The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 SEPTEMBER 23, 2010 VOL. 363 NO. 13

Effect of Screening Mammography on Breast-Cancer Mortality in Norway

Mette Kalager, M.D., Marvin Zelen, Ph.D., Frøydis Langmark, M.D., and Hans-Olov Adami, M.D., Ph.D.

ABSTRACT

BACKGROUND

A challenge in quantifying the effect of screening mammography on breast-cancer mortality is to provide valid comparison groups. The use of historical control subjects does not take into account chronologic trends associated with advances in breast-cancer awareness and treatment.

METHODS

The Norwegian breast-cancer screening program was started in 1996 and expanded geographically during the subsequent 9 years. Women between the ages of 50 and 69 years were offered screening mammography every 2 years. We compared the incidence-based rates of death from breast cancer in four groups: two groups of women who from 1996 through 2005 were living in counties with screening (screening group) or without screening (nonscreening group); and two historical-comparison groups that from 1986 through 1995 mirrored the current groups.

RESULTS

We analyzed data from 40,075 women with breast cancer. The rate of death was reduced by 7.2 deaths per 100,000 person-years in the screening group as compared with the historical screening group (rate ratio, 0.72; 95% confidence interval [CI], 0.63 to 0.81) and by 4.8 deaths per 100,000 person-years in the nonscreening group as compared with the historical nonscreening group (rate ratio, 0.82; 95% CI, 0.71 to 0.93; P<0.001 for both comparisons), for a relative reduction in mortality of 10% in the screening group (P=0.13). Thus, the difference in the reduction in mortality between the current and historical groups that could be attributed to screening alone was 2.4 deaths per 100,000 person-years, or a third of the total reduction of 7.2 deaths.

CONCLUSIONS

The availability of screening mammography was associated with a reduction in the rate of death from breast cancer, but the screening itself accounted for only about a third of the total reduction. (Funded by the Cancer Registry of Norway and the Research Council of Norway.)

From the Cancer Registry of Norway, Oslo (M.K., F.L., H.-O.A.); the Departments of Epidemiology (M.K., H.-O.A.) and Biostatistics (M.Z.), Harvard School of Public Health; and the Dana–Farber Cancer Institute and Harvard Medical School (M.Z., H.-O.A.) — all in Boston; and the Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm (H.-O.A.). Address reprint requests to Dr. Kalager at Oslo University Hospital, Department of Surgery, Montebello, 0310 Oslo, Norway, or at mkalager@hsph.harvard.edu.

N Engl J Med 2010;363:1203-10.

On the basis of several randomized clinical trials,[1-3] the World Health Organization concluded in 2002 that screening mammography for women between the ages of 50 and 69 years reduced the rate of death from breast cancer by 25%.[4] Nevertheless, the use of screening mammography is still debated, chiefly because of concern regarding methodologic limitations in some of the randomized trials.[5] In addition, the benefit of mammography when implemented in a population-based service program remains poorly quantified. Therefore, continued evaluation of breast-cancer screening programs is warranted.[6]

The main challenge in quantifying the reduction in mortality from nonrandomized screening programs is to provide valid comparison groups. Although historical, prescreening control groups are often used, such a comparison has important limitations because it does not take into account confounding by chronological trends in breast-cancer mortality, reflecting such factors as advances in breast-cancer awareness and treatment. According to a statistical model based on data regarding breast-cancer mortality in the United States from 1975 through 2000, only half the observed reduction in mortality was causally related to the mammographic intervention itself, whereas the other half was attributable to improved management.[7] To establish a valid comparison group, we took advantage of several unique features of the nationwide Breast Cancer Screening Program in Norway, which was implemented by means of gradual geographic expansion over a 9-year period.

METHODS

SCREENING PROGRAM

Norway, with a total population of 4.8 million, has a public health care system. Patients generally receive treatment in their county of residence, and there is no private primary care for breast cancer.[8] The nationwide Cancer Registry of Norway is close to 100% complete.[9,10] Patients are identified in the registry by their individually unique national registration number, which includes the date of birth. The registry runs the Breast Cancer Screening Program, which began as a pilot project in 4 of the 19 Norwegian counties in 1996. Two years later, the government decided to expand the program, and over a period of 9 years, the remaining 15 counties were enrolled in a staggered fashion[11] (Fig. 1). The rollout of the program followed no specific geographic pattern. Since 2005, all women in the country between the ages of 50 and 69 years have been invited to participate in screening mammography every 2 years.

Before enrollment in the program, each county was required to establish multidisciplinary breast-cancer management teams and breast units.[12] As a result, breast-cancer management became centralized for all residents within each county, and dedicated teams of radiologists, radiologic technologists, pathologists, surgeons, oncologists, and nurses managed the care of all patients, regardless of age.

The screening program is organized with 26 stationary and 4 mobile screening units.[13] The Central Population Registry of Norway identifies eligible women on the basis of their national registration number. Invitations are mailed to each eligible woman, suggesting a time for an appointment.[14] Overall, 77% of all women who are invited participate in the program.[15] In accordance with European guidelines, mammograms are obtained in two views, which are independently read by two radiologists.[12]

STUDY GROUPS

From Statistics Norway we retrieved information on the Norwegian female population, according to county, from January 1, 1986, through December 31, 2005.[16] From the Cancer Registry, we retrieved data on all women who had received a diagnosis of invasive breast cancer, including age, tumor stage, date and county of residence at diagnosis, date and cause of death, and information on whether the diagnosis had been made before or after the implementation of the screening program.

By comparing two current groups on the basis of whether screening mammography was available in the county, we would avoid confounding by factors such as improvements in treatment and heightened awareness, temporal changes that may be associated with a reduction in breast-cancer mortality. However, we could not make direct comparisons between these two groups because of the nonconstant risk of death from breast cancer according to the time since diagnosis and differences in rates of death from breast cancer between counties before implementation of the screening program.[15] To adjust for such differences and to achieve equal follow-up time in each county, we

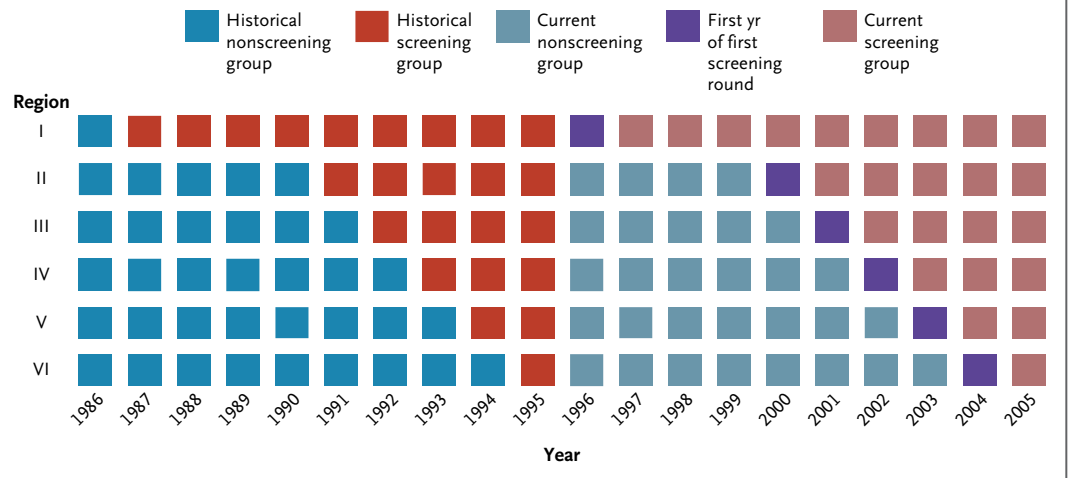

Figure 1. The Four Study Groups, According to Region and Year.

The 19 counties were grouped into six regions according to the date of introduction of the screening program, which was implemented throughout the country in a staggered fashion, starting in 1996. The screening group consisted of women who received a diagnosis of breast cancer after the introduction of the screening program. The nonscreening group consisted of women living in regions where screening was not offered in the same calendar period that screening was offered in other regions. The historical study groups consisted of women residing in the 19 counties in the 10-year period before screening was offered. A screening round lasted for 2 years, and the first year of the first round was included in both the screening and nonscreening groups (purple).

established two historical comparison groups that mirrored the implementation of the screening program during the 10-year period preceding the screening program.

Thus, we defined four groups of women, including those in whom a first invasive breast cancer had been diagnosed: two current groups of women who from 1996 through 2005 were living either in counties in which the screening program had been implemented (screening group) or in counties in which the program had not been implemented (nonscreening group), and two historical-comparison groups that from 1986 through 1995 mirrored the county residence of the current groups before the implementation of the screening program (Fig. 1) (see the Supplementary Appendix, available with the full text of this article at NEJM.org).

As pointed out, each county was required to establish multidisciplinary breast-cancer management teams and breast units before enrollment in the national screening program. As a result, the screening program consists of two components: screening mammography and care from multidisciplinary teams. For women between the ages of 50 and 69 years who were invited to participate in the program, the change in mortality after the introduction of the screening program can be related to both the introduction of screening mammography and the establishment of multidisciplinary teams. However, for women who were outside the age range that was eligible for the screening program (i.e., those between the ages of 20 and 49 years and those between the ages of 70 and 84 years) in the counties in which screening was available, the change in mortality could be related only to the establishment of multidisciplinary teams, since these women were not invited to undergo mammography.

STUDY OVERSIGHT

The Norwegian Social Science Data Services approved the study, which was funded by the Cancer Registry of Norway and the Research Council of Norway. The study was conducted in accordance with the protocol, which is available at NEJM.org.

STATISTICAL ANALYSIS

We obtained information on breast cancer as the underlying cause of death through regular linkage between the Cancer Registry and the Cause of Death Registry at Statistics Norway. To isolate the

effect of the breast-cancer screening program, our calculation of mortality in the screening group includes only deaths from breast cancer in women who received the diagnosis after the screening program was implemented (so-called incidence-based mortality).[17-19] The use of incidence-based mortality avoids the inclusion of breast-cancer deaths that occurred after implementation of the screening program but reflected diagnoses that were made before the program was implemented. So as not to bias our comparisons, we calculated the rate of death in all groups using the incidence-based method. All women in whom breast cancer was diagnosed and who died of breast cancer after implementation of the screening program were included in the screening group, regardless of whether they received the diagnosis at a screening or a diagnostic examination.

On the basis of the date of implementation of the screening program in each county, we grouped the 19 counties into six regions; each county within a given region entered the program at approximately the same time (see the Supplementary Appendix). We compared the rates of death separately for each region. Thus, the regional comparisons have the same follow-up time. This grouping tended to reduce random variation resulting from small numbers and permitted the evaluation of changes in mortality in the same region over a period of time. First, we compared women in the nonscreening group with their historical counterparts to determine the temporal change in mortality that was not attributable to the introduction of the screening program and that was likely to reflect improved treatment and earlier clinical diagnosis. Then, we compared women in the screening group with their historical counterparts to determine the change in mortality after implementation of the screening program. In this second comparison, the difference in the rate of death between the two groups can be attributed both to the screening program and to temporal trends in mortality that were unrelated to the screening program. Thus, the reduction in mortality that was related to the screening program was the difference between the rate ratio for death among women in the screening group as compared with their historical counterparts and the rate ratio for death among women in the nonscreening group as compared with their historical counterparts.

We estimated rates of death from breast cancer in the four study groups according to the age at diagnosis (20 to 49 years, 50 to 69 years, and 70 to 84 years). All tests of statistical significance were one-sided, and a P value of less than 0.05 was considered to indicate statistical significance. (For additional details on the statistical analysis plan, see the Supplementary Appendix.)

RESULTS

SUBJECTS

A total of 40,075 women received a diagnosis of breast cancer between 1986 and 2005. During the follow-up period, 4791 of these women (12%) died from breast cancer. Of the women who died, 423 (9%) had received the diagnosis after the introduction of the screening program. The total follow-up time for the study was 31,613,529 person-years, with an average of 2.2 years and a maximum of 8.9 years of follow-up for women with breast cancer. Among women between the ages of 50 and 69 years, 6967 received a diagnosis of breast cancer between 1986 and 1995, as compared with 12,056 who received the diagnosis between 1996 and 2005. In the latter group, 7975 women (66%) had been invited to participate in screening mammography. In the first screening round, a total of 454,331 women had been invited.

Among women between the ages of 50 and 69 years in the screening group, the rate of death was 18.1 per 100,000 person-years, as compared with 25.3 per 100,000 person-years among their historical counterparts, for a difference of 7.2 deaths per 100,000 person-years (rate ratio, 0.72; 95% confidence interval [CI], 0.63 to 0.81; P<0.001), a relative reduction of 28% (Table 1 and Fig. 2). Among women in the nonscreening group, the rate of death was 21.2 per 100,000 person-years, as compared with 26.0 per 100,000 person-years among their historical counterparts, for a difference of 4.8 deaths per 100,000 person-years (rate ratio, 0.82; 95% CI, 0.71 to 0.93; P<0.001), a relative reduction of 18% (Table 1 and Fig. 2). Given the reduction in mortality among women in the nonscreening group, as compared with their historical counterparts, the relative reduction among women in the screening group was 10% (95% CI, −4 to 24; P=0.13). Since the differences between the current groups and historical groups were 7.2 deaths per 100,000 person-years in the screening group and 4.8 deaths per 100,000 person-years in the nonscreening group, only the overall between-

Table 1. Rates of Death from Breast Cancer, According to Study Group and Age.*

Age Group and Mortality Data	Nonscreening Groups		Screening Groups		Difference		
	Historical Group	Current Group	Historical Group	Current Group	Nonscreening Groups†	Screening Groups‡	Nonscreening Groups vs. Screening Groups§
50–69 Yr							
No. of deaths	494	396	555	423			
No. of person-yr	1,898,989	1,866,741	2,197,469	2,337,323			
No. of deaths/100,000 person-yr	26.0	21.2	25.3	18.1	4.8	7.2	2.4±4.1
Rate ratio for death (95% CI)					0.82 (0.71–0.93)	0.72 (0.63–0.81)	0.10
20–49 Yr							
No. of deaths	238	183	332	267			
No. of person-yr	3,842,740	4,030,443	5,134,212	5,357,163			
No. of deaths/100,000 person-yr	6.2	4.5	6.5	5.0	1.7	1.5	–0.2±4.4
Rate ratio for death (95% CI)					0.73 (0.63–0.92)	0.77 (0.65–0.90)	–0.04
70–84 Yr							
No. of deaths	429	386	623	465			
No. of person-yr	1,101,019	1,173,624	1,349,967	1,318,004			
No. of deaths/100,000 person-yr	39.0	32.9	46.1	35.3	6.1	10.8	4.7±6.9
Rate ratio for death (95% CI)					0.84 (0.74–0.97)	0.76 (0.68–0.86)	0.08

* Only women between the ages of 50 and 69 years were invited to participate in screening mammography. All women in this group were also eligible for treatment by the multidisciplinary teams that are part of the screening program.

† For the nonscreening groups, the value shown is the difference between the rate of death in the historical group and that in the current group. This difference represents changes in mortality over time as a result of increased breast-cancer awareness, improved therapy, and more sensitive diagnostic tools.

‡ For the screening groups, the value shown is the difference between the rate of death in the historical group and that in the current group. This difference represents changes in mortality both over time and after introduction of the breast-cancer screening program.

§ For the comparison of the nonscreening groups with the screening groups, the value shown is the difference between the two rate-of-death differences. This value represents the effect of introducing the breast-cancer screening program. Plus–minus values are 95% confidence intervals.

group difference — 2.4 deaths per 100,000 person-years (95% CI, –1.7 to 6.5) — can be attributed to the screening program alone, representing a third of the total estimated reduction in mortality (2.4 of 7.2).

Among women between the ages of 50 and 69 years in the screening group, those with stage I tumors had a relative reduction in mortality of 16%, as compared with their historical counterparts (rate ratio, 0.84; 95% CI, 0.63 to 1.11); among women in the nonscreening group, the corresponding reduction was 13% (rate ratio, 0.87; 95% CI, 0.62 to 1.23). Among women with stage II tumors, those in the screening group had a marked 29% reduction in mortality, as compared with their historical counterparts (rate ratio, 0.71; 95% CI, 0.58 to 0.86); among women in the nonscreening group, the reduction was 7% (rate ratio, 0.93; 95% CI, 0.76 to 1.12). Among women with stage III or IV tumors, the improvement in prognosis was similar with and without the screening program (rate ratio for death in both groups, 0.70; 95% CI, 0.57 to 0.86 for the screening group and 0.56 to 0.87 for the nonscreening group).

Among women who were not eligible for screening because they were younger than 50 years of age or older than 69 years of age, there was also a significant reduction in the rate of death from breast cancer, as compared with their historical counterparts (Table 1). Women in these age groups who were in the screening group but were not eligible for the screening program had the benefit of the multidisciplinary breast-cancer management teams. Among women under the age of 50 years, there was a nonsignificant relative increase in mortality of 4% (P=1.00) after the introduction of the

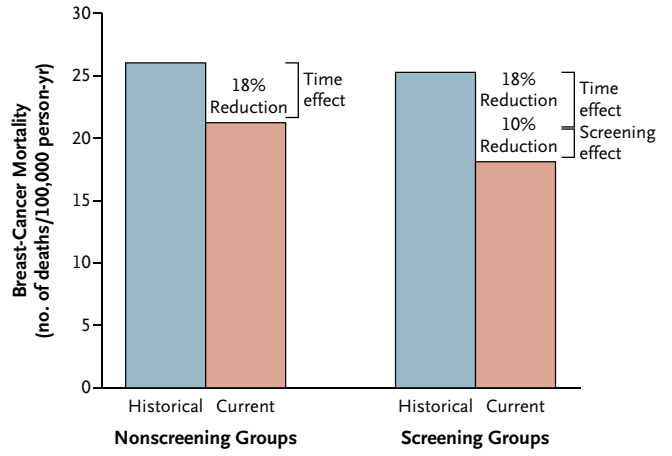

Figure 2. Rates of Death among Women between the Ages of 50 and 69 Years in the Four Study Groups.

Among women in the nonscreening group, there was an 18% reduction in the rate of death from breast cancer, as compared with the preceding 10-year period, presumably as a result of increased breast-cancer awareness, improved therapy, and the use of more sensitive diagnostic tools. Among women in the screening group, there was a 28% reduction in mortality from breast cancer during the same period. Thus, the relative reduction in mortality that was causally related to the screening program alone was 10%.

screening program (Table 1). Among women who were 70 years of age or older, the relative reduction in mortality of 8% (P=0.09) could be attributed to the establishment of multidisciplinary teams in the screening program (Table 1 and Fig. 3).

DISCUSSION

In our study, the rate of death from breast cancer was reduced by the introduction of a breast-cancer screening program. However, when we took into account temporal trends in breast-cancer mortality caused by other factors, the apparent effect was considerably smaller than expected. Indeed, the take-home message is that breast-cancer screening was associated with an absolute reduction of 10 percentage points in the rate of death from breast cancer. However, the screening program accounted for only one third of the total reduction in mortality among women who were invited to participate in the program. For women between the ages of 50 and 69 years, it was impossible to determine whether the reduction in mortality resulted from earlier diagnoses associated with screening mammography or from the management of treatment by an interdisciplinary team. To our surprise, the reduction in breast-cancer mortality among women between the ages of 70 and 84 years was largely the same as that in the screening group. Although none of the older women were invited to undergo mammography, they were all treated by multidisciplinary teams specializing in breast-cancer care.

The fundamental prerequisite for our analysis was the staggered implementation of the Norwegian Breast Cancer Screening Program. This structure provided the opportunity to identify a nonscreening group in order to reduce or perhaps eliminate confounding as a result of temporal changes in breast-cancer mortality attributable to factors other than screening. Additional strengths of our study include its nationwide design, the large size, the high proportion of women participating in the screening program (77%), and the complete follow-up. The incidence-based approach for calculating rates of death also reduced the likelihood that results were obscured by deaths from breast cancers that were diagnosed before the screening program was implemented.

Is it possible that the lead time created a bias in calculating incidence-based mortality? We counted the rate of death from breast cancer only if the death and diagnosis occurred in that group. For example, in the screening group, a death would be attributed to breast cancer only if the disease was diagnosed early by means of screening mammography or if the disease was clinically diagnosed while the woman was in the group. However, for women in whom an early diagnosis was made at screening and who later died of breast cancer, the diagnosis would have been made clinically at an unknown time within the study period. Thus, the lead time plays no role in the calculation of the rate of death, and we believe that the mortality calculations for all groups are free of this bias.

Our study also has limitations. First, the maximum follow-up time of 8.9 years may be too short to show the full potential of the screening program. However, in randomized, controlled trials, there was a reduction in mortality after 4 years, with an increasing effect up to 10 years.[20] In our study, the reduction in mortality was seen mainly in the first 4 years of follow-up (data not shown). Second, since the screening program was implemented gradually in the counties, diagnoses were

made more recently in the screening group than in the nonscreening group (Fig. 1) and there may be an overestimation of the mortality benefit associated with the screening program. Third, some of the women in the nonscreening group may have actually undergone mammography (opportunistic screening), potentially resulting in an underestimation of the benefit of screening. Unfortunately, we have no precise information about the numbers of such examinations. However, several circumstances provide reassuring evidence against contamination by opportunistic screening as an important source of bias. Before the implementation of the screening program, access to mammography was limited, especially in the predominantly rural areas of the country, and the reduction in mortality was of similar magnitude in urban and rural areas (data not shown). Also, the public health care system provides no financial incentives for offering screening mammography. Finally, the organized screening mammography entailed a substantial increase in diagnosed cases of breast cancer, with no similar trends in counties before they joined the program.

Our finding that only about one third of the reduction in mortality can be directly attributed to breast-cancer screening is in line with evidence from the National Health Service screening program in the United Kingdom.[21] Other studies have shown a relative reduction in the rate of death from breast cancer of 6.4 to 25% with follow-up periods of 10 years or less.[18,19,21-25] However, most of these studies have compared current breast-cancer mortality with mortality in a period preceding the introduction of screening mammography, with no ability to account for the confounding effect of temporal trends.[18,21,23-25] As our data show, such confounding may entail a considerable overestimation of the mortality benefit of mammography.[23-25]

The implementation of multidisciplinary breast-cancer management teams was intended to provide comprehensive and integrated optimization of breast-cancer care. As a corollary, it is not possible to attribute the reduction in mortality to any specific component of such a change in health care, although increased breast-cancer awareness, higher sensitivity of diagnostic techniques, and improvements in treatment can all be conducive to a lower rate of death. The greatest reduction in the death rate associated with mammography was observed among women with stage II tumors. This finding might be explained by selective stage migration among screening participants[26] as a result of more sensitive staging techniques (including the use of sentinel-node biopsy, which increased from virtually no use in 1998 to a 65% rate of use in 2004[15]) and improvements in treatment.

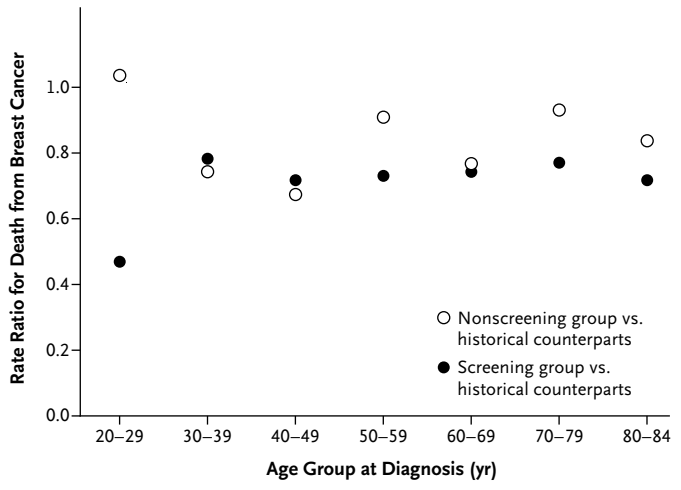

Figure 3. Incidence-Based Rate Ratios for Death from Breast Cancer, According to Age Group.

Shown are the differences in breast-cancer mortality among women living in counties in which breast-cancer screening had been implemented, as compared with their historical counterparts, and corresponding values for women living in counties in which screening had not been implemented, as compared with their historical counterparts. Only women between the ages of 50 and 69 years were invited to participate in mammographic screening.

We conclude that our results support the evidence that screening mammography reduces the rate of death from breast cancer. However, the magnitude of this benefit seems modest in the high-attendance, nationwide screening program we evaluated. Most important, the apparent benefit conveyed by optimized patient care may be missed unless breast-cancer screening is integrated into a well-functioning health care system that is available to the entire population.

Supported by the Cancer Registry of Norway and the Research Council of Norway.

No potential conflict of interest relevant to this article was reported.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank Michael Bretthauer, M.D., Ph.D., of Oslo University Hospital for valuable help with editing and scientific discussion during the preparation of the manuscript.

REFERENCES

1. Chu KC, Smart CR, Tarone RE. Analysis of breast cancer mortality and stage distribution by age for the Health Insurance Plan clinical trial. J Natl Cancer Inst 1988;80:1125-32.
2. Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years. CMAJ 1992;147:1477-88. [Erratum, CMAJ 1993; 148:718.]
3. Nyström L, Rutqvist LE, Walls S, et al. Breast cancer screening with mammography: overview of Swedish randomised trials. Lancet 1993;341:973-8. [Erratum, Lancet 1993;342:1372.]
4. Vainio H, Bianchini F, eds. IARC handbook of cancer prevention. Vol 7. Breast cancer screening. Lyon, France: IARC Press, 2002.
5. Gøtzsche PC, Nielsen M. Screening for breast cancer with mammography. Cochrane Database Syst Rev 2006;4: CD001877.
6. Nyström L. How effective is screening for breast cancer? BMJ 2000;321:647-8.
7. Berry DA, Corin KA, Plevritis SK, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. N Engl J Med 2005;353:1784-92.
8. Consumer report 4/2002. Oslo: The Consumer Council of Norway. (In Norwegian.) (Accessed August 30, 2010, at http://forbrukerportalen.no/filearchive/4_2002_s_1_48.pdf.)
9. Larsen IK, Småstuen 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.
10. Tingulstad S, Haldorsen T, Norstein J, Hagen B, Skjeldestad FE. Completeness and accuracy of registration of ovarian cancer in the Cancer Registry of Norway. Int J Cancer 2002;98:907-11.
11. Kalager M, Haldorsen T, Bretthauer M, Hoff G, Thoresen SO, Adami HO. Improved breast cancer survival following introduction of an organized mammography screening program among both screened and unscreened women: a population-based cohort study. Breast Cancer Res 2009;11:R44.
12. Perry N, Broeders M, de Wolf C. European guidelines for quality assurance in breast cancer screening and diagnosis. 2nd ed. Luxembourg: Office for Official Publications of the European Communities, 1996.
13. Erzaas A. Quality assurance manual of the Norwegian Breast Cancer Screening Program. Oslo: The Cancer Registry of Norway, 2003. (In Norwegian.) (Available at http://www.kreftregisteret.no.)
14. Wang H, Kåresen R, Hervik A, Thoresen S. Mammography screening in Norway: results from the first screening round in four counties and cost effectiveness of a modeled nationwide screening. Cancer Causes Control 2001;12:39-45.
15. Kalager M, Kåresen R, Wist E. Survival after breast cancer — differences between Norwegian counties. Tidsskr Nor Laegeforen 2009;129:2595-600. (In Norwegian.)
16. Statistics Norway home page. (Accessed August 30, 2010, at http://statbank.ssb.no/statistikkbanken/Default_FR.asp?PXSid=0&nvl=true&PLanguage=0&tilside=selecttable/MenuSelS.asp&SubjectCode=02.)
17. The 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.
18. Tabar L, Duffy SW, Yen MF, et al. All-cause mortality among breast cancer patients in a screening trial: support for breast cancer mortality as an end point. J Med Screen 2002;9:159-62.
19. Hakama M, Pukkala E, Heikkilä M, Kallio M. Effectiveness of the public health policy for breast cancer screening in Finland: population based cohort study. BMJ 1997;314:864-7.
20. Nyström L, Andersson I, Bjurstam N, Frisell J, Nordenskjöld B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomized trials. Lancet 2002;359:909-19. [Erratum, Lancet 2002;360:724.]
21. Blanks RG, Moss SM, McGahan CE, Quinn MJ, Babb PJ. Effect of NHS breast screening programme on mortality from breast cancer in England and Wales, 1990-8: comparison of observed with predicted mortality. BMJ 2000;321:665-9.
22. Olsen AH, Njor SH, Vejborg I, et al. Breast cancer mortality in Copenhagen after introduction of mammography screening: cohort study. BMJ 2005;330:220.
23. Tabar L, Yen MF, Vitak B, Chen HHT, Smith R, Duffy SW. Mammography service screening and mortality in breast cancer patients: 20-year follow-up before and after introduction of screening. Lancet 2003; 361:1405-10.
24. Otto SJ, Fracherboud J, Looman CWN, 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.
25. Jonsson H, Nyström L, Törnberg S, Lenner P. Service screening with mammography of women aged 50-69 years in Sweden: effects on mortality of breast cancer. J Med Screen 2001;8:152-60.
26. Feinstein AR, Sosin DM, Wells CK. The Will Rogers phenomenon: stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. N Engl J Med 1985;312: 1604-8.

EDITORIALS

Screening Mammography — A Long Run for a Short Slide?

H. Gilbert Welch, M.D., M.P.H.

No screening test has ever been more carefully studied than screening mammography. In the past 50 years, more than 600,000 women have participated in 10 randomized trials, each involving approximately 10 years of follow-up. Given this extraordinary research effort, it is ironic that screening mammography continues to be one of the most contentious issues within the medical community.

The juxtaposition of such a charged medical debate in the face of such an exhaustive scientific investigation is in itself instructive. For context, one trial involving fewer than 150 men who were followed for less than 2 years was sufficient to convince physicians of the value of treating severe hypertension.[1] That physicians are still debating the relative merits of screening mammography despite the wealth of data suggests that the test is surely a close call, a delicate balance between modest benefit and modest harm.

In this issue of the *Journal*, Kalager et al.[2] provide additional data that the benefit of mammography is modest. Making use of the opportunity provided by the staggered implementation of a national screening program in Norway, the investigators were able to isolate the benefit of the screening program from other factors that may have changed over time, including increased breast-cancer awareness and improvements in treatment. They report that the benefit of the Norwegian screening program was disappointingly small: a 10% reduction in breast-cancer mortality among women between the ages of 50 and 69 years.

Moreover, this reduction in mortality reflected the combined effect of the two interventions that make up the Norwegian screening program: screening mammography and multidisciplinary teams instituted to better treat breast cancer. Kalager et al. provide data that the latter may be the more important of the two factors, since women over the age of 70 years, who were exposed to the program's multidisciplinary teams but were not invited to undergo mammography, had an 8% reduction in breast-cancer mortality. Thus, the relative reduction in mortality due to screening mammography alone could be as low as 2%.

Clinicians who follow the mammography debate will reasonably wonder why the benefit estimated by Kalager et al. is so much smaller than the reduction in mortality of 15 to 23% estimated by the U.S. Preventive Services Task Force.[3] The easiest explanation would be that the Kalager estimate is wrong. Although the task force uses data from randomized trials, the Norwegian data are observational — and as with all observational data, the primary threat to validity is the comparability of the comparison groups.

But the staggered cohort design that was used by Kalager et al. mitigates the concern that the women in the four study groups are somehow different, since many of the women in the study actually contributed data to each group at different points in their life. Contamination is a more relevant concern. If the women in the nonscreening groups were exposed to opportunistic mammography screening or began to benefit from the multidisciplinary teams, which had to be in place before the screening program was initiated, then the background effect of time may have been overestimated. This would have led to an underestimation of the benefit of the screening program. Furthermore, the follow-up period may be too short to fully capture the benefits of screening. The authors argue that these effects are small.

So another explanation must be considered:

the estimates of both the task force and Kalager et al. are correct. But where the randomized trials reflect the world before 1990, the observational data reflect the world after 1990. It is quite plausible that screening mammography was more effective in the past than it is now. If women with new breast lumps now present earlier for evaluation, the benefit of screening will be less. If treatment of clinically detected breast cancer (i.e., tumors that are detected by means other than screening) has now improved, the benefit of screening will be less. Thus, the increased awareness about the importance of promptly seeking care for overt breast abnormalities (there is no debate about diagnostic mammography) and the widespread use of adjuvant therapy have probably combined to make screening now less important.[4,5]

Nevertheless, the public widely perceives screening mammography to be one of the most important services provided by modern medicine. The perception is largely the product of well-crafted public health messaging, such as the American Cancer Society's print campaign in the 1980s that featured the headline "If you haven't had a mammogram, you need more than your breasts examined." Given current data, such messaging must become more balanced.

If we assume that mammography screening is associated with a 10% reduction in the rate of death from breast cancer (making the optimistic assumption that all the benefit comes from screening mammograms), the 10-year risk of breast-cancer death for a 50-year-old woman in the United States is now about 4 per 1000 women.[6] If we assume that this risk already incorporates the benefit of screening mammography, the risk estimate without mammography would be about 4.4 per 1000 women.

Because we are all subject to framing effects, it is important to consider the reverse frame. The number of women who will not die from breast cancer rises from 995.6 to 996 per 1000 women with the addition of screening mammography. Although readers may each respond differently to these frames, both reflect the same absolute benefit: 0.4 per 1000 women. In other words, 2500 women would need to be screened over a 10-year period for 1 to avoid death from breast cancer (Table 1).

What happens to the other 2499 women who had to undergo screening to achieve this benefit is also relevant. Estimates of harm vary considerably. In the United States, more than 1000 women would be expected to have at least one false positive result,[7] a number that would be considerably lower in Europe.[8] Less frequent but more worrisome is the problem of overdiagnosis. Somewhere between 5 and 15 women would be expected to be needlessly treated for a condition that was never going to bother them, with all the accompanying harms.[9,10]

Table 1. Estimated Benefits and Harms Associated with a 10-Year Course of Screening Mammography for 2500 Women Who Are 50 Years of Age.*

Benefit	Harm
One woman will avoid dying from breast cancer.	Up to 1000 women will have at least one "false alarm," about half of whom will undergo biopsy.
	Breast cancer will be overdiagnosed in 5 to 15 women, who will be treated needlessly with surgery, radiation, chemotherapy, or a combination.

* The assumed benefit of screening mammography is a reduction of 10% in the rate of death from breast cancer, as reported by Kalagar et al.[2]

Screening mammography has become one of the most prominent measures of health care performance. Since the inception of health care report cards, such evaluations have focused on ensuring that all women undergo the test.[11] There were practical reasons for this: it was easily measured, easy to understand, and hard to argue against. But by highlighting that the mortality benefit is modest, Kalager et al. help confirm that the decision about whether to undergo screening mammography is, in fact, a close call. Many observers will argue that because it is a delicate decision — involving trade-offs among noncomparable outcomes — it must be left to informed individuals to decide. Others will argue that physicians should continue to persuade women to undergo screening and that the modest benefit is worth the associated harms.

But no one can argue that screening mammography is one of the most important services we provide in medicine. The time has come for it to stop being used as an indicator of the quality of our health care system.

Disclosure forms provided by the author are available with the full text of this article at NEJM.org.

From the Dartmouth Institute for Health Policy and Clinical Practice, Dartmouth Medical School, Lebanon, NH.

1. Effects of treatment on morbidity in hypertension: results in patients with diastolic blood pressures averaging 115 through 129 mm Hg. JAMA 1967;202:1028-34.
2. Kalager M, Zelen M, Langmark F, Adami H-O. Effect of

screening mammography on breast-cancer mortality in Norway. N Engl J Med 2010;363:1203-10.
3. Mandelblatt JS, Cronin KA, Bailey S, et al. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. Ann Intern Med 2009;151:738-47.
4. Rostgaard K, Vaeth M, Rootzén H, Lynge E. Why did the breast cancer lymph node status distribution improve in Denmark in the pre-mammography screening period of 1978-1994? Acta Oncol 2010;49:313-21.
5. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 2005;365:1687-717.
6. The risk of developing/dying from cancer based on November 2009 SEER data submission. Bethesda, MD: National Cancer Institute. (Available at http://seer.cancer.gov/faststats/.)
7. Elmore JG, Barton MB, Moceri VM, Polk S, Arena PJ, Fletcher SW. Ten-year risk of false positive screening mammograms and clinical breast examinations. N Engl J Med 1998;338:1089-96.
8. Breast cancer screening. In: IARC handbooks of cancer prevention. Vol. 7. Lyon, France: International Agency for Research on Cancer, 2002:135-44.
9. Zackrisson S, Andersson I, Janzon L, Manjer J, Garne JP. Rate of over-diagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study. BMJ 2006;332:689-92.
10. Jørgensen KJ, Gøtzsche PC. Content of invitations for publicly funded screening mammography. BMJ 2006;332:538-41.
11. McIntyre D, Rogers L, Heier EJ. Overview, history, and objectives of performance measurement. Health Care Financ Rev 2001;22(3):7-21.

Superficial Phlebitis and Phase 3.5 Trials

Lee Goldman, M.D., and Jeffrey Ginsberg, M.D.

In this issue of the *Journal*, Decousus et al.[1] report on the efficacy and safety of fondaparinux for the treatment of superficial-vein thrombosis in the legs. The results of their carefully conducted, placebo-controlled trial show that treatment with fondaparinux, at a dose of 2.5 mg once daily for 45 days, as compared with placebo, reduced the probability that superficial-vein thrombosis in the legs would progress to deep-vein thrombosis or pulmonary embolism (1.3% with placebo vs. 0.2% with fondaparinux), without an increase in bleeding or other serious adverse events. The probability that patients would undergo surgery for superficial-vein thrombosis was reduced from 3.8% to 0.7%. Two patients in the fondaparinux group and one in the placebo group died, but none of the deaths were apparently the result of a pulmonary embolism. This study adds to previous work describing the natural history of superficial-vein thrombosis,[2-5] although it did not address which patients might be at an increased risk because of previously undiagnosed thrombophilia.[4,5]

To put the rates of deep-vein thrombosis and pulmonary embolism — the most important outcomes — into perspective, it is useful to consider the generally "acceptable" failure rates in strategies to diagnose venous thromboembolism. In the study by Decousus et al., the rate at which symptomatic deep-vein thrombosis or pulmonary embolism developed in untreated patients during follow-up (1.3%) was similar to the rate with widely accepted strategies for diagnosing deep-vein thrombosis and pulmonary embolism. For example, among patients who are evaluated for suspected deep-vein thrombosis but have normal results on a contrast venogram[6] or duplex ultrasonography,[7] about 1.3% and 0.6% of patients, respectively, will return with symptomatic deep-vein thrombosis or pulmonary embolism over the course of long-term follow-up. Similarly, among patients who have a suspected pulmonary embolism but then have normal results on a conventional pulmonary angiogram[8] or a computed tomographic pulmonary angiogram,[9] about 1.7% and 1.2%, respectively, will return with symptomatic deep-vein thrombosis or pulmonary embolism. These historical comparisons and the extremely low mortality among untreated patients with superficial-vein thrombosis support an initial "no anticoagulant treatment" approach, unless conservative measures fail to resolve symptoms or deep-vein thrombosis develops. It is also clear from the stringent inclusion and exclusion criteria in the study by Decousus et al. that treatment with fondaparinux for 45 days is clinically reasonable for patients with severe symptoms, thrombosis in the proximal saphenous vein, or recurrent disease.

Agents such as fondaparinux, low-molecular-weight heparins, and perhaps oral direct factor Xa inhibitors (apixaban, rivaroxaban) and thrombin inhibitors (dabigatran) have better risk profiles than do unfractionated heparin and warfarin, and the favorable risk-to-benefit ratio associated with them could lead to an extension

BMC Cancer

BioMed Central
The Open Access Publisher

An Investigation of the Apparent Breast Cancer Epidemic in France: Screening and incidence trends in birth cohorts

BMC Cancer 2011, **11**:401 doi:10.1186/1471-2407-11-401

Bernard Junod (junod.bernard@wanadoo.fr)
Per-Henrik Zahl (Per-Henrik.Zahl@fhi.no)
Robert M. Kaplan (rmkaplan@ucla.edu)
Jorn Olsen (JO@SOCI.AU.DK)
Sander Greenland (Sandelesdomes@ucla.edu)

ISSN 1471-2407

Article type Research article

Submission date 9 May 2011

Acceptance date 21 September 2011

Publication date 21 September 2011

Article URL http://www.biomedcentral.com/1471-2407/11/401

MS 1316613618549646

An Investigation of the Apparent Breast Cancer Epidemic in France: Screening and incidence trends in birth cohorts

Bernard Junod[1], Per-Henrik Zahl[2], Robert M Kaplan[3], Jørn Olsen[4], Sander Greenland[4,5].

[1] Current address: FORMINDEP, Roubaix, France. Previous position: Department of Epidemiology, Ecole des Hautes Etudes en Sante Publique Rennes, France.

[2] Norwegian Institute of Public Health, Oslo, Norway

[3] UCLA Schools of Public Health and Medicine, Los Angeles, USA

[4] Department of Epidemiology, UCLA School of Public Health, Los Angeles, USA

[5] Department of Statistics, UCLA College of Letters and Science, Los Angeles, USA

Corresponding author:

Bernard Junod

La Maison Neuve

35520 La Chapelle des Fougeretz

France

00 33 2 99 66 43 25

E-mail addresses:

Junod.bernard@wanadoo.fr

Per-Henrik.Zahl@fhi.no

rmkaplan@ucla.edu

jo@ucla.edu

lesdomes@ucla.edu

Abstract

Background: Official descriptive data from France showed a strong increase in breast-cancer incidence between 1980 to 2005 without a corresponding change in breast-cancer mortality. This study quantifies the part of incidence increase due to secular changes in risk factor exposure and in overdiagnosis due to organised or opportunistic screening. Overdiagnosis was defined as non progressive tumours diagnosed as cancer at histology or progressive cancer that would remain asymptomatic until time of death for another cause.

Methods : Comparison between age-matched cohorts from 1980 to 2005. All women residing in France and born 1911-1915, 1926-1930 and 1941-1945 are included. Sources are official data sets and published French reports on screening by mammography, age and time specific breast-cancer incidence and mortality, hormone replacement therapy, alcohol and obesity. Outcome measures include breast-cancer incidence differences adjusted for changes in risk factor distributions between pairs of age-matched cohorts who had experienced different levels of screening intensity.

Results: There was an 8-fold increase in the number of mammography machines operating in France between 1980 and 2000. Opportunistic and organised screening increased over time. In comparison to age-matched cohorts born 15 years earlier, recent cohorts had adjusted incidence proportion over 11 years that were 76% higher [95% confidence limits (CL) 67%, 85%] for women aged 50 to 64 years and 23% higher [95% CL 15%, 31%] for women aged 65 to 79 years. Given that mortality did not change correspondingly, this increase in adjusted 11 year incidence proportion was considered as an estimate of overdiagnosis.

Conclusions: Breast cancer may be overdiagnosed because screening increases diagnosis of slowly progressing non-life threatening cancer and increases misdiagnosis among women without progressive cancer. We suggest that these effects could largely explain the reported “epidemic” of breast cancer in France. Better predictive classification of tumours is needed in order to avoid unnecessary cancer diagnoses and subsequent procedures.

Background

Between 1980 and 2005, age-standardized cancer incidence in France increased by 38%, primarily due to increased reported prostate cancer incidence in men and breast and lung cancer among women [1]. The case-fatality rate of breast cancer estimated from incidence and mortality decreased from 39% in 1980 to 23% in 2005. The increase in breast cancer incidence may be related to increasing exposure to causal factors, such as use of hormone replacement therapy (HRT), alcohol, obesity and change in family size, but may also be an artefact of increased screening.

Reports from the International Agency for Research on Cancer (IARC) and from the French National Institute for Health Research (INSERM) considered the distinction between real and artificial increases in cancer frequency in France by emphasizing mortality data over incidence data [2,3]. When comparing the trends between cancer sites, the IARC report hypothesised that the net impact of early detection methods is to increase reported cancer incidence independently of environmental or lifestyle risk factors. Figure 1 shows breast cancer incidence and breast cancer mortality for the period 1980 to 2005, revealing a substantial discrepancy. If the true incidence in breast cancer was not increasing over time, both screening and improvements in treatment should have substantially reduced breast-cancer mortality.

The goal of breast-cancer screening (testing for the disease in asymptomatic patients) is to reduce mortality by diagnosing and treating tumours earlier in the disease process. Initially, screening programs will increase rates of cancer diagnosis because prevalent tumours are detected earlier. After the introduction of screening, when the reservoir of undiagnosed cases is depleted, a decline of incidence is

expected before a new steady state is achieved [4]. However, recent papers suggest that publicly available mammography screening programs are associated with 10% to 50 % overdiagnosis [5,6], where overdiagnosis is defined as the detection, through screening, of disease that would never have been diagnosed in the absence of screening and thus unlikely to have imposed health consequences throughout life [7]. Increase in screening activity also occurs without organized screening program. For example, after careful modelling, overdiagnosis was over 40% for the younger cohorts that had been exposed to mammograms in Catalonia [8].

A Norwegian study suggested that mammography screening leads to a larger increase in detected invasive breast cancer than can be explained by earlier diagnosis or increased exposure to risk factors. The authors suggested that mammography screening detects many tumours that otherwise would spontaneously regress [9].

Most breast cancers are diagnosed by biopsy following identification by self palpation, clinical examination by a physician, or by mammography. Overdiagnosis is inevitable when testing for asymptomatic disease in almost all screening programs. Clinicians use histology for diagnosing a true progressive disease that would metastasise and cause death without treatment if no other health problem interfered with its progression. The validity of testing for true progressive cancer by histology depends on the sensitivity and the specificity of slides from the biopsy. The number of diagnosed cancer cases in an examined population is the sum of women with progressive cancer correctly diagnosed and of women diagnosed with a cancer that would not progress to clinical detection in their lifetime. The number of true progressive cancers detected in a population reflects the frequency of examinations

among women with progressive cancer, the sensitivity of diagnosis procedures before the biopsy, and the sensitivity of examination by histology.

Global sensitivity is the proportion of progressive cancers correctly identified in a population. All nonprogressive tumours diagnosed as cancer by histology are overdiagnoses. They reflect the frequency of examinations among women without a progressive cancer, the specificity of diagnostic procedures before the biopsy, and the proportion of women without a progressive cancer correctly identified when examined by histology. All the cancer-free women not tested contribute to increase *global specificity*: the proportion of women without a true progressive cancer correctly considered as cancer free in the population. Screening increases global sensitivity. But by doing this, it also results in decreasing global specificity, which in turn produces more overdiagnosis.

Overdiagnosis includes all nonprogressive tumours diagnosed as cancer using histology and those progressive cancers that would never cause symptoms or death during a patient's lifetime. Such cases are *functional* overdiagnoses related to a patient's outcome rather than to the physiological or structural causes of overdiagnosis. Functional overdiagnosis depends not only on the cancer but also on competing causes of death and life expectancy. It occurs more frequently when screening is performed among women with a short remaining life expectancy and when global sensitivity is high.

Our study investigates how the increase in mammography screening is associated with increase in the apparent breast-cancer incidence in France. Such information is relevant to the debate about the benefits and side effects of breast-cancer screening [10-15].

Methods

This investigation is restricted to 1980 - 2005 and to French districts of metropolitan France, the European part of the country. Districts of metropolitan France is an English translation for “départements de la France métropolitaine”.

Data

Breast cancer deaths and the populations of women were provided by the Center for Epidemiology of Medical Causes of Death (CepiDC) [16]. The annual number of newly diagnosed invasive breast cancer, and the population of women in France were used to estimate the time trend of breast-cancer frequency. Diagnoses of invasive breast cancer were estimated from population-based cancer registries operating in France [17]. For year 1992 (the middle of the study period) the national estimate was based on 2193 reported cases [18]. Incidence of invasive cancer was provided up to 2005 by the French Institute for Public Health Surveillance [19].

Age and time specific exposure to HRT [20,21], alcohol [22], and obesity [23] in France were estimated from published data. For HRT, age specific prevalence was based on two cohorts available in France: “ESPS-EPAS” (sample from the social security registry) and “3C” (women from Bordeaux, Dijon and Montpellier). The relative risk estimates were obtained from four models used in the report [20]. Each model takes into account three types of HRT use: oestrogen only, oestrogen plus progesterone, and oestrogen plus progestin. Women were considered exposed to alcohol if they drank at least 6 glasses or more on one occasion and/or at least 14 glasses a week. Obesity was defined by a body mass index equal to or larger than 30 kg/m^2. For alcohol and obesity, relative risk estimates were based on international literature [24].

Changes in diagnostic procedure

We used two data sources to evaluate changes in mammography practice. First, we estimated the number of mammography machines registered annually in France, using the same methods from the 1970s to 2000 [25]. The mean number of screening tests performed per mammography machine is available for 1988 [26]. Second, we estimated the implementation of organised breast-cancer screening programs by mammography in France up to 2004 [27]. For two districts we also had the age distribution of women undergoing mammography when tested either by an organised program or by private initiative in 1995 [28].

Change in incidence due to change in risk factor exposure

Breast-cancer incidence attributable to change in exposure to risk factors over time is obtained in age-specific categories. It is computed from incidence in the reference period, available exposure prevalence for each period, and from relative risk estimates. Additional file 1 provides formulas used for these calculations.

Overdiagnosis estimate from change in incidence and breast-cancer mortality

Change in incidence proportion was obtained by comparing age-matched cohorts 15 years apart. Comparisons of incidence within a pair of cohorts submitted to different screening activity were performed separately for middle age women aged 50 to 64 and for elderly women aged 65 to 79. In each pair, the reference cohort was observed at an earlier calendar period when screening activity was less intense. The reference cohort was observed 15 years before the comparison cohort for both age groups of women. For middle-aged women, the reference cohort included women born between 1926 and 1930 and was compared to the cohort of women born in 1941-1945. For elderly women, the reference cohort included women born in 1911-

1915 and was compared to the 1926-1930 birth cohort. For both middle age and elderly women, incidence was observed yearly for 5-years age groups. Change in crude incidence proportion associated with a 15 year change in screening activity is the difference between 11 years incidence proportion within each pair of cohorts. The detailed computation of incidence proportion is given in additional file 2. The same procedure was used for breast-cancer mortality.

Within each pair of cohorts, incidence attributable to change in risk factor prevalence and to change in mortality proportion, if any, was subtracted from crude incidence proportion to get an estimate of overdiagnosis between the two cohorts in the comparison.

Statistical Methods

Confidence limits (CL) were obtained from a normal approximation to the distribution of proportions for the comparison of initial procedure leading to breast-cancer diagnosis over time. Confidence limits for differences between incidence proportions were obtained using French official data and observed cases in French cancer registries operating in 1992 [18]. Confidence limits were not calculated for estimates of the full population. See additional file 2 for formulas.

Results

Time trend for availability and use of mammography screening

The number of mammography machines increased steadily from 308 in 1980, 499 in 1984, 1351 in 1990, 2282 in 1994 to 2511 in 2000. There were about 8 times more mammography machines available in 2000 than in 1980.

There were three districts with an organised screening program in 1989, 13 in 1994 and 31 in 1999. In 2004, all 96 districts had an organised screening program.

Screening began at age 50, and in 1999, the upper age limit for inviting women to be screened every second year was extended from 69 up to 74 years of age. During the whole period, screening practices were not restricted to the women included in organised programs.

In two districts with an organised program in 1995, the mammography rate before 50 or after 69 years of age was 59% of the mammography rate in the organized screening program for women aged 50 to 69 [28]. In 1988, the mean number of screening mammography per mammography machine amounted to 1050 per year [26,28].

Time trend of exposure to risk factor

Changes in exposure to risk factors are summarised in Table 1. In comparison to 1980-1990, there was an increased prevalence of HRT use and obesity by 1995-2005, whereas alcohol consumption in women decreased.

Changes over time in breast-cancer incidence and breast-cancer mortality

Figure 2 shows the age-specific increase of breast-cancer incidence over time. The largest increase occurred for women 45 to 74 years of age. For women aged 50 to 69, the incidence in 2005 is twice the incidence in 1980. The largest increase occurred in 2005 for the 60 to 64 age group. In 2005, the breast-cancer incidence decreased after age 74 compared to rates for women aged 60-69; the shape of the age-specific incidence rate changed from being non-declining with age to being bell shaped.

In cohorts that had more intensive breast screening, we might expect a reduction of breast-cancer incidence after age 74 since slow growing tumours should

have been detected but this was not seen. Within each pair of cohorts, the observed increase in incidence is even larger at the end of the period of comparison than 11 years before. This is visible in figure 3 for age groups 65-69 to 75-79. Age-specific breast-cancer mortality was similar in the two pairs of cohorts. In the pair of cohorts of middle age women (50 to 64), the cumulative breast-cancer mortality rates were 6.7/1000 from 1980 to 1990 and 6.6/1000 from 1995 to 2005. In the cohorts of elderly women (65-79), the cumulative breast-cancer mortality rates were 9.9/1000 from 1980 to 1990 and 10.7/1000 from 1995 to 2005.

Estimates of overdiagnosis

Incidence rates observed in the cohorts are given in Figure 3. For women aged 50 to 64, the 11-years incidence proportion was 20/1000 in the reference cohort observed from 1980 to 1990. It increased 80% [95% CL: 72%, 89%] to 37/1000 in the age-matched cohort observed from 1995 to 2005. For women aged 65 to 79, the 11-year incidence proportion was 24/1000 in the reference cohort observed from 1980 to 1990. It increased 27% [95%CL 20%, 34%] to 31/1000 in the age-matched cohort observed from 1995 to 2005.

Overdiagnosis estimates take into account changes in incidence proportions, and adjustments due to change in risk factors prevalence, as given in Table 2. Given that breast-cancer mortality changed less than 0.1/1000 per year, and moved in opposite direction according to age group, it was not taken into account for adjustments. Overdiagnosis estimates are slightly lower than the crude difference between incidence proportions within each pair of cohorts. Adjustment for increase in HRT use and in obesity resulted in a slight reduction in the crude difference between incidence proportions. On the contrary, decrease in alcohol consumption contributed to a small increase in the overdiagnosis estimate for each considered pair.

Overdiagnosis estimates from 1995 to 2005 were 76% (95% CL: 67%; 85%) for women aged 50 to 64 and 23% [95% CL: 15%; 31%] for women aged 65 to 79 .

Discussion

We observed that standardised breast-cancer incidence rates increased steadily from 1980 to 2005 as use of screening tools increased, whereas age standardised breast-cancer mortality rates changed only slightly during this period. These trends might reflect a progressive increase in unknown breast cancer exposure and a decrease in case fatality due to better treatment. However, our results are consistent with other studies that fail to demonstrate a benefit of screening for breast cancer at the population level.

Opinion on the value of screening mammography remains divided. Many investigators, particularly from the radiology community, support population screening [10,12,29]. On the other hand, some but not all meta-analysis of randomised controlled trials fail to document survival benefits [14]. Meta analyses may come to different conclusions because they apply different study exclusion criteria. Those with stricter quality criteria tend to favour the null effects of screening, particularly for women younger than 50 years. However, some have argued that the choice in quality criteria is subjective [12] or due to assumptions [11]. Further, systematic quasi-experimental evaluations in Norway reported poor survival benefit in those screened [30]. Similar studies from Denmark suggest that the decline in breast cancer mortality was greater in regions without screening than it was in areas where screening was phased in earlier [31]. A recent study used WHO data to compare trends in breast cancer mortality in three pairs of European countries: Northern Ireland v. Republic of Ireland, Netherlands v. Belgium, and Sweden v. Norway. Although one member of each county pair had a more aggressive

screening policy, reductions in breast cancer mortality were similar for in all three pairs. These findings are consistent with clinical trials and other quasi-experiments that have failed to show significant reductions in mortality directly attributable to mammography [32].

Well-conducted screening programs should lead to an initial increase of both prevalent cases and lead time, and then to a subsequent decline of observed advanced tumours which was not observed. After 74 years of age, when women are not invited by organised screening programs, the incidence rate should not increase as much as when screening occurred previously in the cohort [33]. Nonetheless, compared to earlier ages, the increase in breast-cancer incidence was even larger for women aged 75-79 in 1995 than in 1990, when screening was less intense up to 74 years of age. The unexpected increase in cancer incidence in older women may reflect overdiagnosis due to greater screening [34-35].

The period of observation was chosen to ensure stability in the recording systems for both death, incidence and in the nomenclatures used. Before 1978, breast-cancer incidence was not available and breast-cancer mortality trend was biased by the fact that death due to breast cancer was also declared as “cancer” without specifying the site of the primary tumour. Statistics are available on the age frequency distribution of surgical interventions performed in 1999 for breast cancer in France [36]. Among women aged 50 to 79, there are 16.5% more interventions than incident cases of invasive cancer. This difference is consistent with inclusion of women operated more than once or for ductal carcinoma in situ.

Our study has several important limitations. Only a small proportion of the French population is included in cancer registries (about 7% in the middle of the study period). In addition, we can not rule out secular changes in other breast-cancer

risk factors like age at first birth, nulliparity or socioeconomic status. Adjustments for changes in HRT, alcohol and obesity prevalence over time are by nature imperfect.

Organised screening programs do not give an unbiased appraisal of actual screening activity in France: they do not include opportunistic screening, which is substantial [28]. An increase over time in the number of mammography machines in France is likely to explain changes of initial procedures leading to breast-cancer diagnosis, as shown in a study conducted in the district of Haute Vienne [37]. During 1986-1989, 80% of the cancers (298 of 372) were discovered by the patient, while this proportion fell to 52% (176 of 341) during 1997-1998. The difference between the two groups is 28.5% [95% CL 21.8%, 35.2%]. This reduction was primarily offset by the increase in the proportion of breast cancer discovered by mammography: 24.5% [95% CL 20.2%; 28.8%]. The shorter duration of the second period indicates increased frequency of breast-cancer diagnosis. This increase observed 10 years apart in the district of Haute Vienne is about 2/3 of the corresponding increase observed 15 years apart in cohorts from 50 to 79 years of age at the national level.

A 1% drop of global specificity would explain the observed increase in breast-cancer incidence in France. Suppose that among 1000 women, 4 have undetected true invasive breast cancer and 996 do not have invasive breast cancer. If the women are not examined, the four cases will eventually be diagnosed if they become symptomatic and the specificity is 100%. If these 1000 women undergo diagnostic procedures with a global sensitivity of 90% and a global specificity of 99%, we would get 90% of 4, that is 3.6 true positives; 99% of 996, that is about 986 true negatives; 10% of 4, or 0.4 false negative and 1% of 996, that is about 10 false positive. The positive predictive value among the 13.6 diagnosed “cancer” is thus about 3.6/13.6, less than 30%. This example illustrates how overdiagnosis may increase with

screening even without changes in the intrinsic sensitivity and specificity of each diagnostic procedure.

Changes in the studied risk factors did not explain much of the increase in breast-cancer incidence over time. The emergence of overdiagnosis as a possible explanation of the incidence trend is related to the long period over which screening intensity has been increasing.

An analogous divergence between the trends in incidence and mortality was observed from 1927 through 1947 in Canada. Confidence in the efficacy of Halsted's radical procedure contributed to increasing early screening by breast self-examination. McKinnon hypothesised that the limitations of diagnosis confirmed by histology, which is "fraught with uncertainty", explained all or part of the apparent improvement in survival of cancers screened at an early stage during this period [38].

Undetected invasive breast cancers exist among women at the time of death due to other causes. Welch and Black used autopsy studies to estimate the size of the "reservoirs" of ductal cancer in situ [39]. These studies also revealed undetected invasive breast cancer among women who died from other causes [40-43]. Other publications report an elevated frequency of slowly or non-progressing lesions [44], some likely to be misdiagnosed as invasive cancer [45]. It is therefore possible that opportunistic screening explains most of the excess of overdiagnosis before age 50 and after age 74. Similar breast-cancer mortality in the cohort observed from 1980 to 1990 and the cohort observed from 1995 to 2005 also indicates overdiagnosis as a possible explanation for the incidence increase from age 40 to 79.

The 2003 report of the French Cancer Commission furnishes a key to interpreting this increased incidence: "Overdiagnosis (diagnosis of tumours at the borderline of malignancy) constitutes a serious problem because it can artificially

increase the incidence of cancer and the result of treatment" [46]. In 2005, a discussion at the French Academy of Medicine suggested that the definition of cancer should change to include evidence of the progression of the tumour over time [47].

Conclusion

In summary, there has been a substantial increase in breast-cancer incidence in France without a corresponding increase in mortality. Although this could be explained by a perfect balance between an increased incidence and improved survival, we think the increased breast-cancer incidence observed in France since 1980 largely reflects an increase in overdiagnosis. The latter includes misdiagnoses and true cancer lesions that would not have had any impact on the health of the women during their normal lifetime. Better predictive classification of tumours is needed in order to avoid unnecessary cancer diagnoses and subsequent procedures.

Competing interests

The authors declare that they have no competing interests.

Authors'contributions

BJ: Choice of the topic of the study. Data gathering, data analysis and contribution to writing. PHZ: Definition of specific research objectives, study design and contribution to writing. RMK: Elaboration of the structure of each section and contribution to writing. JO: Choice of epidemiological methods and of appropriate data. Contribution to writing. SG: Choice of statistical methods and of appropriate analysis. Contribution to writing.

All authors read and approved the final manuscript.

Acknowledgement

To Barron Lerner for the commitment of his book, The Breast Cancer Wars, an illuminating promotion of medical ethics.

Supported in part by the National Institutes of Health, Office of Behavioural and Social Sciences Research.

References

[1] Guerin S, Doyon F, Hill C. The frequency of cancer in France in 2006, mortality trends since 1950, incidence trends since 1980 and analysis of the discrepancies between these trends. Bull Cancer 2009 Feb; 96(1): 51-7.

[2] World Health Oganization. Attributable causes of cancer in France in the year 2000. IARC Working Group Reports. Volume 3. WHO, Geneva, 2007.

[3] Cancers-environnement. Expertise collective. INSERM. Octobre 2008. 907 pages.

[4] Boer, WP, de Koning H, van Oortmarssen G. Extraincidence caused by mammographic screening. The Lancet 1994(343): 979.

[5] Zackrisson S, Andersson I, Janzon L, Manjer J, Garne JP. Rate of over-diagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study. BMJ 2006; 332: 689–92.

[6] Jørgensen KJ, Gøtzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. BMJ 2009; 339; b2587.

[7] Etzioni R, Penson DF, Legler JM, di Tommaso D, Boer R, Gann PH, Feuer EJ. Overdiagnosis due to prostate-specific antigen screening: Lessons from U.S. prostate cancer incidence trends. JNCI 2002; 94(13): 981-90.

[8] Martinez-Alonzo M, Vilaprinyo E, Marcos-Gragera R, Rue M. Breast cancer incidence and overdiagnosis in Catalonia (Spain). Breast Cancer Res. 2010;12(4):R58. Epub 2010 Aug 3.

[9] Zahl PH, Maehlen J, Welch HG. The natural history of invasive breast cancers detected by screening mammography. Arch Intern Med. 2008; 168(21): 2311-6.

[10] Freedman DA, Petitti DB, Robins JM. On the efficacy of screening for breast cancer. Int J Epidemiol 2004; 33: 43-55.

[11] Gøtzsche PC On the benefits and harms of screening for breast cancer. Int J Epidemiol 2004; 33: 56-64.

[12] Freedman DA, Petiti DB, Robins JM. Rejoinder. Int J Epidemiol 2004; 33: 69-73.

[13] Welch HG Screening mammography - A long run for a short slide? N Engl J Med 2010; 363(13): 1276-8.

[14] Gøtzsche PC, Nielsen M. Screening for Breast Cancer with Mammography. Cochrane Database Syst Rev. 2009 Oct 7;(4): CD001877.

[15] Welch HG, Black WC. Overdiagnosis in Cancer. JNCI 2010;102 : 605-13.

[16] Centre d'epidemiologie sur les causes medicales de deces. http://www.cepidc.vesinet.inserm.fr/

[17] Remontet L, Esteve J, Bouvier AM Bouvier AM, Grosclaude P, Launoy G, Menegoz F, Exbrayat C, Tretare B, Carli PM, Guizard AV, Troussard X, Bercelli P, Colonna M, Halna JM, Hedelin G, Macé-Lesec'h J, Peng J, Buemi A, Velten M, Jougla E, Arveux P, Le Bodic L, Michel E, Sauvage

M, Schvartz C, Faivre J. Cancer incidence and mortality in France over the period 1978-2000. Rev Epidemiol Sante Publique 2003 Feb;51:3-30.

[18] Parkin DM, Whelan SI, Ferlay J, Raymond L, Young J. Cancer incidence in five continents. Vol VII. IARC Scientific Publication 143. Lyon IARC 1993.

[19] Evolution de l'incidence et de la mortalité par cancer en France de 1980 à 2005. Estimations à partir des données des registres du réseau FRANCIM et du CepiDCIM. Institut de la veille sanitaire. Janvier 2008. http://www.ecosante.fr/FRANFRA/141.html

[20] Tamborini A. Menopause, THS et osteoporose post-menopausique: quoi de neuf? Réalités en Gynécologie-Obstétrique 2008 130: 1-7.

[21] Traitement hormonal substitutif de la menopause. Afssaps. Sept 2005. 68 pp.

[22] Com-Ruelle L, Dourgnon P, Jusot F, Lengagne P. Les problemes d'alcool en France: quelles sont les populations à risque. Questions d'economie de la sante. IRDS 2008. N° 129: 1-6.

[23] De Saint Pol T. Evolution de l'obésité en France de 1981 à 2003 : Les disparités entre milieux sociaux augmentent. Obésité 2007 June (2): 188-94.

[24] Rochefort H, Rouesse J. Cancers du sein, incidence et prevention. Bull. Acad Natle Med 2008; 192: 161-74.

[25] Laugier A. Annuaire de la cancérologie / radiothérapie et des imageries médicales en France (ACRIM). Paris, 2002.

[26] Dubois G. Le depistage des cancers. Situation et perspectives en France. Acta Endoscopica 1992; 22:115-18.

[27] http://lesrapports.ladocumentationfrancaise.fr/BRP/074000352/0000.pdf

[28] Wait S, Schaffer P, Seradour B, Chollot M, Demay M, Dejouhanet S.Opportunistic breast cancer screening in France. Bull Cancer 1997; 84(6): 619-24.

[29] Kopans DB. The 2009 US Preventive Services Task Force (USPSTF) guidelines are not supported by science: the scientific support for mammography screening. Radiol Clin North Am 2010; 48(5): 843-57.

[30] Kalager M, Zelen M, Langmark F, Adami HO. Effect of screening mammography on breast-cancer mortality in Norway. N Engl J Med 2010; 363(13):1203-10.

[31] Jørgensen KJ, Zahl PH, Gøtzsche PC. Breast cancer mortality in organised mammography screening in Denmark. A comparative study. BMJ 2010; 340: c1241.

[32] Autier, P, Moniol, M, Gavin, A, Natten, LJ. Breast cancer mortality in neighbouring European countries with different levels of screening but similar access to treatment: trend analysis of WHO mortality database. BMJ 2011; 343:d4411

[33] Vainio H, Bianchini F, editors. IARC Handbook of Cancer Prevention Volume 7 Breast Cancer Screening. Lyon . IARC Press. 2002.

[34] Hofvind S, Sørum R, Thoresen S. Incidence and tumour characteristics of breast cancer diagnosed before and after implementation of a population-based screening-program. Acta Oncologica, 2008; 47: 225-31.

[35] Kaplan RM, Saltzstein SL. Reduced mammographic screening may explain declines in breast carcinoma in older women. J Am Geriatr Soc 2005; 53(5): 862-6.

[36] Mouquet MC, Cherie-Challine L, Marescaux C. L'analyse des séjours chirurgicaux au sein du PMSI:un nouvel indicateur pour l'observation des cancers. DRESS 2002;27:1-32.

[37] Aubard Y, Genet D, Eyraud JL, Clavere P, Tubiana-Mathieu N, Philippe HJ. Impact of screening on breast cancer detection. Retrospective comparative study of two periods ten years apart. Eur.Gynaec Oncol J 2002; 23: 37-41.

[38] McKinnon NE. Breast Cancer mortality, Ontario, 1909-1947. The lack of any decline, and its significance. Can J Pub Health 1949 ; 40: 257-69.

[39] Welch HG, Black WC. Using Autopsy series to estimate "reservoir" for ductal carcinoma in situ of the breast: How much more breast cancer can we fin? Ann Intern Med 19997 Dec 1; 127(11):1023-8

[40] Nielsen M, Jensen J, Andersen J. Precancerous and cancerous breast lesions during lifetime and at autopsy. A study of 83 women. Cancer 1984; 54: 612-5.

[41] Bhathal PS, Brown RW, Lesueur GC, Russell IS. Frequency of benign and malignant breast lesions in 207 consecutive autopsies in Australian women. Br J Cancer 1985; 51: 271-8.

[42] Bartow SA, Pathak DR, Black WC, Key CR, Teaf SR.Prevalence of benign, atypical, and malignant breast lesions in populations at different risk for breast cancer. A forensic autopsy study. Cancer 1987; 60: 2751-60.

[43] Nielsen M, Thomsen JL, Primdahl S, Dyreborg U, Andersen JA. Breast cancer and atypia among young and middle-aged women: a study of 110 medicolegal autopsies. Br J Cancer 1987; 56: 814-9.

[44] Zahl PH, Strand GH, Maehlen J. Incidence of breast cancer in Norway and Sweden during introduction of nationwide screening: Prospective cohort study. BMJ 2004; 328: 921-4.

[45] Zakrisson S, Andersson I, Janzon L, Manjer J, Garne JP. Rate of over-diagnosis of breast cancer 5 years after end of Malmo mammographic screening trial: Follow-up study. BMJ 2006; 332: 689-92.

[46] Académie nationale de médecine. Analyse du rapport de la commission d'orientation sur le cancer. Académie de médecine. Paris, 2003.

[47] Philip T, Kasparian C, Fagnani F, Moatti JP, Meunier A, Parodi AL, Mornex R, Paolaggi JB, Godeau P, Vacheron A, Dreux C, Dubois F, Junod B, Picard JD. Le dépistage du cancer du sein en France : bilan et limites. Discussion. Bulletin de l'Académie Nationale de Médecine, 189 (2005), n° 2, 321-39.

Table 1

Change over time in the prevalence of risk factor exposure

Age group	HRT (RR = 1.17 [21][a])		Alcohol (RR = 1.7 [24] [b])		Obesity (RR = 2.0 [24] [c])	
	1980 -1990	1995 - 2005	1980 -1990	1995 -2005	1980 -1990	1995 -2005
50-59	7.9 %	31.6 %	16.7 %	13.5 %	4.1 %	6.4 %
60-69	7.7 %	30.7 %	7.0 %	5.7 %	6.1 %	10.4 %
70-79	2.3 %	9.0 %	3.7 %	3.0 %	6.1 %	10.4 %

[a] RR resulting from the four available models in table six of "AFSSAPS report"[21]. RR = Total of expected exposed cases (3922.15) / Total of expected non-exposed cases (3358.78) = 1.17

Prevalence data restricted to population based samples: "ESPS–EPAS" and "3C" [21].

[b] Interpolation between delivered results (1.5 and 2.0)

[c] Prevalence in age group 60-69 was extrapolated to age group 70-79

Table 2

Breast-cancer incidence proportion from 1980-1990 to 1995-2005 in France

Comparison of 11 year follow-up of age matched cohorts 15 years apart

Age in the cohort	Incidence proportion of breast cancer diagnosis for 11 year periods *Cases per 1000*		Incidence proportion attributable to change in risk factor exposure from 1980-1990 to 1995-2005 *Cases per 1000*			Relative change in adjusted incidence proportion attributable to overdiagnosis *Adjusted relative increase*
	(1) 1980 - 1990	*(2)* 1995 - 2005	*(3)* HRT	*(4)* Alcohol	*(5)* Obesity	*{(2)-(1)-(3)-(4)-(5)} / (1)*
50-64	*20.4*	*36.8*	*.82*	*– .40*	*.48*	*76.0 %* *95% CL:66.7; 85.0*
65-79	*24.3*	*30.9*	*.89*	*– .25*	*.33*	*23.0 %* *95% CL: 15.2 %; 30.9 %*

Figure 1. Age standardised breast-cancer deaths and breast-cancer incidence by calendar year in France. Standard: age structure of women aged 35 and more in 1992.

Figure 2. Breast-cancer incidence rates by age and according to screening activity. France, 1980, 1992 and 2005.

Figure 3. Breast-cancer incidence in birth cohorts subject to different levels of screening activity. Crude incidence proportion increase in %. France, 1980 to 2005.

Additional files

Additional file 1

Title : Formulas for change in incidence after change in exposure prevalence for hormone replacement therapy, alcohol and obesity.

Description : Explanation of the formulas used for estimating changes in breast-cancer incidence after change in risk factor exposure.

Additional file 2

Title : Formulas for incidence proportion and confidence limits in cohorts

Description : Explanation of the formulas used for estimating incidence proportion and confidence limits in cohorts.

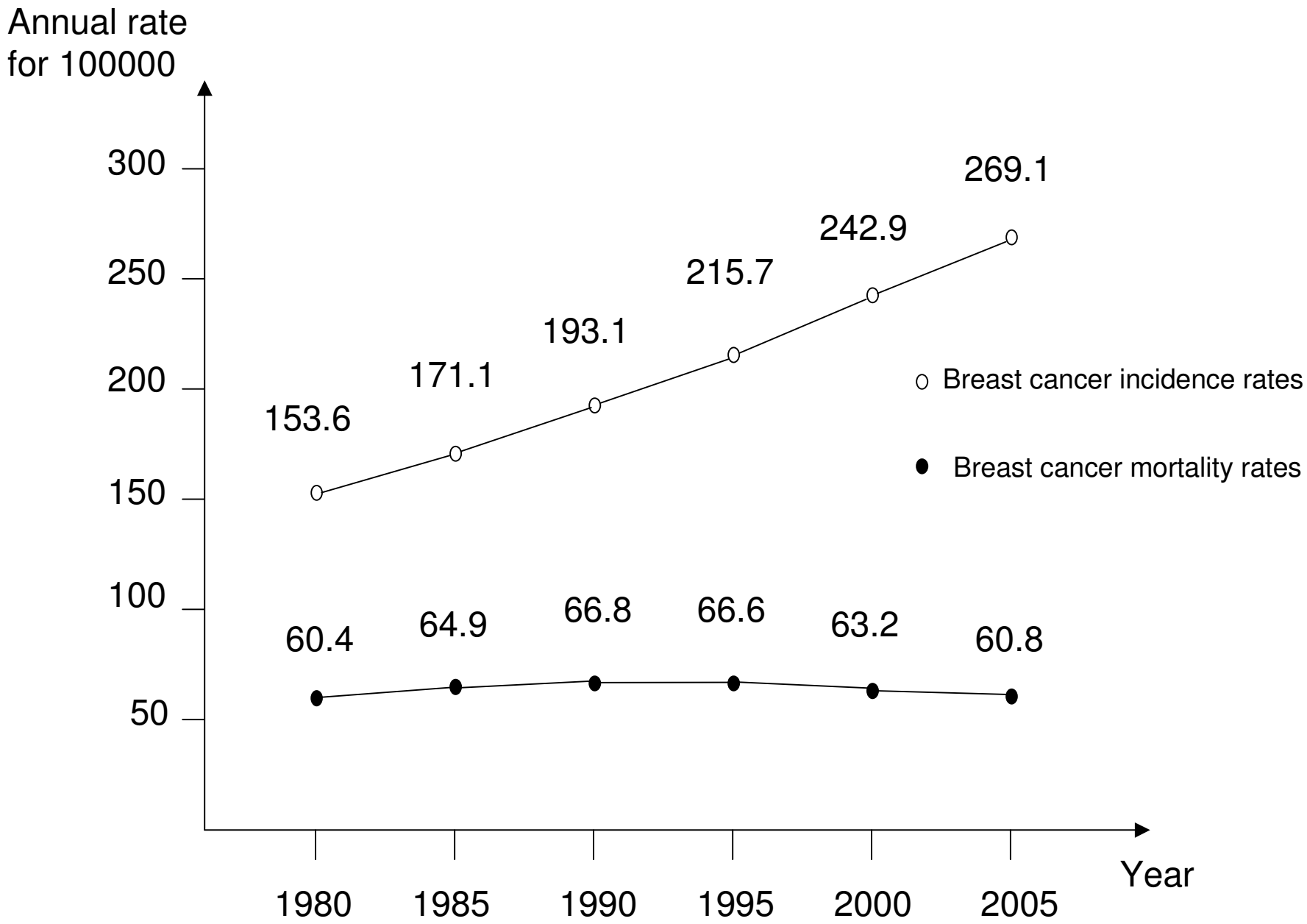
Annual rate
for 100000
300
250
200
150
100
50
269.1
242.9
215.7
193.1
171.1
153.6
60.4
64.9
66.8
66.6
63.2
60.8
Breast cancer incidence rates
Breast cancer mortality rates
1980
1985
1990
1995
2000
2005
Year

Figure 1

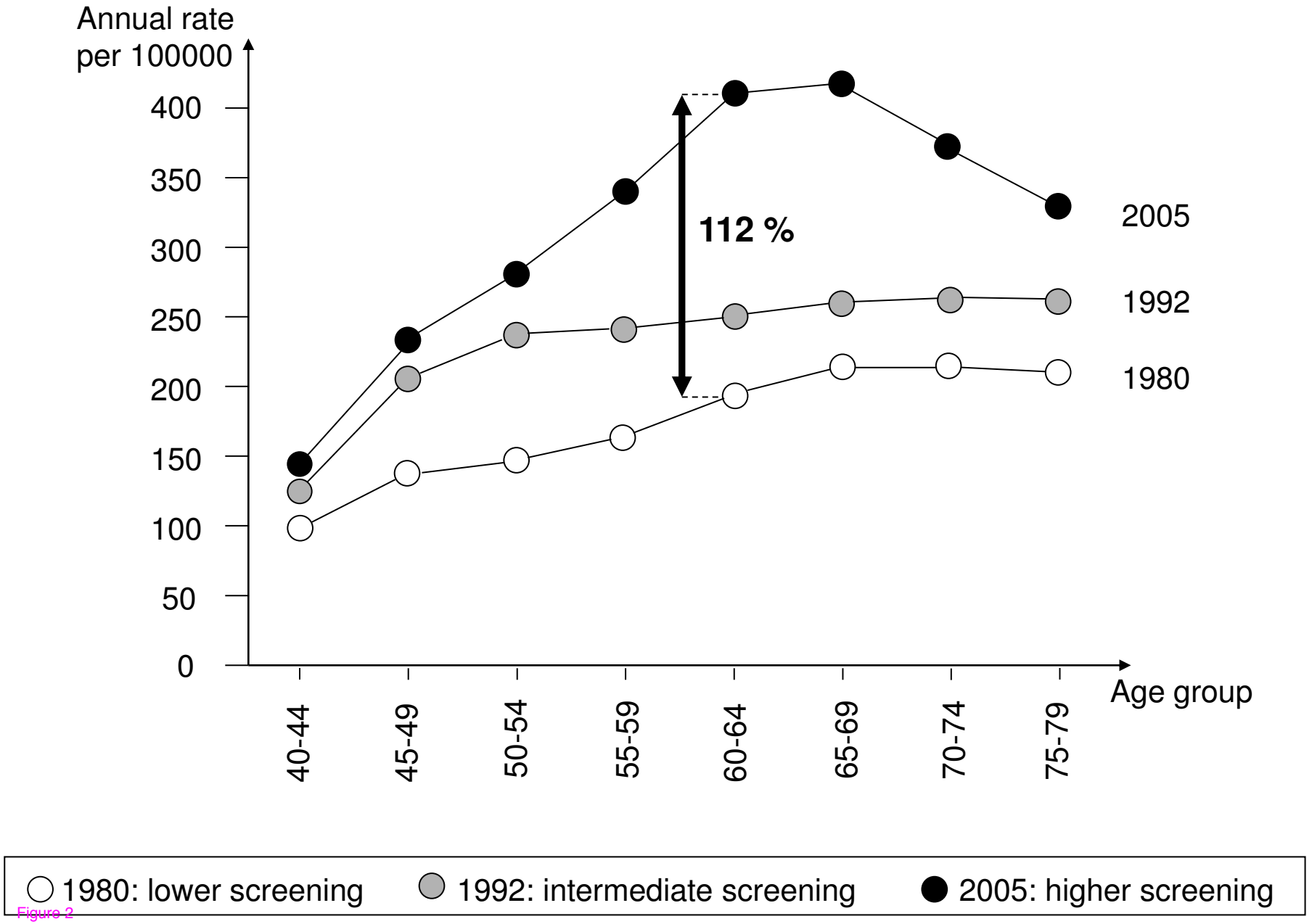
Annual rate
per 100000
400
350
300
250
200
150
100
50
0
112 %
2005
1992
1980
40-44
45-49
50-54
55-59
60-64
65-69
70-74
75-79
Age group
1980: lower screening
1992: intermediate screening
2005: higher screening

Figure 2

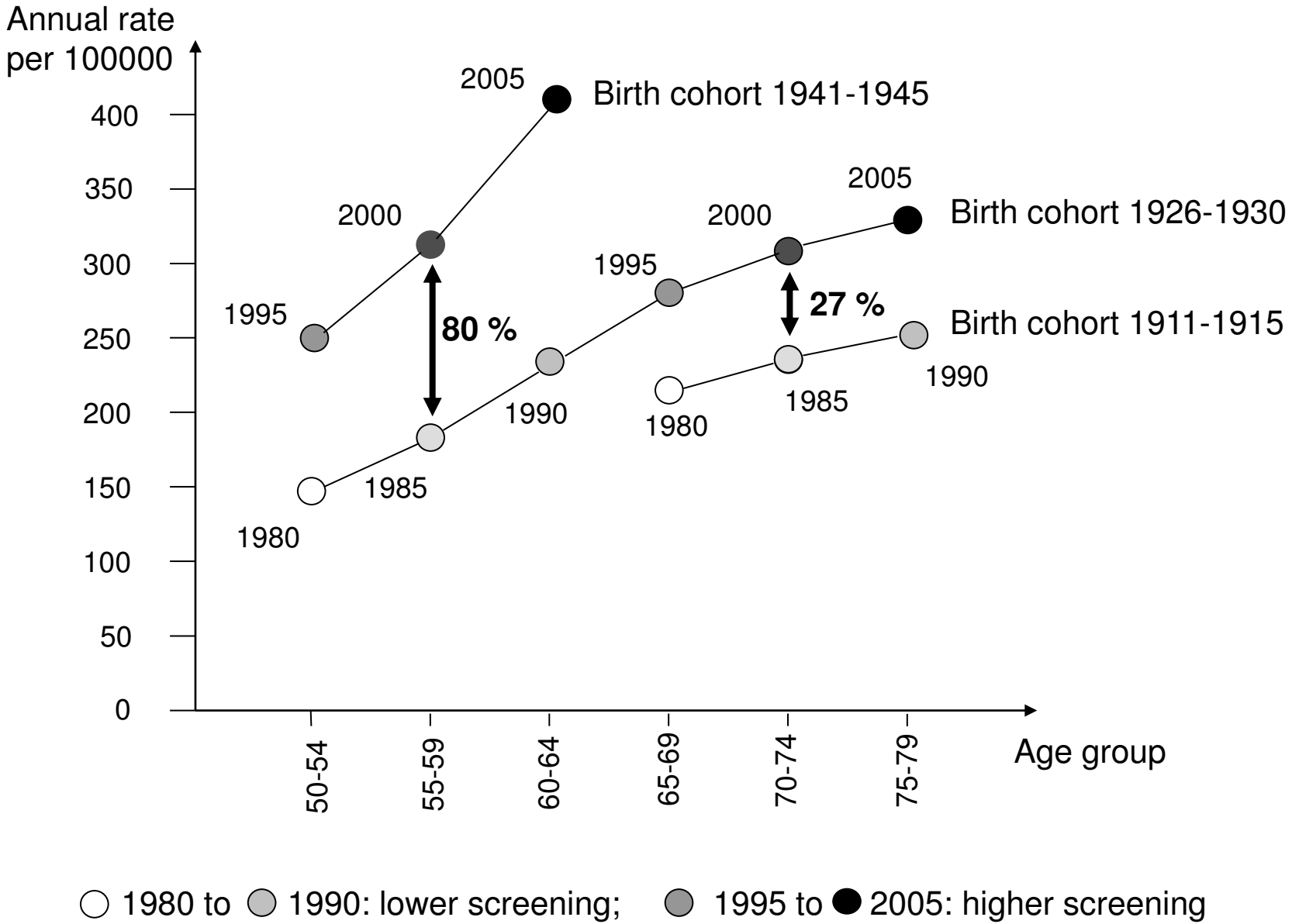
Annual rate per 100000
400
350
300
250
200
150
100
50
0
2005
Birth cohort 1941-1945
2000
1995
80 %
1990
1985
1980
2005
2000
Birth cohort 1926-1930
1995
27 %
Birth cohort 1911-1915
1990
1985
1980
50-54
55-59
60-64
65-69
70-74
75-79
Age group
1980 to 1990: lower screening;
1995 to 2005: higher screening

Figure 3

Additional files provided with this submission:

Additional file 1: JunodAddlFile1.doc, 44K
http://www.biomedcentral.com/imedia/9436701745910153/supp1.doc
Additional file 2: JunodAddFile2.doc, 27K
http://www.biomedcentral.com/imedia/3463407165910152/supp2.doc