Method of primary breast cancer detection and the disease-free interval, adjusting for lead time

Linda de Munck, Ph.D.[1*], Anouk H Eijkelboom, MSc.[1,2*], Johannes DM Otten, Ph.D.[3], Mireille JM Broeders, Ph.D.[3,4], Sabine Siesling, Ph.D.[1,2]

1. Department of Research and Development, Netherlands Comprehensive Cancer Organisation (IKNL), Utrecht, the Netherlands
2. Department of Health Technology and Services Research, Technical Medical Centre, University of Twente, Enschede, the Netherlands
3. Department for Health Evidence, Radboud university medical center, Nijmegen, The Netherlands
4. Dutch Expert Centre for Screening, Nijmegen, The Netherlands

* L. de Munck and A.H. Eijkelboom have contributed equally to this work and share first authorship

Corresponding author

Sabine Siesling

Department of Health Technology and Services Research

University of Twente

Drienerlolaan 5, 7522 NB, Enschede

the Netherlands

s.siesling@utwente.nl

+31 6 13 19 38 06

Abstract

Background. Little is known about the impact of screen-detected breast cancer compared with clinically-detected breast cancer on the disease-free interval (i.e. free from locoregional recurrences, distant metastasis, contralateral breast cancer). Moreover, it is thought that most studies overestimate the beneficial effect of screening, as they do not adjust for lead time. We investigated the association between method of breast cancer detection and disease-free interval, taking lead time into account.

Method. Women, 50-76 years old, diagnosed with breast cancer between 2005-2008 were selected from the Netherlands Cancer Registry. Women diagnosed in 2005 were divided into screen-detected and clinically-detected cancer and had a follow-up of ten-years (2005 cohort). Women diagnosed in 2006-2008 were divided into screen-detected, interval, and non-screen-related cancer, and had a follow-up of five years (2006-2008 cohort). A previously published method was used to adjust for lead time. Analyses were repeated correcting for confounding variables instead of lead time.

Results. The 2005 cohort included 6,215 women. Women with screen-detected cancer had an improved disease-free interval compared to women with clinically-detected cancer (HR: 0.77, 95% CI: 0.68 to 0.87). The 2006-2008 cohort included 15,176 women. Women with screen-detected or interval cancer had an improved disease-free interval compared to women with non-screen-related cancer (HR: 0.76, 95% CI: 0.66 to 0.88; HR: 0.88, 95% CI: 0.78 to 0.99, respectively). Correcting for confounders instead of lead-time did not change associations.

Conclusion. Women with screen-detected cancer had an improved disease-free interval compared to women with a non-screen-related or clinically-detected cancer, after correction for lead time.

Introduction

Previous studies have shown that women with screen-detected breast cancer have a longer overall as well as disease-free survival compared to women with clinically-detected breast cancer [1-5]. However, these studies are susceptible for lead time and length time bias, possibly leading to an overestimation of the beneficial effect of screening. It has been argued that this artificial overestimation is the main reason for favourable survival and therefore detection method of the tumour is currently not used to estimate the risk of cancer recurrence or overall survival [6, 7].

Screening studies are susceptible for two types of biases. First, length time is introduced because more slowly growing tumours have a longer pre-symptomatic screen-detectable phase and are therefore more likely to be screen-detected. Overdiagnosis, the most extreme form of length time bias, means that the patient is diagnosed with a tumour which would not have been diagnosed in the absence of screening. Second, lead time can be defined as the time between the date of detection of a screen-detected cancer and the date it would have been diagnosed without screening. A method to correct for lead time has been described by Duffy et al. [8]. Studies using this method, hereafter referred to as 'Duffy method', found an improved breast cancer-specific survival for patients with a screen-detected cancer compared to patients with an interval breast cancer (a tumour diagnosed after a negative screening result) [9-11]. Another study, adjusting for confounders such as tumour stage, subtype and grade, and adjusting for lead time as well, found no difference in survival between patients with a screen-detected or clinically-detected cancer [12].

Tumour characteristics are usually used as confounding variables in analyses concerning the method of detection and overall and disease-free survival. One might argue that correcting for patient and tumour characteristics can be seen as a (although not perfect) proxy for correcting for lead time. It would be interesting to perform analyses correcting for lead time and correcting for confounders separately, using the same study population, and to compare the results.

The few studies taking lead time into account only investigated the association between the method of detection and (breast cancer-specific) survival. However, as the survival of patients with breast cancer is improving, more patients are at risk of developing recurrent disease. Therefore, more knowledge on recurrent disease is desired. The disease-free interval (i.e. the period of time between the primary tumour and recurrent disease) would be a suitable endpoint in studies into recurrent diseases.

This study aimed to investigate the association between method of breast cancer detection and the disease-free interval, taking lead time into account. Analyses were repeated correcting for confounding

variables (patient and tumour characteristics) only. We hypothesized that screen-detected cancers would have an improved disease-free interval compared to clinically-detected cancers after correction for lead time, as well as after correction for confounding.

Methods

Study population & data collection

Patients were selected from the Netherlands Cancer Registry (NCR). The NCR is a nationwide population-based cancer registry, hosted by the Netherlands Comprehensive Cancer Organisation (IKNL), that includes almost all newly diagnosed cancer patients since 1989. Notifications of newly diagnosed tumours are obtained from the nationwide network and registry of histo- and cytopathology in the Netherlands (PALGA). Trained data managers subsequently register information on patient, tumour and treatment characteristics. For breast cancer patients diagnosed in 2005, additional data on locoregional recurrences or distant metastases diagnosed within 10 years following primary tumour diagnosis were available, as these data were obtained from patient files during previous projects. For patients diagnosed in 2006-2008 data on locoregional recurrences or distant metastases occurring within 5 years following diagnosis were available from previous projects. Data on vital status were obtained by linkage to the Municipal Personal Records database. The study was approved by the Privacy Review Board of the NCR. According to the Dutch Central Committee on Research involving Human Subjects, no ethical approval is needed for this study, as it is a retrospective study, which uses data from the NCR.

Women aged between the 50 and 76 years, diagnosed with invasive non-metastatic breast cancer between January 1st, 2005 and December 31st, 2008, and surgically treated in a Dutch hospital were selected (Figure 1). Patients were ineligible if they had been diagnosed with a previous malignant tumour (breast or other localization) in the past five years or with a synchronous breast tumour (diagnosed within 30 days of each other), developed a locoregional recurrence, distant metastasis, or contralateral breast tumour within 30 days after diagnosis, died within 30 days after diagnosis, or had a macroscopic residue after surgery or a microscopic residue without adjuvant treatment.

Patients were divided into two cohorts based on data availability. The 2005 cohort contains data of patients diagnosed between January 1st, 2005 and December 31st, 2005. The 2006-2008 cohort contains data of patients diagnosed between January 1st, 2006 and December 31st, 2008. Data of patients in this latter cohort had been linked previously on an individual level to the data of the Netherlands Breast

Cancer Screening Programme so data on method of detection was available. This population-based programme has been operational since 1990, initially inviting women aged 49–69 years for a biennial screening examination. From 1998 onwards the age range expanded to 74 years [13]. Permission for use of the data was requested from women when they attended screening. This was based on an opt-out option, which was used by 0.02% of all women screened [14].

Definitions

Women diagnosed in 2005 were divided into two detection groups: screen-detected versus clinically-detected cancer. A screen-detected cancer was defined as a tumour diagnosed within 24 months after a positive screening result. A clinically-detected cancer was defined as a tumour not detected by the screening programme. For women diagnosed between 2006-2008 a more detailed definition of method of detection was derived: screen-detected, interval, and non-screen-related cancer. An interval cancer was defined as a tumour diagnosed within 24 months after a negative screening result. A non-screen-related cancer was defined as a tumour detected more than 24 months after the last screening (i.e., not recently screened) or in a woman who had never attended screening.

The primary outcome of the current study was disease-free interval, defined as the period of time between diagnosis of the primary tumour and diagnosis of the recurrent disease (any locoregional recurrence, distant metastasis, or contralateral invasive breast cancer) or the end of follow-up. A locoregional recurrence was defined as the reappearance of cancer in the ipsilateral breast, chest wall or axillary or supraclavicular lymph nodes. A distant metastasis was defined as the reappearance of breast cancer at a location other than the breast or regional axillary nodes [15]. Patients were censored if they died or at the last date of observation.

A patient's socioeconomic status (SES) was based on scores assigned to the four numbers of each patient's postal code at the time of diagnosis. These scores, based on mean household income, percentage of inhabitants with a low income, percentage of low educatedness and percentage of unemployment, were provided by the Netherlands Institute for Social Research at an aggregated level [16]. Subsequently, these scores were categorized as high, intermediate and low SES. Age was categorized into: <60, 60-69, ≥70 years. Tumours were divided into four subtypes based on the oestrogen receptor (ER), progesterone receptor (PR) and Herceptin receptor (HER2) status: 1) ER+ and/or PR+, and HER2-; 2) ER+ and/or PR+, and HER2+; 3) ER- and PR-, and HER2- and 4) ER- and PR-, and HER2+. Tumour

size and TNM-stage were based on pathological assessment, or on clinical assessment if the patient received neo-adjuvant therapy [17].

Statistical analysis

Descriptive statistics were used to summarize baseline characteristics of the total population and the separate subgroups. Chi-square and Kruskal Wallis tests were used to compare patient, tumour and treatment characteristics.

The crude risk of developing recurrent disease was estimated using the Kaplan-Meier method, and detection-groups were compared using the log-rank test. Cox proportional hazard regression models were used to estimate hazard ratio's (HRs) and 95% confidence intervals (CI) for the association between method of detection and disease-free interval. Women with a clinically-detected cancer (i.e. interval and non-screen-related cancers combined, 2005 cohort) or non-screen-related cancer (2006-2008 cohort) were used as the reference categories. The Duffy method [8] was used to correct for lead time, using a mean sojourn time of 4.3 years to represent the preclinical screen-detectable period (Supplementary Methods) [18]. For the 2005 cohort, this resulted in a corrected follow-up time of $(10-((1-e^{-(1/4.3*10)})/(1/4.3))=)$ 6.1 years for all women with a screen-detected cancer and no recurrent disease after 10-years of follow-up. In the 2006-2008 cohort, the corrected follow-up time of women with a screen-detected cancer and no recurrent disease after 5-years of follow-up was $(5-((1-e^{-(1/4.3*5)})/(1/4.3))=)$ 2.0 years. The corrected disease-free interval was used in the Cox models to adjusted for lead time. All analyses were adjusted for age at diagnosis (continuous).

Additional analyses were performed correcting for confounders (patient and tumour characteristics) only. The following confounders were added to the Cox regression model: age, SES, histology, tumour grade, multifocality, tumour stage, and subtype. Treatment was not added as a confounder to prevent overcorrection, as the type of treatment is mainly based on patient and tumour characteristics. As these analyses were not corrected for lead time, the maximum follow-up was used. To account for missing values, multiple imputation by chained equations was used to impute these values [19]. In 2005, a relatively large percentage of patients had missing values for tumour subtype, because HER2 status had only been routinely collected in the NCR since 2006. Other missing values were related to missing information in the patients' files, and were considered as missing at random. Covariates included in the baseline table were used for imputation. Data was imputed 25 times. Rubin's rule was subsequently

used to pool the estimates and standard errors [20]. The validity of the imputed data was checked by comparing the values of the complete cases with the imputed values. The imputed data was used for the analyses. Complete cases analyses were performed for comparison. Scaled Schoenfeld residuals plots were used to test the proportionality assumptions [21].

Sensitivity analyses were performed to investigate whether results were similar for women diagnosed with a screen-detected or interval cancer at/after an initial or subsequent screening examination.

A two sided p-value <0.05 was considered statistically significant. All data was analysed with STATA version 17 software.

Results

A total of 6,215 women were diagnosed in 2005, of whom 3,467 (55.8%) had a screen-detected and 2,748 (44.2%) a clinically-detected cancer. All baseline characteristics, except multifocality, differed significantly between the detection-groups (Table 1). A total of 15,176 women were diagnosed in 2006-2008, of whom 8,487 (55.9%) had a screen-detected, 3,536 (23.3%) an interval, and 3,153 (20.8%) a non-screen-related cancer. All baseline characteristics differed significantly between the detection-groups (Table 2).

2005 cohort

During the 10-year follow-up period, 84.5% of the women with a screen-detected cancer had a 10-year disease-free interval , compared to 75.1% of the women with a clinically-detected cancer (Figure 2A). After correcting for lead time 84.7% of the women with a screen-detected cancer had a 6.1-year disease-free interval. To compare, 80.6% of women with a clinically-detected cancer had a 6.1-year disease-free interval. The unadjusted analysis showed that women with a screen-detected cancer had an improved 10-year disease-free interval compared to women with a clinically-detected cancer (HR: 0.56, 95% CI: 0.50 to 0.63) (Table 3). This effect remained present after correcting for lead time (HR: 0.77, 95% CI: 0.68 to 0.87) or confounders (HR: 0.72, 95% CI: 0.64 to 0.81) (Table 3, Supplementary Table 1). After adjusting for lead time or confounders, both women with a screen-detected cancer detected at an initial or subsequent screen had an improved disease-free interval compared to women with a clinically-detected tumour (Supplementary Table 2).

2006-2008 cohort

During the 5-year follow-up period, 91.9% of the women with a screen-detected cancer had a 5-year disease-free interval, compared to 85.5% of the women with an interval cancer and 83.8% of the women with a non-screen-related cancer (Figure 2B). After correcting for lead time 94.3% of the women with a screen-detected cancer had a 2-year disease-free interval. To compare, 93.6% of the women with an interval cancer had a 2-year disease-free interval, and 92.8% of the women with a non-screen-related cancer. The unadjusted analyses showed that women with a screen-detected or an interval cancer had an improved 5-year disease-free interval compared to women with a non-screen-related cancer (HR: 0.46, 95% CI: 0.41 to 0.52; HR: 0.88, 95% CI: 0.78 to 0.99, respectively) (Table 3). After correcting for lead time, women with a screen-detected cancer still had an improved 5-year disease-free interval (HR: 0.76, 95% CI: 0.66 to 0.88). After correcting for confounders, and not for lead time, women with a screen-detected or interval cancer both had an improved 5-year disease-free interval compared to women with a non-screen-related cancer (HR: 0.65, 95% CI: 0.57 to 0.73; HR: 0.83, 95% CI: 0.74 to 0.94, respectively) (Table 3, Supplementary Table 2). After adjusting for lead time women with a screen-detected cancer detected at subsequent screens had an improved disease-free interval compared to women with a non-screen-related tumour. A similar trend was observed for women with a screen-detected cancer at an initial screen (Supplementary Table 2). After adjusting for confounders both women with a screen-detected cancer detected at initial or subsequent screens had an improved disease-free interval compared to women with a non-screen-related tumour, as had women with an interval cancer detected after a subsequent screen. A similar trend was observed for women with an interval cancer detected at an initial screen.

Discussion

This study showed that women with a screen-detected cancer and ten years of follow-up had an improved disease-free interval compared to women with a clinically-detected cancer, taking lead time into account. Moreover, women with a screen-detected cancer or an interval cancer and five years of follow-up both had an improved disease-free interval compared to women with a non-screen-related cancer. Our results suggest a positive effect of the screening programme.

Tumour characteristics are well known and usually used as confounding factors, as a proxy for correcting for lead time. We showed that after correcting for confounders only, the disease-free interval of

the different method of detection groups was of similar order of magnitude as the disease-free interval corrected for lead time, in both cohorts. Thereby, our results add more insight to the results of studies that correct for tumour characteristics only. When a reliable estimate of sojourn time is available for a population, we prefer the method correcting for lead time. If this estimate is not available, correcting for confounders (without lead time) can be a good alternative.

In our study, with disease-free interval as outcome, we adjusted for lead time using the Duffy method [8]. This has not been described before. Other studies described an advantage in breast cancer-specific survival after lead time correction, which are in accordance with the advantage in disease-free interval described in this study. A British retrospective study, including about 27,000 patients, corrected for lead-time with a mean sojourn time of 4 years. They found that screen-detected cancer had an increased breast cancer-specific survival (HR: 0.40, 95% CI: 0.37 to 0.44) compared to clinically-detected cancer [9]. A French study from the region Gironde also showed a significant increased net survival for screen-detected cancer compared to clinically-detected cancer (93.0% vs 83.8%, respectively) after adjustment for lead time [22]. Furthermore, we compared our results on disease-free interval correcting for confounders (patient and tumour characteristics) only, with other studies correcting for confounders. Results were similar, for the uncorrected and corrected comparison of screen-detected with clinically-detected cancers [2, 5]. Our results support previous studies that showed that method of detection is an independent prognostic factor [23, 24]. As the results of both methods used in our study are in accordance with previously published other studies, this might be a first indication that our lead time corrected results regarding disease-free interval might be generalisable to other populations. Other studies on method of detection and disease-free interval, correcting for lead time, are needed to support our findings.

Analysing screen-detected cancers diagnosed at initial or at subsequent screening examinations separately, showed that screening seemed to improve the disease-free interval in both situations compared to non-screen-related cancers. This suggests that the cancers detected at initial screens (which often have worse characteristics compared to cancers detected at subsequent screens [14, 25]) still have a better prognosis than non-screen-related cancers. Interval cancers diagnosed after initial or subsequent screening examinations both seemed to have an improved disease-free interval compared to non-screen-related cancers. Even though women are relatively younger at their initial screen compared to at their subsequent screen the positive effect of the screening programme seems to be present in both first and subsequent screens. A possible explanation for the improved disease-free interval of women with an

interval cancer compared to those with a non-screen-related cancers is that women who agree to participate in screening might be more conscious about changes in the breast compared to women who do not participate.

In our study we defined patients with non-screen-related cancer as a reference group. Women with a screen-detected cancer could also be compared to women with an interval cancer, avoiding self-selection bias regarding attending the screening programme. Women with an interval cancer had a significant worse 5-year disease-free interval compared to screen-detected cancers (results not shown), suggesting a positive effect of the screening programme. So far, we found no other studies on the effect of method of detection and disease-free interval. However, the above mentioned British study also found that screen-detected cancer had an increased breast cancer-specific survival (HR: 0.53, 95% CI: 0.49 to 0.59) compared to interval cancer, which supports our results [9].

Previous studies showed an improved breast cancer-specific [26] or overall survival [27, 28] for patients with a longer time period between breast cancer diagnosis and recurrent disease, or no recurrent disease at all, compared to patients with early recurrence. Although we did not study the association between method of detection and (breast cancer-specific) survival, the improved disease-free interval for screen-detected cancers might suggest improved (breast-cancer specific) survival.

We assumed an average sojourn time of 4.3 years over all breast cancers, which was specific for the population in the region Nijmegen of the Netherlands and a biennial screening program [18]. The lead time corrections were based on this average sojourn time. The sojourn time of 4.3 years is the top estimate we found in literature. This relatively high estimate was used, so that any protective effects of screening found would not be due to using a too short sojourn time. When interpreting our results, one should take into account that a screening program with less frequent screening or a different age range included for screening would have a different sojourn time. In addition, the same mean sojourn time was used for every patient, while it is very likely that sojourn time differs per patient.

Lead time and length time bias are a concern when comparing survival between different methods of detection groups. Lead time bias has been shown to be decreased by adjusting for tumour size and lymph node involvement [29], while histology and tumour grade have been shown to decrease the effect of length time bias [2, 30]. Therefore, we performed our analyses correcting for lead time and correcting for confounders separately. Unfortunately, both correction methods used in our study, are not the desired gold standard. We realize that our method using lead time correction can be improved when more

information becomes available on the length and distribution of sojourn time for individual patients. On the other hand, after correction for confounders, residual confounding might remain.

As treatment options have improved since the introduction of the screening programme, the impact of screening on mortality in a recent era remains a point of discussion. A recent meta-analysis showed that most trials and studies found no gain in all-cause mortality due to screening [31]. This might suggest that improvements in treatment may have reduced the impact of screening on mortality. Furthermore, all-cause mortality includes a vast amount of causes of death on which screening has no effect [32] and others have considered it as a misleading endpoint [33, 34]. Our study, set in the Dutch screening situation, showed an improved disease-specific interval, corrected for lead time, suggesting that early detection might remain beneficial.

Screening has a strong interaction with early treatment. It should be acknowledged that the benefits of screening described in this study are partly explained by the benefits of early treatment. However, without screening this early treatment would probably not have been given, and hence there might not have been an improved disease-free interval. An modelling study performed in the United States estimated that, compared to a situation with no screening and no treatment, 37% of the reduction in breast cancer mortality was associated with screening, and 63% with treatment in 2012 [35].

Strengths of this study were its nationwide and population-based design, large sample size, and the availability of data on method of detection (due to linkage with the screening programme). Also, we were able to study the effect of method of detection on disease-free interval in two different cohorts, i.e, with 5-years or 10-years follow-up available. Furthermore, we were able to correct for lead time, but we could also use well known tumour characteristics to correct for confounders. Comparing the results of the two methods can give additional insight in the size of the risk estimates. Limitations of this study are that the exact sojourn time is not known at the individual level, and the average sojourn time of 4.3 years was thought to be the most suitable time according to a previous Dutch study [18]. Finally, length time bias and overdiagnosis might still affect our results.

To conclude, women with a screen-detected cancer had an improved disease-free interval compared to women with a clinically-detected cancer, taking lead time into account. More detailed data on method of detection showed that disease-free interval was also improved for screen-detected compared to non-screen-related cancer. Women with an interval cancer also had an improved disease-free interval compared to non-screen-related cancer, though this was less pronounced. Correcting for confounders led

to results of a similar order of magnitude as correction for lead time. The results of this study suggest that patients with screen-detected breast cancer might have a better prognosis.

Data Availability Statement

The data underlying this article cannot be made publicly available because of the privacy of the individuals included in this study. All data collected for the study, including a data dictionary defining each field in the set, will be made available via the NCR (https://iknl.nl/en/ncr/apply-for-data) upon request and after approval of a proposal from the date of publication. The plan for the statistical analysis will be made available by the corresponding author upon request.

Funding

Not applicable

Conflicts of Interest

S. Siesling receives support from the Beirat Krebsregister Robert Koch Insititut, participates on the Evidencio Advisory board and the Vivica Health and Lifestyle, and is advisor of the Netherlands Epidemiology Society. All other authors have no conflict of interest to declare.

Acknowledgements

The authors thank the registration team of the Netherlands Comprehensive Cancer Organisation (IKNL) for the collection of data for the Netherlands Cancer Registry, and the Netherlands Breast Cancer Screening Organisation for providing screening-related data. The role of the funder is not applicable, as there was no funding for this study. The results of this study have been presented as a poster presentation at the European Breast Cancer Conference in November 2022 and as an oral presentation at ISPOR in November 2023.

Dr Stephen Duffy, (Wolfson Institute for Preventive Medicine, QMUL, London) and Laura Woods (London School of Hygiene and Tropical Medicine, London) provided assistance for the correction of lead time.

References

1. Domingo L, Blanch J, Servitja S*, et al.* Aggressiveness features and outcomes of true interval cancers: comparison between screen-detected and symptom-detected cancers. *European journal of cancer prevention.* 2013;22(1):21-28.
2. Joensuu H, Lehtimäki T, Holli K*, et al.* Risk for distant recurrence of breast cancer detected by mammography screening or other methods. *Jama.* 2004;292(9):1064-1073.
3. Sihto H, Lundin J, Lehtimäki T*, et al.* Molecular subtypes of breast cancers detected in mammography screening and outside of screening. *Clinical Cancer Research.* 2008;14(13):4103-4110.
4. Rayson D, Payne JI, Abdolell M*, et al.* Comparison of clinical-pathologic characteristics and outcomes of true interval and screen-detected invasive breast cancer among participants of a Canadian breast screening program: a nested case-control study. *Clinical breast cancer.* 2011;11(1):27-32.
5. Dong W, Berry DA, Bevers TB*, et al.* Prognostic role of detection method and its relationship with tumor biomarkers in breast cancer: the University of Texas MD Anderson Cancer Center experience. *Cancer Epidemiology and Prevention Biomarkers.* 2008;17(5):1096-1103.
6. Wishart GC, Azzato EM, Greenberg DC*, et al.* PREDICT: a new UK prognostic model that predicts survival following surgery for invasive breast cancer. *Breast Cancer Res.* 2010;12(1):R1.
7. Völkel V, Hueting TA, Draeger T*, et al.* Improved risk estimation of locoregional recurrence, secondary contralateral tumors and distant metastases in early breast cancer: the INFLUENCE 2.0 model. *Breast cancer research and treatment.* 2021;189:817-826.
8. Duffy SW, Nagtegaal ID, Wallis M*, et al.* Correcting for lead time and length bias in estimating the effect of screen detection on cancer survival. *American journal of epidemiology.* 2008;168(1):98-104.
9. Lawrence G, Wallis M, Allgood P*, et al.* Population estimates of survival in women with screen-detected and symptomatic breast cancer taking account of lead time and length bias. *Breast cancer research and treatment.* 2009;116(1):179-185.
10. Morris M, Woods LM, Rogers N*, et al.* Ethnicity, deprivation and screening: survival from breast cancer among screening-eligible women in the West Midlands diagnosed from 1989 to 2011. *British journal of cancer.* 2015;113(3):548-555.
11. Woods LM, Rachet B, O'Connell DL*, et al.* Are international differences in breast cancer survival between A ustralia and the UK present amongst both screen-detected women and non-screen-detected women? survival estimates for women diagnosed in West Midlands and New South Wales 1997–2006. *International journal of cancer.* 2016;138(10):2404-2414.
12. O'Brien KM, Mooney T, Fitzpatrick P*, et al.* Screening status, tumour subtype, and breast cancer survival: a national population-based analysis. *Breast cancer research and treatment.* 2018;172(1):133-142.
13. Fracheboud J, de Koning HJ, Boer R*, et al.* Nationwide breast cancer screening programme fully implemented in The Netherlands. *Breast.* 2001;10(1):6-11.
14. Fracheboud J, van Luijt PA, Sankatsing VDV*, et al. National Evaluation of breast cancer screening in the Netherlands 1990-2011/2012*. Rotterdam: Optima Grafische Communicatie; 2014.
15. Moossdorff M, Van Roozendaal LM, Strobbe LJ*, et al.* Maastricht Delphi consensus on event definitions for classification of recurrence in breast cancer research. *Journal of the National Cancer Institute.* 2014;106(12).
16. van Duin C, Keij I. Sociaal-economische status indicator op postcodeniveau [in Dutch]. *Maandstatistiek van de bevolking.* 2002;50:32-35.
17. Sobin LH, Wittekind C. *TNM Classification of Malignant Tumours. Sixth edition*. New York: Wiley-Liss; 2002.
18. Aarts A, Duffy S, Geurts S*, et al.* Test sensitivity of mammography and mean sojourn time over 40 years of breast cancer screening in Nijmegen (The Netherlands). *Journal of medical screening.* 2019;26(3):147-153.
19. Azur M, Stuart E, Frangakis C*, et al.* Multiple Imputation by chained equation: what is it and how does it work? *International journal of methods in psychiatric research.* 2011;20(1):40-49.
20. White IR, Royston P. Imputing missing covariate values for the Cox model. *Statistics in medicine.* 2009;28(15):1982-1998.
21. Schoenfeld D. Partial residuals for the proportional hazards regression model. *Biometrika.* 1982;69(1):239-241.
22. Poiseuil M, Coureau G, Payet C*, et al.* Deprivation and mass screening: Survival of women diagnosed with breast cancer in France from 2008 to 2010. *Cancer Epidemiol.* 2019;60:149-155.
23. Hofvind S, Holen Å, Román M*, et al.* Mode of detection: an independent prognostic factor for women with breast cancer. *J Med Screen.* 2016;23(2):89-97.
24. Lehtimäki T, Lundin M, Linder N*, et al.* Long-term prognosis of breast cancer detected by mammography screening or other methods. *Breast Cancer Research.* 2011;13(6):1-11.
25. de Munck L, de Bock GH, Reiding D*, et al.* Digital vs screen-film mammography in population-based breast cancer screening: performance indicators and tumour characteristics of screen-detected and interval cancers. *British journal of cancer.* 2016;115(5):517-524.
26. Pedersen RN, Mellemkjær L, Ejlertsen B*, et al.* Mortality After Late Breast Cancer Recurrence in Denmark. *Journal of Clinical Oncology.* 2022;40(13):1450-1463.
27. Tevaarwerk AJ, Gray RJ, Schneider BP*, et al.* Survival in patients with metastatic recurrent breast cancer after adjuvant chemotherapy: little evidence of improvement over the past 30 years. *Cancer.* 2013;119(6):1140-1148.

28. Witteveen A, Kwast AB, Sonke GS*, et al.* Survival after locoregional recurrence or second primary breast cancer: impact of the disease-free interval. *PLoS One.* 2015;10(4).
29. Mook S, Van 't Veer LJ, Rutgers EJ*, et al.* Independent prognostic value of screen detection in invasive breast cancer. *J Natl Cancer Inst.* 2011;103(7):585-97.
30. Redondo M, Funez R, Medina-Cano F*, et al.* Detection methods predict differences in biology and survival in breast cancer patients. *BMC Cancer.* 2012;12(1):604.
31. Bretthauer M, Wieszczy P, Løberg M*, et al.* Estimated Lifetime Gained With Cancer Screening Tests: A Meta-Analysis of Randomized Clinical Trials. *JAMA Internal Medicine.* 2023.
32. Centraal Bureau voor de Statistiek. *StatLine*. https://opendata.cbs.nl/statline/#/CBS/en/.
33. Duffy SW. All-cause mortality in multi-cancer screening trials. In: SAGE Publications Sage UK: London, England; 2022, 1-2.
34. Sasieni PD, Wald NJ. Should a reduction in all-cause mortality be the goal when assessing preventive medical therapies? *Circulation.* 2017;135(21):1985-1987.
35. Plevritis SK, Munoz D, Kurian AW*, et al.* Association of screening and treatment with breast cancer mortality by molecular subtype in US women, 2000-2012. *Jama.* 2018;319(2):154-164.

Tables

Table 1 Patient, tumor and treatment characteristics of the total 2005 cohort and specified by method of detection.

	Total population	Clinically-detected cancer	Screen-detected cancer	P-value[a]
	No. (%)	No. (%)	No. (%)	
Patients	6,215	2,748	3,467	
Recurrence				<0.001
No	4,996 (80.4)	2,065 (75.1)	2,931 (84.5)	
Yes (LRR/DM/contralateral breast cancer)	1,219 (19.6)	683 (24.9)	536 (15.5)	
Age at diagnosis in years Median (IQR)	61 (55-68)	60 (54-67)	62 (56-68)	<0.001
SES				0.02
High	1,948 (31.3)	906 (33.0)	1,042 (30.1)	
Medium	2,523 (40.6)	1,069 (38.9)	1,454 (41.9)	
Low	1,744 (28.1)	773 (28.1)	971 (28.0)	
Screening round				–[b]
Initial screen	–[b]	–[b]	342 (9.9)	
Subsequent screen	–[b]	–[b]	3,125 (90.1)	
Histology				<0.001
Ductal	4,947 (79.6)	2,127 (77.4)	2,820 (81.3)	
Lobular	718 (11.6)	376 (13.7)	342 (9.9)	
Mixed	258 (4.2)	104 (3.8)	154 (4.4)	
Other	292 (4.7)	141 (5.1)	151 (4.4)	
Tumor grade				<0.001
1	1,454 (25.0)	471 (18.7)	983 (29.9)	
2	2,668 (46.0)	1,089 (43.2)	1,579 (48.0)	
3	1,684 (29.0)	958 (38.0)	726 (22.1)	
Unknown	409	230	179	
Multifocality				0.09
Yes	5,156 (85.2)	2,246 (84.3)	2,910 (85.9)	
No	896 (14.8)	418 (15.7)	478 (14.1)	
Unknown	163	84	79	
Tumor size				<0.001
<2 cm	4,108 (67.7)	1,358 (51.8)	2,750 (79.9)	
2 – 5 cm	1,782 (29.4)	1,120 (42.7)	662 (19.2)	
>5 cm	176 (2.9)	146 (5.6)	30 (0.9)	
Unknown	149	124	25	
Positive nodes				<0.001
0	4,065 (65.8)	1,520 (55.8)	2,545 (73.6)	
1 – 3	1,515 (24.5)	790 (29.0)	725 (21.0)	
>3	601 (9.7)	415 (15.2)	186 (5.4)	
Unknown	34	23	11	
Tumor stage				<0.001
I	3,117 (50.2)	934 (34.1)	2,183 (63.1)	
II	2,347 (37.8)	1,274 (46.5)	1,073 (31.0)	
III	740 (11.9)	534 (19.5)	206 (6.0)	
Unknown	11	6	5	
Tumor subtype				<0.001
ER+ and/or PR+ and HER2-	3,756 (74.8)	1,539 (67.7)	2,217 (80.6)	
ER+ and/or PR+ and HER2+	444 (8.8)	226 (9.9)	218 (7.9)	
ER- and PR- and HER2-	300 (6.0)	184 (8.1)	116 (4.2)	
ER- and PR- and HER2+	524 (10.4)	324 (14.3)	200 (7.3)	
Unknown	1,191	475	716	

Type of surgery				<0.001
Breast conserving surgery	3,811 (61.3)	1,378 (50.1)	2,433 (70.2)	
Mastectomy	2,404 (38.7)	1,370 (49.9)	1,034 (29.8)	
Chemotherapy				<0.001
No	4,369 (70.3)	1,587 (57.8)	2,782 (80.2)	
Yes	1,846 (29.7)	1,161 (42.2)	685 (19.8)	
Hormonal therapy				<0.001
No	3,663 (58.9)	1,371 (49.9)	2,292 (66.1)	
Yes	2,552 (41.1)	1,377 (50.1)	1,175 (33.9)	
Targeted therapy				<0.001
No	5,905 (95.0)	2,544 (92.6)	3,361 (96.9)	
Yes	310 (5.0)	204 (7.4)	106 (3.1)	
Radiotherapy				<0.001
No	1,834 (29.5)	929 (33.8)	905 (26.1)	
Yes	4,381 (70.5)	1,819 (66.2)	2,562 (73.9)	
Neo-adjuvant systemic therapy				<0.001
No	6,016 (96.8)	2,582 (94.0)	3,434 (99.0)	
Yes	199 (3.2)	166 (6.0)	33 (1.0)	
Axillary lymph node dissection				<0.001
No	3,457 (55.6)	1,261 (45.9)	2,196 (63.3)	
Yes	2,758 (44.4)	1,487 (54.1)	1,271 (36.7)	

The table contains numbers (percentages) unless otherwise specified. Percentages are calculated on known values only.
DM: distant metastasis, ER: estrogen receptor, IQR: interquartile range, LRR: locoregional recurrence, PR: progesterone receptor, SES: socioeconomic status
[a] Chi-squared and Kruskal Wallis test were used to compare patients in the different method of detection groups. The p-value is calculated on known values only.
[b] Not applicable for this subgroup of patients

Table 2 Patient, tumor and treatment characteristics of the total 2006-2008 cohort and specified by method of detection.

	Total population	Non-screen-related cancer	Screen-detected cancer	Interval cancer	P-value[a]
	No. (%)	No. (%)	No. (%)	No. (%)	
Patients	15,176	3,153	8,487	3,536	
Recurrence					<0.001
No	13,468 (88.7)	2,642 (83.8)	7,802 (91.9)	3,024 (85.5)	
Yes (LRR/DM/contralateral breast cancer)	1,708 (11.3)	511 (16.2)	685 (8.1)	512 (14.5)	
Age at diagnosis in years Median (IQR)	61 (55-68)	61 (54-68)	62 (56-68)	60 (55-66)	<0.001
SES					0.003
High	4,805 (31.7)	970 (30.8)	2,660 (31.3)	1,175 (33.2)	
Medium	6,201 (40.9)	1,244 (39.5)	3,507 (41.3)	1,450 (41.0)	
Low	4,170 (27.5)	939 (29.8)	2,320 (27.3)	911 (25.8)	
Screening round					<0.001
Initial screen	–[b]	–[b]	822 (9.7)	476 (13.5)	
Subsequent screen	–[b]	–[b]	7,667 (90.3)	3,060 (86.5)	
Histology					<0.001
Ductal	12,012 (79.2)	2,465 (78.2)	6,848 (80.7)	2,699 (76.3)	
Lobular	1,684 (11.1)	343 (10.9)	856 (10.1)	485 (13.7)	
Mixed	701 (4.6)	147 (4.7)	373 (4.4)	181 (5.1)	
Other	779 (5.1)	198 (6.3)	410 (4.8)	171 (4.8)	
Tumor grade					<0.001
1	3,750 (26.4)	599 (20.9)	2,575 (31.9)	576 (17.7)	

2	6,514 (45.9)	1,276 (44.5)	3,838 (47.5)	1,400 (43.1)	
3	3,929 (27.7)	991 (34.6)	1,663 (20.6)	1,275 (39.2)	
Unknown	983	287	411	285	
Multifocality					<0.001
Yes	12,729 (85.7)	2,588 (84.1)	7,216 (86.8)	2,925 (84.6)	
No	2,119 (14.3)	489 (15.9)	1,098 (13.2)	532 (15.4)	
Unknown	328	76	173	79	
Tumor size					<0.001
<2 cm	10,101 (68.1)	1,560 (52.3)	6,737 (80.0)	1,804 (52.8)	
2 – 5 cm	4,306 (29.1)	1,257 (42.2)	1,599 (19.0)	1,450 (42.4)	
>5 cm	415 (2.8)	164 (5.5)	87 (1.0)	164 (4.8)	
Unknown	354	172	64	118	
Positive nodes					<0.001
0	9,821 (65.4)	1,702 (54.7)	6,225 (74.1)	1,894 (53.8)	
1 – 3	3,754 (25.0)	956 (30.7)	1,723 (20.5)	1,075 (30.5)	
>3	1,450 (9.7)	451 (14.5)	449 (5.3)	550 (15.6)	
Unknown	151	44	90	17	
Tumor stage					<0.001
I	7,716 (50.9)	1,100 (34.9)	5,407 (63.8)	1,209 (34.2)	
II	5,664 (37.4)	1,444 (45.8)	2,572 (30.3)	1,648 (46.7)	
III	1,776 (11.7)	606 (19.2)	497 (5.9)	673 (19.1)	
Unknown	20	3	11	6	
Tumor subtype					<0.001
ER+ and/or PR+ and HER2-	11,020 (77.3)	2,146 (72.3)	6,559 (82.8)	2,315 (68.9)	
ER+ and/or PR+ and HER2+	1,085 (7.6)	253 (8.5)	536 (6.8)	296 (8.8)	
ER- and PR- and HER2-	744 (5.2)	202 (6.8)	294 (3.7)	248 (7.4)	
ER- and PR- and HER2+	1,403 (9.8)	367 (12.4)	537 (6.8)	499 (14.9)	
Unknown	924	185	561	178	
Type of surgery					<0.001
Breast conserving surgery	9,384 (61.8)	1,512 (48.0)	5,939 (70.0)	1,933 (54.7)	
Mastectomy	5,792 (38.2)	1,641 (52.0)	2,548 (30.0)	1,603 (45.3)	
Chemotherapy					<0.001
No	10,306 (67.9)	1,886 (59.8)	6,603 (77.8)	1,817 (51.4)	
Yes	4,870 (32.1)	1,267 (40.2)	1,884 (22.2)	1,719 (48.6)	
Hormonal therapy					<0.001
No	8,764 (57.7)	1,534 (48.7)	5,494 (64.7)	1,736 (49.1)	
Yes	6,412 (42.3)	1,619 (51.3)	2,993 (35.3)	1,800 (50.9)	
Targeted therapy					<0.001
No	14,121 (93.0)	2,880 (91.3)	8,087 (95.3)	3,154 (89.2)	
Yes	1,055 (7.0)	273 (8.7)	400 (4.7)	382 (10.8)	
Radiotherapy					<0.001
No	4,344 (28.6)	1,114 (35.3)	2,182 (25.7)	1,048 (29.6)	
Yes	10,832 (71.4)	2,039 (64.7)	6,305 (74.3)	2,488 (70.4)	
Neo-adjuvant systemic therapy					<0.001
No	14,524 (95.7)	2,865 (90.9)	8,361 (98.5)	3,298 (93.3)	
Yes	652 (4.3)	288 (9.1)	126 (1.5)	238 (6.7)	
Axillary lymph node dissection					<0.001
No	8,994 (59.3)	1,523 (48.3)	5,753 (67.8)	1,718 (48.6)	
Yes	6,182 (40.7)	1,630 (51.7)	2,734 (32.2)	1,818 (51.4)	

The table contains number (percentages) unless otherwise specified. Percentages are calculated on known values only.
DM: distant metastasis, ER: estrogen receptor, IQR: interquartile range, LRR: locoregional recurrence, PR: progesterone receptor, SES: socioeconomic status;

[a] Chi-squared and Kruskal Wallis test were used to compare patients in the different method of detection groups. The p-value is calculated on known values only.
[b] Not applicable for this subgroup of patients

Table 3 Hazard ratios (HRs) and 95% confidence intervals (CI) for the association between method of detection and disease-free interval in the 2005 and 2006-2008 cohort.

	Unadjusted HR (95% CI)[a]	Lead time adjusted HR (95% CI)	Confounding adjusted HR (95% CI)[a,b]
2005-cohort			
Clinically-detected	1.00 (reference)	1.00 (reference)	1.00 (reference)
Screen-detected	0.56 (0.50 to 0.63)	0.77 (0.68 to 0.87)[c]	0.72 (0.64 to 0.81)
2006-2008 cohort			
Non-screen-related	1.00 (reference)	1.00 (reference)	1.00 (reference)
Screen-detected	0.46 (0.41 to 0.52)	0.76 (0.66 to 0.88)[d]	0.65 (0.57 to 0.73)
Interval	0.88 (0.78 to 0.99)	–[e]	0.83 (0.74 to 0.94)

All analyses are adjusted for age. Disease-free interval: free of locoregional recurrence, distant metastasis or contralateral invasive breast cancer

[a] Using the uncorrected disease-free interval (i.e. time between diagnosis of the primary tumor and diagnosis of the recurrent disease or the end of follow-up).[b] Adjusted for age, social economic status, histology, tumour grade, multifocality, tumour stage, and subtype.

[c] Using the lead time corrected disease-free interval of 6.1 years.

[d] Using the lead time corrected disease-free interval of 2.0 years.

[e] Not applicable for this subgroup of patients

Figure legends

Figure 1 Flowchart of included participants

*Clinically-detected breast cancer includes non-screen-related and interval breast cancer

Figure 2 Disease-free interval of women in the 2005 (A) and 2006-2008 cohort (B), specified by method of detection and with and without lead time correction. Range on y-axis is 0.7-1.0.

Figure 1

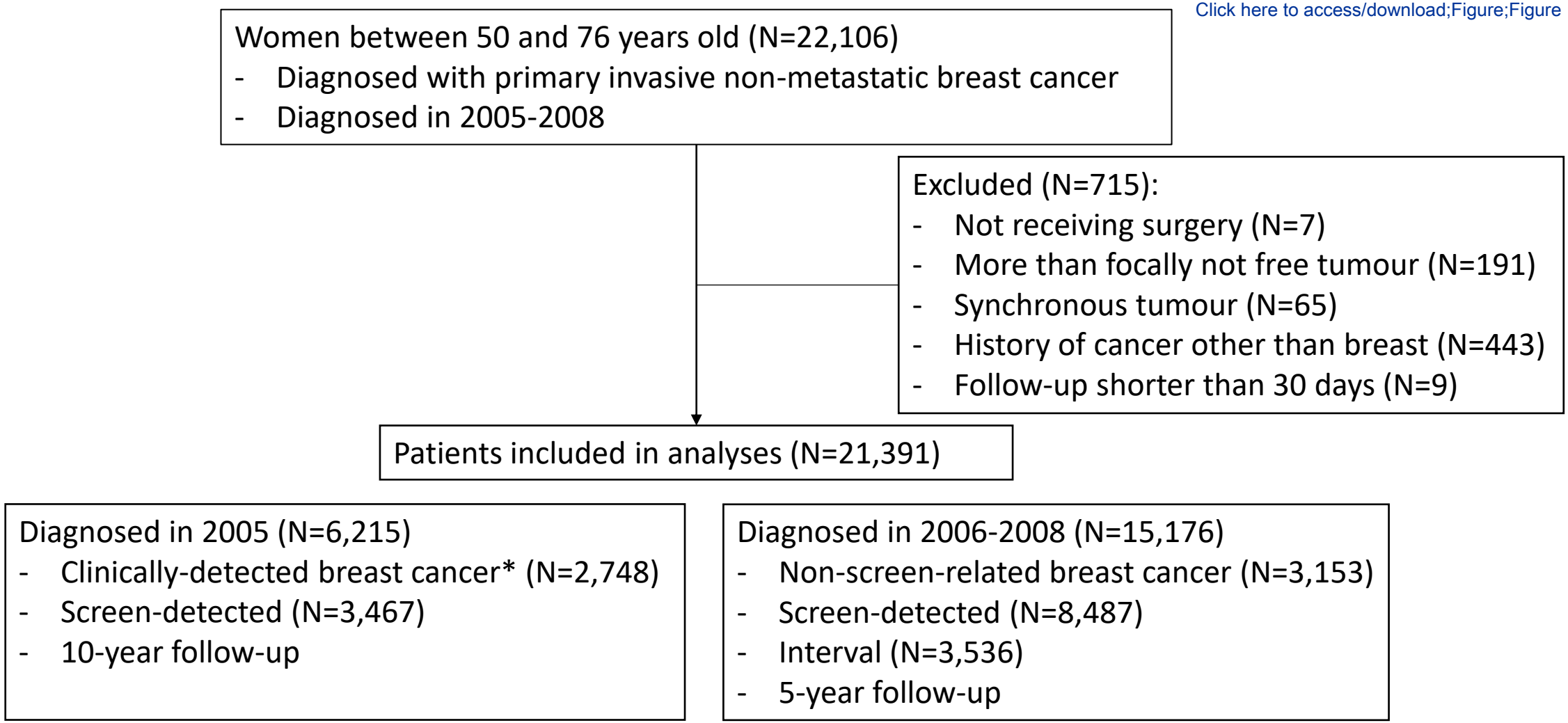

Figure 2A

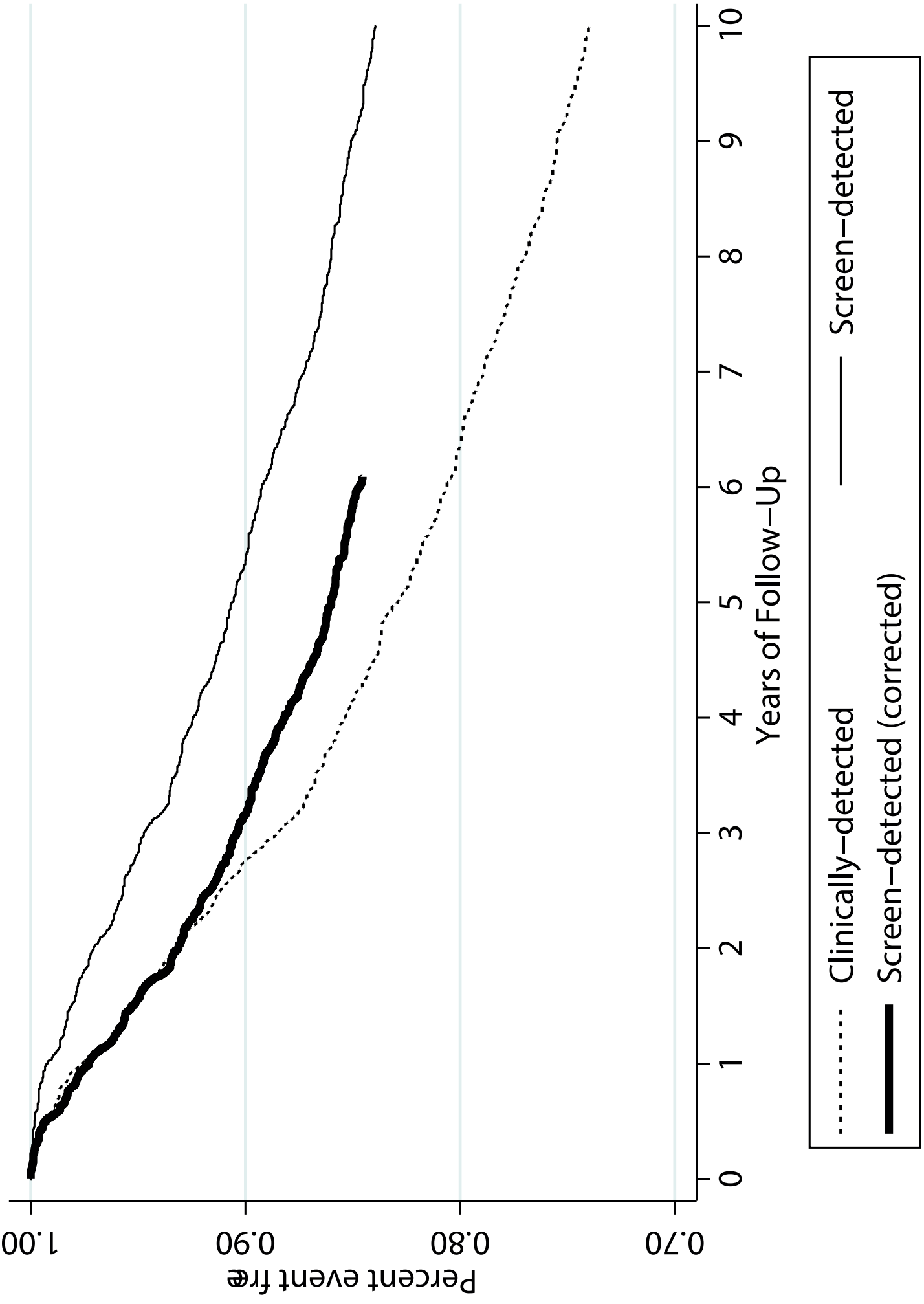

Figure 2B