Breast Cancer Screening: All’s Well That Ends Well, or Much Ado About Nothing?

Myron Moskowitz

The past 25 years of intensive epidemiologic and clinical studies have produced a breakthrough in the management of breast cancer. These studies have shown that approximately 3.5 years can be added to the life of the average woman with breast cancer [1, 2]. The data also suggest that a number of women will be essentially cured of their disease [1, 3–6]. The controlled studies have shown that breast cancer mortality rates can be reduced by 25–30% by appropriate screening with mammography and physical examination [1, 5, 7–14].

Evidence for Mortality Reduction by Periodic Screening

The single most important clinical study in this area to date was completed in New York in the middle 1960s. In this period, Shapiro, W. Venet, Strax, and L. Venet, in cooperation with the Health Insurance Plan of Greater New York (HIP), performed the landmark controlled trial to evaluate screening for breast cancer in 40- to 69-year-old women. In this initial trial, 60,000 women were randomly allocated to study and control groups of about 30,000 each. The women in the study group were offered annual screening with mammography and clinical examination, whereas the women in the control group were allowed to go about their usual health-care routines. Of the 30,000 women offered the screenings (four screenings each [1, 8–10]), only 20,000 accepted at least one; 10,000 of these women did not want any screenings.

The power of this study derives from the measurement of deaths from breast cancer in the entire group offered screening (including those who accepted any screening as well as those who refused all screenings), and from the comparison of this number with the total number of breast cancer deaths among the control group. This approach effectively eliminates the problems of length bias sampling, lead-time bias, self-selection bias, and bias of secular trends. Because no new treatments were available for use in either the study group or the control group, variations in therapy did not occur that might have complicated the analyses.

This study showed conclusively that within 3.5–4 years, mortality was reduced in the study group compared with the control group (Fig. 1A) [15]. By 5 years, the mortality reduction represented about a 30% differential in cumulative mortality. However, the data at that time indicated that the mortality reduction appeared to be limited to women who were 50 years old or older at the time of entry (Figs. 1B and 1C). Later results, however, show parallel trend-lines of mortality reduction for older and younger women (Figs. 1B and 1C). In a recent analysis of the HIP data, Habbema et al. [16] indicated that the value of screening is not affected by age. Breast cancer screening of populations of women age 40 and older, as it was performed in the HIP trial, seems to have the capability of reducing mortality approximately 25–30% across the board. Indeed, 18 years after the start of the study, the percent mortality reduction was greatest for women who were 40–44 years old at time of entry into the program [16].
Researchers (most recently Bailar [17]) have argued that no scientific evidence exists to support the value of screening for this younger group. If we examine the data, we find that the HIP study was set up to reject the null hypothesis that "screening a population of women 40–69 with mammography and clinical exam annually is ineffective in reducing mortality from breast cancer." Because there was no plan to isolate the age-specific effect of screening, the study was designed to enroll a sufficiently large number of women older than 40 so that the resulting statistics would be meaningful. If study of the age-specific effect had been the goal, more women between the ages of 40 and 49 would have been enrolled. As we have seen, the null hypothesis was resoundingly rejected. Not only was a mortality reduction shown for screened women who were older than 40 years, but no scientific evidence was found that screening is ineffective in younger women.

I believe that Dr. Bailar's opinion that screening is ineffective in the younger age group is not founded on fact. The onus is now on those who hold this belief to prove, by scientific prospective studies performed with current medical techniques and at yearly intervals, that screening is ineffective for 40- to 49-year-old women.

The findings of the HIP trial have been verified by a controlled trial in Sweden [12]. Case-control studies in Holland [7, 13, 14, 18] and Italy show that women who undergo screening have a distinctly lower probability of dying of breast cancer than do unscreened women. In the Nijmegen study [13], among the women younger than 50 at entry, the relative risk of dying of breast cancer in screened women vs unscreened women was 0.5; for older women, it was 0.65. In the Swedish and Dutch trials [12, 13, 14, 18], screening was done with single-view mammography and without clinical examination; screening intervals were 2 or 3 years, depending on age at entry.

Although breast screening can reduce mortality in a given population, we must not expect that all cancers so discovered are necessarily cured. Clearly, many patients with invasive, stage 1 carcinomas (and perhaps some with stage 2 carcinomas) have their lives extended by early detection, but a number of these women may ultimately develop metastases. Nevertheless, these screening studies indicate that women...
who undergo screening mammography live longer than those who don’t. Because of breast cancer screening, life has been extended and its quality has been improved for many patients.

Lead Time and Intervals of Screening

The Swedish study [11, 12] in which the screening interval was 2 years, shows that for women who were 40–49 years old at entry, 38% of the number of cancers expected (on the basis of the control group) were seen in the first 11 months after screening. In the next 12–23 months after screening, the number of cancers rose to 68% of the expected rate. In my 1986 study [6], the first 11 months showed a marked drop in the cancer rate for this age group, but by the second year after screening, the rate had returned to normal. These data indicate that, for 40- to 49-year-old women, the lead time gained by screening is 2 years or less. Again, our observations as well as those of others [6, 11, 19] show that the mean detection lead time gained by screening women older than 49 is about 4 years.

Many noninvasive ductal cancers and cancers of 5 mm or less can be found by screening. Detection of these cases, as shown by the controlled Swedish trial, reduces the cumulative number of advanced, stage 2 or higher cancers that will occur in the population [5, 11, 12]. In my study [6], during the screening period, the rate of stage 2 cancers for younger women was 0.4/1000 person-years. In the post-screening period (i.e., after the 2½-year lead time had been consumed), the rate of stage 2 and higher disease was 1.3/1000 person-years. Because this series lacked a control group, I could not conclude with certainty that the total number of stage 2 cancers was reduced by screening. In younger women, the mean detection lead time is relatively short owing to the number of rapidly growing cancers relative to the smaller total annual incidence. Therefore, when equilibrium is reestablished, these rapidly growing cancers predominate in the short term. However, in this series, the mean age of patients at cancer detection before the screening period was 46, and the mean age of patients at cancer detection after screening had been instituted was 47. Therefore, we can presume (1) annual cancer rates for the two groups were the same; (2) if no screening had been done, the total number of 

incident cancers in this group would still have been the same as we actually observed during the screen and post-screen periods (the difference is the stage at which they were detected); (3) if these women had not been screened, 52 advanced cancers probably would have occurred in this period, but only 36 occurred (a 31% reduction). Without a control group, this third conclusion can be taken only as a strong possibility based on observed clinical data. The Swedish data referred to earlier [5] have established that this is no longer a clinical hypothesis but a proved fact. Clearly, screening has resulted in an alteration in the natural history of the disease in younger women.

If the lead times gained by screening are ignored in planning strategies, the potential negative effects are enormous. If the mean detection lead time gained by screening is exceeded, the screening detects only the slower growing tumors. Depending on (1) the lead time gained, (2) the duration of the interval between screenings, and (3) the ratio of slow-growing tumors to those that are growing at an intermediate rate, this strategic error could nullify the impact of screening, certainly in the short term, and possibly in the long term.

In the younger women attending the Cincinnati Breast Cancer Detection Demonstration Project (BCDDP), no deaths have occurred among the patients who have developed breast cancer during the first 11 months after their last screenings. Nor have any advanced cancers occurred in this group. If patients with lobular carcinoma in situ are excluded (a conservative approach) and patients-with-disease rather than breasts-with-disease are counted, the death rate among all women with cancer found either during or within 11 months of the last complete negative screening was 0.65/1000 patient-months.* Among the cases detected 12–59 months after the last complete negative screening, the death rate was 6.2/1000 patient-months. The average age of the patients whose cancers were found during screening was 46 years. The average age of the women whose cancers were found in the post-screening period was 47 years.

For women older than 50 at entry, the death rate among women whose cancers were detected at or within 11 months of a negative screening was 2.6/1000 patient-months. For those cases occurring 12 months or more after the last complete screening, the death rate was 4.7/1000 patient-months.

The comparative rates for both younger and older women are remarkable for the fact that the time of observation for the post-screen cases is very short compared with the time of observation for the screened cases. The rate of death among the post-screen cases will continue to increase considerably over time. When the observation periods of the two groups are equal, the disparity in death rates most likely will be even more dramatic. I believe this will occur because the documented advantages of finding and treating minimal breast cancers should increase about 10–15 years after detection and treatment.

Perhaps the death rates among these screened cases are heavily weighted by length biased sampling and lead time bias. Eliminating the cases of lobular carcinoma in situ has reduced this possible effect. On the other hand, surely not even the most die-hard critics of screening would contend that only 10% of the cases detected in screening were real (or lethal) cancers. The benefit of screening younger women may not be as high as 10:1, but it is certainly not zero either.

If the better survival is due only to length-biased sampling, where are the rapidly growing cancers that would be expected in the first 11 months after screening? They were not seen in the younger or older women in my study nor in the Swedish study [5].

These observations suggest that screening does not detect only the slower growing tumors. Rather, aggressive annual screening and early detection has a great impact on the most active tumors and can reduce even short-term mortality. When the screening is not done, these lesions reach clinical

---

* Patient-months are based on the sum of the total number of months of observation of each woman from time of entry into the appropriate category to the time of last follow-up contact (screen visit, mail, or phone interview).
detection 12–23 months after screening and may go on unchecked to cause rapid death.

Therefore, if we are to maximize the impact of screening on younger women, the best results will be achieved by aggressive, annual screening and by using both two-view mammography and clinical examination.

For older women, screening probably does not need to be so aggressive; for large populations, screening with mammography alone every 3 years has proved to be effective. However, annual or biennial screening with modern mammography and clinical examination may be even more effective.

Certainly, the only reason not to screen more frequently in this older group is the cost.

**Who Should Be Screened?**

If our goal is to reduce the mortality from breast cancer in the population as a whole, we will not be able to do so by limiting our efforts only to those women who have the known or accepted high-risk features. No more than 25% of women who develop breast cancer have these identifiable characteristics [20].

Wolfe’s mammographic risk patterns deserve comment here. For practical purposes, we can presume that patients with certain patterns are at higher risk than those who have other patterns [21–25]. The patterns seem to operate independently of other known risk factors. Also, the magnitude of these risks seems to be much smaller than originally believed [21–23].

In a recent review, Saftlas and Szklo [22] have pointed out that in 11 prospective studies, the relative risk for P2 + Dy pattern, compared with that of P1, ranged from infinity to 21.6. P2 + Dy patterns are assumed to represent those cases with prominent duct pattern occupying 25% or more of the volume of the breast, plus dense breasts containing multiple lobulated, often poorly margined densities. Because these data represent two outcomes of multiple, disparate studies, interobserver consistency of definition cannot be assumed. If one eliminates the study yielding an infinite result and the study showing a relative risk of 21.6, the average relative risk is 3.26. If we eliminate the statistics from Wolfe’s articles, the average relative risk is 2.4.

Failure to demonstrate a large risk may stem from inability to properly classify mammograms as Wolfe does. However, in three series in which interobserver agreement with Wolfe was measured and agreement was based on the combination of high and low risk patterns (N1 + P1 and P2 + Dy), agreement ranged from 0.82 to 1.00 (average, 0.89).

Among 20 case-control or cross-sectional studies in which risk ratios of P2 + Dy were reported, the average relative risk was 1.49 [22]. In this group, when Wolfe’s 1987 series was included, the relative risk was 3.3 [22].

Some risk seems to be associated with certain mammographic patterns. The magnitude of that risk is in line with other known risk factors, except family history. It certainly is unclear to me whether the minimal risk associated with the mammographic patterns reflects abnormal physiology or merely masks current pathology associated with dense breasts.

In my opinion, the most artfully crafted of the trials to evaluate parenchymal patterns was a case-control study conducted by Whitehead et al. [21]. Training, testing, and reinforcement measures were done to ensure inter- and intraobserver consistency. The numbers were large enough to ensure statistical validity of the study, and the analyses were complete. They developed a comprehensive model to estimate the combined risk of parenchymal patterns with all other known risk factors.

The odds ratios for the parenchymal patterns, as reported by Whitehead et al. using P1 as the reference, are 0.5:1:1.8:1.6 for N1:P1:P2:Dy. If N1 is used as the reference baseline, the odds are 1.2:0:3.5:3.1. These data follow closely those recently reported by Gravelle et al. [26]. However, when these data are combined with all other known risk factors and applied to 2003 cases and controls, the resulting histograms for both groups show extensive overlap, and the two distributions cannot be separated satisfactorily on the basis of risk factors [21].

The patient’s age and family history are the only two known risk factors that may help select screening strategies. In women with a family history of breast cancer, the annual risk of developing breast cancer when they are 30 years old is about the same as in other women at age 40 [27]. For example, the usual annual incidence of breast cancer in the 30–34 age range is about 0.23/1000; in the 40–44 age range, it is about 1.03/1000. For 40-year-old women, the annual incidence is about 0.53/1000, whereas for 30-year-old women, it is about 0.08/1000. Therefore, at age 30, women with a family history of breast cancer are seven times more likely to get cancer than are other women their age. Between the ages of 30 and 34, these women are 4.5 times as likely as other women to get breast cancer. In the 10-year interval between ages 30 and 39, the risk for women with a positive family history to get breast cancer can be projected to be about the same as other women in the 40–49 age range (i.e., 1.30% [2]). As the two populations age, a relative excess risk persists, but the magnitude of the excess is significantly reduced.

I think that (1) in the case of a woman whose mother or sister has had breast cancer, annual screening with mammography and clinical examination should begin at age 30; (2) all other women should start annual screening at age 40; and (3) for women who are 50 or older, annual screening can be done as recommended by the American College of Radiology, the American Cancer Society, and the National Cancer Institutes. However, the data suggest that these women could be screened every 2 to 3 years with the same result as shown by HIP.

**What Mammographic Projections Are Needed?**

The mediolateral oblique view was designed by Lundgren and coworkers [28–30] to maximize the amount of breast tissue included in one-view film mammography. He hoped that this would be the only projection needed for screening. However, in my experience [31], only 80% of the cancers detected by screening mammography would have been detected if only the single mediolateral oblique xerogram had...
been used. More recently, Bassett et al. [32] have shown that, although the single-view examination correctly detected 67% of cancers, additional views were recommended in 32% of the normal cases, as well as in 23% of the cancer cases; 7% of the normal cases were recommended for biopsy. When two views were initially obtained, additional views were requested only in 4%, 80% were correctly identified on the initial pass, and only 5% of healthy subjects required additional views.

There seems to be no overriding need to limit mammographic studies to one view. A second view adds only $2.00 to the cost of the examination. In my experience, with highly qualified technologists, a good two-view examination can be done in 6 min, so throughput is not a significant problem. In fact this rate exceeds the heat load capacity of the X-ray tubes, so over a period of time filming at this rate would be counterproductive.

Thus, the use of single mediolateral oblique mammography does not seem to be indicated because it would lead to an excessive number of call-back examinations of healthy patients, which would increase both costs and anxiety.

For those radiologists using film-screen mammography, the mediolateral oblique view (Lundgren projection) is extremely important in order to maximize the amount of breast tissue that is included on the film. A mediolateral, or lateromedial, view with film-screen mammography includes an insufficient amount of glandular tissue. In my opinion, the mediolateral view should be reserved for evaluation of a suspicious area or abnormality. The mediolateral view may be included as part of a routine examination, but a mediolateral oblique projection must always be obtained.

Are Sufficient Personnel Available to Do Screening Mammography?

Some have claimed that current levels of personnel are insufficient to offer mammographic screening for all women at risk. This claim is refuted, however, by a recent report from the Office of Technology Assessment, Congress of the United States [33].

In the United States, approximately 50 million women are 40 years old or older. There are approximately 20,000 radiologists in this country. If all these women had mammograms and all radiologists interpreted the mammograms, each would have to interpret 10 per day. If only 25% of the radiologists were capable, interested, or able to participate in mammographic screening, each would have to interpret 10,000 mammograms per year. On the basis of a 250 work-day year, each radiologist would have to screen 40 examinations per day. For those who have never participated in a screening program, 40 mammograms per day may seem overwhelming. However, once a radiologist has sufficient experience, interpretation of this number of images can be easily handled in about 2 hr. Admittedly, this degree of expertise will not be achieved overnight. It will require some gradual phasing-in, training, and experience. As new radiologists enter into practice from training centers where such volumes are common, this will not be perceived as a major problem. Cases of cancer may be missed or misinterpreted, but this will happen with low volume as well. Double reading has been shown to reduce observer error, but is generally not practical, particularly in a high-volume setting.

It has been proposed that physician assistants (PAs) be trained and used to alleviate the problem of insufficient numbers of physicians for mammographic screening. However, no deficiency exists in the number of physicians available, so this argument does not apply. As far as reducing the cost of mammographic screening (professional interpretation fees of $5–$10 per mammogram are currently charged in many areas), little further cost reduction can be expected by the use of PAs. The time required to train a PA to make the appropriate observations is no shorter than the time required to train physicians to do so.

Hillman et al. [34] have shown that nonphysicians can be trained to review mammograms. I believe that PAs can be of most help in screening, not to screen out the normal cases, but to provide the capability for double reading of all screenings. This should reduce the observer error in a meaningful way and could still be done in a cost-effective fashion.

Quality control is a problem that is being addressed by the American College of Radiology. Although the program will never be perfect, we do not expect it to be worse than any other area of radiologic expertise. There will be good, and there will be bad. In this less-than-perfect world, we hope the good will outweigh the bad. I can see a potentially important application of computer technology to improve mammographic interpretations. Digital imaging may allow information to be captured at lower exposure levels, with reconstruction of the data by image processing. Work is now in progress to develop expert systems to assist the radiologist in the evaluation of troublesome microcalcifications and other low-yield findings. Perhaps the computer will soon allow an in-house consultant to be constantly available, even in remote locations.

Radiation Risk

According to the World Health Organization Commission on Technology Assessment [35], the risk of screening mammography with modern equipment for women who are 40 years of age or older is negligible. Studies involving Japanese data and radiation fluoroscopic studies indicate that although the risk per rad with highly fractionated doses are probably additive and although there is little sparing effect of dose fractionation, women who are 40 years old at exposure are at little or no excess risk compared with the general population [36]. They also state that "another recent estimate on the influence of diagnostic radiography on the incidence of breast cancer and leukemia, which predicted 788 new breast cancers per year induced among 117,000,000 American women (over their lifetime), may have overestimated the risk by a factor of five to ten."
age 35, annual examinations beginning at age 40 for the rest of their lives, and 3 views per breast for each examination. This group would have 93,000 naturally occurring tumors. If film-screen mammography were used, 150 additional cases of breast cancer might be induced. If xeromammography were used, 1000 additional cases might be induced. Over 40 years of exposure, this means that an individual woman might have as many as four chances in one million of developing a new cancer as a result of each year’s screening.

Another way to estimate benefit vs risk is to use the person-years gained by screening as shown in the HIP trial and apply it to the data of Gohagan et al. In the HIP study, each woman destined to get cancer in the screened group had 3.5 years added to her life. If 40% of this benefit was due to mammography alone, about 130,000 added person-years would be gained. If we assume that 3.5 years are lost for each induced cancer, a net gain of 129,675 person-years can be estimated.

Conclusions

In this overview of breast cancer, I have avoided a detailed discussion of the monetary costs. The economics of screening is a morass that probably lies outside the domain of physicians. Our responsibility, as I see it, is to determine the course that is medically best for our patients. If society deems that the course is worth the cost, it will provide the funds.

We do not know yet how to prevent breast cancer. The dietary fat hypothesis, although interesting, is far from proved [38-40] and, in the opinion of some, is untenable [41]. Moreover, we do not know how to treat advanced breast cancer effectively. Our treatments are best for local control. We know that early detection and current treatment can reduce the death rate and add years to life. Mass screening can be done at $40-$60 per examination. Can we afford not to do it?

The data fail to prove that screening work in the past 3 decades have been “much ado about nothing.” Unfortunately, I cannot yet say that “All’s well that ends well,” because this odyssey has barely begun. Despite all the effort, only a small fraction of primary-care practitioners follow the ACR/ACS suggested guidelines for screening [19, 20, 42].

REFERENCES


23. Moskowitz M, Gartside P, McLauglin C. Mammographic patterns as
42. Leads from the MMWR. Screening for cervical and breast cancer—southeastern Kentucky. JAMA 1988;259(8):1141–1142
This article has been cited by:


