

# Results of the Two-County trial of mammography screening are not compatible with contemporaneous official Swedish breast cancer statistics

Per-Henrik Zahl, MD, PhD<sup>1</sup>, Peter C. Gøtzsche, MD, MDS<sup>2</sup>, Jannike Mørch Andersen PhD<sup>1, 3</sup> and Jan Mæhlen MD, PhD<sup>3</sup>

1) Norwegian Institute of Public Health, P.O. Box 4404, Nydalen, N-0403, Oslo, Norway. 2) Nordic Cochrane Centre, H:S Rigshospitalet, DK-2100 København Ø, Denmark. 3) Department of Pathology, Ullevål University Hospital, N-0407 Oslo, Norway.

Correspondence: Per-Henrik Zahl, Norwegian Institute of Public Health, P.O. Box 4404, Nydalen, 0403, Oslo, Norway

E-mail: per-henrik.zahl@fhi.no

Dan Med Bull 2006;53:438-40

## ABSTRACT

**Background:** National mammography screening programmes are based on the results of randomised trials, but the quality of these trials has recently been questioned. The Swedish Two-County trial reported a 31% reduction in breast cancer mortality and was instrumental for the introduction of screening in many countries. In this trial, official Swedish health registries were used to identify breast cancers and breast cancer deaths in the study population.

**Methods:** We used data from the same registries to estimate the numbers of breast cancer cases and breast cancer deaths among the included women.

**Results:** Compared to official Swedish statistics we found that 192 breast cancer cases and 43 breast cancer deaths seem to be missing in the main publication of the Two-County trial; we found similar discrepancies in two updates of the trial. These large differences can hardly be explained by random fluctuations in the cancer occurrence.

**Conclusion:** The data reported for the Two-County trial are incomplete. Other data indicate that the mortality results in a recent report were flawed.

The Two-County trial of screening mammography in Sweden has been instrumental for the introduction of screening in many countries. It reported a 31% reduction in breast cancer mortality [1]. However, the reliability of the study has later been questioned [2, 3]. The mean time of randomisation, as well as the number of women included in the study and the number of breast cancer deaths varied between different papers from the trial, even when the included age groups and follow-up periods were the same.

Another potential problem in the Two-County trial was to identify breast cancers that were not diagnosed during the screening sessions. Such cases included interval cancers in attending women, clinical cancers occurring in invited but non-attending women and all cancers in the control group. Since the researchers used the personal identification number system to link the study population to the Swedish population-based registries for cancer and the causes of death [1,4], the recorded cancer incidence and mortality in the Two-County trial should be identical to those in the official registers for these women.

We have used official Swedish statistics on breast cancer incidence and breast cancer mortality and compared them with the numbers reported by the researchers. We have combined data from the registries and the information on the study population and the average study period in the Two-County trial to obtain likely estimates of

the total number of breast cancers and breast cancer deaths in the trial. As far as we know such an independent quality control of the Two-County trial has not been done previously.

## PATIENTS AND METHODS

The Two-County trial included 134,867 women aged 40-74 years [1]. Randomisation started in October 1977 in Kopparberg and in May 1978 in Östergötland. Since the average length of follow-up was 6.0 years [1] December 31st, 1984, one can infer that the mean time-point for study entry (i.e. randomisation) was close to January 1st, 1979. The Two-County trial used cluster randomisation rather than individual randomisation. The date of entry was taken as the date of randomisation. Only individuals with a previous history of breast cancer were excluded from the study.

Tabar and colleagues [1] recorded 1663 cancers during the follow-up. We compared this number to data in the Swedish cancer register which Tabar and colleagues also used. This register uses personal identification numbers and is considered to be almost complete. In 1979, the population aged 40-74 years in Kopparberg and Östergötland was 134,924, close to the number in the trial.

The trial was closed after the first screening round and new women were not invited. Therefore, we had to identify cases in two steps. First, we recorded cases in the age group 40-74 years in the randomisation period (from study start in October 1977 in Kopparberg and from May 1978 in Östergötland to the end of 1979). Second, we recorded new cases among those aged 40-74 years in 1979 to December 31st, 1984.

## RESULTS

### BREAST CANCER CASES

We found 626 and 1431 breast cancers, respectively, in the two periods giving a total of 2057 new invasive breast cancers in the whole study period, or 394 more than reported in the trial (like Tabar and colleagues [1], we excluded all new cancers in women with first breast cancer diagnosed prior to the study period).

Next, we estimated how many of the 2057 official cancers that could have been excluded from the trial because the diagnosis had been made in the interim period from study start to time of randomisation. The mean ending time of this interim period must have been close to January 1st, 1979 since the mean time-point for randomization was close to January 1st, 1979 (see above). For the whole study population, the number of exposure years in the interim period for the age group 40-74 years was 123,937 (Statistics Sweden). In the 5-year period before the start of the trial, the breast cancer incidence in the age group 40-74 years in Kopparberg and Östergötland was 163 per 100,000 (Swedish national cancer registry). Assuming that the incidence is the same in the interim period we estimated that 202 invasive cancers (123,937 years × 163 per 100,000) were diagnosed in the interim period. By adding these 202 cancers to the 1663 reported by Tabar and colleagues [1], we get 1865 cancers for the entire study period; this is 192 fewer cancers than the 2057 recorded by the cancer registry. The probability that 394 cancers (compared to 202 expected) were diagnosed in the interim is very low ( $p < 0.001$ ). This suggests that the breast cancer incidence in the Two-County trial was incompletely reported.

### BREAST CANCER DEATHS

By the end of 1984, official statistics had recorded 266 breast cancer deaths among the 2057 breast cancers diagnosed in the study period. Of these, 111 had been diagnosed with breast cancer from study start to the end of 1979 and 155 in the period 1980-1984.

Tabar and colleagues [1] reported only 173 breast cancer deaths. If the records of Tabar and colleagues are complete, then the difference of 93 breast cancer deaths must have occurred in patients diagnosed in the interim period. But to the end of 1984, the predicted number of breast cancer deaths among these patients is only approximately 50. The probability that random variation in the breast cancer mor-

**Table 1.** Breast-cancer mortality for the Östergötland part of the Two-County trial (age group 40–74 years, evaluation model).

	Breast cancer deaths		Person-years of follow-up (in 1000s)	Relative risk
	study group	control group		
Nyström et al. [7]. . . . .	177	190	1161	0.90
Tabár et al. [8]. . . . .	167	213	1304	0.76

tality (assuming an incidence of 163 per 100,000) explains an additional 43 deaths is very low ( $p < 0.001$ ). This calculation suggests that the mortality data in the Two-County trial are also incomplete.

We have also compared breast cancer mortality data in two updates of the Two-County trial [4,5] with official statistics. Our analysis yielded results very similar to those reported above; for example, the number of breast cancer deaths for the study population in one of the updates [5] is 97 below those in official record, and even further below in the other update [4].

We have previously used methods similar to those used here to compare the breast cancer incidence recorded in the first four years of the Norwegian screening study with official Norwegian statistics [6]. Our estimate of the incidence in the interim period from the start of the study to the time of screening plus that in non-attending women was almost identical to that observed in the 5-year period before the start of the study. This indicates that our method is sufficiently precise to allow us to conclude that the data on breast cancer incidence and mortality in the Two-County trial appear to be unreliable.

## DISCUSSION

A recent updated overview of the Swedish randomised screening trials by Nyström and colleagues [7] was based on official mortality statistics. Nyström and colleagues reported only 10% reduction in breast cancer mortality for the Östergötland part of the Two-County trial (data were not made available for Kopparberg). In contrast, an update of the Two-County trial from Tabar and colleagues [8] reported a 24% mortality reduction. Compared to the update by Nyström and colleagues, Tabar and colleagues reported 10 fewer deaths from breast cancer in the study group despite the fact that the follow-up was slightly longer and the age group was identical, and 23 more in the control group in Östergötland (Table 1). Such a substantial change is very unlikely to have happened by chance ( $p < 0.001$ ). According to an investigator involved with the Two-County trial [9], other Swedish trialists [7], and an IARC/WHO report [10], cause-of-death assessments were not blind. This might be the reason why the cause of death determination by a local endpoint committee [11, 12] in Tabar and colleagues' update [8] appears to be seriously flawed.

The possibility that some cancers were not included in the trial reports is supported by the low rate of interval cancers after the first screen in the Two-County trial [13]. The rate of interval cancers has been reported to be only 94 per 100,000 [13] which we have confirmed; we found 91 per 100,000 using another publication and the same age group, 50–69 years [4]. In contrast, the rates of interval cancers were 158 and 157 per 100,000 in the national screening programme in the UK and The Netherlands, respectively [13].

It is of note that the differences in the number of breast cancer deaths between the study and control groups in the Two-County trial are small [1]. The mortality reduction would therefore no longer be statistically significant if only a few more breast cancer deaths were added to the study group.

We have used a simple and crude method to compare records of breast cancers and breast cancer deaths in the Two-County trial to the records in official Swedish statistics for the study population. Compared to official Swedish statistics about 192 breast cancer cases and 43 breast cancer deaths seem to be missing in the main publication of the Two-County trial [1]. These differences are large and cannot be explained by random fluctuations in the cancer occurrence. Our analysis of breast cancer cases is based on the assumption that most cases occurring in 1983 and 1984 were included in the

trial report from 1985 [1], while the analysis of breast cancer mortality is less dependent on this assumption. Neither can an imprecise estimate of the mean time of randomisation by Tabar and colleagues [1] change our conclusion that cases are missing in the study report; even if we assume that our estimate of mean time of randomisation was three months longer, our conclusion is still valid ( $p < 0.001$ ). Moreover, if the estimate of the mean time of randomisation is more than three months wrong, then all tables with exposure years in the study [1, 4, 5] are incorrect as well.

The cancer incidence in Sweden [14] and causes of death in Sweden [15] for 1984 were not available in the official version when the Two-County trial was published in April 1985 [1]. This can perhaps explain why there were fewer cancers and deaths in the 1985 report, but this underreporting was not corrected later. In 1992, Tabar and colleagues [4] reported 465 breast cancer deaths in the age group 40–74 years, which is 16 less than the number reported in the overview of the Swedish trials [5], although the follow-up in the overview ended one year earlier. Furthermore, it cannot explain the large discrepancy to official mortality statistics in one of the most recent reports on this trial (Table 1) [8]. We conclude that the data reported for the Two-County trial seem both incomplete and flawed.

Conflict of Interest statement: None declared.

Acknowledgments: We thank Centre for Epidemiology at the National Board of Health and Welfare for linking data sets on breast cancer cases and breast cancer deaths in Kopparberg and Östergötland in the actual time periods. P.-H.Z. and J.M. conceived and implemented the study. J.M.A. and P.C.G. assisted in data analysis and writing. P.-H.Z. is guarantor.

## REFERENCES

1. Tabár L, Fagerberg CJ, Gad A, et al. Reduction in mortality from breast cancer after mass screening with mammography. Randomised trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health and Welfare. *Lancet* 1985;325:829–32.
2. Gøtzsche PC, Mæhlen J, Zahl PH. What is publication? Editorial conflict over screening mammography. *Lancet* 2006;368 (Nov 25) (In press).
3. Gøtzsche PC. On the benefits and harms of screening for breast cancer. *Int J Epidemiol* 2004;33:56–64.
4. Tabar L, Fagerberg G, Duffy SW, Day NE, Gad A, Grøntoft O. Update of the Swedish Two-county program of mammographic screening for breast cancer. *Radiol Clin North Am* 1992;30:187–210.
5. Nyström L, Rutquist LE, Wall S, et al. Breast cancer screening with mammography: overview of the Swedish randomised trials. *Lancet* 1993;341: 973–8.
6. Zahl PH, Mæhlen J. Overdiagnosis in mammography screening. *Tidsskr Nor Laegeforen* 2004;124:2238–9.
7. Nyström L, Andersson I, Bjurstam N, Frisell J, Nordenskjöld B, Rutquist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet* 2002;359:909–19.
8. Tabar L, Vitak B, Chen HH et al. The Swedish Two-County Trial twenty years later. Updated mortality results and new insights from long-term follow-up. *Radiol Clin North Am* 2000;38:625–51.
9. Crewdson J. Swedes doubt mammography trial: disparities found in landmark study. *Chicago Tribune* 2002; March 15. [www.chicagotribune.com/news/chi-0203150264mar15.story](http://www.chicagotribune.com/news/chi-0203150264mar15.story) (accessed 15 March, 2002).
10. Vainio H, Bianchini F. IARC Handbooks of Cancer Prevention. Vol 7: Breast Cancer Screening. IARC Press; Lyon 2002.
11. Tabár L, Smith RA, Duffy SW. Update on effects of screening mammography. *Lancet* 2002;360:337.
12. Nyström L, Andersson I, Bjurstam N, Frisell J, Rutqvist LE. Update on effects of screening mammography (authors' reply). *Lancet* 2002;360: 339–40.
13. Woodman CBJ, Threlfall AG, Boggis CRM, Prior P. Is the three year breast screening interval too long? Occurrence of interval cancers in NHS breast screening programme's north western region. *BMJ* 1995;310:224–6.
14. Cancer incidence in Sweden 1984. Stockholm: National Board of Health and Welfare (Socialstyrelsen), 1987.
15. Dödsorsaker 1984. Stockholm: National Board of Health and Welfare (Socialstyrelsen), 1986.

## EDITOR'S COMMENT

The above paper was first published on the *European Journal of*

Cancer's website on March 9, 2006, but on March 29, the authors were informed by the editor that the article had been removed. It has never reappeared neither on the website nor in print and today the paper is labelled as withdrawn on PubMed (1). The course of events, as seen from the authors perspective, are summarised in the Lancet (2). I share the concerns raised by the authors. First, the process that led to removal of the accepted and published paper was unilateral. Second, a withdrawn or removed paper invariably leaves you with an impression of scientific fraud.

Therefore, DMB has decided to publish the paper.

1. Zahl PH, Gøtzsche PC, Andersen JM, Mæhlen J. WITHDRAWN: Results of the Two-County trial of mammography screening are not compatible with contemporaneous official Swedish breast cancer statistics. *Eur J Cancer*. 2006 Mar 9; [Epub ahead of print].
2. Gøtzsche PC, Maehlen J, Zahl PH. What is publication? Editorial conflict over screening mammography. *Lancet* 2006;368 (25 Nov) (In press).

*Torben V. Schroeder*  
Editor-in-chief  
Danish Medical Bulletin