Articles

Long-term effects of mammography screening: updated overview of the Swedish randomised trials

Lennarth Nyström, Ingvar Andersson, Nils Bjurstam, Jan Frisell, Bo Nordenskjöld, Lars Erik Rutqvist

Summary

Background There has been much debate about the value of screening mammography. Here we update the overview of the Swedish randomised controlled trials on mammography screening up to and including 1996. The Kopparberg part of the Two-County trial was not available for the overview, but the continuation of the Malmö trial (MMST II) has been added. The article also contains basic data from the trials that have not been presented before.

Methods The trials (n=247 010, invited group 129 750, control group 117 260) have been followed up by record linkage to the Swedish Cancer and Cause of Death Registers. The relative risks (RR) for breast cancer death and mortality were calculated for the invited and the control groups. The trial-specific as well as the age-specific effects were analysed. RRs were calculated by the density method, with total person-time experience of the cohort by time interval of follow-up as a basis for estimating mortality rates. We calculated weighted RRs and 95% CI with the Mantel-Haenszel procedure.

Findings The median trial time—the time from randomisation until the first round was completed for the control group or if the control group was not invited, until end of follow-up—was 6·5 years (range 3·0–18·1). The median follow-up time, the time from randomisation, to the end of follow-up, was 15·8 years (5·8–20·2). There were 511 breast cancer deaths in 1 864 770 women-years in the invited groups and 584 breast cancer deaths in 1 688 440 women-years in the control groups, a significant 21% reduction in breast cancer mortality (RR=0·79, 95% CI 0·70–0·89). The reduction was greatest in the age group 60–69 years at entry (33%). Looking at 5-year age groups, there were statistically significant effects in the age groups 55–59, 60–64, and 65–69 years (RR=0·76, 0·68, and 0·69, respectively). There was a small effect in women 50–54 years at randomisation (RR=0·95). The benefit in terms of cumulative breast cancer mortality started to emerge at about 4 years after randomisation and continued to increase to about 10 years. Thereafter the benefit in absolute terms was maintained throughout the period of observation. The age-adjusted relative risk for the total mortality was 0·98 (0·96–1·00).

Department of Public Health and Clinical Medicine, Umeå University (L Nyström PhD); **Department of Diagnostic Radiology, Malmö University** (I Andersson MD); **Department of Diagnostic Radiology, Göteborg University** (N Bjurstam MD); **Department of Surgery, Huddinge University Hospital, Stockholm** (J Frisell MD); **Department of Oncology, Linköping University** (Prof B Nordenskjöld MD); **and Department of Oncology, Huddinge University Hospital, Stockholm, Sweden** (L E Rutqvist MD)

Correspondence to: Dr Lennarth Nyström, Department of Public Health and Clinical Medicine, Umeå University, SE-901 85 Umeå, Sweden
(e-mail: Lennarth.Nystrom@epiph.umu.se)

Interpretation The advantageous effect of breast screening on breast cancer mortality persists after long-term follow-up. The recent criticism against the Swedish randomised controlled trials is misleading and scientifically unfounded.

Introduction

Service-screening for breast cancer occurs in several countries with the aim to decrease breast cancer mortality. The scientific basis for these programmes are the randomised screening trials. There are seven such studies, four from Sweden. The Swedish trials have a similar design: they were all population-based and compared invitation to breast screening with mammography alone versus no invitation. These Swedish trials differed from the other trials (the Health Insurance Plan of Greater New York, the Edinburgh trial, and the Canadian National Breast Screening Study), which all evaluated mammography combined with breast self-examination, clinical breast examination, or both. Moreover, the New York and the Canadian trials were not population-based.

The Swedish Cancer Society initiated an overview of the Swedish trials in the late 1980s. The objective was to validate the results from the individual trials through a method that was common to all trials, including a blind review of all deaths among breast cancer cases by an independent endpoint committee. Another objective was to increase the statistical power. The first results of the overview were published in 1993[1] and an update focusing on the age group 40–49 in 1997.[2] Concerns raised about the validity of the results from the trials.[3] include inappropriate exclusions, poor randomisation, and the excess total mortality in women invited to screening.

Our aim here was to extend the follow-up and to analyse the age-specific and trial-specific effects on breast cancer mortality, to describe the randomisation procedures in more detail, and to assess the quality of the cluster randomisation used in Östergötland. The Kopparberg part of the Two-County (WE) trial was not available for analysis but the continuation of the Malmö trial (MMST II) was added.

Methods

Details of the Swedish mammography screening trials have been published (table 1), and are summarised below.

Swedish trials summary

Malmö trial

The Malmö Mammographic Screening Trial (MMST) included women in the city of Malmö from October, 1976. In the first part (MMST I) women born between 1908 and 1932 were randomised with individual stratification by year of birth. Women were invited to screen-film mammography alone, in the first two rounds with two views (craniocaudal and oblique) and in

	Trial				
	MMST I	MMST II	Östergötland	Stockholm	Göteborg
Randomisation If cluster, type	Individual	Individual	Cluster Municipality, parish	Cluster Day of birth	Individual, cluster Day of birth
Accrual period Invited group Control group	Oct, 1976–Aug, 1978 Oct, 1992–Feb, 1993	Sept, 1978–Nov, 1990 Sept, 1991–April, 1994	May, 1978–March, 1981 April, 1986–Feb, 1988	March, 1981–May, 1983 Oct, 1985–May, 1986	Dec, 1982–April, 1984 Nov, 1987–June, 1991
Birth cohorts	1908–32	1933–45	1903–40	1917–42	1923–44
Age at entry	45–70	43–49	38–75	39–65	39–59
Intervention Number of views	2*	2	1	1	Round 1: 2 Round 2: 1–2§
Number of readers	2	2	1	1	Round 1–3: 1 Round 4–5: 2
Screening interval	18–24	18–24	24,33†	28	18
Number of rounds	1908–17: 6 1918: 7 1919–32: 8	1–7	40–49y: 4 50–69y: 3 70–74y: 2	2	1923–32: 4 1933–44: 5
Attendance rate‡	74%	75–80%	89%	82%	84%

*From round 3 single or two views according to parenchymal pattern. †Average for age groups 39–49 years and 50–75 years, respectively; ‡1st round. §Depending on density of breast.

Table 1: **Overview of Swedish trials included in overview**

subsequent rounds with either two views or the oblique view alone depending on the parenchymal pattern. A single oblique view was used for women whose breasts were mainly fatty on mammography, and two views for women with dense breasts. The endpoint was breast cancer as the underlying cause of death as determined by a blinded independent committee.[4] After MMST I closed in August, 1978, women who reached age 45 were continuously randomised to the Malmö trial, with the same protocol as in MMST I. MMST II comprised all women who were born 1933–45 and were living in Malmö between 1978 and 1990. The women were randomly allocated to receive invitation to screening. The plan was to invite these women when they turned 45, beginning in 1978. Because of limited resources, the plan could not be strictly adhered to, which means that, in some years, no women could be invited, while in other years two or even three birth-year cohorts were randomised and invited to examination. The median age at entry was 45 years. The last birth-year cohort, women born in 1945, was invited in 1990. The first screening round for the control group took place between 1991 and 1994. The endpoint was breast cancer as the underlying cause of death according to the Swedish Cause of Death Registry. A combined analysis of the two trials has been presented.[5]

Two-County trial
The trial included women in two Swedish counties: Kopparberg and Östergötland, with cluster randomisation by geographic area in each. Women aged 40–49, 50–69, and 70–74 years were invited to four, three, and two screening rounds, respectively. The screening interval was 24 months for women below 50 years of age and 33 months for women 50 years and older. The trial has been followed up every second year.[6] The Kopparberg part of the trial was not available for our latest overview.

Stockholm trial
About 60 000 women in the southern part of Stockholm born between 1917 and 1942 were randomised by day of birth to receive an invitation to mammography alone with an oblique single-view scan or no intervention. About 40 000 women were allocated to the invited group and 20 000 to the control group.[7] The screening started in March, 1981, and the screening interval was 28 months

between the first and the second round and 24 months between the second and the third round. After two rounds the control group was invited to screening.[8]

Göteborg trial
Between December, 1982, and April, 1984, all women born between 1923 and 1944 who lived in the city of Göteborg were randomised. Two-view mammography was used unless the observations at the previous screen indicated that single-view mammography would be adequate, depending on the density of the breast.[9,10] To re-invite women every 18 months, the ratio of women randomised to the invited group compared with controls was 1 to 1·2 in the age group 39–49 years and 1 to 1·6 in the age group 50–59 years. Women born in 1923–32 were invited to four screening rounds and women born 1933–44 to five.

Randomisation methods
Individual randomisation was used in the Malmö trials (webtable 1, http://www.lancet.com) and in the second part of the Göteborg trial (women born between 1936 and 1944) (webtable 2). During the first part of the Göteborg trial (women born between 1923 and 1935 and randomised between Dec 21, 1982, and Nov 3, 1983), day of birth was used for randomisation with varying days for each year-cohort.

The Stockholm trial used randomisation by day of birth. Between March, 1981, and April, 1982, women born on day 1–20 of the month between 1917 and 1941 were included, women born on day 1–10 in the invited group and women born on day 11–20 in the control group. Between May, 1982, and May, 1983, women born on day 21–30 between 1918 and 1942 were included in the invited group and women born on day 11–20 in 1942 were included in the controls (webtable 3).

The Two-County trial used cluster randomisation with geographic area (municipalities, parishes, or tax districts) as the unit of randomisation. Logistic problems with mobile mammographic units made individual randomisation unfeasible. The sparsely populated municipalities in Östergötland were grouped pairwise for size of population and geographic characteristics (adjacent municipalities constituted pairs as they were presumed to be similar in most respects). The larger population municipalities of Linköping, Norrköping, and Motala were
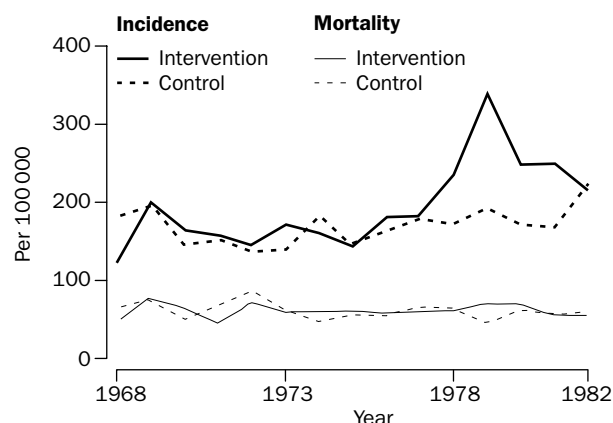
Figure 1: **Breast cancer incidence and mortality per 100 000 in 1968–82 in women randomised to invited group and control group in Östergötland trial**

split up into six, eight, and two clusters, respectively, of similar size, thereby creating three, four, and one pairs to increase the number of clusters (webtable 4). The clusters were randomly allocated to the invited or the control groups by tossing a coin under the supervision of the chairman of the County Council. 92 927 women lived in Östergötland; two of whom had a permanent address outside the county and 53 of whom did not have a permanent address. Thus 92 872 women were randomised.

To assess the comparability of the clusters for breast cancer risks, we studied breast cancer incidence and mortality before the start of the trial among women aged 40–74 years in the invited and control cluster areas in Östergötland (figure 1). Whilst breast cancer mortality was fairly constant during the 15-year period there was a slight increase in breast cancer incidence. During the pre-trial period 1968–77 in the cluster areas that were randomised to intervention the mean annual incidence per 100 000 was 162·4 and in the cluster areas not invited to intervention 162·0 (p=0·99). As expected, the incidence among the invited clusters was significantly higher than in the control clusters during 1978–82 (257·9 *vs* 185·8, p<0·001). In 1982 (that is, after two screening rounds) there was no difference in the breast cancer incidence between the two types of clusters. Breast cancer mortality per 100 000 was 60·6 and 63·4 (p=0·59), respectively, in the invited and control clusters during the pre-trial period and 62·7 and 57·7 (p=0·51) during the trial period. The similarity between the clusters in breast cancer incidence

and mortality before the start of the trial indicates that no significant bias was introduced by the clustering.

Inclusion and exclusion criteria
We excluded from this overview women with a diagnosis of an invasive epithelial breast cancer before randomisation, according to the Swedish Cancer Registry. Women in the Östergötland trial without a permanent address in one of the municipalities also had to be excluded, because they could not be randomised. Further, women born on day 31 of the month in the Stockholm trial were not randomised.

All analyses in the overview were based on exact age at randomisation, despite the fact that most trials, for practical reasons, used year-of-birth cohorts. Hence there are differences between the publications from each trial and the overview in number of women in the invited and control groups. The rationale for using exact age was to achieve uniformity between the trials. In addition, analyses of epidemiological studies are usually based on 5-year or 10-year intervals, and we consequently focused on the age group 40–74 years at entry. This means that a few women younger than 40 years in the Östergötland, Stockholm, and Göteborg trials were excluded, as were women 75 years and older in the Östergötland trial.

Endpoints
The primary endpoint in this overview was breast cancer as underlying cause of death according to the Swedish Cause of Death Registry. Data on deaths from other causes were also retrieved from the same register. We have previously compared breast cancer as the underlying cause with breast cancer present at death as determined by an independent endpoint committee in a blind review based on available clinical records and necropsy protocols with the officially recorded underlying cause of death. That comparison revealed close concordance.[11] Moreover, the estimated benefit associated with invitation to breast cancer screening was almost identical irrespective of endpoint and how cause of death was determined. Therefore, for this overview, we considered it appropriate to use the officially recorded underlying cause of death.

Data retrieval
The original file, selected from the official population register, which served as the basis for the randomisation, was obtained from the principal investigator of each of the participating trials. All individuals were identified through their unique identification number. For each woman, information on date of randomisation and allocation group was added. Records were linked to the six Regional Oncologic Centres to retrieve date of breast cancer

Age at randomisation	MMST I		MMST II		Östergötland		Stockholm		Göteborg		Total	
	IG	CG	IG	CG	IG	CG	IG	CG	IG	CG	IG	CG
≤39	··	··	··	··	··	1296	··	683	··	2 022		4001
40–44	··	··	3857	2923	5187	5354	7569	4519	5691	7141	22304	19937
45–49	3987	4067	5724	5289	5098	5105	6734	3502	5197	6062	26740	24025
50–54	4686	4244	··	··	5578	5552	7354	3907	4842	7456	22460	21159
55–59	4599	5078	··	··	6433	5943	8592	4514	5270	8541	24894	24076
60–64	3595	3601	··	··	5824	5486	8765	4476	··	··	18184	13563
65–69	3925	3914	··	··	5749	5376	125	60	··	··	9799	9350
70–74	296	291	··	··	5073	4859	··	··	··	··	5369	5150
75*	··	··	··	··		959	··	··	··	··		959
>75	··	··	··	··		14 000	··	··	··	··		14 000
Total		42 283		17 793		92 872		60 800		52 222		265 970
40–74	21 088	21 195	9581	8212	38 942	37 675	39 139	20 978	21 000	29 200	129 750	117 260

IG=invited group; CG=control group. Only women 40–74 years were included in overview. *Reached age 75 during the year of invitation.

Table 2: **Number of women randomised by trial study group and 5-year age group**

Trial	Age group	Trial time		Follow-up time					
				Invited group		Control group		Total	
		Median	Range	Median	Range	Median	Range	Median	Range
MMST I	45–70	18·8	13·9–20·2	19·2	18·3–20·2	19·2	18·3–20·2	19·2	18·3–20·2
MMST II	43–49	5·8	3·1–18·1	9·1	6·2–18·3	9·0	5·8–18·3	9·1	5·8–18·3
Östergötland	40–74	7·7	6·5–10·9	17·2	15·8–18·6	17·4	15·8–18·6	17·4	15·8–18·6
Stockholm	40–65	4·4	3·2–4·8	14·7	13·6–15·8	15·1	13·6–15·8	14·9	13·6–15·8
Göteborg	40–59	6·7	4·8–7·5	13·2	12·7–14·0	13·3	12·7–14·0	13·3	12·7–14·0
Overview									
5-year age group	40–44	7·0	3·1–18·1	14·6	7·8–18·6	14·7	5·8–18·6	14·7	5·8–18·6
	45–49	6·5	3·0–13·1	14·7	6·2–19·8	15·1	5·8–19·8	14·9	5·8–19·8
	50–54	4·9	3·2–8·7	15·6	13·3–19·9	15·7	13·3–19·9	15·6	13·3–20·2
	55–59	4·9	3·2–8·7	15·3	13·6–20·2	15·6	13·7–20·2	15·6	13·6–20·2
	60–64	4·5	3·2–8·7	15·8	13·6–20·2	17·1	14·7–19·8	16·8	13·6–19·8
	65–69	7·7	3·4–10·9	18·4	13·6–18·9	18·4	14·7–18·9	18·4	13·6–18·9
	70–74	9·2	6·6–10·9	17·4	15·8–18·6	17·2	15·8–18·6	17·4	15·8–18·6
10-year age group	40–49	6·6	3·0–18·1	14·7	6·2–19·8	14·9	5·8–19·8	14·8	5·8–19·8
	45–54	4·9	3·0–13·1	15·1	6·2–20·2	15·3	5·8–19·9	15·2	5·8–20·2
	50–59	4·9	3·2–8·7	15·6	13·3–20·2	15·6	13·3–20·2	15·6	13·3–20·2
	55–64	4·9	3·2–8·7	15·7	13·6–20·2	16·2	13·7–20·2	15·8	13·6–20·2
	60–69	7·0	3·2–10·9	17·1	13·6–19·8	17·9	14·7–19·8	17·4	13·6–19·8
	65–74	7·8	3·2–10·9	17·9	13·6–18·9	17·9	14·7–18·9	17·9	13·6–18·9
Total	*40–74*	*6·5*	*3·0–18·1*	*15·7*	*5·8–20·2*	*15·8*	*5·8–20·2*	*15·8*	*5·8–20·2*

Median and range. Trial time=length of time from date of randomisation until control group had first round of screenngor until Dec 31, 1996. Women with breast cancer before randomisation excluded). Follow-up until Dec 31, 1996.

Table 3: **Trial time for invited group and follow-up time by age at randomisation and trial**

diagnosis, and to the Swedish Cause of Death Register at Statistics Sweden to obtain date and cause of death. The end-date for the computerised follow-up was Dec 31, 1996.

Statistical methods
Women allocated to the control group were also invited to screening after a varying number of screening rounds in the invited group. However, this process did not apply to women born between 1908 and 1922 in MMST I, and to women aged 70–74 years in the Östergötland trial. To minimise problems related to possible dilution of the effect of screening in the invited group from screening in the control group, we developed two statistical models to analyse the outcome: the evaluation model and the follow-up model.[11] Briefly, the follow-up model includes as an event all diagnoses of breast cancer in women after date of randomisation, who died with breast cancer as the underlying cause before date of follow-up. The evaluation model ignores breast cancer deaths among women whose breast cancer diagnosis was made after the first screening round of the control group was completed.

We analysed data with QUEST, a program for statistical and epidemiological data analysis developed by Lennarth Gustafsson and Stig Wall, Umeå University. Relative risks (RR) were calculated by the density method, with total person-time experience of the cohort by time interval of follow-up as a basis for estimating the mortality rates. We calculated weighted RRs and 95% CI with the Mantel-Haenszel procedure.

Ethical considerations
This overview was approved by the Regional Ethics Committee of the Karolinska Institute, Stockholm.

Age at randomisation	Women-years×1000		Evaluation model				Follow-up model			
			Number of deaths				Number of deaths			
	IG	CG	IG	CG	RR	95% CI	IG	CG	**RR**	**95% CI**
5-year age group										
40–44	320	281	55	57	0·85	0·59–1·23	85	88	0·85	0·64–1·13
45–49	377	338	85	98	0·78	0·59–1·04	135	128	0·95	0·75–1·21
50–54	341	320	99	98	0·95	0·72–1·25	144	146	0·92	0·73–1·16
55–59	368	357	107	137	0·76	0·59–0·98	177	201	0·86	0·70–1·05
60–64	260	201	73	85	0·68	0·50–0·92	128	129	0·79	0·62–1·01
65–69	137	131	60	83	0·69	0·49–0·96	84	119	0·68	0·52–0·89
70–74	62	59	32	26	1·18	0·71–1·97	42	36	1·12	0·73–1·72
10-year age group										
40–49	697	620	140	155	0·80	0·63–1·01	220	216	0·91	0·76–1·09
45–54	718	658	184	196	0·86	0·70–1·05	279	274	0·93	0·78–1·11
50–59	709	677	206	235	0·84	0·70–1·01	321	347	0·88	0·75–1·03
55–64	628	559	180	222	0·73	0·60–0·89	305	330	0·83	0·71–0·97
60–69	397	332	133	168	0·67	0·53–0·84	212	248	0·73	0·61–0·87
65–74	199	190	92	109	0·81	0·61–1·07	126	155	0·78	0·62–0·99
Total										
40–74	1865	1688	511	584	0·79	0·70–0·89	795	847	0·85	0·77–0·94
40–74	··	··	··	··	0·80*	0·71–0·90	··	··	0·85*	0·77–0·94

RR and 95% CI. Follow-up until Dec 31, 1996. *Age-adjusted estimate.

Table 4: **All trials combined, number of 1000 women-years and number of cases with breast cancer as underlying cause of death according to Statistics Sweden by age at randomisation**

Role of funding source
The study sponsor had no role in the conduction of the study or writing of the report.

Results

Number of women randomised, trial time, and follow-up time
Our analysis was based on the follow-up of 247 010 women, 129 750 of whom were invited to mammography screening and 117 260 of whom were controls. 4001 women below the age of 40 and 14 959 women from Östergötland aged 75 and above were excluded. Age distribution by trial is in table 2.

Median trial time and range are in table 3. Trial time was defined as time from date of randomisation until the control groups completed the first round of screening. In trials in which the control groups were not invited to screening before the end of follow-up, trial time was defined as time from date of randomisation until date of follow-up (Dec 31, 1996). The median trial time in the overview was 6·5 years (range 3·0–18·1), varying from 4·4 in the Stockholm trial to 18·8 in MMST I. The median trial time by age at entry varied from 4·9 years in women 45–64 years to 7·8 years in women 65–74 years.

The follow-up time, defined as the time between date of randomisation and the end-date of follow-up (Dec 31, 1996) is also in table 3. The median follow-up time was 15·8 years (range 5·8–20·2), varying from 14·8 years in the 40–49-age group to 17·9 years in the 65–74-age group. The median follow-up time in the trials varied from 19·2 years in MMST I to 9·1 years in MMST II. There was no difference in median follow-up time between the invited group and the control group in the individually randomised trials. In the Göteborg trial, in which about 60% of the women were randomised in clusters, the difference was only 0·1 year (13·2 *vs* 13·3). As expected, the difference was largest in Östergötland , which used cluster randomisation, with a median follow-up time of 17·2 and 17·4 years, respectively, in the invited and control groups.

Breast cancer mortality
With the evaluation model, there were 511 breast cancer deaths in 1 864 770 women-years in the invited groups and 584 breast cancer deaths in 1 688 440 women-years in the control groups, resulting in a 21% significant reduction in breast cancer mortality associated with invitation to mammography screening (RR=0·79, 95% CI 0·70–0·89, table 4). The age-adjusted estimate was almost identical (0·80; 0·71–0·90). Tests of heterogeneity in terms of screening benefit by 5-year age groups (40–44, 45–49 to 70–74) and 10-year age groups (40–49, 50–59, 60–69, 70–74) were not significant (p=0·07 and 0·09, respectively). Trial-specific results are in table 5.

To assess the age-dependency of the effect of screening, RRs were calculated for consecutive 10-year and 5-year age groups (figure 2). In the 10-year age groups, the effect was significant in 12 consecutive age groups, 53–62, 54–63 to 64–73 years. For the 5-year age groups the pattern is unstable, although the effect is less for women 49–53 years and 50–54 years at randomisation (RR=0·97 and 0·95, respectively).

Our data do not support the possiblity that there is a difference in the effect of screening on the breast cancer mortality between the trials (test of heterogeneity, p=0·74).

The RRs for both the evaluation and the follow-up model by 5-year and 10-year age groups are presented in

table 4. As expected the difference between the two analytical models is now greater than in the earlier follow-up, with greater differences for the older trials. Overall, by the follow-up model, the RRs risk for all trials combined was 0·85 (0·77–0·94). This estimate was unaffected by adjustment for age.

Cumulative breast cancer mortality
The cumulative breast cancer mortalities per 100 000 women in the invited groups and the control groups by trial and age at entry are in figures 3 and 4. The curves for the age groups 55–64 and 60–69 years started to diverge earlier than for the age groups 40–49 and 50–59, whereas the curves for the 45–54 year age group hardly diverged at all. The greatest absolute differences were observed at ages 55 years and above. At 18 years after randomisation the absolute reduction for all women 40–74 years at entry was 136 per 100 000.

The absolute difference in cumulative breast cancer mortality between the invited and the control groups at 8, 12, and 16 years is in table 6. In general, the absolute effect increased up to 12 years after randomisation, whereafter it was maintained.

Total mortality
There were 22 398 deaths in 1 864 770 women-years in the invited group and 20 945 deaths in 1 688 440 women-years in the control group, resulting in an RR of 0·98 (0·96–1·00, table 7). Age-adjustment did not have any impact on the estimate. The RR was below 1·00 in all consecutive 10-year age groups except for 40–49 years at entry.
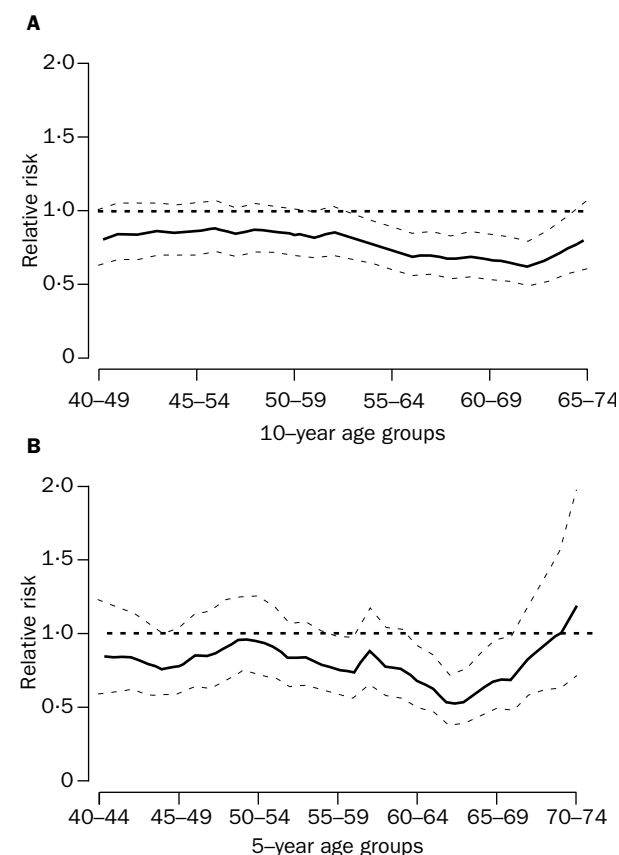

Figure 2: **Relative risk and 95% CI, evaluation model, all trials, follow-up until Dec, 1996**
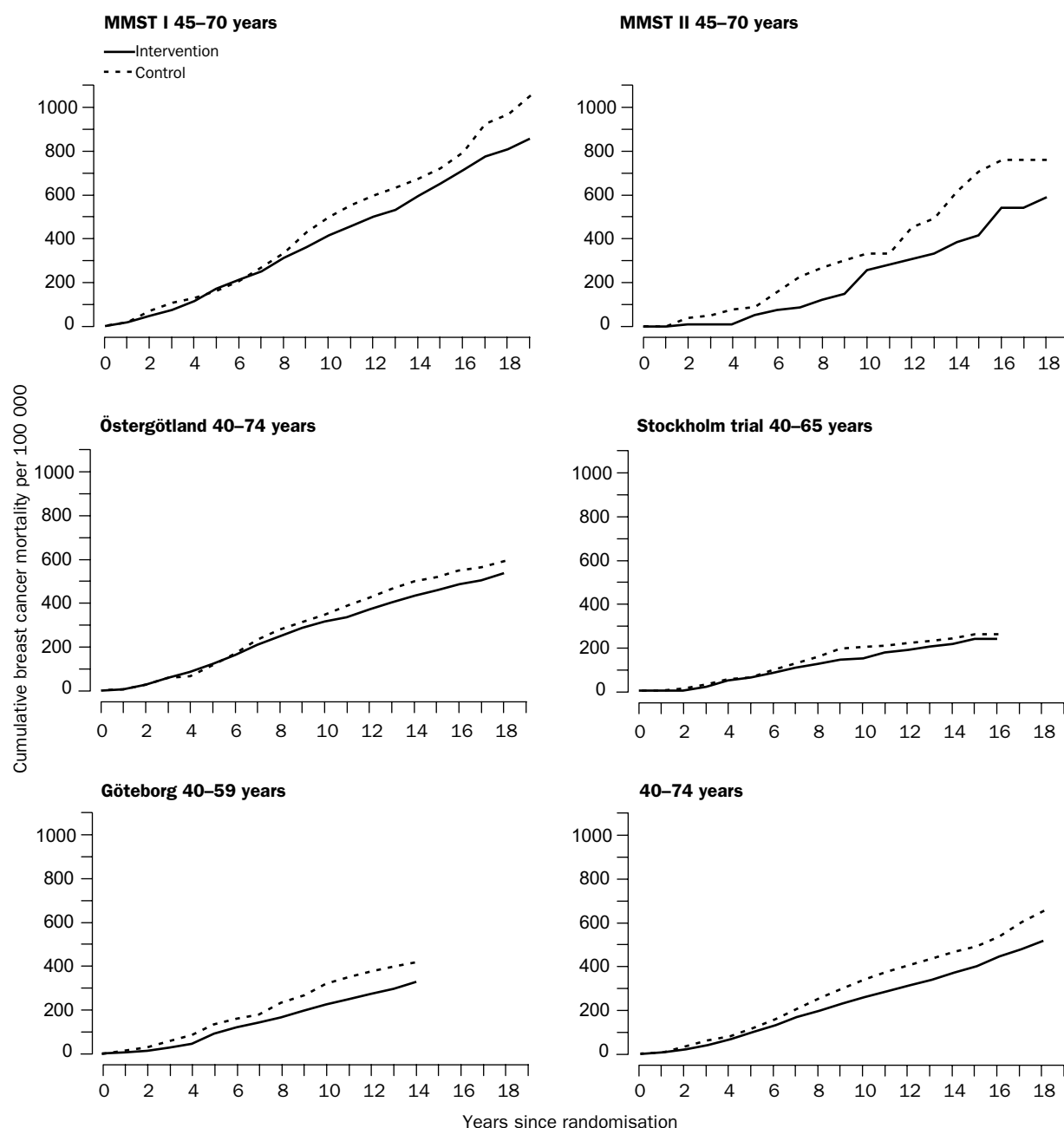(A) consecutive 10-year age groups, (B) consecutive 5-year age groups.

Figure 3: Cumulative breast cancer mortality per 100 000 in invited group and control group by trial and all trials
Evaluation method, follow-up until Dec, 1996.

Discussion

Our aim was to elucidate some issues that have been raised in recent reviews of the Swedish trials. In addition, we wanted to assess the long-term effects on mortality, including age-specific and trial-specific effects. Our latest overview, which is unbiased and unconfounded for study design, confirms and extends previous results. Our main observation was that the benefit of screening was maintained several years after the trials had been closed. In general, the benefit in absolute terms increased up to 12 years after randomisation and thereafter it was maintained.

The Kopparberg part of the Two-County trial was not available for this overview. The unavailability of Kopparberg data was due to a decision not to continue with the collaboration with the Swedish collaborative group by the Kopparberg trialists shortly after the publication of the first overview.[11] We regret this decision.

On the other hand, the continuation of the Malmö trial (MMST II) was added. The reason for not including MMST II in our previous overview was a decision at that time to restrict the analysis to the original Swedish trials.

Randomisation
Recently, concerns have been raised that the randomisation methods used in some of the Swedish trials of mammography screening may have been biased and that inappropriate exclusions in previous publications may have distorted the reported results.[3] Here we have presented in detail the randomisation methods used in the different trials.

MMST used individual randomisation stratified by year of birth. However, because of an administrative error the entire 1934-year birth cohort (n=1341) was invited to screening without randomisation. Also, there was slightly
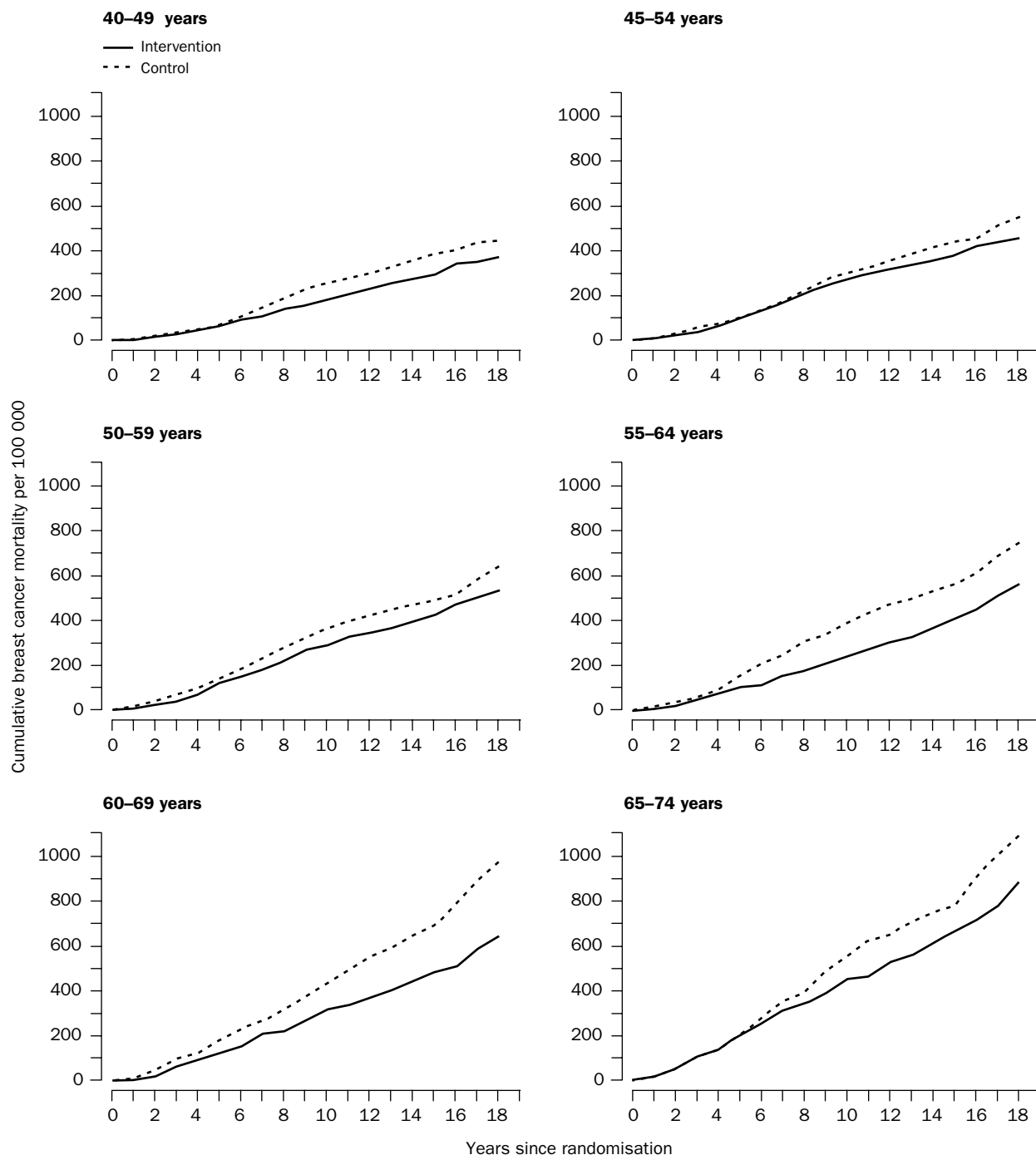
Figure 4: Cumulative breast cancer mortality per 100 000 in invited group and control group in women 40–49, 45–54, 50–59, 55–64, 60–69, and 65–74 years at entry
All trials, evaluation model, follow-up until Dec, 1996.

skewed distribution between invited women and controls in the 1929-year birth cohort. It could be argued that these women should be excluded in the mortality analysis. However, an analysis of MMST II based on the 1933 and 1935–45 year cohort resulted in an RR of 0·65 (95% CI 0·38–1·10). Exclusion of the 1929 cohort in MMST I decreased the RR of invited versus controls to 0·64. We concluded that these aberrations did not in any significant way change the estimated benefit associated with screening in MMST.

Cluster randomisation by day of birth, as used in the Stockholm and part of the Göteborg trial, may introduce

bias if used in a conventional treatment trial because the method implies foreknowledge of the allocated treatment of a potential participant. However, in a population-based trial, the day-of-birth method is unbiased, because there is no reason to assume that day of birth is related to outcome (death due to breast cancer). Moreover, since all women in a defined geographic area are included, there can be no inappropriate exclusions or inclusions on the basis of foreknowledge of allocation.

Cluster randomisation by geographic area may entail bias if the areas exhibit significant differences in pretrial characteristics related to the study outcome, and if, by

Trial	Age at randomisation	Women-years×1000		Number of breast cancer deaths		RR	95% CI
		IG	CG	IG	CG		
MMST I	45–54	158	151	71	78	0·87	0·63–1·20
	50–59	168	168	88	90	0·98	0·75–1·29
	55–64	141	149	63	83	0·80	0·57–1·12
	60–69	117	116	46	72	0·64	0·45–0·92
	70	4	4	3	3	0·98	0·15–6·60
	45–70	*360*	*362*	*161*	*198*	*0·82*	*0·67–1·00*
	..	..	..	..	..	*0·81**	*0·66–1·00*
MMST II	*43–49*	*113*	*86*	*29*	*33*	*0·64*	*0·39–1·06*
	..	..	..	..	..	*0·65**	*0·39–1·08*
Östergötland	40–49	172	176	31	30	1·05	0·64–1·71
	45–54	176	176	44	36	1·22	0·78–1·90
	50–59	194	185	53	54	0·94	0·66–1·35
	55–64	189	177	59	68	0·81	0·57–1·14
	60–69	166	155	64	83	0·72	0·52–1·00
	65–74	136	128	64	72	0·84	0·60–1·18
	40–74	*589*	*572*	*177*	*190*	*0·90*	*0·73–1·11*
	..	..	..	..	..	*0·89**	*0·72–1·09*
Stockholm	40–49	203	117	34	13	1·52	0·80–2·88
	45–54	196	107	25	15	0·89	0·47–1·67
	50–59	217	118	25	24	0·56	0·32–0·97
	55–64	229	122	39	28	0·75	0·46–1·21
	65	1	1	1	0	..	..
	40–65	*535*	*296*	*82*	*50*	*0·91*	*0·65–1·27*
	..	..	..	..	..	*0·90**	*0·63–1·28*
Göteborg	40–49	138	167	22	46	0·58	0·35–0·96
	45–54	128	173	30	47	0·86	0·55–1·36
	50–59	130	206	40	67	0·94	0·62–1·43
	40–59	*268*	*373*	*62*	*113*	*0·76*	*0·56–1·04*
	..	..	..	..	..	*0·78**	*0·57–1·07*

RR and 95% CI. Evaluation model, follow-up until Dec 31, 1996. *Age-adjusted estimate.

Table 5: **Number of 1000 women-years and number of cases with breast cancer as underlying cause of death according to Statistics Sweden in invited and control groups by age at randomisation and trial**

chance, the random allocation of the clusters fails to achieve a balance for these characteristics. The fact that the pretrial breast cancer incidence and mortality were almost identical in the clusters allocated to screening compared with the control clusters in the Östergötland trial indicates that the randomisation was successful.

Trial	Age at randomisation	Years since randomisation		
		8	12	16
MMST I	45–70	23	95	78
MMST II	43–49	148	146	217
Östergötland	40–74	29	52	62
Stockholm	40–65	32	30	22
Göteborg	40–59	70	101	..
Overview 5-year age group	40–44	9	21	51
	45–49	73	92	77
	50–54	−62	−34	−30
	55–59	163	185	114
	60–64	75	135	207
	65–69	117	219	334
	70–74	−90	−74	−161
10-year age group	40–49	49	73	61
	45–54	11	47	29
	50–59	56	81	47
	55–64	132	166	154
	60–69	101	179	270
	65–74	45	125	187
Total	*40–74*	*57*	*92*	*93*

Evaluation model, follow-up until Dec 31, 1996.

Table 6: **Differences in cumulative breast cancer mortality per 100 000 women at 8, 12, and 16 years after randomisation between control and invited groups by trial and age at randomisation**

The cluster randomisation used in the Östergötland, Stockholm, and part of the Göteborg trial may result in slight imbalances in the number of women allocated to the screening and control group as well as minor differences in mean age between the groups. Therefore it is a fallacy to interpret such marginal imbalances as an indication of biased allocation.

Our analyses here were not based on methods that formally take into account the fact that some of the trials were randomised by clusters. The rationale was that use of such methods in an overview of several trials with different methods for randomisation is not straightforward. Moreover, because of the mentioned lack of bias in any of the randomisation techniques that were used, there is no reason to assume that an alternative analytical approach would result in point estimates that differ from those we found. The only difference we would anticipate is slightly wider CIs, as was illustrated in a recent publication based on the Two-County trial.[12] In Östergötland the RR for the age group 40–74 years was 0·79 (0·66–0·96) with the Mantel-Haenszel method and 0·79 (0·64–0·97), 0·79 (0·64–0·97), 0·79 (0·61–1·00), and 0·79 (0·60–0·99), with four different logistic random-effects models. There were, as expected, no differences in the point estimates and the CIs were 0·01–0·05 wider depending on the model.

We are aware of the consequences of cluster randomisation, including the fact that the CI of the point estimate of the effect of the intervention will be wider. However, the Östergötland part of the Two-County trial has been analysed in detail.[12] That report applied different models allowing for different effects due to cluster randomisation. The effect on the CI was only marginal (by 0·01–0·02) under realistic assumptions. Moreover, the slightly excessive weight given to the Östergötland study,

Trial	Age at randomisation	Women—years×1000		Total no of deaths		RR	95% CI
		IG	CG	IG	CG		
MMST I	45–70	360	362	5672	5796	0·99	0·97–1·01
MMST II	43–49	113	86	402	300	1·03	0·89–1·20
Östergötland	40–74	589	572	10357	10036	0·98	0·95–1·01
Stockholm	40–65	534	296	4537	2572	0·99	0·95–1·03
Göteborg	40–59	268	373	1430	2241	0·94	0·88–1·00
Overview							
5-year age	40–44	320	281	921	791	1·03	0·92–1·15
group	45–49	377	338	1701	1534	0·99	0·89–1·10
	50–54	341	320	2393	2295	0·98	0·93–1·03
	55–59	368	357	4005	4169	0·93	0·89–0·97
	60–64	260	201	4850	3899	0·96	0·92–1·00
	65–69	137	131	4787	4643	0·99	0·96–1·02
	70–74	62	59	3741	3614	0·99	0·91–1·07
10-year age	40–49	697	620	2622	2325	1·00	0·95–1·06
group	45–54	718	658	4094	3829	0·98	0·94–1·02
	50–59	709	677	6398	6464	0·95	0·92–0·98
	55–64	628	559	8855	8068	0·98	0·96–1·01
	60–69	397	332	9637	8542	0·94	0·91–0·97
	65–74	199	190	8528	8257	0·99	0·96–1·02
Total	*40–74*	*1865*	*1689*	*22398*	*20945*	*0·98*	*0·96–1·00*
	40–74	*..*	*..*	*..*	*0·98**	*0·96–1·00*	*..*

Follow-up model until Dec 31, 1996 *Age-adjusted estimate.

Table 7: **Total mortality in invited and control groups**

without allowance for the cluster randomisation, only makes the overall estimate of the effect of screening more conservative because Östergötland had the lowest effect. Further, there are no validated well-accepted statistical methods to meta-analyse trials, some of which are individually randomised and others are randomised by clusters, while at the same time allowing for the effects of the cluster randomisation. For the Swedish trials of mammography screening such an analysis would be expected only to produce a marginally wider CI of the point estimate of the effect of screening in one of the trials, and an even smaller effect on the CI for the estimate of the effect based on all trials. The point estimate of the effect in the analysis of all trials would only be marginally more extreme.

Number of randomised women reported
Our latest overview as well as all previous reports of the Swedish overview was based on files from local population registers including all women in the areas covered by the trials. We have presented detailed information on which women were excluded, that is, women with a diagnosis of invasive breast cancer before randomisation and those without permanent address in the Östergötland trial. To avoid any possibility of bias, information on previous breast cancers among women in the invited and the control groups was obtained through computerised record linkage with the Swedish Cancer Registry. The rationale for the exclusion of women with previous invasive breast cancer was that the risk of death due to that breast cancer was considered not to be influenced by an invitation to screening. Such women would, therefore, tend to inappropriately dilute the observed effect of the intervention.

The only other exclusions were the 1154 women born on day 31 in the Stockholm trial and the 55 women in the Östergötland trial who did not have a registered permanent address and therefore could not be allocated to a geographic cluster.

The reported numbers of randomised women have differed slightly in some previous reports from the individual trials. For instance, in publications from the Two-County trial up to and including 1987, the number

of women in the Östergötland part of the study in the invited and control group was 39 034 and 37 936, and from 1989, 38 491 and 37 403. The post-1989 figures represent the number of women after exclusion of cases with a history of breast cancer before randomisation. The figures deviate slightly from the figures in our follow-up (38 942 and 37 675 in the invited and control group, respectively).[1] As we pointed out in our previous publications, this discrepancy is explained by the fact that the trial reports refer to all randomised women according to their birth cohort, whereas the overview figures[1] refer to randomised women aged exactly 40–74 years at randomisation.

In reports from the Stockholm trial, the number of women in the invited group and the control group was, in some reports, approximated at 40 000 and 20 000. However, such approximations were never used in the statistical analysis of the Stockholm trial. The differences in figures between earlier reports and the present overview are due to a difference in the definition of age at entry. The overview[1] of the Swedish trials used only exact age, and consequently women who were 39 years at randomisation (347 in the invited group and 336 controls) were excluded. Thus, in the present overview, the Stockholm trial contributed 39 139 women to the invited group and 20 978 to the control group.

Precision in point estimates
Point estimates are affected by contamination and dilution. The development of the evaluation model was done to minimise the effect of dilution due to invitation of the control group to screening. One possible remaining effect involves the fact that women born between 1908 and 1922 in the control group in MMST I were never invited to screening. Women born between 1908 and 1917 were invited to the sixth and last round in 1986, women born in 1918 were invited to the seventh and last round in 1988, and women born between 1919 and 1922 were invited to the eighth and last round in 1989 at the ages of 70–78, 70, and 67–70 years, respectively. Even if these women were diagnosed with breast cancer and died with breast cancer as the underlying cause of death during the period between the last screening and the time for

follow-up, Dec 31, 1996, they were included in the analysis. For women born between 1908 and 1917, the length of time since the last screening could have been up to 10 years. The reason for not taking this into account involved the problem of introducing lead-time bias. The situation was the same for women 70–74 years at randomisation in the Östergötland trial.

Variation in point estimates by age and trial
Our latest overview confirms the results of the earlier overview.[1] The main finding is that the benefit remained several years after the trials had been closed. The difference between median follow-up time and median trial time was 9·3 years. It was 0·4 years in MMST I, 3·3 years in MMST II, 6·6 years in the Göteborg trial, 9·7 years in the Östergötland trial, and 10·5 years in the Stockholm trial. The figure for MMST I has to be interpreted with caution because of the definition of the concept of trial time. The trial and follow-up time were, for women 45–54 years at entry, 14·6 and 19·2 years, respectively—a difference of 5·6 years—while for women aged 55–70 years the follow-up and trial time were almost identical. The duration from the date when the first round was finished for the control group until date for follow-up was, in Östergötland and Stockholm, around 10 years; thus, if the effect of the intervention levels off, it should be possible to detect such an effect.

Figure 2A and 2B both indicate a difference in effect by age at randomisation. This difference was not supported by the test of heterogeneity; however, the power of the test is low.

The finding of virtually no effect in women in the 5-year age groups 49–53 and 50–54 years at randomisation is surprising. A similar result was presented in the Edinburgh trial for the 50–54 age group.[13] With a median trial time in this age group of 4·9 years these women were 49–58 and 50–59 years, respectively, at invitation to screening. The explanation is not clear: an association with hormonal changes during menopause is possible.

Another question is whether there is an effect due to the intensity of the intervention in the trials—ie, number of screening rounds, screening interval, and number of projections and whether single or double reading. To answer this question properly would have required a specific study design in which different trials of proper size applied different screening modalities throughout. This was not the case with the trials involved in the overview; thus it is not possible to evaluate whether MMST and the Göteborg trial, which both had shorter screening intervals and double reading (the Göteborg trial only at the end of the study) and two-view mammography, or MMST I, which invited women to at least six screening rounds, had a better effect than the Östergötland and Stockholm trial, which invited women to two to four and two rounds, respectively, and used a longer screening interval. However, a comparison of MMST plus the Göteborg trial with the Stockholm and Östergötland trials for the age groups 40–59 and 40–64, where all four trials were fairly well represented, gave age-adjusted RRs of 0·86 versus 0·88 and 0·82 versus 0·85, respectively. Thus the effect seems to be similar.

Tabár et al[14] presented updated mortality results up to and including 1998 from the Two-County trial. The RRs for the Östergötland part were, in the 40–49, 50–59, 60–69, and 70–74 year age groups, 1·06, 0·76, 0·65, and 0·73, respectively. Thus the effect seems to be much larger than in the present follow-up up to and including 1996. The difference may be due to different criteria for cause of death determination and differences in definition

of age. We used official data from Statistics Sweden, while in the Two-County trial a local team of physicians determined the cause of death. The fact that the Two-County team was not blinded for the screening status of the deceased women may have introduced a bias. We have earlier shown[11] that cause of death determination according to Statistics Sweden results in a more conservative estimate than a blind determination of cause of death by an independent endpoint committee (0·80 *vs* 0·77 with the follow-up model). The difference was greater for the Östergötland trial (0·89 *vs* 0·82). We also used exact age, while the Two-County trial used birth cohort, thus their 40–49 year age group also contains women 39 years old, and so on.

In all trials there was an increased relative beneficial effect during the first 4–10 years followed by a few years with constant relative effect and a few years with decreasing relative effect. The absolute effect increased during the first 12 years. The absolute effect at 16 years of follow-up has to be interpreted with caution, because only MMST I and Östergötland and women randomised in 1978 (birth-year cohort 1933–34) in MMST II could contribute to the estimate (table 6). Similarly almost only MMST I contributed to the 18-year follow-up estimate

Total mortality
Another concern raised about the Swedish trials was that the screening cohorts appeared to exhibit a higher total mortality than the controls in the follow-up up to and including 1989. However, this is based on a misunderstanding that age-adjustment in our previous report,[15] which resulted in a non-significant difference in the total mortality between the cohorts, was inappropriate. As we said above, cluster randomisation may result in slight imbalances in the age distribution, which makes age-adjustment necessary and appropriate in analyses of total mortality because age is a strong determinant. When such an adjustment was made there was, as expected, no significant difference between the invited and the control groups.

In our overview, there was a 2% lower total mortality in the invited cohorts. This estimate was not changed by adjustment for age. Mortality from breast cancer in the age groups 50–59, 60–69, 70–79, and 80–84 years in Sweden in 1990 constituted 13·2%, 7·2%, 3·2%, and 2·1%, respectively, of the total female mortality. Because the median age at death in women 40–49, 50–59, 60–69, and 70–74 years at entry was 55, 65, 75, and 82 years, respectively, and the relative effect according to the follow-up model was 9%, 12%, 17%, and –12%, respectively, the expected effect of the intervention on the total mortality was 2·3%, which is in accordance with our results.

Conclusion
The effect of breast screening in terms of breast cancer mortality reduction persists after long-term follow-up. The effect is age-dependent: highest effect in women aged 55–69 years at randomisation and lowest in women aged 50–54 years at randomisation. Further, we conclude that the recent criticism against the Swedish randomised controlled trials is misleading and scientifically unfounded.

the overview group and coordinated the work. He also contributed to the design of the study and drafting of the manuscript. BN has represented the Östergötland study and contributed to the drafting of the manuscript.

References

1 Nyström L, Rutqvist LE, Wall S, et al. Breast cancer screening with mammography: overview of Swedish randomised studies. *Lancet* 1993; **341:** 973–78.

2 Larsson L-G, Andersson I, Bjurstam N, et al. Updated overview of the Swedish randomised trials on breast cancer screening with mammography: age group 40–49 at randomisation. *J Natl Cancer Inst Mono* 1997; **22:** 57–61.

3 Götzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet* 2000; **355:** 129–34.

4 Andersson I, Aspegren K, Janzon L, et al. Mammographic screening and mortality from breast cancer: the Malmö Mammographic Screening Trial. *BMJ* 1988; **297:** 943–48.

5 Andersson I, Janzon L. Reduced breast cancer mortality in women under age **50:** updated results from the Malmö Mammographic Screening Program. *J Natl Cancer Inst Mono* 1997; **22:** 63–67.

6 Tabár L, Fagerberg CJG, Gad A, et al. Reduction in mortality from breast cancer after mass screening with mammography: randomised trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health and Welfare. *Lancet* 1985; **1:** 829–32.

7 Frisell J, Glas U, Hellström L, Somell A. Randomised mammographic screening for breast cancer in Stockholm: design, first round results and comparison. *Breast Cancer Res Treat* 1986; **8:** 45–54.

8. Frisell J, Eklund G, Hellström L, Lidbrink E, Rutqvist LE, Somell A. Randomised study of mammography screening in the Stockholm trial. *Breast Cancer Res Treat* 1991; **18:** 49–56.

9 Bjurstam N, Björneld L, Duffy SW, et al. The Gothenburg Breast Screening Trial: first results on mortality, incidence, and mode of detection for women ages 39–49 years at randomisation. *Cancer* 1997; **80:** 2091–99.

10 Bjurstam N, Björneld L, Duffy SW, et al. The Gothenburg Breast Cancer Screening Trial: preliminary results on breast cancer mortality for women aged 39–49. *J Natl Cancer Inst Mono* 1997; **22:** 53–55.

11 Nyström L, Larsson L-G, Rutqvist LE, et al. Determination of cause of death among breast cancer cases in the Swedish randomised mammography screening trials: a comparison between official statistics and validation by an endpoint committee. *Acta Oncol* 1995; **34:** 145–52.

12 Nixon R, Prevost TC, Duffy SW, Tabár L, Vitak B, Chen HH. Some random-effects models for the analysis of matched-cluster randomised: application to the Swedish two-county trial of breast-cancer screening. *J Epidemiol Biostat* 2000; **5:** 349–58.

13 Alexander F, Anderson TJ, Brown HK, et al. 14 years of follow-up from the Edinburgh randomised trial of breast-cancer screening. *Lancet* 1999; **353:** 1903–08.

14 Tabár L, Vitak B, Chen H-H, 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.

15 Nyström L, Larsson L-G, Wall S, et al. The overview of the Swedish randomised mammography trials: the total mortality pattern and the representivity of the study cohorts. *J Med Screen* 1997; **3:** 85–87.

Uses of error

A breathtaking patient

Maarten Boers

The error that still haunts me occurred when I was a resident in internal medicine. I was on call over the weekend when I was paged to the emergency room to see a man in his early twenties whom I had seen only a few nights before. He had presented with typical signs and symptoms of hyperventilation, including sweating, shortness of breath, and a sensation of central chest pressure. I couldn't find anything abnormal on physical examination and I don't remember whether I had ordered arterial blood gases (or any other lab tests) at that time. I did remember that he had been somewhat recalcitrant, so I was irritated at the prospect of having to see him again. I ordered an electrocardiogram over the telephone and went to the emergency room to shoo him out before the serious cases started to arrive. As I entered, the nurse told me he wasn't well at all and asked me to look at him straight away. I only glanced at the electrocardiogram and did not register anything. We went into his room, saw the patient take a few very deep breaths, stop breathing and lose consciousness. Confidently I told the nurse not to worry, that it was possible to wash out so much CO_2 that one lost consciousness. I started counting silently, and after 10 he was still apnoeic. I felt for his femoral pulse, and found none! We immediately started resuscitation, and he responded. The electrocardiogram showed (and lab tests later confirmed), a large anterior myocardial infarction. I had refused to look at the electrocardiogram because I assumed it would be normal, and I intended to use it to convince the patient to leave me alone. This patient taught me that although pattern recognition may be essential for an efficient delivery of care, it must be balanced by a strong reflex of self doubt, and a clear-headed appraisal of every bit of evidence.

Department of Clinical Epidemiology and Biostatistics, Vrije Universiteit Medical Center, 1007 MB Amsterdam, Netherlands (Prof M Boers MD)