Autier, Philippe and Boniol, Mathieu and Smans, Michel and Sullivan, Richard and Boyle, Peter (2015) Statistical analyses in Swedish randomised trials on mammography screening and in other randomised trials on cancer screening : a systematic review. Journal of the Royal Society of Medicine, 108 (11). pp. 440-450. ISSN 0141-0768 , http://dx.doi.org/10.1177/0141076815593403

Revised version

N=~~3242~~ ~~→~~ ~~3000~~2961

Statistical analyses in Swedish randomized trials on mammography screening

and ~~A comparison with statistical analyses of~~ in other randomized trials on cancer

screening: a systematic review

Short title: Revisiting Swedish mammography trials

Philippe Autier* [1,2]; Mathieu Boniol [1,2]; Michel Smans [2]; Richard Sullivan [3]; Peter Boyle [1,2]

[1] University of Strathclyde Institute of Global Public Health at iPRI, International Prevention Research Institute, Espace Européen, Building G, Allée Claude Debussy, 69130 Ecully Lyon, France

[2] International Prevention Research Institute (iPRI), 95 Cours Lafayette, 69006, Lyon, France

[3] Institute of Cancer Policy, Kings Health Partners Cancer Centre, Guy's Campus, London SE1 9RT, United Kingdom

***Corresponding author and reprint requests**: Dr Philippe Autier, University of Strathclyde Institute of Global Public Health, Espace Européen, Building G, Allée Claude Debussy, 69130 Ecully, France, philippe.autier@i-pri.org

Competing interests: Authors have no conflict of interest to disclose.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethical approval: Not needed

Guarantor: Philippe Autier.

Philippe Autier accepts full responsibility for the work and/or the conduct of the study, had access to the data, and controlled the decision to publish.

Formatted: Underline

Formatted: Left: 0.98", Right: 0.98", Top: 0.98", Bottom: 0.98", Width: 8.27", Height: 11.69"

Contributorship: Philippe Autier (PA, Mathieu Boniol (MB), Michel Smans (MS), Richard Sullivan (RS), Peter Boyle (PB)

PA and MS undertook the literature search. PA, MB and MS were responsible for the figures and data extraction. PA was responsible for the study design and study coordination. The data analysis was done by PA and MB. All co-authors participated in the data interpretation. ~~Finally,~~ PA wrote the first version of the article and PA, RS and PB finalized ~~wrote the~~ the article submitted to the journal.

Acknowledgements: No acknowledgements

Keywords: Breast cancer; screening; randomized trials; statistical analyses

Abstract

Objectives: Mammography screening seems not to play a major role in breast cancer mortality reductions observed in many populations. However the overview of Swedish trials of 2002 reported relative risks of 0.79 (95% CI: 0.70; 0.89) for the risk of breast cancer death associated with mammography screening. We compared investigated how calculations of relative risks of cancer death made in Swedish mammography trials and compared to calculation in other cancer screening trials.

Setting: Randomized trials on cancer screening.

Design: For each trial, Within the follow-up period of each trial, we identified the intervention period, when screening was offered to screening groups and not to control groups, and the post-intervention period, when screening (or absence of screening) was the same in screening and control groups. We then examined which cancer deaths had been used for the computation of relative risk of cancer death.

Main outcome measures: Relative risk of cancer death.

Results: In 17 non-breast screening trials, deaths due to cancer diagnosed during the follow-up periods were used for relative risk calculations. In the 5 Swedish trials, relative risk calculations used deaths due to breast cancers found during intervention periods, but deaths due to breast cancer found at first screening of control groups were added to these groups. After re-allocation of the added breast cancer deaths to post-intervention periods of control groups, relative risks of 0.86 (0.76; 0.97) were obtained for cancers found during intervention periods and 0.83 (0.71; 0.97) for cancers found during post intervention periods, indicating constant reduction in the risk of breast cancer death during follow-up, irrespective of screening.

Conclusions: The use of unconventional statistical methods in Swedish trials has led to over-estimation of risk reduction in breast cancer death attributable to mammography screening. The constant risk reduction observed in screening groups was probably due to the trial design that optimized awareness and medical management of women allocated to screening groups ~~compared to women allocated to control groups~~.

Introduction

Between 1977 and 1996, five randomized trials on mammography screening were conducted in Sweden, including women aged 40-74 at trial start. An overview of these trials published in 2002 reported that that 2 to 4 rounds of mammography screening could decrease breast cancer risk by 21 % (relative risk of 0.79; 95% CI: 0.70; 0.89)(1). This meta-analysis is considered the strongest evidence proving the efficacy of periodic mammography screening [2].

Mammography screening works through finding non-clinically detectable breast cancer before progression into advanced cancer with metastatic spread in lymph nodes and distant organs. Since reduction in cancer deaths due to reduction in the incidence of advanced cancer is not influenced by treatment efficacy, it was concluded from Swedish trials that decreases in the incidence of advanced breast cancer after screening introduction would provide the best indication that mammography screening reduces breast cancer mortality (2).

However, in communities where screening participation was high for more than ten years, only modest or no declines in the incidence of advanced breast cancer were observed (3-12). This e situation of breast cancer screening is in sharp contrast with that of colorectal and cervical cancer screening because in communities where screening for cervical and colorectal cancers is widespread, marked declines in the incidence of these types of cancers at an advanced stage have been observed, which indicates a substantial contribution of these screening modalities (13, 14) because

. Randomized trials have shown that screening for the latter two cancers reduced the risk of advanced cancer and of cancer death [16 17]. Iin communities where screening for cervical and colorectal cancers is widespread, marked declines in the incidence of these types of cancers at an advanced stage have been observed, which indicates a substantial contribution of these screening modalities (13, 14).

Breast screening trials were initiated at a time (1980's) when there was limited experience for designing, conducting and analyzing cancer screening trials (15). We therefore postulate that the contrasts between breast and cervical or colorectal cancers could be due to differences in the way randomized trials were conducted and analyzed. In this study, we re-examine the mortality data used and the way risks of breast cancer death were computed in Swedish trials in the light of study design and statistical analyses performed in screening trials on cancers other than breast cancer.

Designs of randomized trials for the evaluation of cancer screening tests

These trials are typically composed of two successive periods (Figure 1a): the intervention period that extends from randomization to termination of the last screening round in the screening group, and the post-intervention period that extends from the end of the last screening round in the screening group to the date of last check of vital status of subjects that were included in the trial. The follow-up period is the total of the intervention and the post-intervention periods. Depending on the number of screening rounds and follow-up extent, intervention and post-intervention periods may be of variable duration. Randomized trials evaluating cancer screening methods may consist of a single intervention of short duration including invitation to screening, the screening test itself and possible work up procedures in case of suspicious screening result. In other trials, the intervention period lasts for several years because the screening test is repeated every year or every two years. After the last screening round in the screening group, screening may be interrupted. Alternatively, screening may be pursued in the screening group and implemented in the control group, when for instance, decision is taken to launch a population screening program.

Relative risks of cancer death associated with screening are computed by dividing the cancer death rate in the screening group by the cancer death rate in the control group (Box). Cancer death rates can be calculated using deaths due to cancers found during the follow-up period as numerator (follow-up method), or using deaths due to cancers found during the intervention period as numerator (evaluation method). Denominators are the same in both methods. If in a trial, there is no post-intervention period, then the evaluation and follow-up periods coincides. During post-intervention periods, because screening (or absence of screening) activities are similar in the screening and in the control group, cancer detection rates in the two groups (i.e., Dsp/Ns and Dcp/Nc in Box) are also similar. In the follow-up method, growing numbers of deaths due to cancers found during steadily longer post-intervention periods will progressively narrow (or dilute) the difference in cancer death rates between the two groups. In this regard, reduction in the risk of cancer death calculated according to the follow-up method may be smaller than when calculated according to the evaluation method. For instance, in the fecal-occult-blood-test (FOBT) trial in England, the relative risk of colorectal cancer death after 7.7 years of follow-up (6.7 years of intervention and 1 year of post-intervention) was 0.85 (95% CI: 0.74;0.94) and 0.91 (95% CI: 0.84;0.98) after 20 years of follow-up (6.7 years of intervention and 13.3 years of post-intervention)(16)For instance, in the fecal-occult-blood-test (FOBT) trial in England, the relative risk of colorectal cancer death after 7.7 years of follow-up (6.7 years of intervention and 1 year of post-intervention) was 0.85 (95% CI: 0.74;0.94)[20]. After 20 years of follow-up (6.7 years of intervention and 13.3 years of post-intervention) the relative risk was 0.91 (95% CI: 0.84;0.98)[19].

7

Cause of death assessment and statistical analysis in trials on screening for cancer other than breast cancer

We retrieved publications on 17 cancer screening trials other than breast cancer in which main trial results were presented (see eTable 1 in the Supplement). In 14 trials, cause of death assessment was done by committees unaware of the screening status of subjects that decided on likely causes of death using all available information. In all 17 trials, the relative risk of cancer-specific death associated with screening was calculated using deaths due to target cancers found during follow-up periods (follow-up method).

Cause of death assessment and statistical analysis in breast cancer screening trials

Committees for cause of death assessment independent of trial conduct and blinded as to the screening status of deceased women were implemented in the HIP (17) and in the Canadian trials (18-20)(Table 1)(1, 18, 21-23). The Two-County trial used causes of death established by local endpoint committees or a Joint Review Committee, both of which included trial investigators (24). Swedish trials included in the overview of 2002 and in the Age trial used causes of death reported on death certificates (1, 23).

~~To avoid dilution of risk reductions caused by breast cancer deaths of cancers found after the intervention period, main results in all breast screening trials were based on the evaluation method.~~ All ~~B~~breast screening trials ~~conducted in the USA, Canada and England~~ calculated relative risks of breast cancer death associated with screening using deaths due to breast cancers found during the intervention period of the screening and of the control groups (Evaluation method)(Table 1). ~~In contrast, t~~However, the Swedish trials and their ~~he~~ overview ~~of Swedish trials~~ used a different selection of breast cancer deaths for control goups, as one sentence in the statistical section of the 2002 overview makes clear, "The evaluation [method] ignores breast cancer deaths among women whose breast cancer diagnosis was made after the first screening round of the control group was completed"(1). This means that the breast cancer deaths in the control group that were used for calculating the relative risk included breast cancer deaths related to cancer cases found at first screening of this group (R_{C1} in Figure 1b). In the Two-County, Malmö and in the Stockholm trials, tThis first screening of the control group generally took place in years following the last screening round in the screening group (25-27). In the Goteborg trial, about half of first screening was done at the time of the last screening round of the screening group, and about half was done 3 to 8 months after the last screening round (28). Hence, most breast cancers found at first screening of the control group were in fact part of the post-intervention period, and if screening of the control group had not taken place, these cancers would have been diagnosed during the post-intervention period. Thus, this incorporation approach was thus equivalent to transferring to the intervention period a number of cancers and associated deaths that were part of the post-intervention period. It is important to note that this approach was applied to the control group only. As a consequence, publications reported more cancers per women in control groups than per women in screening groups (28-30). Translating this incorporation approach in equations displayed in Box gives: $RR_{EM/ST} = (D_{SI}/N_S)/[(D_{CI} + D_{RC1})/N_C]$, where $RR_{EM/ST}$ stands for the evaluation method specific to Swedish trials. D_{RC1} are deaths due to breast cancers found at first screening of the control group that pertain to the post-intervention period, (i.e., D_{CP} in Box) and not to the intervention period, (i.e., D_{CI} in Box)

The Two-County and the Stockholm trials reported numbers and stage of cancers found at first screening of control groups, showing that the incorporation approach resulted in adding 72 advanced (i.e., 20 mm size or more) cancers to the 434 advanced cancers diagnosed in the control group during the intervention period of the Two-County trial(25) and 30 advanced cancers (i.e., stage 2 or more) to the 173 advanced cancers diagnosed in the control group during the intervention period of the Stockholm trial(31). Because of their high fatality rate, these extra advanced cancers led to a substantial number of extra cancer deaths i.e., D_{RC1}. Thus the greater the value of D_{RC1}, the smaller the value of $RR_{EM/ST}$ and thus the greater the apparent reduction in the risk of breast cancer death associated with mammography screening.

Alternative calculation of results of Swedish trials

We estimated a relative risk according to the evaluation method that would not incorporate deaths due to cancers found at first screening of control groups, that is, we estimated D_{CI} and D_{RC1} of the $RR_{EM/ST}$ equation. In Swedish trials, the ratio between breast cancer mortality rates in the screening and control groups remained relatively equivalent after 10 to 12 years of follow-up (1, 22). Furthermore, the Two-County trial reported that after 29 years of follow-up, 10% of breast cancer deaths in the control group were associated with cancers found during the first screening of control women (22). The 10% figure is plausible because follow-up of the additional cancers was shorter than for cancers found during intervention periods. We thus inferred that 10% represented a valid estimate of the proportion of extra deaths added to intervention periods of control groups in the overview of 2002.

~~Table 2 displays the main results of the overview of 2002~~ [1]~~.~~

The evaluation method specific to Swedish trials found a relative risk of 0.79 while the follow-up method found a relative risk of 0.85, reflecting dilution of effect over time (Table 2). ~~Reduction of the risk of breast cancer death is smaller with the follow-up method because of the dilution by the addition of breast cancer deaths related to breast cancers found during the post-intervention period, when screening activities in both groups were identical.~~

Trial-specific data on breast cancer deaths are displayed in Table 3.

In the central column of this tTable 3, we estimated breast cancer deaths linked to cases found at first screening of control women by multiplying by 10% the number of breast cancer deaths in control groups of the Östergötland, Goteborg, and Stockholm trials. The 10% hypothesis was probably excessive for the Malmö I trial because first screening of control group concerned women born in 1923-32, and not women born in 1908-22[36], i.e., about 45% of the total number of control women included in the trial[37]. We thus set the estimate to 4.5% for Malmö I because first screening of control group concerned about 45% of the total number of control women included in the trial(32). In Malmö II, we set estimates to 7.5% because the follow-up period lasted 9.1 years(1), i.e., about three-quarter of 12 years. Therefore, we set estimates to 7.5% for Malmö II. We obtained an estimate of 46 breast cancer deaths related to breast cancers found at first screening of control groups. In the two right-hand columns, we re-allocated to post-intervention periods the 46 breast cancer deaths associated with cases found during first screening of control groups.

We then re-worked results of the overview of 2002 (1) in Table 2 using numbers of breast cancer deaths in control groups we estimated in Table 3. The relative risk of breast cancer death over the follow-up period remained unchanged, but the relative risk of breast cancer death for the evaluation method was 0.86 instead of 0.79. For breast cancers diagnosed during the post-intervention period, the relative risk of breast cancer death dropped to 0.83. Sensitivity analysis using 8 or 12% for re-working numbers of breast cancer deaths in control groups of the Östergötland, Goteborg, and Stockholm trials did not change much the corrected relative risk estimates (data not shown).

So, proper allocation of breast cancer deaths to the intervention and post-intervention periods led to an equalization of relative risks found for the intervention, post-intervention, and follow-up periods, with a risk of breast cancer death that remained about 15% lower in the screening group throughout the entire trial duration.

Discussion

Computations performed by the overview of Swedish mammography trials incorporated deaths of breast cancers found at first screening of the control group as if these cancers were part of intervention periods (1). The consequence of this incorporation approach was the overestimation of rates of breast cancer death in the control groups, which ended up in the overestimation of the protection conferred by mammography screening against breast cancer death. Other authors raised similar concerns, estimating that the evaluation method adopted by Swedish trials resulted in including in the control groups many cancers that would not have been found in the screening group, which biased results in favor of screening (33).

Non-Swedish breast screening trials and trials on screening for cancer other than breast cancer never used the incorporation approach and . In contrast, whenever possible, Swedish trials had recourse to the incorporation approach. But we found practically no methodological justification for this approach. The second publication on Swedish trials overview just provided an ethical justification [39]. The 2002 overview [1] did not comment on the incorporation approach. The Goteborg trial investigators argued that there was a need to compensate for the extra number of cancer found by screening that are included for follow-up to death in the screening group (28, 34). However all extra screen-detected invasive cancers in screening groups were early cancers, i.e., tumors less than 20 mm diameter or stage 1 (25, 29, 31, 32). Hence, the conceivable need to compensate for screen detection of extra numbers of early cancer could not justify the transfer to intervention periods of substantial numbers of advanced cancers found at first screening of control groups. Substantial numbers of extra cancers were also found in screening groups of trials of prostate and lung cancer. However, none of these trials resorted to screening the control group after termination of the intervention and to transfer these cancers to the intervention period. The compensation argument invoked by Swedish trial investigators (28, 34) is thus not tenable.

Our re-calculations of Swedish trial revealed that risks of breast cancer death were similar for cancers found during the intervention and the post-intervention periods, indicating that reductions in the risk of breast cancer death also applied to cancer cases diagnosed when screening (or absence of screening) was the same in both screening and control groups. Such result is compatible with an effect of being allocated to the screening or to the control group on the risk of breast cancer death (allocation effect), but not with an effect of mammography screening (screening effect) on that risk.

Two reasons could explain a lower risk of breast cancer deaths independent of mammography screening. First, the HIP (21), Age (23)and all Swedish trials (1, 22, 28, 30, 32, 35) that found decreased risk of breast cancer death associated with mammography screening adopted a “left-to-nature” design. Typically, parallel group randomized trials first recruit a group of eligible subjects that are informed on trial objectives, on potential health benefits and probable side effects. Subjects agreeing to participate must first sign an informed consent form after which they are randomized in an intervention or in a control group. In left-to-nature trials, only women invited to participate in breast screening knew they were part of a clinical trial. Women allocated to control groups were never contacted, did not sign an informed consent and were completely ignorant they were part of a trial. Health professionals knew or could detect which women were invited to screening but did not know which women were allocated to control groups. Imbalance between the two groups probably led to increased awareness and better information (e.g., on early breast symptoms) and medical management of women in screening groups. Women invited to screening had probably quicker access to specialized care than women in control groups. The Two-County trial provides the best evidence for factors other than mammography screening influencing breast cancer mortality. Besides mammography screening, the intervention also encompassed enhancing breast cancer awareness, breast self-examination, and rapid referral of women presenting at screening with breast symptoms, all factors that would have, according to investigators, reduced patient delay and led to earlier detection of interval cancers and their treatment (36). In addition, the Two-County trial randomized women by geographical cluster, each cluster comprising about 2,700 women in Dalarna (Kopparberg) county and about 3,200 women in Östergötland county (25, 37). This large cluster randomization scheme is likely to have exacerbated differences between screening and control groups with respect to information, awareness and medical management. Finally, some data indicate different management of breast cancer patients according to randomization group: the histological grade of cancers found during the Two-County trial was unknown for 19% of patients in the control group vs. 10% in the screening group ($p<0.0001$)(25). Lymph node status was missing for 5.0% of patients in the screening group and 7.3% of patients in the control group ($p=0.0396$)(25).

It seems likely that Swedish mammography screening trials have departed from the "*ceteris paribus*" principle by which an experiment evaluating the effect of one action must make sure that all other things remain equal and will not interfere with study results.

In contrast, the Canadian trials that found no reduction in the risk of breast cancer death associated with mammography screening, adopted the typical parallel group randomized trial design. All enrolled women were volunteers who signed an informed consent form before randomization and received the same information and medical attention (18-20).

A second reason for the persistent lower risk of breast cancer death for cancers found in the intervention and post-intervention periods could be biased attribution of causes of death. Of the 8 major breast screening trials, only the HIP and the Canadian trial implemented endpoint committees unaware of the screening status of deceased women. In left-to-nature trials, health professionals completing death certificates of being part of local endpoint committees may have known or guessed which women have been invited to screening but had no idea regarding women allocated to control groups. To circumvent this problem, the overview of 2002 used death certificates for cause of death assessment because the overview of 1993 found that causes reported on certificates correlated well with causes established by an independent endpoint committee that had access to all medical and necropsy information(1, 38). However, in the 2002 overview, there were nearly twice as many breast cancer deaths for the Malmö, Östergötland, Stockholm and Goteborg trials than in the 1993 overview (39) and it is unknown up to which point the reliability of death certificates was maintained over time.

In conclusion, unconventional computation of the relative risk of breast cancer death impacted on the reported results of the Swedish trials on mammography screening. This led to an intrinsic bias in favor of screening. If calculations of relative risks had been carried out using similar methodological approaches to other cancer screening trials conducted in the more recent era, the Swedish trials would not have found a 20% reduction of breast cancer death due to mammography screening. This conclusion can be verified through a re-analysis of Swedish trial original data according to methods used in other cancer screening trials.

Supplementary materials:

Formatted: Highlight

Declaration of conflicting interests

The Authors declare that there is no conflict of interest.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

1. Nystrom L, Andersson I, Bjurstam N, Frisell J, Nordenskjold B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. Lancet 2002;359:909-919
2. Smith RA, Duffy SW, Gabe R, Tabar L, Yen AM, Chen TH. The randomized trials of breast cancer screening: what have we learned? Radiol Clin North Am 2004;42:793-806, v
3. Esserman L, Shieh Y, Thompson I. Rethinking screening for breast cancer and prostate cancer. JAMA 2009;302:1685-1692
4. Autier P, Boniol M, Middleton R, Dore JF, Hery C, Zheng T, et al. Advanced breast cancer incidence following population-based mammographic screening. Ann Oncol 2011;22:1726-1735
5. Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. N Engl J Med 2012;367:1998-2005
6. Autier P, Koechlin A, Smans M, Vatten L, Boniol M. Mammography screening and breast cancer mortality in Sweden. J Natl Cancer Inst 2012;104:1080-1093
7. Nederend J, Duijm LE, Voogd AC, Groenewoud JH, Jansen FH, Louwman MW. Trends in incidence and detection of advanced breast cancer at biennial screening mammography in The Netherlands: a population based study. Breast Cancer Res 2012;14:R10
8. Kalager M, Adami HO, Bretthauer M, Tamimi RM. Overdiagnosis of invasive breast cancer due to mammography screening: results from the Norwegian screening program. Ann Intern Med 2012;156:491-499
9. Lousdal ML, Kristiansen IS, Moller B, Stovring H. Trends in breast cancer stage distribution before, during and after introduction of a screening programme in Norway. Eur J Public Health 2014;24:1017-1022
10. Autier P. [Screening for breast cancer: worries about its effectiveness]. Rev Prat 2013;63:1369-1377
11. de Glas NA, de Craen AJ, Bastiaannet E, Op 't Land EG, Kiderlen M, van de Water W, et al. Effect of implementation of the mass breast cancer screening programme in older women in the Netherlands: population based study. BMJ 2014;349:g5410
12. Autier P, Boniol M. The incidence of advanced breast cancer in the West Midlands, United Kingdom. Eur J Cancer Prev 2012;21:217-221
13. Sigurdsson K, Sigvaldason H. Effectiveness of cervical cancer screening in Iceland, 1964-2002: a study on trends in incidence and mortality and the effect of risk factors. Acta Obstet Gynecol Scand 2006;85:343-349
14. Edwards BK, Ward E, Kohler BA, Eheman C, Zauber AG, Anderson RN, et al. Annual report to the nation on the status of cancer, 1975-2006, featuring colorectal cancer trends and impact of

interventions (risk factors, screening, and treatment) to reduce future rates. Cancer 2010;116:544-573
15. Biller-Andorno N, Juni P. Abolishing mammography screening programs? A view from the Swiss Medical Board. N Engl J Med 2014;370:1965-1967
16. Scholefield JH, Moss SM, Mangham CM, Whynes DK, Hardcastle JD. Nottingham trial of faecal occult blood testing for colorectal cancer: a 20-year follow-up. Gut 2012;61:1036-1040
17. Miller AB. Screening for breast cancer with mammography. Lancet 2001;358:2164; author reply 2167-2168
18. Miller AB, Wall C, Baines CJ, Sun P, To T, Narod SA. Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National Breast Screening Study: randomised screening trial. BMJ 2014;348:g366
19. Miller AB, To T, Baines CJ, Wall C. Canadian National Breast Screening Study-2: 13-year results of a randomized trial in women aged 50-59 years. J Natl Cancer Inst 2000;92:1490-1499
20. Miller AB, To T, Baines CJ, Wall C. The Canadian National Breast Screening Study-1: breast cancer mortality after 11 to 16 years of follow-up. A randomized screening trial of mammography in women age 40 to 49 years. Ann Intern Med 2002;137:305-312
21. Shapiro S. Periodic screening for breast cancer: the HIP Randomized Controlled Trial. Health Insurance Plan. J Natl Cancer Inst Monogr 1997:27-30
22. Tabar L, Vitak B, Chen TH, Yen AM, Cohen A, Tot T, et al. Swedish two-county trial: impact of mammographic screening on breast cancer mortality during 3 decades. Radiology 2011;260:658-663
23. Moss SM, Cuckle H, Evans A, Johns L, Waller M, Bobrow L. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. Lancet 2006;368:2053-2060
24. Holmberg L, Duffy SW, Yen AM, Tabar L, Vitak B, Nystrom L, et al. Differences in endpoints between the Swedish W-E (two county) trial of mammographic screening and the Swedish overview: methodological consequences. J Med Screen 2009;16:73-80
25. Tabar L, Fagerberg G, Duffy SW, Day NE, Gad A, Grontoft O. Update of the Swedish two-county program of mammographic screening for breast cancer. Radiol Clin North Am 1992;30:187-210
26. Frisell J, Lidbrink E, Hellstrom L, Rutqvist LE. Folluwup after 11 years - update of mortality results in the Stockholm mammographic screening trial. Breast Cancer Res Treat 1997;45:263-270
27. Andersson I, Janzon L. Reduced breast cancer mortality in women under age 50: updated results from the Malmo Mammographic Screening Program. J Natl Cancer Inst Monogr 1997:63-67
28. Bjurstam N, Bjorneld L, Warwick J, Sala E, Duffy SW, Nystrom L, et al. The Gothenburg Breast Screening Trial. Cancer 2003;97:2387-2396

29. Tabar L, Fagerberg G, Day NE, Duffy SW, Kitchin RM. Breast cancer treatment and natural history: new insights from results of screening. Lancet 1992;339:412-414

30. Frisell J, Lidbrink E, Hellstrom L, Rutqvist LE. Followup after 11 years--update of mortality results in the Stockholm mammographic screening trial. Breast Cancer Res Treat 1997;45:263-270

31. Frisell J, Eklund G, Hellstrom L, Lidbrink E, Rutqvist LE, Somell A. Randomized study of mammography screening--preliminary report on mortality in the Stockholm trial. Breast Cancer Res Treat 1991;18:49-56

32. Andersson I, Aspegren K, Janzon L, Landberg T, Lindholm K, Linell F, et al. Mammographic screening and mortality from breast cancer: the Malmo mammographic screening trial. BMJ 1988;297:943-948

33. Berry DA. Benefits and risks of screening mammography for women in their forties: a statistical appraisal. J Natl Cancer Inst 1998;90:1431-1439

34. Bjurstam N, Bjorneld L. Author reply to AB Miller et al. RE The Gothenburg breast screening trial. Cancer 1998;83:188-190

35. Tabar L, Fagerberg CJ, Gad A, Baldetorp L, Holmberg LH, Grontoft O, 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;1:829-832

36. Tabar L, Akerlund E, Gad A. Five-year experience with single-view mammography randomized controlled screening in Sweden. Recent Results Cancer Res 1984;90:105-113

37. Tabar L, Gad A. Screening for breast cancer: the Swedish trial. Radiology 1981;138:219-222

38. Nystrom L, Larsson LG, Rutqvist LE, Lindgren A, Lindgvist M, Ryden S, et al. Determination of cause of death among breast cancer cases in the Swedish randomized mammography screening trials. A comparison between official statistics and validation by an endpoint committee. Acta Oncol 1995;34:145-152

39. Nystrom L, Rutqvist LE, Wall S, Lindgren A, Lindqvist M, Rydén S, et al. Breast cancer screening with mammography: overview of Swedish randomised trials. Lancet 1993;341:973-978

CAPTIONS

Figure 1 – Design of randomized trials for the evaluation of cancer screening methods (R : screening round). Intervention periods are the continuous lines and the post-intervention periods are the dashed lines. (a) Typical design; (b) design specific to Swedish trials on breast cancer screening.

Box – Computation of relative risk (RR) of cancer death in randomized trials on cancer screening

Table 1 - Data used for relative risks calculation in randomized trials on breast cancer screening

Table 2 - Breast cancer deaths in the Swedish trials included in the 2002 overview

Table 3 - Breast cancer deaths in Swedish mammography trials

Revised version

N= 2961

Statistical analyses in Swedish randomized trials on mammography screening and in other randomized trials on cancer screening: a systematic review

Short title: Revisiting Swedish mammography trials

Philippe Autier* [1,2]; Mathieu Boniol [1,2]; Michel Smans [2]; Richard Sullivan [3]; Peter Boyle [1,2]

[1] University of Strathclyde Institute of Global Public Health at iPRI, International Prevention Research Institute, Espace Européen, Building G, Allée Claude Debussy, 69130 Ecully Lyon, France

[2] International Prevention Research Institute (iPRI), 95 Cours Lafayette, 69006, Lyon, France

[3] Institute of Cancer Policy, Kings Health Partners Cancer Centre, Guy's Campus, London SE1 9RT, United Kingdom

***Corresponding author and reprint requests**: Dr Philippe Autier, University of Strathclyde Institute of Global Public Health, Espace Européen, Building G, Allée Claude Debussy, 69130 Ecully, France, philippe.autier@i-pri.org

Competing interests: Authors have no conflict of interest to disclose.

Funding: This research was funded by the International Prevention Research Institute (iPRI), Lyon, France. The research did not receive specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethical approval: Not needed because of the use of data extracted from published articles.

Guarantor: Philippe Autier.

Philippe Autier accepts full responsibility for the work and/or the conduct of the study, had access to the data, and controlled the decision to publish.

Contributorship: Philippe Autier (PA), Mathieu Boniol (MB), Michel Smans (MS), Richard Sullivan (RS), Peter Boyle (PB)

PA was responsible for the study design and study coordination. PA and MS undertook the literature search. PA, MB and MS were responsible for the figures and data extraction. The data analysis was done by PA and MB. All co-authors participated in the data interpretation. PA wrote the first version of the article and PA, RS and PB finalized the article submitted to the journal.

Acknowledgements: No acknowledgements

Keywords: Breast cancer; screening; randomized trials; statistical analyses

Supplementary materials: Etable - Data used for relative risks calculation in randomized trials on cancer screening other than breast cancer screening

Abstract

Objectives: We compared calculations of relative risks of cancer death in Swedish mammography trials and in other cancer screening trials.

Setting: Randomized trials on cancer screening.

Design: For each trial, we identified the intervention period, when screening was offered to screening groups and not to control groups, and the post-intervention period, when screening (or absence of screening) was the same in screening and control groups. We then examined which cancer deaths had been used for the computation of relative risk of cancer death.

Main outcome measures: Relative risk of cancer death.

Results: In 17 non-breast screening trials, deaths due to cancers diagnosed during the follow-up periods were used for relative risk calculations. In the 5 Swedish trials, relative risk calculations used deaths due to breast cancers found during intervention periods, but deaths due to breast cancer found at first screening of control groups were added to these groups. After re-allocation of the added breast cancer deaths to post-intervention periods of control groups, relative risks of 0.86 (0.76; 0.97) were obtained for cancers found during intervention periods and 0.83 (0.71; 0.97) for cancers found during post intervention periods, indicating constant reduction in the risk of breast cancer death during follow-up, irrespective of screening.

Conclusions: The use of unconventional statistical methods in Swedish trials has led to over-estimation of risk reduction in breast cancer death attributable to mammography screening. The constant risk reduction observed in screening groups was probably due to the trial

design that optimized awareness and medical management of women allocated to screening groups.

Introduction

Between 1977 and 1996, five randomized trials on mammography screening were conducted in Sweden. An overview of these trials published in 2002 reported that that 2 to 4 rounds of mammography screening could decrease breast cancer risk by 21 % (1). Mammography screening works through finding non-clinically detectable breast cancer before progression into advanced cancer with metastatic spread in lymph nodes and distant organs. Since reduction in cancer deaths due to reduction in the incidence of advanced cancer is not influenced by treatment efficacy, it was concluded from Swedish trials that decreases in the incidence of advanced breast cancer after screening introduction would provide the best indication that mammography screening reduces breast cancer mortality (2).

However, in communities where screening participation was high for more than ten years, only modest or no declines in the incidence of advanced breast cancer were observed (3-5). This situation is in sharp contrast with that of colorectal and cervical cancer screening, because in communities where screening for cervical and colorectal cancers is widespread, marked declines in the incidence of these types of cancers at an advanced stage have been observed, which indicates a substantial contribution of these screening modalities (6, 7). Breast screening trials were initiated at a time when there was limited experience for designing, conducting and analyzing cancer screening trials. We therefore postulate that the contrasts between breast and cervical or colorectal cancers could be due to differences in the way randomized trials were conducted and analyzed. In this study, we re-examine the mortality data used and the way risks of breast cancer death were computed in Swedish trials in the light of study design and statistical analyses performed in screening trials on cancers other than breast cancer.

Designs of randomized trials for the evaluation of cancer screening tests

These trials are typically composed of two successive periods (Figure 1a): the intervention period that extends from randomization to termination of the last screening round in the screening group, and the post-intervention period that extends from the end of the last screening round in the screening group to the date of last check of vital status of subjects that were included in the trial. The follow-up period is the total of the intervention and the post-intervention periods. Depending on the number of screening rounds and follow-up extent, intervention and post-intervention periods may be of variable duration. Randomized trials evaluating cancer screening methods may consist of a single intervention of short duration including invitation to screening, the screening test itself and possible work up procedures in case of suspicious screening result. In other trials, the intervention period lasts for several years because the screening test is repeated every year or every two years. After the last screening round in the screening group, screening may be interrupted. Alternatively, screening may be pursued in the screening group and implemented in the control group, when for instance, decision is taken to launch a population screening program.

Relative risks of cancer death associated with screening are computed by dividing the cancer death rate in the screening group by the cancer death rate in the control group (Box). Cancer death rates can be calculated using deaths due to cancers found during the follow-up period as numerator (follow-up method), or using deaths due to cancers found during the intervention period as numerator (evaluation method). Denominators are the same in both methods. If in a trial, there is no post-intervention period, then the evaluation and follow-up periods coincides. During post-intervention periods, because screening (or absence of screening) activities are similar in the screening and in the control group, cancer detection rates in the two groups (i.e., Dsp/Ns and Dcp/Nc in Box) are also similar. In the follow-up method, growing numbers of deaths due to cancers found during steadily longer post-intervention periods will progressively narrow (or dilute) the difference in cancer death rates between the two groups. In this regard, reduction in the risk of cancer death calculated according to the follow-up method may be smaller than when calculated according to the evaluation method. For instance, in the fecal-occult-blood-test (FOBT) trial in England, the relative risk of colorectal cancer death after 7.7 years of follow-up (6.7 years of intervention and 1 year of post-intervention) was 0.85 (95% CI: 0.74;0.94) and 0.91 (95% CI: 0.84;0.98) after 20 years of follow-up (6.7 years of intervention and 13.3 years of post-intervention)(8).

Cause of death assessment and statistical analysis in trials on screening for cancer other than breast cancer

We retrieved publications on 17 cancer screening trials other than breast cancer in which main trial results were presented (see eTable in the Supplement). In 14 trials, cause of death assessment was done by committees unaware of the screening status of subjects that decided on likely causes of death using all available information. In all 17 trials, the relative risk of cancer-specific death associated with screening was calculated using deaths due to target cancers found during follow-up periods (follow-up method).

Cause of death assessment and statistical analysis in breast cancer screening trials

Committees for cause of death assessment independent of trial conduct and blinded as to the screening status of deceased women were implemented in the HIP (9) and in the Canadian trials (10)(Table 1). The Two-County trial used causes of death established by local endpoint committees or a Joint Review Committee, both of which included trial investigators (11). Swedish trials included in the overview of 2002 and in the Age trial used causes of death reported on death certificates (1, 12).

All breast screening trials calculated relative risks of breast cancer death associated with screening using deaths due to breast cancers found during the intervention period of the screening and of the control groups (Evaluation method)(Table 1). However, the Swedish trials and their overview used a different selection of breast cancer deaths for control groups, as one sentence in the statistical section of the 2002 overview makes clear, “The evaluation [method] ignores breast cancer deaths among women whose breast cancer diagnosis was made after the first screening round of the control group was completed”(1). This means that the breast cancer deaths in the control group that were used for calculating the relative risk included breast cancer deaths related to cancer cases found at first screening of this group (R_{C1} in Figure 1b). This first screening of the control group generally took place in years following the last screening round in the screening group (13-16). Hence, if screening of the control group had not taken place, these cancers would have been diagnosed during the post-intervention period. This incorporation approach was thus equivalent to transferring to the intervention period a number of cancers and associated deaths that were part of the post-intervention period. It is important to note that this approach was applied to the control group only. As a consequence, publications reported more cancers per women in control groups than per women in screening groups (16-18).

Translating this incorporation approach in equations displayed in Box gives:

$RR_{EM/ST} = (D_{SI}/N_S)/[(D_{CI} + D_{RC1})/N_C]$, where $RR_{EM/ST}$ stands for the evaluation method specific to Swedish trials. D_{RC1} are deaths due to breast cancers found at first screening of the control group that pertain to the post-intervention period, (i.e., D_{CP} in Box) and not to the intervention period, (i.e., D_{CI} in Box)

The Two-County and the Stockholm trials reported numbers and stage of cancers found at first screening of control groups, showing that the incorporation approach resulted in adding

72 advanced (i.e., 20 mm size or more) cancers to the 434 advanced cancers diagnosed in the control group during the intervention period of the Two-County trial (13) and 30 advanced cancers (i.e., stage 2 or more) to the 173 advanced cancers diagnosed in the control group during the intervention period of the Stockholm trial (19). Because of their high fatality rate, these extra advanced cancers led to a substantial number of extra cancer deaths i.e., D_{RC1}. Thus the greater the value of D_{RC1}, the smaller the value of $RR_{EM/ST}$ and thus the greater the apparent reduction in the risk of breast cancer death associated with mammography screening.

Alternative calculation of results of Swedish trials

We estimated a relative risk according to the evaluation method that would not incorporate deaths due to cancers found at first screening of control groups, that is, we estimated D_{CI} and D_{RC1} of the $RR_{EM/ST}$ equation. In Swedish trials, the ratio between breast cancer mortality rates in the screening and control groups remained relatively equivalent after 10 to 12 years of follow-up (1, 20). Furthermore, the Two-County trial reported that after 29 years of follow-up, 10% of breast cancer deaths in the control group were associated with cancers found during the first screening of control women (20). The 10% figure is plausible because follow-up of the additional cancers was shorter than for cancers found during intervention periods. We thus inferred that 10% represented a valid estimate of the proportion of extra deaths added to intervention periods of control groups in the overview of 2002.

The evaluation method specific to Swedish trials found a relative risk of 0.79 while the follow-up method found a relative risk of 0.85, reflecting dilution of effect over time (Table 2).

In the central column of Table 3, we estimated breast cancer deaths linked to cases found at first screening of control women by multiplying by 10% the number of breast cancer deaths in control groups of the Östergötland, Goteborg, and Stockholm trials. We set the estimate to 4.5% for Malmö I because first screening of control group concerned about 45% of the total number of control women included in the trial (21). In Malmö II, we set estimates to 7.5% because the follow-up period lasted 9.1 years (1) We obtained an estimate of 46 breast cancer deaths related to breast cancers found at first screening of control groups. In the two right-hand columns, we re-allocated to post-intervention periods the 46 breast cancer deaths associated with cases found during first screening of control groups.

We then re-worked results of the overview of 2002 (22) in Table 2 using numbers of breast cancer deaths in control groups we estimated in Table 3. The relative risk of breast cancer death over the follow-up period remained unchanged, but the relative risk of breast cancer death for the evaluation method was 0.86 instead of 0.79. For breast cancers diagnosed during the post-intervention period, the relative risk of breast cancer death dropped to 0.83. Sensitivity analysis using 8 or 12% for re-working numbers of breast cancer deaths in control groups of the Östergötland, Goteborg, and Stockholm trials did not change much the corrected relative risk estimates (data not shown).

So, proper allocation of breast cancer deaths to the intervention and post-intervention periods led to an equalization of relative risks found for the intervention, post-intervention, and follow-up periods, with a risk of breast cancer death that remained about 15% lower in the screening group throughout the entire trial duration.

Discussion

Computations performed by the overview of Swedish mammography trials incorporated deaths of breast cancers found at first screening of the control group as if these cancers were part of intervention periods (1). The consequence of this incorporation approach was the overestimation of rates of breast cancer death in the control groups, which ended up in the overestimation of the protection conferred by mammography screening against breast cancer death. Other authors raised similar concerns, estimating that the evaluation method adopted by Swedish trials resulted in including in the control groups many cancers that would not have been found in the screening group, which biased results in favor of screening (23).

Non-Swedish breast screening trials and trials on screening for cancer other than breast cancer never used the incorporation approach and we found practically no methodological justification for this approach. The Goteborg trial investigators argued that there was a need to compensate for the extra number of cancer found by screening that are included for follow-up to death in the screening group (16, 24). However all extra screen-detected invasive cancers in screening groups were early cancers, i.e., tumors less than 20 mm diameter or stage 1 (13, 17, 19, 25). Hence, the conceivable need to compensate for screen detection of extra numbers of early cancer could not justify the transfer to intervention periods of substantial numbers of advanced cancers found at first screening of control groups. Substantial numbers of extra cancers were also found in screening groups of trials of prostate and lung cancer. However, none of these trials resorted to screening the control group after termination of the intervention and to transfer these cancers to the intervention period. The compensation argument invoked by Swedish trial investigators (16, 24) is thus not tenable.

Our re-calculations of Swedish trial revealed that risks of breast cancer death were similar for cancers found during the intervention and the post-intervention periods, indicating that reductions in the risk of breast cancer death also applied to cancer cases diagnosed when screening (or absence of screening) was the same in both screening and control groups. Such result is compatible with an effect of being allocated to the screening or to the control group on the risk of breast cancer death (allocation effect), but not with an effect of mammography screening (screening effect) on that risk.

Two reasons could explain a lower risk of breast cancer deaths independent of mammography screening. First, the HIP (26), Age (12) and all Swedish trials (1, 16, 18, 20, 25, 27) that found decreased risk of breast cancer death associated with mammography screening adopted a “left-to-nature” design. Typically, parallel group randomized trials first recruit a group of eligible subjects that are informed on trial objectives, on potential health benefits and probable side effects. Subjects agreeing to participate must first sign an informed consent form after which they are randomized in an intervention or in a control group. In left-to-nature trials, only women invited to participate in breast screening knew they were part of a clinical trial. Women allocated to control groups were never contacted, did not sign an informed consent and were completely ignorant they were part of a trial. Health professionals knew or could detect which women were invited to screening but did not know which women were allocated to control groups. Imbalance between the two groups probably led to increased awareness and better information (e.g., on early breast symptoms) and medical management of women in screening groups. Women invited to screening had probably quicker access to specialized care than women in control groups.

The Two-County trial provides the best evidence that factors other than mammography screening influenced breast cancer mortality. Besides mammography screening, the intervention also encompassed enhancing breast cancer awareness, breast self-examination, and rapid referral of women presenting at screening with breast symptoms, all factors that would have, according to investigators, reduced patient delay and led to earlier detection of interval cancers and their treatment (28). In addition, the Two-County trial randomized women by geographical cluster, each cluster comprising about 2,700 women in Dalarna

(Kopparberg) county and about 3,200 women in Östergötland county (13). This large cluster randomization scheme is likely to have exacerbated differences between screening and control groups with respect to information, awareness and medical management. Finally, some data indicate different management of breast cancer patients according to randomization group: the histological grade of cancers found during the Two-County trial was unknown for 19% of patients in the control group vs. 10% in the screening group (p<0.0001)(13). Lymph node status was missing for 5.0% of patients in the screening group and 7.3% of patients in the control group (p=0.0396)(13).

It seems likely that Swedish trials have departed from the "*ceteris paribus*" principle by which an experiment evaluating the effect of one action must make sure that all other things remain equal and will not interfere with study results.

In contrast, the Canadian trials that found no reduction in the risk of breast cancer death associated with mammography screening, adopted the typical parallel group randomized trial design. All enrolled women were volunteers who signed an informed consent form before randomization and received the same information and medical attention (10).

A second reason for the persistent lower risk of breast cancer death for cancers found in the intervention and post-intervention periods could be biased attribution of causes of death. Of the 8 major breast screening trials, only the HIP and the Canadian trial implemented endpoint committees unaware of the screening status of deceased women. In left-to-nature trials, health professionals completing death certificates of being part of local endpoint committees may have known or guessed which women have been invited to screening but had no idea regarding women allocated to control groups. To circumvent this problem, the overview of 2002 used death certificates for cause of death assessment because the overview of 1993 found that causes reported on certificates correlated well with causes established by an independent endpoint committee that had access to all medical and necropsy information (1). However, in the 2002 overview, there were nearly twice as many breast cancer deaths for the Malmö, Östergötland, Stockholm and Goteborg trials than in the 1993 overview (29) and it is unknown up to which point the reliability of death certificates was maintained over time.

In conclusion, unconventional computation of the relative risk of breast cancer death impacted on the reported results of the Swedish trials on mammography screening. This led to an intrinsic bias in favor of screening. If calculations of relative risks had been carried out using similar methodological approaches to other cancer screening trials conducted in the more recent era, the Swedish trials would not have found a 20% reduction of breast cancer death due to mammography screening. This conclusion can be verified through a re-analysis of Swedish trial original data according to methods used in other cancer screening trials.

References

1. Nystrom L, Andersson I, Bjurstam N, Frisell J, Nordenskjold B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. Lancet 2002;359:909-919
2. Smith RA, Duffy SW, Gabe R, Tabar L, Yen AM, Chen TH. The randomized trials of breast cancer screening: what have we learned? Radiol Clin North Am 2004;42:793-806, v
3. Autier P, Boniol M, Middleton R, Dore JF, Hery C, Zheng T, et al. Advanced breast cancer incidence following population-based mammographic screening. Ann Oncol 2011;22:1726-1735
4. de Glas NA, de Craen AJ, Bastiaannet E, Op 't Land EG, Kiderlen M, van de Water W, et al. Effect of implementation of the mass breast cancer screening programme in older women in the Netherlands: population based study. BMJ 2014;349:g5410
5. Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. N Engl J Med 2012;367:1998-2005
6. Sigurdsson K, Sigvaldason H. Effectiveness of cervical cancer screening in Iceland, 1964-2002: a study on trends in incidence and mortality and the effect of risk factors. Acta Obstet Gynecol Scand 2006;85:343-349
7. Edwards BK, Ward E, Kohler BA, Eheman C, Zauber AG, Anderson RN, et al. Annual report to the nation on the status of cancer, 1975-2006, featuring colorectal cancer trends and impact of interventions (risk factors, screening, and treatment) to reduce future rates. Cancer 2010;116:544-573
8. Scholefield JH, Moss SM, Mangham CM, Whynes DK, Hardcastle JD. Nottingham trial of faecal occult blood testing for colorectal cancer: a 20-year follow-up. Gut 2012;61:1036-1040
9. Miller AB. Screening for breast cancer with mammography. Lancet 2001;358:2164; author reply 2167-2168
10. Miller AB, Wall C, Baines CJ, Sun P, To T, Narod SA. Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National Breast Screening Study: randomised screening trial. BMJ 2014;348:g366
11. Holmberg L, Duffy SW, Yen AM, Tabar L, Vitak B, Nystrom L, et al. Differences in endpoints between the Swedish W-E (two county) trial of mammographic screening and the Swedish overview: methodological consequences. J Med Screen 2009;16:73-80
12. Moss SM, Cuckle H, Evans A, Johns L, Waller M, Bobrow L. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. Lancet 2006;368:2053-2060
13. Tabar L, Fagerberg G, Duffy SW, Day NE, Gad A, Grontoft O. Update of the Swedish two-county program of mammographic screening for breast cancer. Radiol Clin North Am 1992;30:187-210
14. Frisell J, Lidbrink E, Hellstrom L, Rutqvist LE. Folluwup after 11 years - update of mortality results in the Stockholm mammographic screening trial. Breast Cancer Res Treat 1997;45:263-270
15. Andersson I, Janzon L. Reduced breast cancer mortality in women under age 50: updated results from the Malmo Mammographic Screening Program. J Natl Cancer Inst Monogr 1997:63-67
16. Bjurstam N, Bjorneld L, Warwick J, Sala E, Duffy SW, Nystrom L, et al. The Gothenburg Breast Screening Trial. Cancer 2003;97:2387-2396
17. Tabar L, Fagerberg G, Day NE, Duffy SW, Kitchin RM. Breast cancer treatment and natural history: new insights from results of screening. Lancet 1992;339:412-414
18. Frisell J, Lidbrink E, Hellstrom L, Rutqvist LE. Followup after 11 years--update of mortality results in the Stockholm mammographic screening trial. Breast Cancer Res Treat 1997;45:263-270
19. Frisell J, Eklund G, Hellstrom L, Lidbrink E, Rutqvist LE, Somell A. Randomized study of mammography screening--preliminary report on mortality in the Stockholm trial. Breast Cancer Res Treat 1991;18:49-56

20. Tabar L, Vitak B, Chen TH, Yen AM, Cohen A, Tot T, et al. Swedish two-county trial: impact of mammographic screening on breast cancer mortality during 3 decades. Radiology 2011;260:658-663
21. Zackrisson S, Andersson I, Janzon L, Manjer J, Garne JP. Rate of over-diagnosis of breast cancer 15 years after end of Malmo mammographic screening trial: follow-up study. BMJ 2006;332:689-692
22. Nyström L, Andersson I, Bjurstam N, Frisell J, Nordenskjöld B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. The Lancet 2002;359:909-919
23. Berry DA. Benefits and risks of screening mammography for women in their forties: a statistical appraisal. J Natl Cancer Inst 1998;90:1431-1439
24. Bjurstam N, Bjorneld L. Author reply to AB Miller et al. RE The Gothenburg breast screening trial. Cancer 1998;83:188-190
25. Andersson I, Aspegren K, Janzon L, Landberg T, Lindholm K, Linell F, et al. Mammographic screening and mortality from breast cancer: the Malmo mammographic screening trial. BMJ 1988;297:943-948
26. Shapiro S. Periodic screening for breast cancer: the HIP Randomized Controlled Trial. Health Insurance Plan. J Natl Cancer Inst Monogr 1997:27-30
27. Tabar L, Fagerberg CJ, Gad A, Baldetorp L, Holmberg LH, Grontoft O, 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;1:829-832
28. Tabar L, Akerlund E, Gad A. Five-year experience with single-view mammography randomized controlled screening in Sweden. Recent Results Cancer Res 1984;90:105-113
29. Nystrom L, Rutqvist LE, Wall S, Lindgren A, Lindqvist M, Rydén S, et al. Breast cancer screening with mammography: overview of Swedish randomised trials. Lancet 1993;341:973-978

CAPTIONS

Figure 1 – Design of randomized trials for the evaluation of cancer screening methods (R : screening round). Intervention periods are the continuous lines and the post-intervention periods are the dashed lines. (a) Typical design; (b) design specific to Swedish trials on breast cancer screening.

Box – Computation of relative risk (RR) of cancer death in randomized trials on cancer screening

Table 1 - Data used for relative risks calculation in randomized trials on breast cancer screening

Table 2 - Breast cancer deaths in the Swedish trials included in the 2002 overview

Table 3 - Breast cancer deaths in Swedish mammography trials

Table 1 - Data used for relative risks calculation in randomized trials on breast cancer screening

Trial No.	First author, year of publication*	Country, Study acronym	Screening method (as compared to the control group)	Follow-up period (years)		Cause of death assessment	Cancer-specific deaths used for calculation of the main relative risk associated with screening	RR	95% CI
				Intervention period	Post-intervention period				
1	Shapiro et al., 1997 [25]	USA, Greater New-York Health Insurance Plan (HIP)	MMS+BCE every 12 months, 4 rounds	5	13	0	Cancer-specific deaths of cancers found during the intervention period	0.77	NR
2	Tabar et al., 2011 [26]	Sweden, Two-County trial †	MMS, 2 to 4 rounds	7	22	Local committee	Cancer-specific deaths of cancers found during the intervention period plus, for the control group, incorporation of cancer-specific deaths of cancers found at first screening of this group	0.69	0.56;0.85
	id.	id.	id.	id.	id.	Joint review committee ‡	Cancer-specific deaths of cancers found during the intervention period plus, for the control group, incorporation of cancer-specific deaths of cancers found at first screening of this group	0.73	0.59;0.89
2	Nyström et al., 2002 # [1]	Sweden, Ostergotland §	MMS, 2 to 4 rounds	7.7	9.7	Death certificates	Cancer-specific deaths of cancers found during the intervention period plus, for the control group, incorporation of cancer-specific deaths of cancers found at first screening of this group	0.90	0.73;1.11
3	Nyström et al., 2002 [1]	Sweden, Malmö I	MMS every 18-24 months, 6 to 8 rounds	15	5	Death certificates	Cancer-specific deaths of cancers found during the intervention period plus, for the control group, incorporation of cancer-specific deaths of cancers found at first screening of this group	0.82	0.67;1.00

Trial No.	First author, year of publication*	Country, Study acronym	Screening method (as compared to the control group)	Follow-up period (years)		Cause of death assessment	Cancer-specific deaths used for calculation of the main relative risk associated with screening	RR	95% CI
				Intervention period	Post-intervention period				
4	Nyström et al., 2002 [1]	Sweden, Malmö II	MMS every 18-24 months, 1 to 7 rounds	5.8	3.3	Death certificates	Cancer-specific deaths of cancers found during the intervention period plus, for the control group, incorporation of cancer-specific deaths of cancers found at first screening of this group	0.64	0.39;1.06
5	Nyström et al., 2002 [1]	Sweden, Stockholm	MMS every 24-28 months, 2 rounds	4.4	10.5	Death certificates	Cancer-specific deaths of cancers found during the intervention period plus, for the control group, incorporation of cancer-specific deaths of cancers found at first screening of this group	0.91	0.65;1.27
6	Nyström et al., 2002 [1]	Sweden, Göteborg	MMS, 3 to 5 rounds	7 (women 39-49) and 5 (women 50-59)	7 (women 39-49) and 9 (women 50-59)	Death certificates	Cancer-specific deaths of cancers found during the intervention period plus, for the control group, incorporation of cancer-specific deaths of cancers found at first screening of this group	0.76	0.56;1.04
7	Miller et al., 2014 [22]	Canada, NBSS I and II	MMS every year, 4 to 5 rounds	5	20	Committee unaware of screening status	Cancer-specific deaths of cancers found during the intervention period	1.05	0.85;1.30
8	Moss et al., 2006 [27]	England, Age trial	MMS every 12 months, 4 to 6 rounds	5	6	Death certificates	Cancer-specific deaths of cancers found during the intervention period	0.83	0.66;1.04

BC: breast cancer; BCE: breast physical examination; CG: control group; IG: intervention group; MMS: mammography screening; RR: relative risk. NBSS: National Breast Screening Study; TCT: Two-County trial (Dalarna [formerly Kopparberg] and Ostergötland counties).

* The most recent publication reporting on main trial results is displayed in the Table.

† This trial was done in the counties of Dalarna (formerly Kopparberg) and Ostergötland.

‡ The Joint Review Committee included Two-County trial investigators (Holmberg et al., 2009 [28]) and has to be distinguished from the Independent Endpoint Committee set up by Swedish trial overviews (Nyström et al., 1993, 1995 [44,39])

§ The Ostergötland county trial was part of the Two-County trial, but results specific to the Ostergötland trial were published in Nyström et al., 2002. [1]

Table 2 - Breast cancer deaths in the Swedish trials included in the 2002 overview *

Group	No. women 40-74 included in trials †	Person-years of follow-up (thousand) ‡	No. BC deaths related to:			RR (95% CI) of BC death for BCs detected:		
			BC found during the intervention period	BC found during the post-intervention period	BC found during the follow-up period	During the intervention period (evaluation model) §	During the post-intervention period §	During the follow-up period (follow-up model) §
			As reported in the overview					
Screening	129750	1865	511	284	795	0.79	0.98	0.85
Control	117260	1688	584	263	847	(0.70; 0.89)	(0.83; 1.15)	(0.77; 0.94)
			After re-allocation to the post-intervention period of 10% breast cancer deaths found at first screening of the control group #					
Screening	129750	1865	511	284	795	0.86	0.83	0.85
Control	117260	1688	538	309	847	(0.76; 0.97)	(0.71; 0.97)	(0.77; 0.94)

BC: breast cancer: PY: person-year; RR: relative risk.

* Nyström et al., 2002 [1]; trials included in the overview are listed in Table 3.

† Data from table 2 of Nyström et al., 2002 [1]

‡ Data from table 4 of Nyström et al., 2002 [1]

§ RR computed using No. of women 40-74 as denominator

See Table 3 for computation of BC deaths in the control group.

Table 3 - Breast cancer deaths in Swedish mammography trials

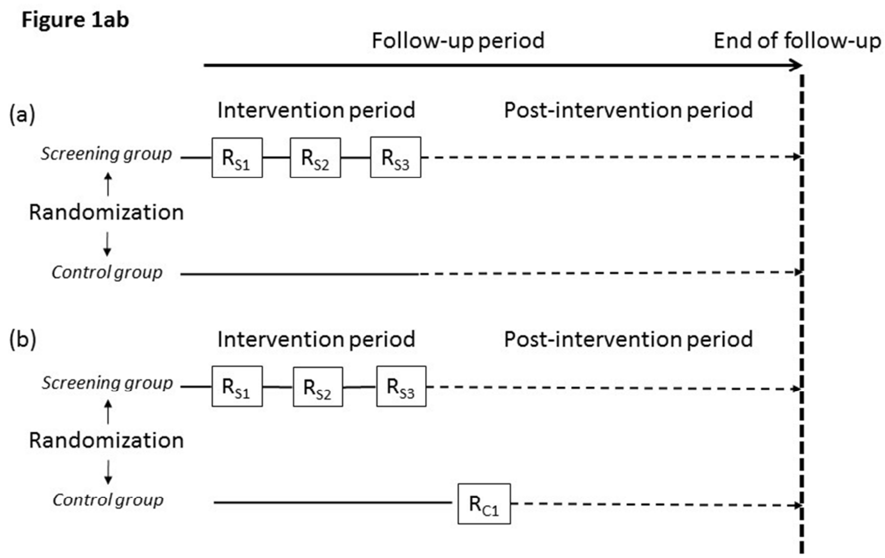
Breast cancer deaths of:	Screening group*	Control group*	BC deaths of BCs found at first screening of the control group (10% hypothesis†)	Re-allocation of BC deaths found at first screening of the control group	Corrected numbers of BC deaths in control groups
BC found during intervention periods					
Malmö I	161	198	9	198 - 9 =	189
Malmö II	29	33	2	33 - 2 =	31
Ostergötland	177	190	19	190 - 19 =	171
Stockholm	82	50	5	50 - 5 =	45
Göteborg	62	113	11	113 - 11 =	102
All five trials	**511**	**584**	**46**	**584 - 46 =**	**538**
BC found during post-intervention periods‡	**284**	**263**		**263 + 46 =**	**309**
BC found during follow-up periods	**795**	**847**			**847**

BC: breast cancer

* From table 4 of Nystrom et al., 2002[1]

† For Malmö I, the hypothesis was 4.5% and for Malmö II, the hypothesis was 7.5%.

‡ Numbers of BCs in each trial during the post-intervention period were not provided.



Design of randomized trials for the evaluation of cancer screening methods (R : screening round). Intervention periods are the continuous lines and the post-intervention periods are the dashed lines. (a) Typical design; (b) design specific to Swedish trials on breast cancer screening.
254x190mm (96 x 96 DPI)

Computation of relative risk (RR) of cancer death in randomized trials on cancer screening

For the screening group: D_{SI} is the number of cancer deaths related to cancers found during the intervention period, D_{SP} is the number of cancer deaths related to cancers found during the post-intervention period, and N_S is the number of subjects included in the screening group.

For the control group: D_{CI} is the number of cancer deaths related to cancers found during the intervention period, D_{CP} is the number of cancer deaths related to cancers found during the post-intervention period, and N_C is the number of subjects included in the control group.

Computations of RR of cancer death are:

Evaluation method: $RR_{EM} = (D_{SI}/N_S)/(D_{CI}/N_C)$

Follow-up method: $RR_{FUM} = [(D_{SI}+D_{SP})/N_S]/[D_{CI}+D_{CP})/N_C]$

Note: Numbers of person-years of follow-up (i.e., years spent by each subject in the trial) may be used instead of numbers of subjects. The two types of denominators provide similar results because in trials, duration of follow-up of subjects in screening and in control groups is the same.

Computation of relative risk (RR) of cancer death in randomized trials on cancer screening
338x190mm (96 x 96 DPI)

Supplementary materials to:

Statistical analyses in Swedish randomized trials on mammography screening and in other randomized trials on cancer screening: a systematic review

Running title: Revisiting Swedish mammography trials

Philippe Autier* [1,2]; Mathieu Boniol [1,2]; Michel Smans [2]; Richard Sullivan [3]; Peter Boyle [1,2]

Etable 1 - Data used for relative risks calculation in randomized trials on cancer screening other than breast cancer screening

Trial No.	First author, year of publication reference	Country, study acronym (if provided)	Target cancer	Screening method (as compared to the control group)	Follow-up period		Cause of death assessment	Cancer-specific deaths used for calculation of the main relative risk associated with screening
					Intervention period	Post-intervention period (years)		
1	Shaukat et al., 2013 (1)	USA, MCCCS	Colorectum	FOBT (rehydrated) every year or every two years	13 years	0 to 17	Committee unaware of screening status	Cancer-specific deaths of cancers found during the follow-up period
2	Scholefield et al., 2012 (2)	England	Colorectum	FOBT every two years	6.7 years	1 to 12.8	One to 3 investigators unaware of screening status	Cancer-specific deaths of cancers found during the follow-up period
3	Jorgensen et al., 2002 (3)	Denmark	Colorectum	FOBT every two years	10 to 13 years	0	Committee unaware of screening status	Cancer-specific deaths of cancers found during the follow-up period
4	Atkin et al., 2010 (4)	England	Colorectum	Once only flexible sigmoidoscopy	Few weeks	11.2 (median)	(i) Death certificates and (ii) cause of death assessed by independent coder unaware of screening status	Cancer-specific deaths of cancers found during the follow-up period
5	Segnan et al., 2011 (5)	Italy	Colorectum	Once only flexible sigmoidoscopy	Few weeks	11.4 (median)	Committee unaware of	Cancer-specific deaths of cancers found during

Trial No.	First author, year of publication reference	Country, study acronym (if provided)	Target cancer	Screening method (as compared to the control group)	Follow-up period		Cause of death assessment	Cancer-specific deaths used for calculation of the main relative risk associated with screening
					Intervention period	Post-intervention period (years)		
							screening status	the follow-up period
6	Schoen et al., 2012 (6)	USA, PLCO Cancer Screening Trial	Colorectum	Flexible sigmoidoscopy, two rounds	3 or 5 years	6.9 or 8.9	Committee unaware of screening status	Cancer-specific deaths of cancers found during the follow-up period
7	Thiis-Evensen et al., 2013 (7)	Norway, TPS	Colorectum	One round of flexible sigmoidoscopy at year 1 and one round of colonoscopy or simoidoscopy in the intervention and in the control group at year 14	14 years	12	Death certificates	Cancer-specific deaths of cancers found during the follow-up period
8	Holme et al., 2014 (8)	Norway	Colorectum	Once only flexible sigmoidoscopy with or without iFOBT	Few weeks	11.2 (median)	Death certificates	Cancer-specific deaths of cancers found during the follow-up period
9	Sankaranarayanan et al., 2007 (9)	India	Cervix	Once only visual inspection	Few weeks	7	Cancer registry staff unaware of screening status	Cancer-specific deaths of cancers found during the follow-up period
10	Sankaranarayanan et al., 2009 (10)	India	Cervix	Once only visual inspection or once only cytology, or once only HPV detection	Few weeks	8	Cancer registry staff unaware of screening status	Cancer-specific deaths of cancers found during the follow-up period

Trial No.	First author, year of publication reference	Country, study acronym (if provided)	Target cancer	Screening method (as compared to the control group)	Follow-up period		Cause of death assessment	Cancer-specific deaths used for calculation of the main relative risk associated with screening
					Intervention period	Post-intervention period (years)		
11	Andriole et al., 2009 (11)	USA, PLCO Cancer Screening Trial	Prostate	Digital rectal examination and serum PSA level every year	6	1 to 4	Committee unaware of screening status	Cancer-specific deaths of cancers found during the follow-up period
12	Schröder et al., 2014 (12)	Europe, ERSPC	Prostate	Serum PSA level every 4 years	8.8 to 13 years (average)	0	Causes of death were evaluated in a blinded fashion and according to a standard algorithm	Cancer-specific deaths of cancers found during the follow-up period
13	Oken et al., 2011 (13)	USA, PLCO Cancer Screening Trial	Lung	Chest X-ray every year	3 years	10	Committee unaware of screening status	Cancer-specific deaths of cancers found during the follow-up period
14	Pastorino et al., 2012 (14)	Italy	Lung	Low-dose CT-Scan every, annual or biennial	5 years	0	Vital status of participants was traced blindly, without knowing the random allocation	Cancer-specific deaths of cancers found during the follow-up period
15	Church et al., 2013 (15)	USA, NLST	Lung	Low-dose CT-Scan every year	2 years	4.5 (median);	Committee unaware of	Cancer-specific deaths of cancers found during

Trial No.	First author, year of publication reference	Country, study acronym (if provided)	Target cancer	Screening method (as compared to the control group)	Follow-up period		Cause of death assessment	Cancer-specific deaths used for calculation of the main relative risk associated with screening
					Intervention period	Post-intervention period (years)		
						7.4 (maximum)	screening status	the follow-up period
16	Sankaranarayanan et al., 2013 (16)	India, Kerala	Mouth	Three and 4 rounds of triennial visual inspection of the mouth	10	0 to 5	3 doctors unaware of screening status	Cancer-specific deaths of cancers found during the follow-up period
17	Buys et al., 2011 (17)	USA, PLCO Cancer Screening Trial	Ovary	Annual serum CA 125 for 6 years and annual TVU for 4 years.	6	6.4 (median)	Committee unaware of screening status	Cancer-specific deaths of cancers found during the follow-up period

FOBT: fecal occult blood test based on guaiac reaction; iFOBT: immunological fecal occult blood test; RR: relative risk; TVU: transvaginal ultrasonography.
ERSPC: European Randomized Study of Screening for Prostate Cancer; MCCCS: Minnesota Colon Cancer Control Study; NLST: National Lung Screening Trial ; PLCO: Prostate, Lung, Colorectal and Ovarian; TPS: Telemark Polyp Study.
*The most recent publication reporting on main trial results is displayed in the Table.

References

1. Shaukat A, Mongin SJ, Geisser MS, Lederle FA, Bond JH, Mandel JS, et al. Long-term mortality after screening for colorectal cancer. N Engl J Med 2013;369:1106-1114
2. Scholefield JH, Moss SM, Mangham CM, Whynes DK, Hardcastle JD. Nottingham trial of faecal occult blood testing for colorectal cancer: a 20-year follow-up. Gut 2012;61:1036-1040
3. Jorgensen OD, Kronborg O, Fenger C. A randomised study of screening for colorectal cancer using faecal occult blood testing: results after 13 years and seven biennial screening rounds. Gut 2002;50:29-32
4. Atkin WS, Edwards R, Kralj-Hans I, Wooldrage K, Hart AR, Northover JM, et al. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. Lancet 2010;375:1624-1633
5. Segnan N, Armaroli P, Bonelli L, Risio M, Sciallero S, Zappa M, et al. Once-only sigmoidoscopy in colorectal cancer screening: follow-up findings of the Italian Randomized Controlled Trial--SCORE. J Natl Cancer Inst 2011;103:1310-1322
6. Schoen RE, Pinsky PF, Weissfeld JL, Yokochi LA, Church T, Laiyemo AO, et al. Colorectal-cancer incidence and mortality with screening flexible sigmoidoscopy. N Engl J Med 2012;366:2345-2357
7. Thiis-Evensen E, Kalager M, Bretthauer M, Hoff G. Long-term effectiveness of endoscopic screening on incidence and mortality of colorectal cancer: A randomized trial. United European Gastroenterol J 2013;1:162-168
8. Holme O, Loberg M, Kalager M, Bretthauer M, Hernan MA, Aas E, et al. Effect of flexible sigmoidoscopy screening on colorectal cancer incidence and mortality: a randomized clinical trial. JAMA 2014;312:606-615
9. Sankaranarayanan R, Esmy PO, Rajkumar R, Muwonge R, Swaminathan R, Shanthakumari S, et al. Effect of visual screening on cervical cancer incidence and mortality in Tamil Nadu, India: a cluster-randomised trial. Lancet 2007;370:398-406
10. Sankaranarayanan R, Nene BM, Shastri SS, Jayant K, Muwonge R, Budukh AM, et al. HPV screening for cervical cancer in rural India. N Engl J Med 2009;360:1385-1394
11. Andriole GL, Crawford ED, Grubb RL, 3rd, Buys SS, Chia D, Church TR, et al. Mortality results from a randomized prostate-cancer screening trial. N Engl J Med 2009;360:1310-1319
12. Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Zappa M, Nelen V, et al. Screening and prostate cancer mortality: results of the European Randomised Study of Screening for Prostate Cancer (ERSPC) at 13 years of follow-up. Lancet 2014;384:2027-2035
13. Oken MM, Hocking WG, Kvale PA, Andriole GL, Buys SS, Church TR, et al. Screening by chest radiograph and lung cancer mortality: the Prostate, Lung, Colorectal, and Ovarian (PLCO) randomized trial. JAMA 2011;306:1865-1873
14. Pastorino U, Rossi M, Rosato V, Marchiano A, Sverzellati N, Morosi C, et al. Annual or biennial CT screening versus observation in heavy smokers: 5-year results of the MILD trial. Eur J Cancer Prev 2012;21:308-315
15. Church TR, Black WC, Aberle DR, Berg CD, Clingan KL, Duan F, et al. Results of initial low-dose computed tomographic screening for lung cancer. N Engl J Med 2013;368:1980-1991
16. Sankaranarayanan R, Ramadas K, Thara S, Muwonge R, Thomas G, Anju G, et al. Long term effect of visual screening on oral cancer incidence and mortality in a randomized trial in Kerala, India. Oral Oncol 2013;49:314-321
17. Buys SS, Partridge E, Black A, Johnson CC, Lamerato L, Isaacs C, et al. Effect of screening on ovarian cancer mortality: the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Randomized Controlled Trial. JAMA 2011;305:2295-2303