2014.10.019.

[24] Goossens M, De Brabander I, De Grève J, et al. Flemish breast cancer screening programme: 15 years of key performance indicators (2002-2016). BMC Cancer 2019;19(1). https://doi.org/10.1186/s12885-019-6230-z.

[25] Greuter MJW, Jansen-van der Weide MC, Jacobi CE, et al. The validation of a simulation model incorporating radiation risk for mammography breast cancer screening in women with a hereditary-increased breast cancer risk. Eur J Cancer 2010;46(3):495–504. https://doi.org/10.1016/j.ejca.2009.10.030.

[26] Koleva-Kolarova RG, Daszczuck AM, de Jonge C, et al. A modelling study to evaluate the costs and effects of lowering the starting age of population breast cancer screening. Maturitas 2018;109:81–8. https://doi.org/10.1016/j.maturitas.2017.12.009.

[27] Du Y, Sidorenkov G, Heuvelmans MA, et al. Cost-effectiveness of lung cancer screening with low-dose computed tomography in heavy smokers: a microsimulation modelling study. Eur J Cancer 2020;135:121–9. https://doi.org/10.1016/j.ejca.2020.05.004.

[28] De Bock GH, Vermeulen KM, Jansen L, et al. Which screening strategy should be offered to women with BRCA1 or BRCA2 mutations? A simulation of comparative cost-effectiveness. Br J Cancer 2013;108(8):1579–86. https://doi.org/10.1038/bjc.2013.149.

[29] Belgian cancer Registry - welcome to the Belgian cancer Registry. http://kankerregister.org/. [Accessed 25 November 2020].

[30] Statbel. The Belgian statistical office | Statbel. https://statbel.fgov.be/en. [Accessed 22 February 2022].

[31] Xiii N. National evaluation of breast cancer screening in the national evaluation of breast cancer screening in The Netherlands national evaluation team for breast cancer screening. 2014. https://www.lrcb.nl/resources/uploads/2017/02/LETB-XIII-English.pdf. [Accessed 25 November 2020].

[32] Jeffreys M, Harvey J, Highnam R. Comparing a new volumetric breast density method (Volpara$^{TM}$) to Cumulus. In: Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics), vol. 6136. LNCS; 2010. p. 408–13. https://doi.org/10.1007/978-3-642-13666-5_55.

[33] Isheden G, Humphreys K. Modelling breast cancer tumour growth for a stable disease population. Stat Methods Med Res 2019;28(3):681–702. https://doi.org/10.1177/0962280217734583.

[34] Lehman Constance D, Emily W, Susan P, Mariann JD, Urban N. Effect of age and breast density on screening mammograms with false-positive findings. AJR Am J Roentgenol 1999;173(6): 1651–5. https://doi.org/10.2214/AJR.173.6.10584815.

[35] Kerlikowske K, Grady D, Barclay J, Sickles EA, Ernster V. Effect of age, breast density, and family history on the sensitivity of first screening mammography. JAMA 1996;276(1):33–8. https://doi.org/10.1001/JAMA.1996.03540010035027.

[36] Cm C, Je C, Fr S, J L, H T. The relationship of mammographic density and age: implications for breast cancer screening. AJR Am J Roentgenol 2012;198(3). https://doi.org/10.2214/AJR.10.6049.

[37] Council NR. Health risks from exposure to low levels of ionizing radiation: BEIR VII phase 2. National Academies Press; 2006. https://doi.org/10.17226/11340.

[38] De Gelder R, Heijnsdijk EAM, Van Ravesteyn NT, Fracheboud J, Draisma G, De Koning HJ. Interpreting overdiagnosis estimates in population-based mammography screening. Epidemiol Rev 2011; 33(1):111–21. https://doi.org/10.1093/epirev/mxr009.

[39] Puliti D, Zappa M, Miccinesi G, Falini P, Crocetti E, Paci E. An estimate of overdiagnosis 15 years after the start of mammographic screening in Florence. Eur J Cancer 2009;45(18):3166–71. https://doi.org/10.1016/J.EJCA.2009.06.014.

[40] Riihimäki M, Thomsen H, Brandt A, Sundquist J, Hemminki K. Death causes in breast cancer patients. Ann Oncol 2012;23(3): 604–10. https://doi.org/10.1093/ANNONC/MDR160.

[41] Afifi AM, Saad AM, Al-Husseini MJ, Elmehrath AO, Northfelt DW, Sonbol MB. Causes of death after breast cancer diagnosis: a US population-based analysis. Cancer 2020;126(7): 1559–67. https://doi.org/10.1002/CNCR.32648.

[42] Duffy SW, Vulkan D, Cuckle H, et al. Effect of mammographic screening from age 40 years on breast cancer mortality (UK Age trial): final results of a randomised, controlled trial. Lancet Oncol 2020;21(9):1165–72. https://doi.org/10.1016/S1470-2045(20)30398-3/ATTACHMENT/78B7104D-3A69-4063-9858-BDF6B0C6889A/MMC1.

[43] Neal CH, Helvie MA. Overdiagnosis and risks of breast cancer screening. Radiol Clin North Am 2021;59(1):19–27. https://doi.org/10.1016/J.RCL.2020.09.005.

[44] Cowell CF, Weigelt B, Sakr RA, et al. Progression from ductal carcinoma in situ to invasive breast cancer: revisited. Mol Oncol 2013;7(5):859–69. https://doi.org/10.1016/J.MOLONC.2013.07.005.