The Excess of Patients with Advanced Breast Cancer in Young Women Screened with Mammography in the Canadian National Breast Screening Study

Robert E. Tarone, Ph.D.

Background. An unexpected excess of patients with breast cancer with four or more positive lymph nodes was observed in mammographically screened women who were younger than the age of 50 years at enrollment into the Canadian National Breast Screening Study (NBSS). It has been suggested that this excess is consistent with prior screening trials evaluating mammography.

Methods. A quantitative evaluation of the distribution of patients with breast cancer with four or more positive lymph nodes in the NBSS was undertaken, and the percentages of patients with breast cancer who were at an advanced state at diagnosis in the NBSS and in previous randomized screening trials were compared. The validity of mortality analyses after eliminating advanced cases detected by physical examination at the initial screening visit is examined.

Results. The excess of patients with cancer with four or more positive lymph nodes in the 40-49-year mammography age group of the NBSS was statistically significant, even when expressed as a percentage of all invasive cancers diagnosed. Such an excess is inconsistent with published data on extent of disease at diagnosis from previous studies. Analysis of NBSS mortality data after eliminating advanced cases detected by physical examination at the initial visit should result in minimal, if any, bias.

Conclusions. Mortality analyses eliminating advanced cases detected by physical examination at the initial screening visit may be less susceptible to bias caused by possible nonrandom allocation of study participants and should be considered in future evaluations of the NBSS cohort after longer follow-up periods and in meta-analyses that may include the NBSS in assessments of the efficacy of mammography. Cancer 1995; 75:997-1003.

Key words: breast cancer, mammography, pseudo-variable, randomized trial.

The Canadian National Breast Screening Study (NBSS) is the largest randomized study to date designed to investigate the effectiveness of mammography in women younger than the age of 50 years. Because of the size and recency of the NBSS, it has received considerable weight in deliberations on mammography guidelines for younger women. Of women younger than the age of 50 years at enrollment into the NBSS, half were assigned to be screened annually using mammography and physical breast examination, and half were assigned to be screened only at enrollment, and then only by physical breast examination. The observation after 7 years of follow-up of 10 more deaths due to breast cancer in the group screened annually compared with the group screened only at enrollment intensified an existing debate on the effectiveness of mammography in young women. The excess breast cancer mortality in the women screened with mammography even led to speculation in press articles about how mammography might have been directly responsible for the increased mortality.

An unexpected result of the NBSS was a substantial excess of patients with advanced disease diagnosed at the initial screening visit in the mammography group for women younger than the age of 50 years at enrollment. This disproportionate assignment of patients with a poor prognosis to the NBSS 40-49-year mammography group raises questions about the relevance of the subsequent mortality excess in this group to determinations of the efficacy of mammography. Indeed, the initial excess of patients with advanced disease assigned to the mammography group was identified as the main reason for the increased breast cancer
mortality in young women screened by mammography in the NBSS.\(^1\)

It has been argued that the excess of advanced disease in the mammography group is consistent with prior randomized studies.\(^1,12,18\) It also has been suggested that the availability of screening mammograms in subsequent evaluations at NBSS review clinics of women in whom abnormalities were detected by physical examination during a screening visit may have led to an increased diagnosis of breast cancers, including lymph node positive cancers, in the mammography group.\(^1,12,18\) Both of these rationales attempting to explain the initial excess of advanced disease in the NBSS mammography group are investigated. In addition, the assertion is examined that inferences are not valid if patients with advanced disease detected by physical examination at the initial screening visit are removed before mortality analyses.\(^12\)

**Materials and Methods**

**NBSS Design and Data**

The comparison groups in the NBSS varied by age at recruitment into the trial. For women 40–49 years of age at recruitment, 25,214 patients were assigned to receive annual mammography and physical examination (MP), whereas 25,216 were assigned to the control group to receive a single physical examination at enrollment into the study and usual medical care thereafter (UC).\(^1\) For women 50–59 years of age at recruitment, 19,711 were assigned to receive annual mammography and physical examination (MP), whereas 19,694 were assigned to receive an annual physical examination alone (PO).\(^19\) The numbers of invasive breast cancers and the extent of disease at the time of diagnosis were obtained from Table 7 of each of the publications reporting the 7-year follow-up results of the NBSS.\(^1,19\) The total numbers of breast cancers, including in situ breast cancers, were obtained from Table 10 of each publication.\(^1,19\)

**Statistical Analyses**

Statistical analyses comparing the number of breast cancers in each mammography group with the number in its corresponding control group were based on the locally most powerful test assuming that the number of cancers followed a Poisson distribution.\(^20\) That is, if there were \(X_1\) breast cancers of some type in the mammography group and \(X_2\) breast cancers in the control group, then the \(P\) value for the comparison of \(X_1\) and \(X_2\) was calculated using the binomial distribution with a sample size of \(X_1 + X_2\) and a probability of success equal to one-half. The proportions of breast cancers diagnosed at a particular advanced state of disease in the mammography and control groups were compared using Fisher's exact test. Two-tailed \(P\) values were reported for all analyses.

**Results**

Table 1 summarizes the distribution of all patients with invasive breast cancer diagnosed in the NBSS as well as patients with four or more positive lymph nodes at diagnosis by age at entry into the study, and by year and detection method. Eliminating patients with breast cancers detected by mammography only, there were 17 with four or more positive lymph nodes detected in the MP group at the initial screening in the 40–49-year age group compared with only 5 in the tgroup \((P = 0.017)\).\(^1\) During the first 7 years of follow-up in this age group, there were 47 patients with breast cancer with four or more positive lymph nodes in the MP group compared with 23 in the UC group \((P = 0.006)\). Most of this excess of patients with a poor prognosis was evident within 12 months of entry into the study (i.e., 25 patients in the MP group compared with 7 in the UC group, \(P = 0.002\)). In women 50–59 years old at recruitment, there were no significant differences in the number of patients diagnosed with breast cancer with four or more positive lymph nodes either at the initial screening or during the first 7 years of follow-up \((P > 0.5)\).

Mammography potentially can detect tumors before they become apparent symptomatically or are palpable, and one thus would expect the proportion of cancers detected at an advanced stage to be smaller in groups screened by mammography than in groups who were not. Contrary to this expectation, of the 86 patients with invasive cancer diagnosed at the initial screening in the MP group for women 40–49 years of age, 22% (19) had four or more positive lymph nodes compared with only 8% observed (Table 1) in the UC group \((P = 0.043)\). In the first 7 years of follow-up, in the 40–49-year age group, 14% of patients diagnosed in the MP group had four or more positive lymph nodes compared with 8% in the UC group \((P = 0.037)\). Even including in situ tumors in the analysis, the proportions remained higher in the MP group than in the UC group at the initial screening \((19/98 [19\%] \text{ vs. } 5/62 [8\%], P = 0.068)\) and after 7 years \((47/397 [12\%] \text{ vs. } 23/296 [8\%], P = 0.097)\).

In the NBSS 50–59-year age group, the expected lower percentage of patients with advanced disease was observed in the group screened by mammography, despite the fact that women in the PO group, unlike those
Table 1. Proportion* (%) of Patients With Invasive Breast Cancer Diagnosed With Four or More Positive Lymph Nodes in the Canadian National Breast Screening Study by Year and Method of Detection

<table>
<thead>
<tr>
<th>Age 40-49 yr</th>
<th>Age 50-59 yr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mammogram and physical examination (%)</td>
<td>Usual care (%)</td>
</tr>
<tr>
<td>Screen 1:Pt</td>
<td>17/65 (26)</td>
</tr>
<tr>
<td>Screen 1:MO</td>
<td>2/21 (10)</td>
</tr>
<tr>
<td>&lt; 12 Months</td>
<td>6/16 (38)</td>
</tr>
<tr>
<td>Screens 2–5:Pt</td>
<td>8/73 (11)</td>
</tr>
<tr>
<td>Screens 2–5:MO</td>
<td>1/46 (2)</td>
</tr>
<tr>
<td>Years 2–5</td>
<td>10/65 (15)</td>
</tr>
<tr>
<td>≥ Year 6</td>
<td>3/45 (7)</td>
</tr>
<tr>
<td>Total</td>
<td>47/331 (14)</td>
</tr>
</tbody>
</table>

* Denominator is the number of patients with invasive breast cancer diagnosed, and numerator is the number with four or more positive lymph nodes.
† Pt: detected by physical examination.
‡ MO: detected by mammography only.
§ Interval or incident cases.

in the UC group, received annual physical breast examinations. Of 119 patients with invasive breast cancer diagnosed at the initial screening in the 50–59-year MP group, 8% (9) had four or more positive lymph nodes compared with 17% observed (Table 1) in the PO group (P = 0.079). In the first 7 years of follow-up in the 50–59-year age group, 9% of patients diagnosed in the MP group had four or more positive lymph nodes compared with 11% in the PO group (P = 0.36). Including in situ cancers resulted in a significantly lower proportion of advanced disease in the MP group at the initial screening (9/142 [6%] vs. 11/68 [16%], P = 0.041) but not for the 7-year follow-up period (32/438 [7%] vs. 34/333 [10%], P = 0.16).

Table 2 presents the percentage of patients with high-stage disease among all patients with invasive breast cancer diagnosed in ongoing and previous randomized trials designed to evaluate breast cancer screening with protocols including mammography. The classification of advanced cancers as Stage II or higher was available for all other studies, and, thus, this classification was used to provide a common measure of disease severity for the largest number of studies. The Swedish two-county data are from patients diagnosed in the first 8 years of follow-up. The Stockholm data are from patients diagnosed in the first 5 and 6 years after enrollment, because the control group was offered mammographic screening during the sixth year. The Malmo data are from patients diagnosed during 10 years of mammographic screening at 18- to 24-month intervals. The Malmo data by age at enrollment (i.e., younger than 55 years of age and 55 years of age and older) were calculated from published tumor-stage-distribution relative-frequency data. The Edinburgh data are from patients diagnosed in the first 7 years. The data for the Health Insurance Plan of Greater New York study are from patients diagnosed in the first 6 years after enrollment and are given by age at enrollment (i.e., younger than 50 years of age and 50 years of age and older). Each of these trials, the percentage of Stage II or higher disease was smaller in the group screened with mammography than in the corresponding control group (Table 2), consistent with the expected earlier diagnosis of tumors by mammography. Furthermore, in keeping with the expected increased detection of early stage disease due to mammography, the percentage of advanced disease in the control group of the Stockholm study was reduced considerably by the mammographic screening in the sixth year. The lower percentage of advanced disease in the group screened using mammography was apparent in younger women in both trials for which age-stratified data were available, although the difference between screened and control groups was greater in older women in both trials.

Staging data have not been published for the NBSS study. For comparison purposes, Table 2 presents the proportion of NBSS patients with any lymph node involvement. All such patients would be classified as having Stage II disease or higher. Of the studies summarized in Table 2, only the NBSS study had a higher percentage of advanced disease in the mammography group than in the control group, a result that is due to
the 40–49-year age group. Removal of the 93 MP cases (38 in the 40–49-year age group and 55 in the 50–59-year age group) and the 95 control subjects (49 in the UC group and 46 in the PO group) with unknown lymph node status from the denominators of the NBSS proportions in Table 2 does not change the qualitative findings.

Table 3 presents the percentage of patients classified as having Stage III or IV disease in previous randomized trials. Published information on such patients is more limited than that for patients classified as having Stage II or higher disease, but the patients with Stage III and IV disease are of interest because they are likely more comparable with the NBSS patients with four or more positive lymph nodes. As was the case with Stage II and higher disease, the percentage of patients with Stage III and IV disease was lower in the group screened with mammography than in the corresponding control group in every study.

It has been suggested that the availability of screen-

ing mammograms in the follow-up evaluation of women in whom abnormalities were discovered by physical examination during a screening visit may have led to increased diagnosis of breast cancer in the NBSS MP groups, contributing to the excess of advanced disease in the 40–49-year MP group. Table 1 shows, however, that the numbers of diagnosed cancers which were detectable by physical examination at the initial screening visit were similar in the MP and control groups for the 40–49-year age group (i.e., 65 in MP and 60 in UC, \( P = 0.72 \)) and the 50–59-year age group (i.e., 71 in MP and 64 in PO, \( P = 0.61 \) ). Thus, any increase in the diagnosis of palpable tumors due to the availability of screening mammograms in the MP groups was small. Also, contrary to an underdiagnosis of patients with four or more positive lymph nodes in the UC group at the initial screening is the fact that only two such cancers were diagnosed in the 12 months after the first screening visit in the UC group compared with six in the MP group.

**Discussion**

Claims that the excess of patients with four or more positive lymph nodes in the MP group in young women in the NBSS is consistent with other randomized trials of breast cancer screening\(^{1,12,18}\) are not justified by published data. The Swedish trials referred to in these claims observed an initially increased incidence of advanced breast cancer in the groups screened with mammography; however, the control groups in these trials were not contacted initially or observed by personnel...
involved in the trials, and breast cancer among control subjects was diagnosed in the course of usual medical care and ascertained through cancer registries. Thus, it is expected that the screened groups in these trials will have an initially increased incidence of breast cancer of any stage or degree of severity compared with their corresponding control groups, and this expectation was met for patients with advanced breast cancer in the first years after enrollment.  

Despite the initial increased number of patients with advanced breast cancer in the mammography groups of the Swedish studies, the cumulative incidence of advanced disease in the mammography groups of these studies decreased below the cumulative incidence of advanced disease in their corresponding control groups by the fifth year after enrollment. In contrast, the excess of patients with advanced disease persisted in the 40-49-year MP group of the NBSS for at least 7 years (Tables 1 and 2). The statistically significant excess of patients in the 40-49-year MP group with four or more positive lymph nodes is contrary to the lower cumulative incidence of advanced disease in the mammography groups of previous trials. Thus, claims that the data on extent of disease at diagnosis from the 40-49-year age group of the NBSS are consistent with published data from previous mammography studies must be questioned.

The reason for the excess of advanced disease in the 40-49-year MP group is unknown. Although analyses presented above suggest that it is unlikely that such an imbalance is due to an unlucky randomization, this possibility cannot be excluded. Questions on the randomization process persist, however, partly because of the deficient randomization scheme used in the NBSS. Despite the fact that the initial physical breast examination preceded group assignment and symptomatic women were not to be excluded from the study, group assignments were made by local center coordinators using lists with preprinted identification numbers and group designations. Such nonblinded randomization could have allowed some women to be assigned preferentially to the mammography group based on adverse signs discovered during the physical examination. Evidence has been presented to refute claims that symptomatic women were assigned differentially to the mammography group; in particular, the MP and UC groups have similar distributions for a variety of demographic variables and for factors or self-reported symptoms that may have suggested an increased risk of breast cancer.

Axillary lymph node involvement can be detected by physical examination, however, and, thus, the significant excess of patients with cancer with extensive lymph node involvement in the 40-49-year MP group could, in itself, be evidence of nonrandom allocation. If there was such nonrandom allocation, it was not extensive, because only a small, nonsignificant excess of cancers detectable by physical examination was observed in the MP group. The nonrandom assignment of even a few advanced cases to the MP group could, however, have a marked effect on assessing the efficacy of mammography, particularly in the early years of follow-up.

The reason for the excess of advanced disease diagnosed in the MP group at Screening 1 in women 40-49 years of age likely will never be known. Mortality analyses that include these patients with a poor prognosis will provide a fair assessment of the efficacy of mammography, however, only if the initial excess of advanced disease was a result of the mammography. Thus, it is important to consider the possibility that mammography is directly responsible for the excess of advanced disease in the 40-49-year MP group. There is no evidence of a harmful effect of mammography from prior studies. The fact that the excess of patients with cancer with four or more positive lymph nodes in the NBSS is primarily due to patients diagnosed within 12 months of the initial screening visit challenges a causal role for NBSS mammography. Also contradicting a causal role for mammography in the NBSS is the absence of a similar excess of advanced disease in the 50-59-year MP group at the initial mammographic examination. If the NBSS mammography was inherently deleterious, it operated through a mechanism that increased the risk of advanced disease only among young women, with most of the damage manifest within 12 months of the initial exposure. Lacking such a plausible mechanism, there is no convincing evidence that NBSS mammography caused more advanced disease.

Because of the greater number of patients with a poor prognosis in the 40-49-year MP group, an excess of breast cancer deaths would be expected in the MP group even if there was no change in survival due to mammography. In fact, eight of the MP patients diagnosed at the initial screening visit as having four or more positive lymph nodes died in the first 7 years of NBSS follow-up compared with only one such patient in the UC group. A small reduction in mortality due to mammography would be obscured by the excess mortality assured by the initial unequal assignment of advanced disease to the MP group. Accordingly, it would seem that a mortality analysis omitting all patients with advanced disease detected by physical examination at the initial screening visit is warranted to assess the impact of the initial imbalance of advanced disease. Furthermore, such analyses would seem appropriate in future evaluations of the NBSS cohort after longer follow-up periods and in future metaanalyses that may include...
the NBSS in summary assessments of the effectiveness of mammography.

It has been argued that such a mortality analysis would be invalid, because the determination of the lymph node status of a patient with cancer may depend on the screening modality.\textsuperscript{12} A variable such as disease stage at diagnosis, which can be differentially affected by the screening modalities being compared, has been termed a pseudo-variable.\textsuperscript{29} It is clear that, in general, adjustment for pseudo-variables can lead to invalid inferences.\textsuperscript{29} An example would be a cancer for which only patients diagnosed as having Stage I disease receive any benefit from treatment. A screening modality for this cancer only could reduce mortality by detecting more cases at disease Stage I, and an effective modality detecting many more cases in Stage I may show no survival benefit if mortality analyses were performed after stratifying for stage at diagnosis.

The concept of disease severity as a pseudo-variable is more complex for the NBSS initial screening visit, however, and is particularly important because symptomatic women were not excluded from the study. The lymph node status of patients with breast cancer detected by physical examination at the initial screening would not seem to be dependent on the screening modality because the MP and the UC groups received the same thorough breast examination. It has been argued\textsuperscript{13,18} that the availability of screening mammograms may have led to increased diagnosis of cancer in women from the MP group in whom abnormalities had been detected by physical examination. As noted above, there is little support for this conjecture in the NBSS data. Moreover, it seems unlikely that such a mechanism would lead to the preferential diagnosis of advanced disease found at the initial screening in the NBSS. A few more patients with palpable tumors were diagnosed with cancer at the initial screening visit in the MP group in both age groups, and, thus, a slight diagnostic benefit may have resulted from having the screening mammogram at subsequent examinations in the MP group. Such a benefit cannot, however, explain the substantial disparity in disease severity between the MP and UC groups.

Because the patients with breast cancer diagnosed with four or more positive lymph nodes were unlikely to receive much benefit from their earlier detection due to screening, mortality analyses excluding such patients detected by physical examination at the initial screening visit deserve consideration. To the extent that a few such advanced cancers diagnosed in the MP groups might have been missed in the absence of a screening mammogram, such an analysis would be biased in favor of finding a beneficial effect due to mammography.

Previous analyses of the 40–49-year cohort of the NBSS have been biased against finding a benefit of mammography, however, because of the excess of patients with prevalent, poor prognosis cancers in the mammography group. Mortality analyses omitting advanced cases detected by physical examination at the initial visit should be considered in all future investigations involving the NBSS, to permit an impartial evaluation of mammography.

References

21. Tabar L, Faberberg G, Day NE, Holmberg L. What is the optimum interval between mammographic screening examina-


