**MDMP&P**
Policy & Practice

$\bigcirc$SAGE

# Interpreting Breast Cancer Mortality Trends Related to Introduction of Mammography Screening: A Simulation Study

**Torunn Heggland**📍, **Lars Johan Vatten, Signe Opdahl, and Harald Weedon-Fekjær**

**Abstract**
**Background.** Several studies have evaluated the effect of mammography screening on breast cancer mortality based on overall breast cancer mortality trends, with varied conclusions. The statistical power of such trend analyses is, however, not carefully studied. **Methods.** We estimated how the effect of screening on overall breast cancer mortality is likely to unfold. Because a screening effect is based on earlier treatment, screening can affect only new incident cases after screening introduction. To evaluate the likelihood of detecting screening effects on overall breast cancer mortality time trends, we calculated the statistical power of joinpoint regression analysis on breast cancer mortality trends around screening introduction using simulations. **Results.** We found that a very gradual increase in population-level screening effect is expected due to prescreening incident cases. Assuming 25% effectiveness of a biennial screening program in reducing breast cancer mortality among women 50 to 69 y of age, the expected reduction in overall breast cancer mortality was 3% after 2 y and reached a long-term effect of 18% after 20 y. In common settings, the statistical power to detect any screening effects using joinpoint regression analysis is very low ($<50\%$), even in an artificial setting of constant risk of baseline breast cancer mortality over time. **Conclusions.** Population effects of screening on breast cancer mortality emerge very gradually and are expected to be considerably lower than the effects reported in trials excluding women diagnosed before screening. Studies of overall breast cancer mortality time trends have too low statistical power to reliably detect screening effects in most populations. **Implications.** Researchers and policy makers evaluating mammography screening should avoid using breast cancer mortality trend analysis that does not separate pre- and postscreening incident cases.

**Highlights**

- Population-level mammography screening effects on breast cancer mortality emerge gradually following screening introduction, resulting in very low statistical power of trend analysis.
- Researchers and policy makers evaluating mammography screening should avoid relying on population-wide breast cancer mortality trends.
- Expected mammography screening effects at population level are lower than those from screening trials, as many cases of breast cancer fall outside the screening age range.

**Corresponding Author**
Torunn Heggland, Oslo Centre for Biostatistics and Epidemiology (OCBE), Research Support Services, Oslo University Hospital, Postboks 4950 Nydalen, Oslo, 0424, Norway; (torunn.heggland@studmed.uio.no).

## Introduction

In the 1980s and 1990s, public mammography screening was introduced in many Western countries,[1–3] based on results of randomized trials showing a 20% to 30% reduction in breast cancer mortality among screened women.[4,5]

Following the introduction of population-based screening programs, several observational studies assessing overall time trends have shown no clear effect of screening on breast cancer mortality,[6–10] whereas studies looking exclusively for an effect of screening among women without prescreening diagnoses (refined breast cancer mortality) have shown results in line with the randomized trials.[11–16]

As screening improves the prognosis by opening for earlier treatment, screening will not affect breast cancer cases already diagnosed, as these cases are already under treatment. Even with a screening program that perfectly prevents breast cancer deaths, we would still see a considerable number of breast cancer deaths overall due to cases diagnosed before screening introduction. When breast cancer mortality studies do not separate cases diagnosed before and after the introduction of screening, the effect of screening will be diluted.[17]

How strong the dilution effect is, and whether changes in breast cancer mortality trend related to screening can be distinguished from random fluctuations, is not well known. We here perform a simulation study to assess whether studies of breast cancer mortality trends not separating cases diagnosed before and after screening introduction can be expected to detect any effect of screening.

Oslo Centre for Biostatistics and Epidemiology, Research Support Services, Oslo University Hospital, Oslo, Norway (TH, HWF); Oslo Centre for Biostatistics and Epidemiology, Department of Biostatistics, Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway (TH); Department of Public Health and Nursing, Faculty of Medicine and Health Science, Norwegian University of Science and Technology, Trondheim, Norway (LJV, SO).

## Methods

Studying how the effect of screening on breast cancer mortality is likely to appear at population level, we simulated breast cancer mortality in a given population around the time of screening introduction. We studied the population breast cancer mortality rate, defined as number of breast cancer deaths divided by woman-years under study. Postscreening breast cancer deaths could originate from diagnosis both pre and post screening introduction, while only the breast cancer cases diagnosed after the introduction of screening could benefit from screening through earlier treatment. In our simulations, we start with assuming no screening effect, before adding a potential screening effect to the cases incident after screening introduction. To estimate the expected number of new postscreening incident cases in the absence of screening effects, we applied the distribution of times from primary breast cancer diagnosis to death for women who died of breast cancer in Norway prescreening. We used 1990 to 1994 as the prescreening period, the most recent 5-year prescreening period.

We assumed the effectiveness of screening in preventing breast cancer deaths among women with no prescreening diagnosis to be 33% (Table 1). We applied an attendance rate in organized screening of 75%, and for the sake of simplicity, we assumed that women who participated in the program attended all screening exams to which they were invited. This implies a screening program effectiveness in preventing breast cancer deaths of 25% among invited women with no prescreening diagnosis (75% of 33%), which is on par with reductions in breast cancer mortality seen in randomized trials and in population-based studies of refined breast cancer mortality.[4,5,13,18] For age groups that are no longer in the target group for program screening, we assumed a decreasing screening effectiveness, using weights based on simulations from the Cancer Intervention and Surveillance Modelling Network (CISNET) Stanford breast cancer simulation model.[3,19–22] The population-level screening effect, including both women with or without a breast cancer diagnosed before screening introduction, is the main outcome of this study (Table 1).

In Europe, a common and recommended mammography schedule is to screen women for breast cancer every other year between the ages of 50 and 69 y,[23] and we aimed to reproduce that schedule here.

For reasons of resources and capacity, screening is typically introduced at different times across a country's regions. Hence, in addition to introducing a screening program over 2 y, we used the introduction design of

**Table 1** Different levels of observing the impact of screening on mortality

| | Applied Term | Interpretation | Assumptions |
|---|---|---|---|
| I | Screening effectiveness | Reduction in breast cancer mortality among screened women with no prescreening diagnosis | 33% |
| II | Screening program effectiveness | Reduction in breast cancer mortality among invited women with no prescreening diagnosis | 25% (based on 75% screening attendance) |
| III | Population-level screening effect | Observable effect of screening on population-wide breast cancer mortality | Based on I and II[a] |

[a]Calculated in this work based on I and II.

BreastScreen Norway as an example of a realistic gradual screening introduction.

BreastScreen Norway, organized by the Cancer Registry of Norway, invites women aged 50 to 69 y living in Norway to mammography screening every other year.[24] In practice, women are invited by birth cohorts to countywide screening rounds,[24] with age typically ranging from 48 to 53 y at first invitation and from 66 to 71 y at last invitation. The program started between November 1995 and February 1996 as a pilot study in 4 large counties accounting for approximately 40% of all eligible women in Norway. The remaining 15 counties were included in a stepwise manner between 1999 and 2004 (Figure 1). The pilot study was designed to offer the women at least 2 screening exams, which led to some women being invited up to 73 y of age. Attendance in the program has remained fairly stable over the years at about 76%,[24] which is close to the 75% attendance rate applied in this study. At initiation of public mammography screening, Norway had a total of 4.4 million inhabitants.[25]

For the simulations, we used the number of female breast cancer deaths and residents in Norway aged 0 to 84 y, by county, birth cohort, and single calendar year (1985–2018). Reporting of cancer cases to the Cancer Registry of Norway is mandatory, and diagnostic information is obtained separately from clinicians, pathologists, and death certificates, with only 0.2% of all cancers ascertained only from death certificates.[26] The unique 11-digit national identity number of each citizen allows reliable follow-up.

## Statistical Analysis

We first estimated the population-level screening effect, assuming that all women eligible for screening in a given population were invited on the same day. To understand the potential of screening in lowering overall breast cancer mortality, we started by considering a screening
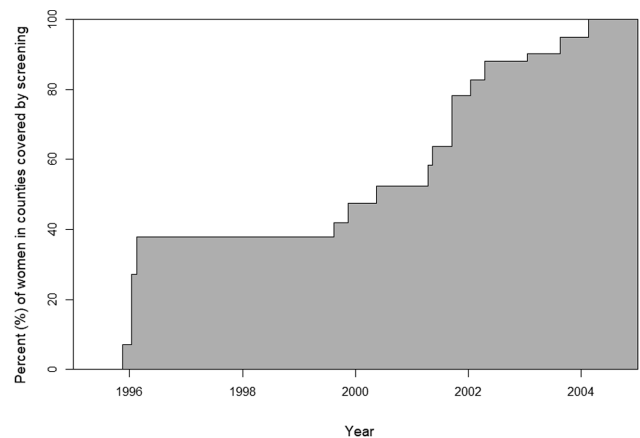


**Figure 1** Proportion of Norwegian women aged 50 to 69 y living in counties with screening. Because the screening program is biennial, women are typically screened between 0 and 2 y after initiation of screening in their county.

program with no screening effectiveness. After the program was introduced, a proportion of breast cancer deaths would arise from new postscreening incident cases, whereas the remaining deaths would arise from old prescreening incident cases. To estimate the proportion of breast cancer deaths based on new postscreening incident cases, in the absence of screening effects, we used the distribution of times from diagnosis to death for all breast cancer deaths that occurred in Norway from 1990 to 1994, when no deaths were averted by program screening. As an example, the proportion of the prescreening period deaths that occurred within 1 year after the time of breast cancer diagnosis is an estimate of the proportion of deaths based on new postscreening incident cases 1 year after screening initiation, with a potential for screening effects. Hence, we calculated the proportion of deaths that occurred within 1, 2, . . . , 419,

420 + mo after the time of breast cancer diagnosis, for the age groups 50 to 59, 60 to 69, 70 to 79, and 80 to 89 y (Supplementary Figure A1).

We then applied the 25% screening program effectiveness on the estimated proportion of breast cancer deaths based on new postscreening incident cases. We calculated the effect for 1-y age groups for every month following the program initiation, accounting for age and time since screening initiation for the given cohort. For age groups no longer in the target group for screening, we applied weights deduced from the CISNET Stanford model to estimate the decline in the screening effectiveness after passing the program's upper age limit with no new additional screening exams (Supplementary Appendix A2). Furthermore, we added screening introduction over 2 y by letting 1/24 of women currently in the target age group for screening be invited each month (Supplementary Appendix A3).

When calculating population-level screening effects for different age groups (50–54, 55–59, 60–64, 65–69, 50–69, 55–74 y and all ages), we weighted the estimated age-specific population-level screening effects by the observed age distribution of breast cancer deaths in Norway from 1990 to 1994 (prescreening). The long-term population-level screening effect for a given age group reflects the effect when all women have been invited to screening since 50 or 51 y of age.

To illustrate the expected population-level screening effect on the breast cancer mortality rate in Norway, we calculated the counterfactual breast cancer mortality rate in the absence of screening. This counterfactual breast cancer mortality rate was estimated by adjusting the observed breast cancer mortality rate for the estimated expected population-level effects of screening (Supplementary Appendix A4 and A5).

Furthermore, we used simulated scenarios to evaluate whether it is possible to separate any effect of screening from random fluctuations. The baseline variations in breast cancer mortality due to risk factors and treatment are mostly unknown. Hence, we chose to test whether a screening effect is possible to detect even in the artificially stable scenario of only random (Poisson distributed) fluctuations: the expected breast cancer mortality rate was set to be constant, at the mean 1990 to 1994 rate in Norway of 60 deaths per 100 000 woman-years for women aged 50 to 69 y and 31 deaths per 100 000 woman-years for women up to 84 y of age. On top of this, we added the always present random variation and our estimated expected population-level screening effect based on the introduction of BreastScreen Norway. We simulated the number of breast cancer deaths, based on

approximately 400 000 women aged 50 to 69 y, for the period 1985 to 2015.

When evaluating the simulated breast cancer mortality rates searching for the screening effect, we applied joinpoint regression analysis.[27–30] Joinpoint is a robust, well-recognized, and well-documented package for identification of distinct changes in trends and has been used in several breast cancer mortality trend studies.[6–8,10,17] A $k$-joinpoint model consists of $k + 1$ linear segments, continuously connected at change points (joinpoints). We searched for up to 2 joinpoints in the simulated mortality rates. For any effect of screening to be regarded as identifiable, a joinpoint must be identified near the time of program initiation. We also performed simulations with an assumed screening program effectiveness of 10% and 33%, simulations with screening introduced over 2 y from 1996 ("rapid introduction"), and simulations with the gradual BreastScreen Norway introduction combined with doubling and 5 times the Norwegian female population. Last, we restricted the study periods to 1985 to 2005 and 1985 to 2010.

We used version 4.9.0.0 of the National Cancer Institute's Joinpoint Regression Program (National Cancer Institute, Bethesda, MD, USA).[30] Searching for up to 2 joinpoints in the mortality rate means that 3 models are fitted. To select among the joinpoint models, we let the joinpoint program use the weighted Bayesian information criteria (BIC), which combines the BIC and a version of BIC with a harsher penalty term (BIC3), as this method is less conservative (joinpoints more easily detected) than the default choice of permutation tests.[31] Otherwise, we used the default settings in the joinpoint regression program, which includes logarithmic transformation of rates. All other statistical analyses and plots were conducted using the R statistical package (version 4.0.3, R Foundation for Statistical Computing, Vienna, Austria).[32]

The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report.

## Results

We found that the population-level screening effect is likely to emerge very gradually (Figure 2; Supplementary Table A1). Given a screening program effectiveness of 25%, the expected population-level effect of screening on breast cancer mortality for women aged 50 to 69 y was estimated to 3% after 2 y, 11% after 5 y, and 16% after 10 y of screening when screening is introduced over 2 y.
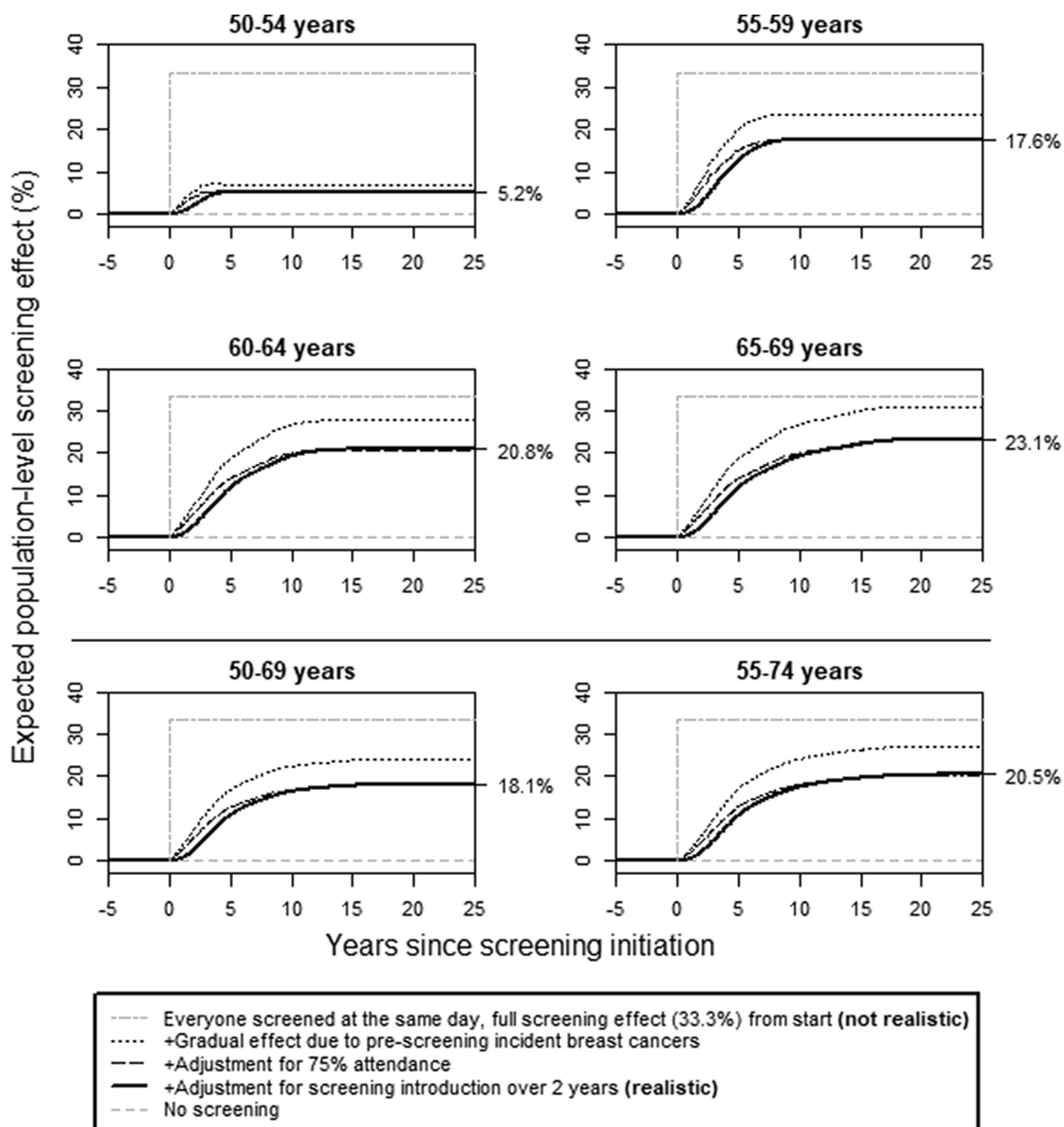
**Figure 2** Expected population-level screening effect on breast cancer mortality, based on an assumed 25% screening program effectiveness (33.3% among attending women and 75% attendance). The expected effect after 25 y of invitations is marked on the right axis.

We estimated the long-term effect for this age group, reached after 20 y, to be 18%. Based on the introduction of BreastScreen Norway, we estimated the expected population-level screening effect in 2018 to be 17% for women between 50 and 69 y of age (Supplementary Table A2). The observed breast cancer mortality rate in Norway, and the corresponding counterfactual rate in the absence of screening, showed a very gradual separation over time (Figure 3).

For the simulations, we defined the statistical power to detect the screening effect as the percentages of simulated mortality rates with at least 1 jointpoint identified

**Table 2** Statistical power to detect changes in breast cancer mortality rates, 50-69 years of age, using joinpoint analysis in a setting like the Norwegian screening introduction. Simulations are performed for an artificially stable scenario of constant rates in the absence of a screening effect. Due to additional variations in risk factors and treatment, real-life statistical power is expected to be substantially lower. Calculations are based on 10 000 simulated breast cancer mortality rates.

| Population Size | Screening Program Effectiveness (Assumed Prerequisite), % | Most Common Year for Joinpoints in Simulated Rates | Simulations with Joinpoints within 1993–2003 ("Statistical Power"), % | Simulations with Joinpoints within 1996–2003, % | Simulations with Joinpoints within 1985–2015, % |
|---|---|---|---|---|---|
| Norway: "baseline setting" | 0 | 2013 | 2 | 1 | 7 |
| (with approximately | 10 | 1987 | 4 | 3 | 9 |
| 400 000 women in the | 25 | 1997 | 25 | 18 | 34 |
| 50–69 y age group) | 33 | 1996 | 42 | 28 | 53 |
| 2 × "baseline setting" | 25 | 1997 | 51 | 37 | 61 |
| 5 × "baseline setting" | 25 | 1996 | 90 | 67 | 96 |

between 1993 and 2003. Based on 10 000 simulations for women aged 50 to 69 y, with a 25% screening program effectiveness, we found a statistical power of 25% (Table 2). By doubling the size of the population, we found a power of 51%. A screening program effectiveness of 10% gave a statistical power of 4%, while an effectiveness of 33% (equaling a 44% effect among attendees) resulted in a statistical power of 42% (Table 2). Limitation to the study period 1985 to 2010 resulted in somewhat higher chances of joinpoints (Supplementary Table A3), but the statistical power was still low. Extension to 0 to 84 y for long-term follow-up data increased the statistical power somewhat as we also covered the postscreening age groups (Supplementary Table A4), but the effect is moderate.

The effect of screening on breast cancer mortality will always stabilize at some time point, ending the declining trend. Longer follow-up could then dilute change points in trend, so a model free of joinpoints fits the rate better. With a more rapid screening introduction, the trend is stabilized earlier, increasing this phenomenon. With a rapid screening introduction and long follow-up, there is also a tendency for a screening-related joinpoint to occur before program initiation, especially in a 1-joinpoint model (Figure 4).

## Discussion

In this study, we found that an effect of screening on breast cancer mortality could not be detected in population-wide breast cancer mortality trend studies for the most realistic settings. The low statistical power makes analyses of overall breast cancer mortality trends nonconclusive. We found that the population-level
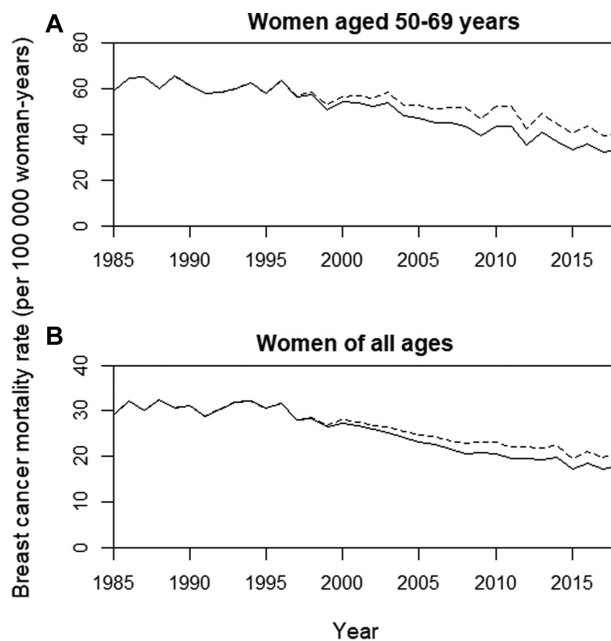


**Figure 3** Observed Norwegian breast cancer mortality rate (solid line) and the expected corresponding counterfactual breast cancer mortality rate in the absence of screening effect, assuming a 25% screening program effectiveness (dotted line), for (A) women aged 50 to 69 y and (B) women aged up to 84 y (age standardized).

screening effect emerges very gradually and that the population-level effect is considerably lower than the effect reported in screening trials. The reason for the low
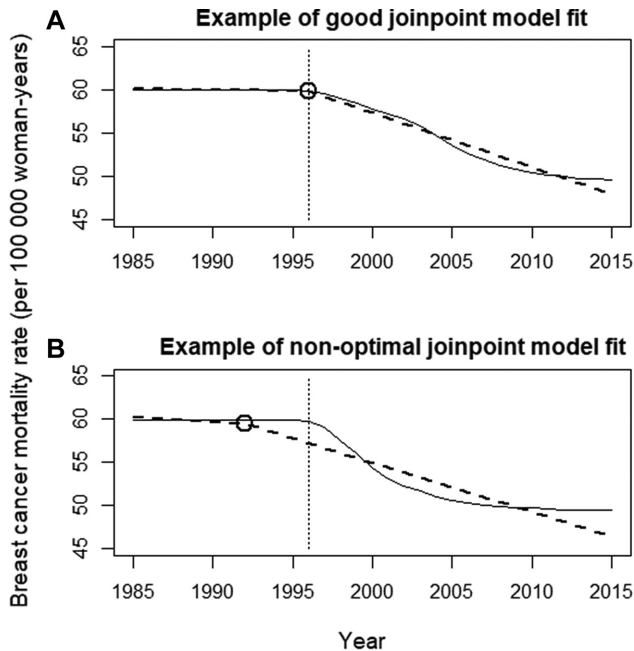
Figure 4 Changes in the expected breast cancer mortality rate (solid line) due to (A) a screening introduction a la Norway and (B) a 2-y introduction. The dashed line shows the fitted 1-joipoint model, with the joinpoint highlighted (circle). The vertical dotted line shows the time of screening initiation.

statistical power of jointpoint and other methods of linear trend analysis is 2-fold: first, a screening effect emerges so gradually that trend analysis using joinpoint regression has a very low chance of detecting the effect, and second, the methods are well suited for detecting linear trends, but screening effects do not appear in a linear fashion (Figure 2). A cohort of approximately 2 million women in the age range 50 to 69 y was necessary for the statistical power to be satisfactory, even in the artificially stable scenario of constant baseline breast cancer risk over time. This is 5 times the number of women aged 50 to 69 y in Norway.

It has been discussed at which time point breast cancer mortality rates may start showing a downward trend that can be attributed to mammography screening.[17,33] Population-based studies have sometimes detected turning points within the first 5 y after screening initiation,[33–37] but turning points have also been found before screening was introduced.[38] Some studies that identified joinpoints before screening started concluded that the change could not be attributed to screening.[6,8,17] However, using joinpoint analysis, we found that due to random fluctuations and

nonoptimal fit to the joinpoint line segments, a break downward may well be observed before screening started. In practice, the joinpoint model aims to fit both the initial decline when screening first starts to have an impact and the flatter, long-term trend that occurs some years later. Balancing the fit to both the decline and the flatter part of the trend, the joinpoint may sometimes be shifted backward in time (Figure 4). Generally, we find that a turning point in the trend should appear around the time screening is introduced: the joinpoint routine typically highlights joinpoints early after screening introduction, as this coincides with the initial decline (Table 2). In practice, fitting a simulation model with different levels of screening effectiveness will be better for studying trend changes than joinpoint analysis, as simulation models provide a more realistic change in mortality trends.[39] However, with data combining deaths from breast cancers diagnosed before and after screening, all analyses will have a low statistical power to identify any effect of screening.

The distribution of times from diagnosis to death from breast cancer (Supplementary Appendix A1) is a key input for our analysis. In practice, we had only the time interval since first diagnosis of breast cancer. This could overrate time to death and bias our results, but second breast cancers are rare.[40] Regardless of screening, this distribution might change somewhat over time, due to increased awareness, changes in risk factors, and improved treatment. However, these factors could lead to both shorter and longer time intervals between diagnosis and death, and historical data showed only small variation by time (data not shown).

The assumed screening program effectiveness of 25% was based on results from high-quality studies.[11–15] The real efficiency of modern mammography screening is not precisely known,[41] and some studies suggested smaller[42] or larger effectiveness.[43,44] Regardless of the actual effectiveness, the screening would influence the population-wide breast cancer mortality rate just as gradually as shown here, only with a different scaling. The power to detect the effect of screening on breast cancer mortality is affected by the screening effectiveness, but even with an attendance rate of 75% and 44% mortality reduction among screened women, a very large population with similar screening startup times is needed.

We assumed the screening effectiveness to be present and constant already from program initiation among new breast cancer cases. Screening typically shifts the time of diagnosis forward by several years.[18,45,46] Hence, some might argue that screening will have no influence on breast cancer deaths the first years after program

start. However, many of the breast cancer deaths with a short time from clinical diagnosis to death are probably fast-growing tumors reaching high stages at clinical diagnosis. These fast-growing tumors probably also have a considerable potential for an improved prognosis with just slightly earlier diagnosis through screening. The Swedish randomized trials overview indicates a relatively larger impact of screening on breast cancer mortality the first years after screening introduction, but the numbers were too small to draw firm conclusions or determine statistical significance.[18] Overall, it is not known precisely how the effectiveness of screening varies; however, smaller variations in the initial effectiveness among the few breast cancer cases diagnosed soon after screening introduction would have limited impact on our results.

The down-weighting of the screening effectiveness at higher ages was based on simulations from the CISNET Stanford model. As this model is calibrated according to American screening practices, there are likely some deviations that may somewhat influence our results. This does, however, relate only to the 70 + y age group and would not affect our main results.

In practice, who is attending screening might vary somewhat between screening rounds. This might increase the screening effect compared with our simulated scenario with regular attendance. The increased statistical power is, however, limited, as even with 100% screening attendance, the statistical power is 42%. The attendance rate level affects the statistical power of trend analysis methods. Hence, the real statistical power is likely to be even lower in many countries that do not have the high attendance rate seen in Norway. Also, inaccurate registration of breast cancer deaths could lead to more noisy data, making it even more difficult to find an effect of screening.

In the real world, both breast cancer risk factors and available treatment change over time, and the true statistical power of trend analyses is probably considerably lower than demonstrated in this study.

## Conclusion

Studies of breast cancer mortality trends that do not separate cases diagnosed before and after screening introduction cannot be expected to detect any effect of screening. Screening influences the overall breast cancer mortality rate very gradually. Hence, the population-level effect of screening will usually not be possible to reliably separate from random variations in trend analysis.

Public health officials should not rely on breast cancer mortality trend analysis mixing cases diagnosed before

and after screening introduction and should be aware that the population-level screening effect is expected to be lower than the effect reported in screening trials.

## ORCID iD

Torunn Heggland (iD) https://orcid.org/0000-0003-3165-8086

## Data Availability

The data are available for research projects from the legal administrator of the data, the Cancer Registry of Norway. For data requests, contact due@kreftregisteret.no.

## References

1. International Agency for Research on Cancer. *Cancer Screening in the European Union. Report on the Implementation of the Council Recommendation on Cancer Screening*. Brussels: European Commission; 2017.
2. Hofvind S, Vacek PM, Skelly J, et al. Comparing screening mammography for early breast cancer detection in Vermont and Norway. *J Natl Cancer Inst*. 2008;100:1082–91. DOI: 10.1093/jnci/djn224
3. Berry DA, Cronin KA, Plevritis SK, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. *N Engl J Med*. 2005;353:1784–92. DOI: 10.1056/NEJMoa050518
4. Nystrom L, Rutqvist LE, Wall S, et al. Breast cancer screening with mammography: overview of Swedish randomised trials. *Lancet*. 1993;341:973–8.
5. Tabar L, Fagerberg CJ, Gad A, et al. Reduction in mortality from breast cancer after mass screening with mammography: randomised trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health and Welfare. *Lancet*. 1985;1:829–32. DOI: 10.1016/s0140-6736(85)92204-4
6. Autier P, Boniol M, Gavin A, et al. Breast cancer mortality in neighbouring European countries with different levels of screening but similar access to treatment: trend analysis of WHO mortality database. *BMJ*. 2011;343: d4411. DOI: 10.1136/bmj.d4411

7. Autier P, Koechlin A, Smans M, et al. Mammography screening and breast cancer mortality in Sweden. *J Natl Cancer Inst*. 2012;104:1080–93. DOI: 10.1093/jnci/djs272

8. Mukhtar TK, Yeates DR, Goldacre MJ. Breast cancer mortality trends in England and the assessment of the effectiveness of mammography screening: population-based study. *J R Soc Med*. 2013;106:234–42. DOI: 10.1177/0141076813486779

9. Jørgensen KJ, Zahl PH, Gøtzsche PC. Breast cancer mortality in organised mammography screening in Denmark: comparative study. *BMJ*. 2010;340:c1241. DOI: 10.1136/bmj.c1241

10. Burton R, Stevenson C. Assessment of breast cancer mortality trends associated with mammographic screening and adjuvant therapy from 1986 to 2013 in the State of Victoria, Australia. *JAMA Netw Open*. 2020;3:e208249. DOI: 10.1001/jamanetworkopen.2020.8249

11. Olsen AH, Njor SH, Vejborg I, et al. Breast cancer mortality in Copenhagen after introduction of mammography screening: cohort study. *BMJ*. 2005;330:220. DOI: 10.1136/bmj.38313.639236.82

12. Broeders M, Moss S, Nystrom L, et al. The impact of mammographic screening on breast cancer mortality in Europe: a review of observational studies. *J Med Screen*. 2012;19(suppl 1):14–25. DOI: 10.1258/jms.2012.012078

13. Weedon-Fekjaer H, Romundstad PR, Vatten LJ. Modern mammography screening and breast cancer mortality: population study. *BMJ*. 2014;348:g3701. DOI: 10.1136/bmj.g3701

14. Beau A-B, Andersen PK, Vejborg I, et al. Limitations in the effect of screening on breast cancer mortality. *J Clin Oncol*. 2018;36:2988–94. DOI: 10.1200/jco.2018.78.0270

15. Sebuødegård S, Botteri E, Hofvind S. Breast cancer mortality after implementation of organized population-based breast cancer screening in Norway. *J Natl Cancer Inst*. 2020;112:839–46. DOI: 10.1093/jnci/djz220

16. Dibden A, Offman J, Duffy SW, et al. Worldwide review and meta-analysis of cohort studies measuring the effect of mammography screening programmes on incidence-based breast cancer mortality. *Cancers (Basel)*. 2020;12(4):976. DOI: 10.3390/cancers12040976

17. Moss SM, Nystrom L, Jonsson H, et al. The impact of mammographic screening on breast cancer mortality in Europe: a review of trend studies. *J Med Screen*. 2012;19(suppl 1):26–32. DOI: 10.1258/jms.2012.012079

18. Nystrom L, Andersson I, Bjurstam N, et al. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet*. 2002;359:909–19. DOI: 10.1016/s0140-6736(02)08020-0

19. Cancer Intervention and Surveillance Modeling Network. 2021. Available from: https://cisnet.cancer.gov/. Accessed May 20, 2021.

20. Plevritis SK, Sigal BM, Salzman P, et al. A stochastic simulation model of U.S. breast cancer mortality trends from 1975 to 2000. *J Natl Cancer Inst Monogr*. 2006;36:86–95. DOI: 10.1093/jncimonographs/lgj01

21. Alagoz O, Berry DA, de Koning HJ, et al. Introduction to the Cancer Intervention and Surveillance Modeling Network (CISNET) breast cancer models. *Med Decis Making*. 2018;38:3s–8s. DOI: 10.1177/0272989x17737507

22. 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;319:154–64. DOI: 10.1001/jama.2017.19130

23. European Commission Initiative on Breast Cancer. Recommendations from the European breast guidelines. 2020. Available from: https://ecibc.jrc.ec.europa.eu/recommendations/. Accessed October 28, 2020.

24. Hofvind S, Tsuruda K, Mangerud G, et al. The Norwegian Breast Cancer Screening Program, 1996-2016: celebrating 20 years of organised mammographic screening. In: *Cancer in Norway 2016—Cancer Incidence, Mortality, Survival and Prevalence in Norway*. Oslo: Cancer Registry of Norway; 2017.

25. Statistics Norway. Table 07459: population, by sex and one-year age groups (M) 1986-2021. Available from: https://www.ssb.no/statbank/table/07459/. Accessed May 21, 2021.

26. Larsen IK, Smastuen M, Johannesen TB, et al. Data quality at the Cancer Registry of Norway: an overview of comparability, completeness, validity and timeliness. *Eur J Cancer*. 2009;45:1218–31.

27. Kim HJ, Fay MP, Feuer EJ, et al. Permutation tests for joinpoint regression with applications to cancer rates [published correction appears in *Stat Med*. 2001;2020:2655). *Stat Med*. 2000;19:335–1. DOI: 10.1002/(sici)1097-0258(20000215)19:3<335::aid-sim336>3.0.co;2-z

28. Kim HJ, Yu B, Feuer EJ. Selecting the number of change-points in segmented line regression. *Stat Sin*. 2009;19:597–609.

29. Kim J, Kim HJ. Consistent model selection in segmented line regression. *J Stat Plann Infer*. 2016;170:106–16. DOI: 10.1016/j.jspi.2015.09.008

30. Statistical Methodology and Applications Branch Surveillance Research Program. *Joinpoint Regression Program, Version 4.9.0.0*. Bethesda (MD): National Cancer Institute; 2021.

31. National Cancer Institute. *Joinpoint Help Manual 4.8.0.1*. 2020. Available from: https://surveillance.cancer.gov/help/joinpoint/. Accessed September 6, 2021.

32. R Developer Core Team. *R: A Language and Environment for Statistical Computing*. 2020. Vienna (Austria): R Foundation for Statistical Computing. Available from: https://www.R-project.org/

33. Sankatsing VDV, van Ravesteyn NT, Heijnsdijk EAM, et al. The effect of population-based mammography screening in Dutch municipalities on breast cancer mortality: 20 years of follow-up. *Int J Cancer*. 2017;141:671–7. DOI: 10.1002/ijc.30754

34. Otto SJ, Fracheboud J, Looman CW, et al. Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality: a systematic review. *Lancet*. 2003;361:1411–7. DOI: 10.1016/s0140-6736(03)13132-7

35. Iwamoto Y, Kaucher S, Lorenz E, et al. Development of breast cancer mortality considering the implementation of mammography screening programs—a comparison of Western European countries. *BMC Public Health*. 2019;19: 823. DOI: 10.1186/s12889-019-7166-6

36. Otten JD, Broeders MJ, Fracheboud J, et al. Impressive time-related influence of the Dutch screening programme on breast cancer incidence and mortality, 1975-2006. *Int J Cancer*. 2008;123:1929–34. DOI: 10.1002/ijc.23736

37. Ascunce EN, Moreno-Iribas C, Barcos Urtiaga A, et al. Changes in breast cancer mortality in Navarre (Spain) after introduction of a screening programme. *J Med Screen*. 2007;14:14–20. DOI: 10.1258/096914107780154558

38. Taylor R, Gregory M, Sexton K, et al. Breast cancer mortality and screening mammography in New Zealand: incidence-based and aggregate analyses. *J Med Screen*. 2019;26:35–43. DOI: 10.1177/0969141318776039

39. Trentham-Dietz A, Alagoz O, Chapman C, et al. Reflecting on 20 years of breast cancer modeling in CISNET: recommendations for future cancer systems modeling efforts. *PLoS Comput Biol*. 2021;17:e1009020. DOI: 10.1371/journal.pcbi.1009020

40. Qian X, Jia H, Zhang Y, et al. Risk factors and prediction of second primary cancer in primary female non-metastatic breast cancer survivors. *Aging (Albany NY)*. 2020;12: 19628–40. DOI: 10.18632/aging.103939

41. Gøtzsche PC. Mammography screening is harmful and should be abandoned. *J R Soc Med*. 2015;108:341–5. DOI: 10.1177/0141076815602452

42. Kalager M, Zelen M, Langmark F, et al. Effect of screening mammography on breast-cancer mortality in Norway. *N Engl J Med*. 2010;363:1203–10. DOI: 10.1056/NEJMoa 1000727

43. Duffy SW, Tabár L, Yen AM, et al. Beneficial effect of consecutive screening mammography examinations on mortality from breast cancer: a prospective study. *Radiology*. 2021;299:541–7. DOI: 10.1148/radiol.2021203935

44. Swedish Organised Service Screening Evaluation Group. Reduction in breast cancer mortality from organized service screening with mammography: 1. Further confirmation with extended data. *Cancer Epidemiol Biomarkers Prev*. 2006;15:45–51. DOI: 10.1158/1055-9965.Epi-05-0349

45. Paci E, Duffy SW. Modelling the analysis of breast cancer screening programmes: sensitivity, lead time and predictive value in the Florence District Programme (1975-1986). *Int J Epidemiol*. 1991;20:852–8.

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