Perspective

# The Canadian National Breast Screening Study: A Critical Review

Daniel B. Kopans[1] and Stephen A. Feig[2]

**Public health planners around the world had been awaiting the preliminary results of the randomized, controlled trial of breast cancer screening performed by the Canadian National Breast Screening Study (CNBSS) during the 1980s. It had been hoped that this large study would answer, with statistical validity, many of the questions concerning breast cancer screening that had not been answered satisfactorily by previous studies. Among the major unresolved issues is the desire to establish "absolute" proof that mammographic screening can benefit women who are between 40 and 49 years old.**

Previous studies, such as the Breast Cancer Detection Demonstration Project (BCDDP) conducted in the United States in the 1970s, have demonstrated that cancers can be detected with mammography at a smaller size and earlier stage than with clinical examination in these women. Comparisons between women in the BCDDP and those in the general population show a marked improvement in survival rates among women who participated in the screening project. These data imply that a mortality benefit should be expected [1]. The Kopparberg trial in Sweden further established the fact that properly performed mammography could lead to down-staging of cancers among women 40–49 years old, and down-staging has been shown to result in decreased mortality [2]. Nevertheless, most authorities require a randomized, controlled trial for "absolute" proof.

In reality, scientific proof is a statistical measure. It is generally accepted that a screening benefit exists if the difference in mortality between the screened group and the control group has a less than 5% probability of being due to chance alone. This in turn requires enough women in the trial to provide a sufficient number of cancers, and, since a reduction in breast cancer mortality is the desired end-point, a sufficient number of deaths from breast cancer among the control group to provide statistical significance. The smaller the anticipated benefit, the larger must be the number of women in the trial.

None of the randomized, controlled trials undertaken before the CNBSS were prospectively designed to evaluate the subgroup of women 40–49 years old. All analyses stratified by age have been done retrospectively, and since subgroup analysis was not part of the original study design for these trials, none of these randomized, controlled trials had the statistical power or were performed with the appropriate screening methodology to conclusively demonstrate a benefit. The CNBSS was expected to provide this information, but major mistakes made in the design and implementation of the trial cast serious doubts on the applicability of its results.

## Goals and Basic Design of the CNBSS

The CNBSS was undertaken to determine the efficacy of screening women 40–49 years old with mammography and physical examination, and to assess the incremental contribution of mammography over physical examination in the screening of asymptomatic women 50–59 years old. The trial began in 1980 and involved self-selected women (volunteers) between the ages of 40 and 59 who were recruited to 15 centers across Canada. Once a woman consented to participate, she was given a clinical breast examination by a trained nurse or physician and was then assigned to either the study group or the control group.

Received January 26, 1993; accepted after revision June 12, 1993.

[1]Department of Radiology, Massachusetts General Hospital and Harvard Medical School, 32 Fruit St., Boston, MA 02114. Address correspondence to D. B. Kopans.

[2]Department of Radiology, Thomas Jefferson University Hospital and Jefferson Medical College, 111 S. 11th St., Philadelphia, PA 19107.

*AJR* **1993;161:755–760** 0361–803X/93/1614–0755 © American Roentgen Ray Society

This perspective focuses on the group of women who were 40–49 at the time of accrual. While both groups were instructed in breast self-examination, women in the study group were offered annual mammography and physical examination whereas women in the control group received an initial physical examination and then were followed up by mail to ascertain their subsequent health status. The investigators apparently linked the CNBSS data base with national registries; however, it is suggested in their publication that the national linkage was maintained only through the end of 1988. Thus, although subsequent provincial linkages were provided, it is unclear how complete the documentation is as to the number of breast cancers and breast cancer deaths among the control group for the additional months of follow-up reported.

Epidemiologists have suggested that screening trials should exclude women with clinical symptoms of breast cancer, such as lumps [3], but these women were allowed to participate in the CNBSS. Inclusion of these women, who are unlikely to benefit from breast screening, increases the number of cancers in the trial (improving the statistical power) but dilutes any assessment of the effects of screening. Benefit may also be masked if women with late-stage cancer are disproportionately assigned to the study group.

### Lack of Statistical Power to Prove Benefit

Investigators have tried to use previous trials in the United States, Sweden, and the United Kingdom to retrospectively analyze the subgroup of women 40–49 years old [3]. Retrospective stratification is fraught with statistical weakness. Since the trials were not prospectively designed to analyze subgroups, they do not have the requisite statistical power (sufficient numbers of women and of cancer deaths among women in the control group) and cannot be used to prove or disprove efficacy. It is misleading and fallacious to analyze the results of these trials to seek proof of substantial benefit since none of the studies included sufficient numbers of women ages 40–49 to be able to provide proof of an expected benefit of 20–25%.

When a clinical trial is designed to evaluate the benefit of mammographic screening, the investigators must determine the minimum number of women who must participate in the trial so that the study will have the statistical power to demonstrate the anticipated benefit. The usual requirement is that the sample size be large enough to have an 80% chance of demonstrating the expected benefit with 95% confidence that the benefit is not due to chance. The power calculation thus provides the minimum number of women who must be enrolled to establish the expected benefit. A larger benefit may be demonstrated with the resulting sample size, but a smaller benefit cannot be demonstrated with statistical validity unless the sample size is commensurately increased. To prove screening efficacy, the key to the sample size is the number of women in the control group who die of breast cancer, since it is a reduction in this mortality among those screened that is a measure of the benefit of screening.

Many of the assumptions made by the CNBSS investigators in their calculation of the necessary sample size have proved to be incorrect. This deficiency is compounded by the fact that significant details of trial design were also overlooked, preventing the trial from answering the question it purportedly was designed to answer. As outlined in the preliminary publication [4] and in a previous summary of the study design [5], a population size was chosen on the basis of "feasibility" that if could be performed. It was expected that it would have an 80% power to show a 40% or greater benefit 5 years after the first screening at a level of .05. Even by the investigators' own calculations, and assuming all parameters were met, the trial was designed with an insufficient number of women to prove anything less than a 40% benefit. Although the investigators may have made reasonable assumptions based on the available data in the late 1970s, a 40% benefit is greater than most would expect even if there was 100% compliance among those offered screening (which there was not). Based on the data from the 1960s' Health Insurance Plan (HIP) project in New York, a 25% reduction is more realistic [6]. To demonstrate this level of benefit, the CNBSS would have probably needed far more women between the ages of 40 and 49 than the 50,000 (25,000 in each arm of the trial) that were included. It has been estimated that 250,000 women would be required in each arm of a trial to prove a 25% benefit at 5 years if the control group experiences an 80% 5-year survival. The CNBSS had one tenth the necessary number of women.

The investigators based their power calculation on the expectation that there would be 212 breast cancer deaths per 100,000 women at 5 years among the entire control group of women less than 50 years old [5]. In fact, the control group had only 111 breast cancer deaths per 100,000 women over the supposed 7 years of follow-up (28 deaths among 25,216 women). Not only was the death rate lower than the background mortality from breast cancer among Canadian women aged 40–49 but it did not even reach the level upon which their power calculation was based.

For this reason, the CNBSS did not have the statistical power to begin with to demonstrate a 40% benefit, let alone a more realistic 25% benefit. This inadequate power was further diluted by the fact that 26% of the "unscreened" women in the control group had mammography outside the trial. The investigators have argued that these were "diagnostic," not screening, mammograms. This argument might be supportable if either all these women had breast cancer or no clinically occult cancers were detected on the "diagnostic" mammograms. However, the suggestion by the investigators that these were inconsequential diagnostic studies merely displays a lack of understanding of the capabilities of mammography. All mammograms "screen" the breasts. Even among symptomatic women, clinically occult ("screen-detected") cancers are found by mammography when the palpable lump proves to be benign (most palpable lumps are benign). The investigators have suggested that they were aware that "contamination" would occur, implying that they planned for it; yet there is no evidence of this in their power calculation. Proper study design requires that this type of crossover or contamination be accounted for, and additional women must be added to the trial to make up for dilution of the control group. Using their own power calculation (which estimated an inadequate sample size from the start), the CNBSS investigators would have had to increase the total number of women in the

trial who were less than 50 years old by almost 40,000 to compensate for this amount of contamination: (estimated sample × (1/[1–the rate of contamination]$^2$) [7]. This would necessitate increasing the total number of women in the age 40–49 part of the trial from 50,000 to almost 90,000.

The preliminary results are further compromised by too short a follow-up period. The HIP trial, as well as the Swedish trials, suggests that mortality reduction for these women begins to appear approximately 8–10 years after screening begins [6, 8]. From the statistical assumptions of the investigators, as well as from the actual experience in the trial, it is unlikely that this trial could prove a benefit at all, and certainly not so soon.

## The CNBSS Results Cannot Be Generalized

Even if the CNBSS contained enough women ages 40–49 in each arm of the study (which it did not), the fact that the control group has enjoyed an unprecedented survival of greater than 90% at 5 years would mean there would have to have been virtually no breast cancer deaths among the screened women to produce a statistically significant benefit. This would have been impossible, since the randomization process allocated significantly more women with advanced cancers (four or more positive axillary nodes) to the screened group in the prevalence year (the first year of screening). This greater number of prognostically poor tumors guaranteed more deaths from breast cancer among the screened women and ensured that a benefit could not be shown for screening with only 7 years of follow-up. Regardless, since contemporary women ages 40–49 in Canada and the United States had a less than 80% 5-year survival rate [9], the results of the CNBSS are certainly not applicable to the general population.

The unexpectedly high survival rate among the women in the control group may in part be a result of the randomization process, which had the effect of shifting the allocation of some late-stage cancers from the control group to the screened population. This is reflected in the fact that there were fewer than expected cancer deaths among the control women and a greater number of cancer deaths than expected among the screened women ages 40–49.

It is likely that the overall high survival rates in the CNBSS are due to selection bias. It is a common phenomenon that persons who volunteer for screening programs enjoy better than average health. This is why the CNBSS investigators estimated that they would have fewer deaths among the women in the trial than among the background Canadian population. The actual death rate among the control population was even lower than they had anticipated. This result was likely due in part to the effect of selection bias. The women who participated in the CNBSS were a selected population defined by the fact that they were self-referred volunteers. According to the CNBSS data, "Only 6.6% of the participants in the CNBSS heard about the study from their physicians" [4]. In the United States, however, it has been repeatedly demonstrated that women do not "self-refer," but are screened because their physicians refer them [10, 11]. Thus, American women likely represent a different population from the CNBSS subjects, arguing strongly against applying the CNBSS results to the general population.

## A Test of Poor-Quality Mammography

Since some of the basic trial design assumptions of the CNBSS were incorrect, and since the sample size was too small to have statistical validity, additional analysis should be superfluous. Unfortunately, many of those reviewing the trial have looked only at the results and not at the flaws in the underlying design and statistical assumptions [12]. They have also ignored flaws in the execution of the trial, particularly deficiencies in mammographic technique and interpretation.

Despite the fact that the most significant goal of the CNBSS was to test the benefit of mammographic screening in younger women, a decision apparently was made at the outset that the quality of the mammography was not to be of major importance.

The screening centers were permitted to use whatever mammographic equipment they had available, since funding was not provided to purchase modern equipment. There was no special training for the technologists performing the mammograms, or for the radiologists interpreting the studies. Thus, instead of evaluating the efficacy of high-quality mammography for screening, the designers of the trial decided to test the validity of mammographic screening as they assumed it was then being generally practiced at the time the study was forming (Miller AB, personal communication). In addition, despite concerns voiced by early advisers to the CNBSS, little effort was made to improve the quality of the mammography; indeed, two advisers (W. Wende Logan and Stephen A. Feig) resigned in protest over this issue.

In contrast, the quality of the physical breast examinations in the CNBSS was given a high priority. The breast examination involved a series of maneuvers that were much more extensive and took longer than those used in routine clinical practice. The physical examinations were primarily performed by nurses who received special training in the performance of the protocol [13] and who were closely monitored throughout the study. As the principal investigator in the CNBSS has acknowledged, the intensive type of clinical breast examination performed in the CNBSS is unlikely to be practical for general screening (Miller AB, presented at the National Cancer Institute Workshop on Breast Cancer Screening, February 1993).

As a result of concerns raised by consultants to the program and others familiar with the CNBSS, the CNBSS investigators organized an objective outside review of the quality of the mammograms, the results of which were published in 1990 [14, 15]. Although the quality of the mammograms gradually improved over time, the external review clearly showed that, for the majority of the trial (years 1–4), more than 50% of the mammograms were judged as poor or completely unacceptable. Two unpublished reviews conducted at the end of the fifth year and beginning of the sixth year of the trial by Stephen Feig, then adviser to the CNBSS, found that poor-quality mammograms were common in more than 75% of the centers. Only 33% of the mammograms reviewed from 1985 were judged as technically acceptable. This figure actually dropped to only 15% in a similar review in 1986. The principal investigators have argued that the worst mammography occurred in the first 2 years of the study, when only a small number of women were screened (some centers had

not yet begun screening), but their own figures demonstrate that poor image quality dominated through the fourth year. More than 50% of women had been recruited by this time (the mid-point of recruitment was 1983) [16]. This was not merely due to the failure to use the mediolateral oblique projection (thus effectively excluding from evaluation the tail of the breast, where some cancers develop), as the investigators have argued, but also to the overall poor quality of the CNBSS images, ranging from poor positioning to poor contrast and image sharpness. To the credit of the program managers, and as a result of the repeated concerns raised by advisers to the program, as well as the lower sensitivity of cancer detection found on internal review by the program's own reference radiologist [17], the quality of the mammography was improved over time. However, for a study with this importance, quality should have been high at the outset.

Although the CNBSS investigators have defended their quality control program as an indicator of good mammographic practice, they have acknowledged that low radiation dose was the primary requisite in the program. Image quality was not considered as important, so to maintain low dose, scatter-reduction grids were not used until late in the trial. A recent statement [18] by the physicist responsible for the quality control of the mammography in the CNBSS confirms the significant deficiencies in the mammography performed in the trial:

> In my work as reference physicist to the CNBSS, [I] identified many concerns regarding the quality of mammography carried out in some of the CNBSS screening centers. That quality [in the CNBSS] was far below state of the art, even for that time (early 1980's). Problems in the quality of mammography resulted not only from inadequate equipment in some cases, but also from inappropriate imaging technique and lack of availability of specialized training for technologists and radiologists.

### Suboptimal Mammographic Interpretation

An internal review conducted by the CNBSS's own reference radiologist revealed that 42% of the cancers that were missed at screening and became palpable in the interval between screenings were visible on a previous mammogram and had been missed by the interpreting radiologist [17]. A total of 25% (143/575) of the cancers diagnosed in the screened population were identified by the reference radiologist on mammograms taken at least 1 year before their ultimate detection, having been missed by the interpreting radiologist. The number of cancers not even visible because of poor-quality mammography was not reported in this review, but the investigators themselves stated that "a frequent problem was inadequate positioning which caused areas of breast tissue not to be visualized."

### Nonblinded Randomization—More Advanced Cancers Allocated to the Study Group

An additional problem with the CNBSS lies in the fact that, in the first round of screening, significantly more women with advanced, lymph node–positive breast cancer were randomized to the screening arm than to the control group. As stated earlier, a test of the efficacy of screening will be diluted by including women with palpable masses, who will not benefit from screening when these masses represent cancer. Nevertheless, the CNBSS permitted women with clinical signs of breast cancer to participate. Among the women aged 40–49, in the first round of screening, 33 with cancer in the axillary nodes were assigned to the screening group whereas only 21 such women were placed in the control group. Among these 33 women, 30 had palpable cancers. For women with particularly poor prognosis (four or more cancerous nodes) the ratio between those assigned to the screening group and those to the control group was 19:5, and 17 of the 19 had palpable breast cancer. This latter fact is important since the randomization was not blinded, contrary to the practice in most randomized, controlled trials. Women who volunteered were first given a clinical breast examination and then were assigned to the mammography or the control group. Logistical efficiency has been cited as prompting the decision to perform the clinical breast examination before randomization. Although there is no proof that this occurred, nonblinded randomization can compromise the randomization process and with it the statistical objectivity of the trial. The investigators argue that there is no evidence of a systematic flaw in the randomization, since the demographic features of the women in the screened and control groups are symmetrical. What is overlooked is the fact that as many as 100 or more women could be shifted between the two groups without producing a demographic imbalance, and that shifting of a much smaller number of women with advanced cancer would substantially affect the relative number of deaths from breast cancer. Whether by chance or through a flaw in the randomization process, more women who already had prognostically poor tumors were, indeed, placed in the screening group. This ensures a poorer survival and higher mortality among the screened women in the first years after screening, when the prevalent cancers predominate. If a substantial number of advanced cancers are present at the outset and are disproportionately assigned to the screening arm, then more deaths from cancer would be expected among these women a priori, and screening may never show an influence on mortality.

Mammographic screening appears to be effective because it detects cancers at a smaller size and earlier stage than does usual care. Statistically, and as demonstrated in the CNBSS, it is not cancers found by mammography that kill. The survival rates for cancers detected by mammography alone in the CNBSS was 95% at 7 years whereas it was only 89.3% for those detected by physical examination alone. If there is an excess of deaths from cancer among screened women it must, in part, come from cancers missed by mammography, again implicating the poor quality of the mammograms in this program.

It has been argued that prevalence screening (the first year of screening) reveals more cancers of all stages, and that more advanced cancers in the prevalence year would be expected among screened women. However, the fact that 91% of the cancers that had spread to the lymph nodes

among women assigned to the screening group were palpable, and that all the women (in both groups) had a clinical breast examination before randomization, raises uncertainty about the assignation process. If randomization had been successful, there should have been the same number of advanced cancers in the control group that perhaps remained initially undetected in the absence of screening. If this were the case, then one would expect that these cancers would surface in later years and that the relative excess of advanced cancers among the screened women would disappear. In fact, the opposite occurred. Not only was there no diminution in the excess of advanced cancers, but the excess increased among the screened group over the period of the trial. This suggests that there was a probable failure in the randomization process. The number of cancers that had spread to the lymph nodes among the screened women increased from 33 in the first year to a total of 102 at the end of the screening period (an increase of 69 cancers). Among the control group, this number went from 21 to a total of 68 (an increase of only 47 cancers). The fact that there was not only a lack of equilibration in the number of advanced cancers between the screened women and the control group over the course of the trial, but indeed a further separation, suggests that delays in diagnosis occurred among the screened women. This probably resulted from the false reassurance that some women derived from a normal finding at the screening study, and false reassurance was compounded by poor-quality mammography. Other hypotheses are, at best, on the scientific fringe.

## False Reassurance

Delays in diagnosis can occur when women are falsely reassured from a normal finding on a screening study. Unfortunately, in convincing women to be screened, it is often incorrectly implied that mammography can exclude breast cancer and that normal results on screening evaluation are somehow protective. In the CNBSS it is likely that some women who underwent screening and had normal results developed a false sense of security. Believing that normal findings conferred a degree of protection, a woman who incurred a problem in her breast during the interval between screenings, assuming that the change could not be significant, would be more likely to defer immediate evaluation, or she might defer evaluation knowing that she would be screened again in the future. In contrast, her counterpart in the control group would have no similar false reassurance and would be more likely to seek immediate evaluation when she noted a breast abnormality. Thus, the screened woman would delay the diagnosis of her potential breast cancer, which would be more likely to be diagnosed at a later stage. This problem would be amplified by poor-quality screening that converted potentially early lesions at detection into later stage and fatal cancers as a result of delayed diagnosis. This delay from false reassurance among screened women has been documented in a Finnish study [19]. Delay in diagnosis from false reassurance is a likely reason for the increase in the excess of advanced cancers found among the screened group over the course of the CNBSS.

## Inconsistent Thresholds for Intervention

There were no agreed-upon parameters for intervention in the CNBSS, and no requirements for biopsy when concern was raised by the screening. If a mammogram was interpreted as abnormal, the case was referred to a center surgeon who would determine whether a biopsy was indicated. This recommendation was then conveyed to the patient's nonprogram physician, who made the final decision whether or not to intervene. The surgeons had no agreed-upon guidelines, and the investigators have acknowledged that there were some mammographically detected cancers that were not cared for promptly. It was possible for center radiologists to raise concern, but the surgeon might disagree. According to the director of the CNBSS, at least one physician refused to do a biopsy of a lesion unless it was palpable. In fact, almost 25% of the needle localizations and biopsies that were recommended were not performed [4].

## Conclusion

It is admittedly very difficult to acknowledge that a study that required an enormous amount of work and resources cannot answer the important questions it set out to answer because of faulty design and execution. It would be more unfortunate, however, to compound these mistakes by basing any significant medical policy decisions on this study.

Important lessons can be learned from the CNBSS and the other screening programs. Perhaps the most significant is the demonstration that poor-quality mammography, and the poor integration of mammographic screening with diagnostic intervention, reduces the effectiveness of screening. Rather than withdrawing this potentially lifesaving approach, we must strive to provide high-quality screening and, through physician education, to better integrate detection and diagnosis.

A great deal of confusion has resulted from the, at best, premature analysis of the CNBSS and from the retrospective analyses of the other screening trials with regard to benefit for women ages 40–49. Unfortunately, reviewers have merely looked at the results of the studies, without critically reviewing the design and performance of the trials that generated the results. In fact, none of the trials, including the CNBSS, has sufficient statistical power to be able to absolutely prove a benefit among the subgroup of women ages 40–49. Subgroup analysis of data from trials that were not designed or performed to legitimately analyze, with statistical validity, the subgroup can be greatly misleading. In fact, virtually all the trials have shown a benefit in screening the entire population of women for which they were designed. Therefore, they have all shown a benefit for screening women beginning in their forties. Subgroup analysis with insufficient numbers of women in the subgroup further dilutes any statistical power. As an example of the weakness of subgroup analysis, one could argue that there are no data that prove an absolute benefit for screening women age 57. There are simply not enough women of this age in the trials to provide statistical significance; yet most agree that there is definite benefit for screening women in their fifties. It is a specious argument that there

is no "scientific proof of benefit" among a subgroup when no studies have been done that could provide the proof.

In the absence of an appropriate study, one must rely on inferential data to guide clinical recommendations. The HIP trial showed a 25% mortality reduction for women 40–49, although this is not statistically significant since the trial did not include a sufficient number of women to be able to provide absolute proof. The overview of the Swedish trials shows an overall benefit of 13% at 10 years that appears to be increasing with time [14], and the Edinburgh trial shows a 14% benefit for women 40–49 [20], but once again these trials were not designed to have sufficient numbers of women to be able to prove benefit. Although the BCDDP cannot provide direct mortality data because it was not a randomized trial, it clearly showed increased survival for women less than 50 years old. Large-scale demonstration projects of mammographic screening performed in the current era, after completion of screening in the randomized, controlled trials, show that screen-detected cancers in women 40–49 years old have the same if not more favorable prognoses as cancers detected in older women (Sickles EA, presented to the President's Panel on Breast Cancer, March 1993).

Women 40–49 years old should not be misled by incomplete reports stating that there is no scientific proof of benefit. They should know that no study has been performed that could provide scientific proof. Access to screening should not be withdrawn simply because investigators have not performed the appropriate scientific studies or because the studies that have been performed have not been performed properly. If "absolute" proof is needed, then a randomized, controlled trial is required that would involve at least 500,000 women. Since it is likely to be difficult or impossible to have sufficient numbers of women willing to be randomized to the unscreened control group, a trial in which the randomization is by interval between screens (6 months, 12 months, 24 months) has been suggested. This would require a massive number of women, but may be feasible in light of the fact that several million women aged 40–49 are probably already being screened.

It is important for physicians and women alike to understand that no "science" that is available at the present time can guide our recommendations with certainty. There is good reason to believe that screening should be useful in women under 50, since it is clearly beneficial in women over 50. Since breast cancer is more common among women over 50, and more older women die of breast cancer, a trial would require more younger women to demonstrate the same benefit as among older women. The distribution of women in the trials has been just the opposite. Only one third of the women in the trials have been less than 50 years old. The much larger number of older women accounts for the statistically significant benefit seen for women age 50 and over. There is no fundamental reason to believe that properly conducted screening of women aged 40–49 would not provide the same benefit it already does in older women. Tabar [21] has recently analyzed the survival rates for women at different ages based on the size, nodal status, and histologic grade of disease. His results clearly show that women with tumors of equal size, nodal status, or grade, have very similar outcomes, irrespective of age.

Decisions about screening for breast cancer must be based on a thorough understanding of all the available information so as to provide women with the best medical advice. We should not lose sight of the fact that 40% of the years of life lost to deaths from breast cancer are from women in whom the disease was diagnosed when they were less than 50 years old [6]. We believe that the available data support the use of annual mammographic screenings for all women beginning at age 40.

## REFERENCES


1. Feig SA, Hendrick RE. Risk, benefit, and controversies in mammographic screening. In: Haus AG, Yaffe MJ, eds. *A categorical course in physics: technical aspects of breast imaging.* Chicago: Radiological Society of North America, **1992:**103–118
2. Tabar L, Gad A, Holmberg L, Ljungquist U. Significant reduction in advanced breast cancer, results of the first seven years of mammography screening in Kopparberg, Sweden. *Diagn Imag Clin Med* **1985**;54:158–164
3. Elwood JM, Cox B, Richardson AK. The effectiveness of breast cancer screening by mammography in younger women. *Online J Curr Clin Trials* **1993**;32
4. Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study. 1. Breast cancer detection and death rates among women aged 40–49. *Can Med Assoc J* **1992**;147:1459–1476
5. Miller AB, Howe GR, Wall C. The National Study of Breast Cancer Screening. *Clin Invest Med* **1981**;4:227–258
6. Shapiro S, Venet W, Strax P, Venet L. *Periodic screening for breast cancer: the Health Insurance Plan project and its sequelae. 1963–1986.* Baltimore: Johns Hopkins University Press, **1988**
7. Lachin JM. Introduction to sample size determination and power analysis for clinical trials. *Controlled Clin Trials* **1981**;2:93–113
8. Nystrom L, Rutqvist LE, Wall S, et al. Breast cancer screening with mammography: overview of Swedish randomised trials. *Lancet* **1993**;341:973–978
9. Ries LAG, Miller BA, Hartman AM, Edwards BK. *Cancer Statistics review 1973–1988.* National Cancer Institute Publication 91-2789, **1991**: Table IV-12
10. The NCI Breast Cancer Screening Consortium. Screening mammography: a missed clincal opportunity? *JAMA* **1990**;264:54–58
11. Fox S, Klos DS, Tsou CV, Baum JK. Breast cancer screening recommendations: current status of women's knowledge. *Fam Commun Health* **1987**;10:39–50
12. Fletcher SW, Fletcher RH. The breast is close to the heart. *Ann Intern Med* **1992**;117:969–971
13. Miller AB, Baines CJ, Turnbull C. The role of the nurse-examiner in the National Breast Screening Study. *Can J Public Health* **1991**;82:162–167
14. Baines CJ, Miller AB, Kopans DB, et al. Canadian National Breast Screening Study. Time-related changes in mammographic technical quality: an external review. *AJR* **1990;**155:743–747
15. Kopans DB. The Canadian Screening Program: a different perspective. *AJR* **1990;**155:748–749
16. Miller AB, Baines CJ, Wall C. The Canadian National Breast Screening Study. In: *Cancer screening: international union against cancer.* Cambridge: Cambridge University Press, **1991:**45–55
17. Baines CJ, McFarlane DV, Miller AB. The role of the reference radiologist: estimate of inter-observer agreement and potential delay in cancer detection in the National Breast Screening Study. *Invest Radiol* **1990**;25:971–976
18. Yaffe MJ. Correction: Canada study (letter). *J Natl Cancer Inst* **1993**;85:435
19. Joensu H, Klemi PJ, Tuominen J, Rasanen O, Parvinen I. Breast cancer found at screening and previous detection by women themselves (letter). *Lancet* **1992**;339:315
20. Fletcher SW, Black W, Harris R, Rimer BK, Shapiro S. *Report of the International Workshop on Screening for Breast Cancer, February 24–25.* Bethesda, MD: National Cancer Insititute, **1993**
21. Tabar L. New Swedish breast cancer detection results: women aged 40–49. *Cancer* (in press)


The reader's attention is directed to the companion to this article, which appears on pages 761-763