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

Lennarth Nyström, Ingvar Andersson, Nils Bjurestam, 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).

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.

Lancet 2002; 359: 909–19

See Commentary page 904

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 and an update focusing on the age group 40–49 in 1997. Concerns raised about the validity of the results from the trials 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 (cranio-caudal and oblique) and in
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. 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. 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.

**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. 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 randomised between Dec 21, 1982, and Nov 3, 1983, 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
intervention group in breast cancer incidence and mortality per 100,000 in
Figure 1: 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 the mean annual incidence per 100,000
increase was 162.0 (p=0.99). As expected, the incidence
was 185.8, vs 162.4 and in the cluster areas not invited to
intervention the mean annual incidence per 100,000
was 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
was added. Records were linked to the six Regional
register, which served as the basis for the randomisation,
participating trials. All individuals were identified through
the original file, selected from the official population
register, which served as the basis for the randomisation,
therefore, for this overview, we considered it appropriate
to use the officially recorded underlying cause of death.

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

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

<table>
<thead>
<tr>
<th>Age at randomisation</th>
<th>MMST I IG</th>
<th>MMST I CG</th>
<th>MMST II IG</th>
<th>MMST II CG</th>
<th>Östergötland IG</th>
<th>Östergötland CG</th>
<th>Stockholm IG</th>
<th>Stockholm CG</th>
<th>Göteborg IG</th>
<th>Göteborg CG</th>
<th>Total IG</th>
<th>Total CG</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;39</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
</tr>
<tr>
<td>40–44</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
</tr>
<tr>
<td>45–49</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
</tr>
<tr>
<td>50–54</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
</tr>
<tr>
<td>55–59</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
</tr>
<tr>
<td>60–64</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
</tr>
<tr>
<td>65–69</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
</tr>
<tr>
<td>70–74</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
</tr>
<tr>
<td>75*</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
</tr>
<tr>
<td>&gt;75</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
</tr>
<tr>
<td>Total</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
<td>3526</td>
</tr>
</tbody>
</table>

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

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. 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
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,
among 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.
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 1918 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 Lennarh 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.

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

Median and range. Trial time—length of time from date of randomisation until control group had first round of screening 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

<table>
<thead>
<tr>
<th>Age at randomisation</th>
<th>Women-years ×1000</th>
<th>Evaluation model</th>
<th>Number of deaths</th>
<th>RR</th>
<th>95% CI</th>
<th>Follow-up model</th>
<th>Number of deaths</th>
<th>RR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-year age group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40–44</td>
<td>320</td>
<td>281</td>
<td>55</td>
<td>57</td>
<td>0.85</td>
<td>0.59–1.23</td>
<td>85</td>
<td>88</td>
<td>0.85</td>
</tr>
<tr>
<td>45–49</td>
<td>377</td>
<td>338</td>
<td>85</td>
<td>98</td>
<td>0.78</td>
<td>0.59–1.04</td>
<td>135</td>
<td>128</td>
<td>0.95</td>
</tr>
<tr>
<td>50–54</td>
<td>341</td>
<td>320</td>
<td>99</td>
<td>98</td>
<td>0.95</td>
<td>0.72–1.25</td>
<td>144</td>
<td>146</td>
<td>0.92</td>
</tr>
<tr>
<td>55–59</td>
<td>368</td>
<td>357</td>
<td>107</td>
<td>137</td>
<td>0.76</td>
<td>0.59–0.98</td>
<td>177</td>
<td>201</td>
<td>0.86</td>
</tr>
<tr>
<td>60–64</td>
<td>260</td>
<td>201</td>
<td>73</td>
<td>85</td>
<td>0.68</td>
<td>0.50–0.92</td>
<td>128</td>
<td>129</td>
<td>0.79</td>
</tr>
<tr>
<td>65–69</td>
<td>137</td>
<td>131</td>
<td>60</td>
<td>83</td>
<td>0.69</td>
<td>0.49–0.96</td>
<td>84</td>
<td>119</td>
<td>0.68</td>
</tr>
<tr>
<td>70–74</td>
<td>62</td>
<td>59</td>
<td>32</td>
<td>26</td>
<td>1.18</td>
<td>0.71–1.97</td>
<td>42</td>
<td>36</td>
<td>1.12</td>
</tr>
<tr>
<td>10-year age group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40–49</td>
<td>697</td>
<td>620</td>
<td>140</td>
<td>155</td>
<td>0.80</td>
<td>0.63–1.01</td>
<td>220</td>
<td>216</td>
<td>0.91</td>
</tr>
<tr>
<td>45–54</td>
<td>718</td>
<td>658</td>
<td>184</td>
<td>196</td>
<td>0.86</td>
<td>0.70–1.05</td>
<td>279</td>
<td>274</td>
<td>0.93</td>
</tr>
<tr>
<td>50–59</td>
<td>709</td>
<td>677</td>
<td>206</td>
<td>235</td>
<td>0.84</td>
<td>0.70–1.01</td>
<td>321</td>
<td>347</td>
<td>0.88</td>
</tr>
<tr>
<td>55–64</td>
<td>628</td>
<td>559</td>
<td>180</td>
<td>222</td>
<td>0.73</td>
<td>0.60–0.89</td>
<td>305</td>
<td>330</td>
<td>0.83</td>
</tr>
<tr>
<td>60–69</td>
<td>397</td>
<td>332</td>
<td>133</td>
<td>168</td>
<td>0.67</td>
<td>0.53–0.84</td>
<td>212</td>
<td>248</td>
<td>0.73</td>
</tr>
<tr>
<td>65–74</td>
<td>199</td>
<td>190</td>
<td>92</td>
<td>109</td>
<td>0.81</td>
<td>0.61–1.07</td>
<td>126</td>
<td>155</td>
<td>0.78</td>
</tr>
<tr>
<td>Total</td>
<td>40–74</td>
<td>1865</td>
<td>1688</td>
<td>511</td>
<td>584</td>
<td>0.79</td>
<td>0.70–0.89</td>
<td>795</td>
<td>847</td>
</tr>
<tr>
<td>40–74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| 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

For personal use. Only reproduce with permission from The Lancet Publishing Group.
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.

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.

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 possibility 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.
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.5 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
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

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.
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.11 In Östergötland the RR for the age group 40–74 years was 0·79 (0·64–0·97), 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 than 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.11
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 37 075 in the invited and control group, respectively). 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 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 figures refer to randomised women aged exactly 40–74 years at randomisation.

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 well represented, gave age-adjusted RRs of 0·86 versus 0·85, respectively. Thus 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 70–74 years at randomisation. Further, we conclude that the recent criticism against the Swedish randomised controlled trials is misleading and scientifically unfounded.

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. We have earlier shown3 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.

Contributors

LN contributed to the design of the study, and performed the record linkage and statistical analysis as well as the drafting of the preliminary manuscript. IA, NB, and JP are the main trialists of the Malmö, Göteborg, and Stockholm studies, respectively, and contributed to the design of the study and the drafting of the manuscript. LR is chairman of
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.

Conflict of interest statement
None declared.

Acknowledgments
We thank Dr Lars Hellström, who initiated the Stockholm Mammography Screening Trial, and Dr Gunnar Fagerberg, who was the radiologist responsible for the Östergötland part of the Two-County trial, for making this overview possible. We also thank Prof Gunnar Eklund for important information about the planning phase of the Swedish trials. The Swedish Cancer Society supported this study.

Tables with number of women randomised to the invited group and the control group by year of birth and date of randomisation in the Malmö trial, by year of birth, date of randomisation, and method of randomisation in the Göteborg trial, by day of birth and date of randomisation in the Stockholm trial, and by date of randomisation and cluster in the Östergötland part of the Two-County trial can be requested by e-mail from the first author.

References

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 shoosh 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 CO2 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.