An Overview of the Breast Cancer Screening Controversy

*Daniel B. Kopans**

Randomized controlled studies show that screening mammograms are as important for women aged 40–49 as for women 50 years old and above. It was the improper use of retrospective, unplanned, sub-group analysis to advise women and their physicians that caused the controversy over mammograms for women under 50. Furthermore, arbitrarily grouping women into two groups leads to the incorrect conclusion that the age of 50 is a significant break point when it is not. The data demonstrates that none of the parameters of screening change abruptly at age 50. The recall rates (an abnormal mammogram) and the rate at which biopsies are recommended are virtually the same, regardless of age. Breast cancer is not a trivial problem for women in their forties. More than 30% of the years of life lost to breast cancer are from women diagnosed while in their forties. Because of changing demographics, in 1995 and 1996, there were actually more women diagnosed with breast cancer in their forties than for women in their fifties. The data clearly show that screening women for breast cancer, on an annual basis, beginning by age 40, can reduce the death rate by approximately 24%. It is important to separate medical and scientific analyses from the economic considerations. ''Society'' may decide that it is too expensive to screen women for breast cancer, but women should be provided with the scientific and medical information so that they can participate in the discussion of whether screening is ''worthwhile'' and decide whether or not to avail themselves of its benefit. The economics should not be used to influence the scientific and medical analysis of benefit. [Monogr Natl Cancer Inst 1997; 22:1–3]

There is now clear proof of benefit for screening women ages 40–49 for breast cancer. Not only have the randomized, controlled trials demonstrated a statistically significant mortality reduction of 18%, (*1*), but the Gothenburg trial has demonstrated a 44% mortality reduction that is statistically significant, by itself, and the Malmö trial has demonstrated a statistically significant reduction of 35% (presented to the NIH Consensus Development Conference, January 21–23, 1997). The data are now as strong as the results for women ages 50 and over, among whom only two trials are significant by themselves.

The benefit is even higher since the National Breast Screening Study (NBSS) of Canada should not be included in the analysis. Not only was it a trial of volunteers that differed from the 7 other trials that were trials by invitation, but the control group was screened by clinical breast examination unlike the unscreened controls in the other trials. Of greater concern is the fact that women with signs and symptoms of breast cancer were knowingly permitted to participate in the trial. This resulted in a major randomization problem (*2*,*3*) since the randomization was not blinded. All the women were first given a clinical breast examination and then were allocated to be screened, or to act as unscreened controls, based on open lists rather than blinded assignment. There were more women with lymph node positive cancers in the screened group than the controls. This has never equilibrated, as would be expected, suggesting an allocation imbalance. It resulted in 19 women with advanced breast cancer (4 or more positive nodes) being allocated to the screening arm, whereas there were only 5 women with advanced cancers allocated to the control arm. These are women who, not only could not be helped by screening, but who were likely to have died in the early years of follow-up. The explanation that the control women with breast cancer were treated in community hospitals and had fewer and less extensive axillary dissections than the screened women not only does not explain the imbalance, but it suggests a worrisome treatment asymmetry, as well, that could influence the results. The effort by MacMahon and Bailar to review the allocation process (*4*) was, unfortunately, inadequate since only a few centers were reviewed, and individuals who were involved in the allocation were never interviewed. The NBSS has yet to explain the excess of deaths that persist in the longer follow-up of the trial. Its results, by all estimates, make it a major outlier among the screening trials.

Why Has There Been a Controversy?

The randomized, controlled trials of breast cancer screening have actually, for many years, shown a statistically significant benefit for mammographic screening beginning by the age of 40. It was the inappropriate use of unplanned subgroup analysis that caused the confusion. The controversy over mammographic screening for women in their forties was not based on scientific analysis, but the incorrect use of data. With the exception of the NBSS, *none* of the RCTs were designed to evaluate women ages 40–49 as a separate group. *None* of the trials individually, or even collectively, had sufficient numbers of women in this decade of life to permit an expected benefit of 25% to be statistically significant in the early years of follow-up. In order to have an 80% power to demonstrate a 25% mortality reduction at five years (assuming a five-year survival of 75%), the trials would have had to involve almost 500,000 women split evenly into

**Affiliation of author:* Associate Professor of Radiology, Harvard Medical School, Cambridge, MA; Director of Breast Imaging, Massachusetts General Hospital, Boston.

Correspondence to: Daniel Kopans, M.D., Massachusetts General Hospital, Department of Radiology, Ambulatory Care Center, 15 Parkman Street, Second Floor, Room 219, Boston, MA 02114.

study and control groups (*5*). In addition to the fact that the trials were not designed to evaluate women ages 40–49 as a separate group (the screening intervals and techniques were not optimized) there were actually only 175,000 women under the age of 50 in all of the trials put together. Since it was mathematically impossible for an expected benefit of 25% to be statistically significant in the early years of follow-up, it was specious to suggest that there was no benefit when the benefits that did appear failed to reach significance (*6*). Advising women based on subgroup analysis of data from trials that lacked the statistical power to permit such analysis has been, at best, inappropriate, and the justification for this has never been provided. When analyzed as they were designed, however, the trials have, for many years, demonstrated a statistically significant benefit for screening beginning by the age of 40 (*7*). It is only the improper use of retrospective, unplanned, subgroup analysis to advise women and their physicians that caused the controversy.

Dichotomous Analysis Is Misleading

The confusion was compounded by reviews that purported to show abrupt changes in the parameters of screening occurring at the age of 50 (*8*). This was the result of data grouping that compared women ages 40–49 (as if they were a uniform group) to *all other women* ages 50 and over (as if they were a uniform group). This type of dichotomous grouping, making the age of 50 the point of analysis, leads to the fallacious interpretation and incorrect conclusion that the age of 50 is a significant break point when it is not. The data, in fact, when analyzed by smaller age groups, or individual age, demonstrate that the recall rates (an abnormal mammogram) are virtually the same, regardless of age and the rate at which biopsies are recommended is the same, regardless of age. The only thing that varies is the yield of cancer, and this changes gradually with increasing age, with no abrupt change at the age of 50, reflecting the prior probability of cancer in the population (*9*).

Despite the fact that the trials were not designed for sub-group analysis, with longer follow-up and more deaths, the trials now demonstrate statistically significant benefit, even when women ages 40–49 are analyzed separately. The most recent overview of the seven trials with similar design shows a 24% mortality reduction for women ages 40–49, that is significant. Even with the addition of the flawed NBSS data, the benefit is significant (*1*).

The Benefit Is Not Due to Women Reaching the Age of 50

The argument should be moot, but it has been suggested that this benefit is due to women reaching the age of 50 and screening suddenly becoming effective. Not only is this biologically not supportable, but RCT data cannot legitimately be analyzed by age at diagnosis. Age at diagnosis is a pseudovariable that is influenced by the intervention. Its use will, *a priori,* bias an analysis against cancers detected among younger women in the screened groups (*10*). RCT divide women into two groups. If the numbers involved are large enough, and the assignment is truly random, then every woman in the screened group will have a twin in the common group. For every woman in the screened group who develops a cancer there will be a woman in the control group whose cancer will behave in the same fashion. Using the age at diagnosis will bias the conclusions against the younger screened women. For example, assume that woman A (in the screened group) has her cancer detected when she is in her forties, and, as a consequence, she will not die from breast cancer. Her ''twin,'' patient B (in the control group), does not have her cancer diagnosed until she is in her fifties. If the age at diagnosis is used, the avoidance of death by ''A'' will not have any control group counterpart, and there will be no apparent mortality benefit for women screened in their forties. The death of woman ''B'' will be attributed to women over the age of 50. Thus, analyzing the data using the age at diagnosis will be misleading and will bias the results against screening the younger women. Nevertheless, even if the rules of RCT analysis are ignored and age at diagnosis is used, in the three trials that have performed such analyses, the benefit has been shown to be primarily for women whose cancers were diagnosed while they were still in their forties in the HIP trial (*11*), the Kopparberg trial (*12*), and in the Gothenburg trial (*1*).

The Benefit Is Actually Greater Than Indicated by the RCTs

What is often forgotten is that the RCTs underestimate the benefit of screening due to noncompliance and contamination. With the exception of the Canadian trial, which involved volunteers (a separate problem), the seven trials first randomized a population and then invited them to be screened. Women allocated to be screened who refused the invitation (noncompliance) are still counted as having been screened, and if they die of breast cancer their deaths are attributed to the screened group. Similarly, women who had mammograms on their own, outside of the screening program, and whose lives were saved as a result, are still counted as unscreened controls. The benefit of screening is likely higher than the trial results would indicate.

The ''Harms'' of Screening Do Not Change Suddenly at Age 50

Some analysts have raised the issue of ''harms'' from screening. These include anxiety from the process as well as biopsies that prove to be for a benign reason (termed unnecessary). Not only are these ''harms'' not equivalent to dying from breast cancer, but they are true for women at all ages, and do not change abruptly at the age of 50. As noted above, the recall rate for an abnormal mammogram is fairly constant across all ages, as is the ''biopsy recommended'' rate. The yield of breast cancer increases steadily with increasing age and merely reflects the prior probability of breast cancer in the population with no abrupt change at any age (*13*).

Breast Cancer Is Not a Trivial Problem for Women in Their Forties

Finally it has been suggested that breast cancer is not a major problem for women in their forties. In fact, more than 30% of the years of life lost to breast cancer are from women diagnosed while in their forties (*11*). Although the incidence of breast cancer increases steadily with increasing age, there are so many women in their forties, that, in 1995 and 1996, there were actually more women diagnosed with breast cancer in their forties than among women in fifties (*14*). It is also often forgotten that

many cancers that are diagnosed after the age of 50 have been growing for several years, and could have been diagnosed while the woman was in her forties.

A Delayed Benefit Does Not Mean No Benefit

Opponents have implied that, since the trials took longer for a benefit to appear among younger women than older women, that the benefit is not important. This is incorrect. To begin with, there is no biological reason to expect an immediate benefit. Given the parameters of the screening trials, a ''delayed'' benefit makes biological sense.

Most of the RCTs used a screening interval that was too long for younger women (two or more years between screens). Faster growing tumors were not interrupted. The benefit from interrupting the more moderate-growth cancers among the screened women cannot appear until the women in the control group succumb to their cancers. This is likely to not occur for five or more years after the cancers among the screened women were detected. Since most cancers are not detected in the first year of screening (the date from which the benefit is measured) and many women live for many years, even with breast cancer that will, ultimately, be lethal, the result is the appearance of a ''delayed'' benefit. Trials that screened at a shorter interval (Gothenburg and HIP) showed an earlier divergence of the mortality curves (years 5–7). Nevertheless, a ''delayed'' benefit does not lessen the value. As Feig has pointed out, a woman whose cancer is diagnosed at age 42 and consequently lives beyond age 52 derives as much if not more benefit than a woman whose cancer is found at age 55 such that she lives beyond age 60 (she had already lived beyond age 52).

The Determination of Medical Benefit Should Be Separated from Economics

It is important to separate the medical and scientific analysis from the economic considerations. ''Society'' may decide that it is too expensive to screen women for breast cancer, but women should be provided with the scientific and medical information, so that they can participate in the discussion of whether screening is ''worthwhile'' and decide whether or not to avail themselves of its benefit. The economics should not be used to influence the scientific and medical analysis of benefit.

Summary

The age of 50 has no biological significance, yet women and their physicians have been led to believe from data grouping and improper data analysis, that it represents a true threshold. There are no parameters of screening that change abruptly at age 50, or any other age. As with any test, there are false-negative examinations and false-positive examinations. Women at all ages should be provided with information concerning the ''risks'' and benefits of screening, so that they can make informed decisions.

The data clearly show that annually screening women for breast cancer, beginning by age 40, can reduce the death rate by approximately 24%. The benefit is likely even higher (*15*). Since there are no known ''risks'' that relate to an annual screening interval, women should know that the only reason to go to a longer interval between screens is economic. There is probably little or no radiation risk for women by the time they reach the age of 40 (*16*). Since the lead time for detecting cancer by mammography is approximately two years for younger women (it is not clear where ''younger'' ends and ''older'' begins) (*17,18*), screening at this interval, or longer, will not add much to the health care without screening. They should be screened at an interval that is less than two years (*19*). It may be possible to go to a longer interval among older women, since the lead time appears to be longer for them, but the age at which this can be done safely has not been determined. Since a 30% benefit has been shown for women over the age of 49 who were screened with intervals of almost three years, a much greater benefit will likely occur with more frequent screening.

References

(*1*) Tabar L, Larson LG, Andersson I, Duffy SW, Nystrom L, Rutqvist LE, et al. Breast cancer screening with mammography in women aged 40–49. Report of the organizing committee and collaborators, Falun Meeting, Falun, Sweden. March 21–22 1996. Int J Cancer 1996;68:693–9.

(*2*) Kopans DB, Feig SA. The Canadian National Breast Screening Study: a critical review. AJR AM J Roentgenol 1993;161:755–60.

(*3*) Tarone RE. The excess of patients with advanced breast cancers in young women screened with mammography in the Canadian National Breast Screening Study. Cancer 1995;75:997–1003.

(*4*) Bailar JC, MacMahon B. Randomization in the Canadian National Breast Screening Study: a review for evidence of subversion. Can Med Assoc J 1997;156:193–9.

(*5*) Kopans DB, Halpern E, Hulka CA. Statistical power in breast cancer screening trials and mortality reduction among women 40–49 with particular emphasis on the National Breast Screening Study of Canada. Cancer 1994;74:1196–1203.

(*6*) Fletcher SW, Black W, Harris R, Rimer BK, Shapiro S. Report of the International Workshop on Screening for Breast Cancer. J Natl Cancer Inst 1993;85:1644–56.

(*7*) Shapiro S. Screening: assessment of current studies. Cancer 1994;74: 231–8.

(*8*) Kerlikowske K, Grady D, Barclay J, Sickles EA, Eaton A, Ernster V. Positive predictive value of screening mammography by age and family history of breast cancer. JAMA 1993;270:2444–50.

(*9*) Kopans DB, Moore RH, McCarthy KA, Hall DA, Hulka CA, Whitman GJ, et al. The positive predictive value of mammographically initiated breast biopsy: there is no abrupt change at age 50 years. Radiology 1996;200:357–60.

(*10*) Prorok PC, Hankey BF, Bundy BN. Concepts and problems in the evaluation of screening programs. J Chron Dis 1981;34;159–71.

(*11*) Shapiro S, Venet W, Strax P, Venet L. Periodic Screening for Breast Cancer: The Health Insurance Plan Project and its Sequelae, 1963–1986. Baltimore (MD): Johns Hopkins University Press, 1988.

(*12*) Tabar L, Duffy SW, Chen HH. Re: quantitative interpretation of age-specific mortality reductions from the Swedish Breast Cancer Screening Trials. J Natl Cancer Inst 1996;88:52–3.

(*13*) Kopans DB, Moore RH, McCarthy KA, Hall DA, Hulka CA, Whitman GJ, et al. Biasing the interpretation of mammography screening data by age grouping: nothing changes abruptly at age 50. Presented at the Radiological Society of North America Meeting 1995, Chicago (IL).

(*14*) American Cancer Society statistics based on 1979–1993 SEER incidence rates and U.S. Census Projections provided by SEER. American Cancer Society, Atlanta (GA).

(*15*) Feig S. Estimation of currently attainable benefit from mammographic screening of women aged 40–49 years. Cancer 1995;75:2412–9.

(*16*) Mettler FA, Upton AC, Kelsey CA, Rosenberg RD, Linver MN. Benefits versus Risks from mammography: a critical assessment. Cancer 1996;77: 903–9.

(*17*) Moskowitz M. Breast cancer: age-specific growth rates and screening strategies. Radiology 1986;161:37–41.

(*18*) Tabar L, Faberberg G, Day NE, Holmberg L. What is the optimum interval between screening examinations? An analysis based on the latest results of the Swedish Two-county Breast Cancer Screening Trial. Br. J. Cancer 1987;55:547–51.

(*19*) Kerlikowske K, Grady D, Barclay J, Sickles EA, Ernster V. Likelihood ratios for modern screening mammography: risk of breast cancer based on age and mammographic interpretation. JAMA 1996;276:39–43.