
TARGET ARTICLE

Mammography Screening: Prospects and Opportunity Costs

Ana M. Navarro and Robert M. Kaplan

*Department of Family and Preventive Medicine
University of California, San Diego*

Health care costs in the United States now consume nearly 15% of the gross domestic product. Continued expansion of health expenditures may have serious economic consequences, including reduction in the standard of living. Health care reform must include cost control without consequent detrimental effects on health status. As a case example, we consider the controversy surrounding mammography screening for premenopausal women. Several literature reviews of published studies suggest that screening of women less than 50 years of age does not statistically significantly reduce mortality from breast cancer. These results are not explained by screening interval, recency of study, or patient compliance to screening. We conclude that screening is effective in decreasing mortality from breast cancer for women older than 50 years. For women less than 50, mammography screening programs displace resources that could have a greater benefit in women's health status if used for other purposes.

Key words: health policy, mammography, opportunity cost, cost-effectiveness analysis

Three important problems in American health care are affordability, access, and accountability (Kaplan, 1993). The affordability problem results from the inability to pay for all health services that are desired. Health care costs in the United States

have grown exponentially since 1940 and the rate of increase has continued to accelerate through the 1990s. Health care in the United States now consumes 14.5% of the gross domestic product, whereas no other country in the world spends more than 10% (Health Care Financing Administration, 1994).

Despite high expenditures on health care, 58 million Americans have no health insurance for part of each year and 38 million are uninsured throughout the entire year (Levit, Olin, & Letsch, 1992). Programs designed to serve the poor, such as Medicaid, have had to redefine eligibility to exclude the majority of low-income people. The costs of Medicaid programs have grown dramatically and all states must now consider cost-cutting strategies (Merritt & Demkovich, 1991). The third problem is accountability. Despite the fact that we spend more on health care than any other country in the world, we have a great deal of difficulty demonstrating that our high expenditures result in health benefits (Kaplan, 1993).

The three problems (affordability, access, and accountability) are connected. Access has become limited because health care is unaffordable, and care may be too expensive because there is poor accountability. Better accountability may preserve resources that could be used to provide access to a larger number of people.

OPPORTUNITY COST PROGRAM

With limited resources, health plans cannot include all programs. To some extent, political factors and lobbying influence what programs are selected. Successful lobbying to obtain reimbursement for a specific service may necessarily mean that another service is excluded. This is often called the opportunity cost problem. Opportunity costs are the foregone opportunities that are surrendered as a result of using resources to support a particular decision. With a fixed budget, the decision to increase spending for a program includes a decision to spend less money elsewhere.

When confronted with the choice between two programs, many people think the solution is easy—do them both. The difficulty is that it is expensive to offer multiple programs. The cost of programs is represented in the fees for health insurance or the cost of health care to taxpayers. We can choose to offer as many health programs as we want, but we need to pay for them. Employees do not want the fees for their health insurance to rise and taxpayers do not want tax increases. The goal of formal decision models is to get higher quality health care at a lower cost (Eddy, 1994). In this article, we argue that success in obtaining support for some programs may have unintended consequences. As a result, resources may not be used most efficiently in order to deliver a public health benefit. The campaign to achieve mammography for all women will be used as a case study.

SERVICES FOR WOMEN

An increasing percentage of Americans under the age of 65 have no health insurance. The age-adjusted percentage of uninsured Americans increased from 12.5% in 1980 to 17.2% in 1992 (National Center for Health Statistics [NCHS], 1994). Nonelderly women are more likely to be insured than men (85% versus 81%), partly because more women than men are covered by Medicaid (8% versus 3%). Overall, one third of all women in the lowest income bracket are uninsured and in greatest need of help, compared with 16% of the nonelderly U.S. population as a whole (Collins, Romero, Drummond, & Shannon, 1993).

Although the proportion of insured women is similar to the proportion of insured men, women may have many disadvantages in using the system. In Michigan, Bashshur, Homan, and Smith (1994) found that 20% of women had access problems compared to 14% of men. Other findings from the analysis confirmed inequity of access to health care. Particularly disadvantaged were members of certain ethnic minorities, women laborers and the poor, and those in poorer health (Bashshur et al., 1994). Furthermore, using a method to simulate out-of-pocket cost, Sofaer and Abel (1990) showed that Medicare provides better coverage for illnesses that predominate among men than those that predominate among women. The authors also found that women on Medicare who supplement their basic coverage by purchasing a typical private insurance Medigap policy do not receive as much of an advantage from their purchases as men do (Sofaer & Abel, 1990).

Preventive science has advanced to the point where preventive interventions are expected to improve health. However, the health care system has done a poor job of delivering some important services to women. For example, since 1980 the percentage of mothers receiving early prenatal care has remained stable at 79% for White mothers and 60% to 62% for Black mothers (NCHS, 1994). Furthermore, there is an important gap between knowledge in how to prevent morbidity and mortality of diseases and the implementation of prevention measures in areas such as cardiovascular diseases that constitute the principal cause of death among women. Thus, approximately 36% of cardiovascular disease morbidity in women over 44 years of age can be attributed to elevated serum cholesterol. In addition, glucose intolerance is of major importance for women and accounts for 36%–37% of cardiovascular mortality (White, Tolsma, Haynes, & McGee, 1987). Prevalence of glucose intolerance could be reduced through reduction in rates of obesity. Data from the National Health Interview Surveys suggest that the age-adjusted percentage of adults who were overweight increased from 25% to 33% between 1960–1962 and 1988–1991. The prevalence of overweight persons increased for all population subgroups, and more women were classified as overweight than men (NCHS, 1994; Piani & Schoenborn, 1993). In a variety of different problem areas, the financially

crippled system has had difficulty delivering necessary care. Clearly, many potentially valuable services for women are not being delivered because of limited resources. At the same time significant resources have been devoted to services that may be unnecessary. In the next sections, methods for achieving more efficient use of resources are reviewed.

MAMMOGRAPHY

Mammography screening programs have become the central focus of women's health advocacy campaigns. The American Cancer Society (ACS) recommends that screening mammography begin at age 40 (American Cancer Society, 1994), whereas many other organizations suggest that screening mammograms are unnecessary before age 50 (Fletcher, Black, Harris, Rimer, & Shapiro, 1993). This difference of opinion highlights some important problems in public decision making. Two points of view must be considered. First, mammography coverage has become the focal point for women's health policy despite the need to provide a variety of other services for women. Another constituency group is providers who have made significant profit offering these tests. There is a legitimate controversy over whether screening mammography should be advocated for to women less than 50 years of age who do not have other risk factors for breast cancer. The medical establishment's position is best exemplified in statements by a past president of the ACS of California, who stated, "I don't know if mammograms are effective under the age of 50, but I don't see any reason not to have them. Nobody is going to get hurt by them" (Duerksen, 1994).

Is it true that women are not hurt by these policies? The issue is not whether mammograms are dangerous. Rather, the problem is one of opportunity cost. Devoting resources to mammography is harmful when it detracts from the opportunity to use the resources for other services that may be necessary to enhance the health of women.

MAMMOGRAPHY ADVOCACY

In their testimony on health care reform, the ACS argued for mammography every 1 to 2 years for women age 40 to 49. These are the guidelines currently supported by the ACS. The ACS believes published studies support the use of screening mammography for women 40 to 49 years of age. The ACS explains the difference between their own position and that taken by the National Cancer Institute (NCI) on the basis of the types of studies considered (Mettlin & Smart, 1994). The ACS considers that the NCI places too much emphasis on evidence of mortality reduction derived from randomized clinical trials. The ACS argues that earlier detection

provides opportunities for more conservative treatment. Studies do show a relation between the use of mammography and the stage of cancer at time of detection. Because women are more likely to survive if cancer is detected early, it stands to reason that early detection through mammography will result in better outcomes. In addition, the ACS argues that the NCI did not pay enough attention to descriptive studies, trends in breast cancer incidence and mortality, and large nonrandomized studies such as the Breast Cancer Detection Demonstration Project (BCDDP), which support the value of screening 40 to 49-year-old women (Mettlin & Smart, 1994). Although compelling, we side with the NCI and consider randomized clinical trials as the most persuasive evidence for the benefits of screening. The reason for this position derives from the difficult issues of lead time and length bias that we address in the next section.

EPIDEMIOLOGIC BIASES AND RESEARCH METHODOLOGY

To understand this controversy, it is necessary to consider two biases: lead time bias and length bias.

Lead Time Bias

Cancer screening may result in early detection of disease. Survival is typically calculated from the date that disease is documented until death. Because screening is associated with earlier disease detection, the interval between detection and death is longer for screened cases than for unscreened cases. Epidemiologists refer to this as *lead time bias*. Figure 1 illustrates this bias.

Imagine that two women each develop breast cancer in 1980 and died in 1995. Hypothetically, the progression of the cancer is identical in these two women. The woman illustrated on the top line of Figure 1 was screened in 1982 and the cancer was detected. After this diagnosis, she lived 13 additional years before her death in 1995. The woman shown on the lower line did not receive screening and detected a lump herself in 1992. After this, she lived 3 additional years. Survival for the woman on the top appears to be much longer than that for the woman on the bottom, even though the interval between developing cancer and dying is exactly the same. Observational (nonrandomized) studies are unable to separate lead time bias from treatment effect and it has been suggested that increased survival associated with screening can be attributed to lead time and not to early detection and treatment (Eddy, 1989; Fletcher et al., 1993; Kerlikowske, Grady, Rubin, Sandrock, & Ernster, 1995). The only way to eliminate lead time bias is to perform clinical trials in which women are randomly assigned to either treatment or control groups and

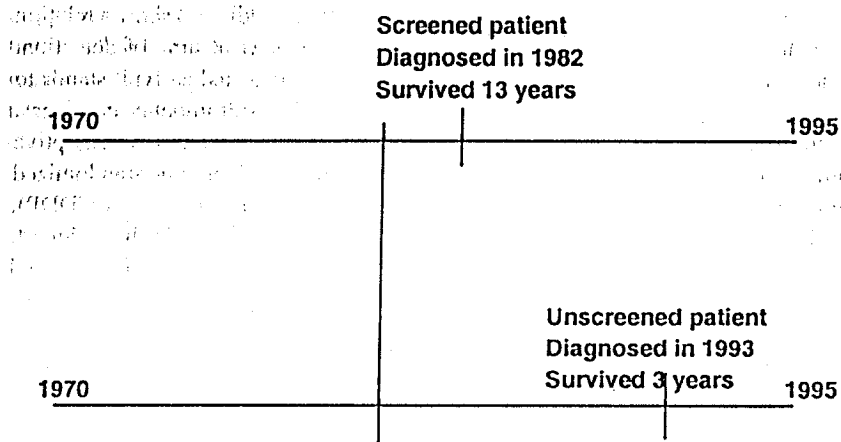


FIGURE 1 Example of lead time bias. Both patients developed cancer in 1980 and died in 1995.

followed for many years. These trials have many methodological problems, but they remain our best way of determining the value of screening.

Length Bias

Tumors progress at different rates. Some cancers are very slow-moving whereas other tumors progress very rapidly. Some cases may regress, remain stable, or progress so slowly that they never produce a clinical problem during an ordinary lifetime. These cases might be described as pseudo-disease because they are not clinically important (Black & Welch, 1993). The probability that disease is detected through screening is inversely proportional to the rate of progression. For example, with rapidly progressing disease, early detection may not produce a clinical benefit because cases are detected too late. On the other hand, diseases with very long preclinical phases are more likely to be detected by screening. However, diseases that are progressing extremely slow may never cause clinical problems. Ironically, advances in screening technology have a greater likelihood of detecting cases for which a clinical manifestation will never materialize (Black & Welch, 1993).

It is possible that some of the apparent benefits of screening and treatment for cancer are actually attributable to lead time and length bias. If this were true, then the greater incidence of detected disease would not be reflected in reduced mortality rates. This appears to be the case for breast cancer. Current data suggest that, despite increases in screening, rates for breast cancer mortality have remained constant over the last two decades (Kaplan, 1993). The same holds for ovarian cancer, colon

cancer, and most other malignancies (except lung cancer). The center of the controversy is a set of experimental trials that evaluate the benefits of screening. These trials are important because they eliminate lead time and length bias. Thus, we focus our review on these experimental trials.

META-ANALYSIS

We recently reviewed trials on the efficacy of screening mammography published as of 1992. Only published studies allowing comparisons between intervention (i.e., mammography screening) and control group were included in our review. These include the Health Insurance Plan (HIP) of New York (Aron & Prorok, 1986; Shapiro, Venet, Strax, & Venet, 1988a, 1988b; Shapiro, Venet, Strax, Venet, & Roeser, 1982), the Malmö Study (Andersson et al., 1988), the Swedish Two-County Study (Fagerberg & Tabár, 1988; Tabár, Fagerberg, Day, & Holmberg, 1987; Tabár et al., 1985), the Nijmegen Study (Verbeek, Hendriks, Holland, Mravunac, & Sturmans, 1985; Verbeek et al., 1984; Verbeek, Straatman, & Hendriks, 1988), the Florence Study (Palli et al., 1986), the Canadian Breast Cancer Screening Study (Baines, McFarlane, & Miller, 1988, 1990; Baines, Miller, et al., 1990; Miller, 1988; Miller, Baines, To, & Wall, 1992a, 1992b), the Periodic Multiphasic Health Checkup Study (Dales, Friedman, & Collen, 1979), and the United Kingdom Study (Roberts et al., 1990; UK Trial of Early Detection of Breast Cancer Group, 1988a, 1988b).

These studies were identified by a computer assisted MEDLINE search as well as a review of references of the identified published articles on effectiveness of mammographic screening. All studies were abstracted following a coding system developed to include the following variables: population target (i.e., geographic location, members of a particular health maintenance organization [HMO]), design of the study (i.e., randomized trial versus case-control study), sample selection, screening procedures (i.e., type of mammography, physical breast exams, breast self-examination), the frequency of breast cancer screening, compliance with the experimental treatment, the nature of the control group, breast cancer screening rates in the control group, the quality of the mammogram, and breast cancer mortality. In addition, the data were systematically evaluated according to the age at entry of the women who were screened and the age at which breast cancer was diagnosed. Characteristics of the reviewed studies that were systematically reported in the respective publications are summarized in Table 1.

Figure 2 summarizes the odds ratio of breast cancer death by age at entry for the eight studies included in our review. Results include mortality from breast cancer exclusively because data on quality of life were not available in any of the published studies. Mortality data included in the analyses constitute the latest follow-up that was reported in the respective published studies. Odds ratios less than 1.0 imply

TABLE 1
Characteristics of the Studies Included in the Meta-Analysis

<i>Study</i>	<i>Country</i>	<i>Duration of the Study</i>	<i>Length of Follow-Up (in Years)</i>	<i>Age of Subjects</i>	<i>No. of Cases</i>	<i>Design</i>	<i>Frequency of Screening</i>	<i>Physical Breast Exam</i>	<i>Compliance With Screening</i>	<i>Type of Control Group</i>
Health Insurance Plan	United States	1963-1981	18	40-64	30,131	Experimental	12 months	Annual	80%	Usual care
Malmö	Sweden	1976-1986	8.8 (mean)	44-68	21,088	Experimental	18-24 months	No	74%	Usual care
Swedish Two-County Study	Sweden	1977-1985	7	40-75	78,085	Experimental	40-49 every 2 years, 50-75 every 33 months	No	89.2%	Usual care
Nijmegen	The Netherlands	1975-1982	7	35-65	23,205	Case-control	24 months	No	84.8%	Usual care matched for age
Florence	Italy	1970-1984	7 (median)	40-70	24,813	Case-control	30 months	In selected cases	60%	Usual care
National Breast Screening Study (NBSS)	Canada	1980-1990	8.5 (mean)	40-59	89,835	Experimental	12 months	Annual	88%	BSE; physical breast exam (only once for women younger than 50)
Periodic Multiphasic Health Checkups (MHC)	United States	1964-1975	11	48-54	5,156	Experimental	24 months	Uncertain	60%	Not urged to use MHC
U.K. Trial	United Kingdom	1979-1988	7	45-64	30,473	Nonrandomized prospective study	24 months	Annual	66.5%	BSE; usual care

Note. BSE = breast self-exam.

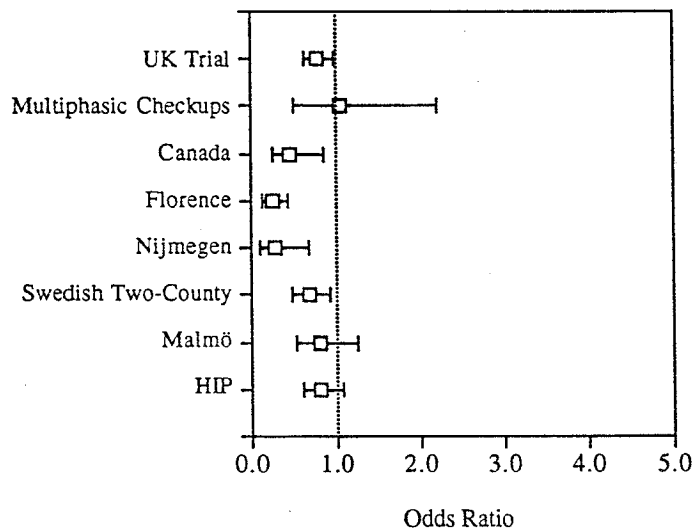
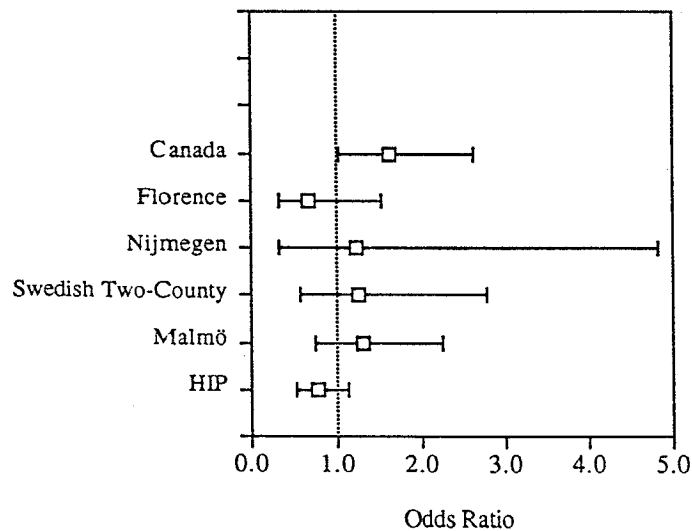


FIGURE 2 Odds ratio (odds ratio and 95% confidence interval) of breast cancer mortality by age at entry.

that mortality from breast cancer is lower in the experimental group (i.e., mammography screening) than in the control group. Conversely, odds ratios greater than 1.0 imply that mortality from breast cancer is greater in the experimental than in the control group. However, only six of the studies presented the results broken down by age. For these six studies the results are shown separately for women under age 50 and women over 50. Age 50 was chosen because it approximates the average age at menopause. As Figure 2 shows, there appears to be no statistically significant benefit of breast cancer screening for women less than 50 years old. The HIP study did suggest some benefit among younger women who were followed for a longer period of time (18 years). Also, the Florence Study reported lower breast cancer mortality rates among women younger than 50 who had received mammography screening as compared to those who had not. However, these differences were not statistically significant.

Assuming a fixed effects model, the summary odds ratio for experimental randomized studies was calculated following the Mantel and Haenszel statistic (Fleiss, 1981). The summary odds ratio for women 50 years old and older was of .75 ($p < .001$) and the 95% confidence interval was between .54 and 1.04. For women younger than 50 years, the summary odds ratio was of 1.09 ($p = .21$) and the pertinent 95% confidence interval fell between .67 and 1.77. The homogeneity test statistics (Breslow & Day, 1980) for women over 50 years old and for women younger than 50 years were of 4.26 ($p = .04$) and 6.52 ($p = .01$), respectively. Given the heterogeneity of the odds ratios in the reviewed studies, we examined the relation between the odds ratios and some of the key variables that were available in all published studies. These included screening interval, date of study, and compliance rates.

Figure 3 summarizes the relation between screening interval and benefit. It might be argued that more frequent screening should produce the greatest benefit. The interval of screening is important because more frequent screening significantly increases costs. However, the figure shows that there is no systematic relation between screening interval and outcome for women less than 50 years of age. Although women older than age 50 seem to benefit from mammography, the interval of screening does not appear to provide systematic advantage.

One argument refuting these results is that the equipment used in mammography screening studies is out of date. Thus, it might be argued that more recent studies may be more likely to show the benefit of screening. However, averaged across all women, there was no systematic relation between the first year of study implementation publication and benefit (see Figure 4). For women less than 50 there was a slight trend, but in the opposite direction. It was the early (i.e., HIP and Florence studies), rather than the later studies, that tended to suggest a benefit of mammography.

Another argument is that some studies on mammography screening must be discounted because compliance rates were low. For example, in some cases women

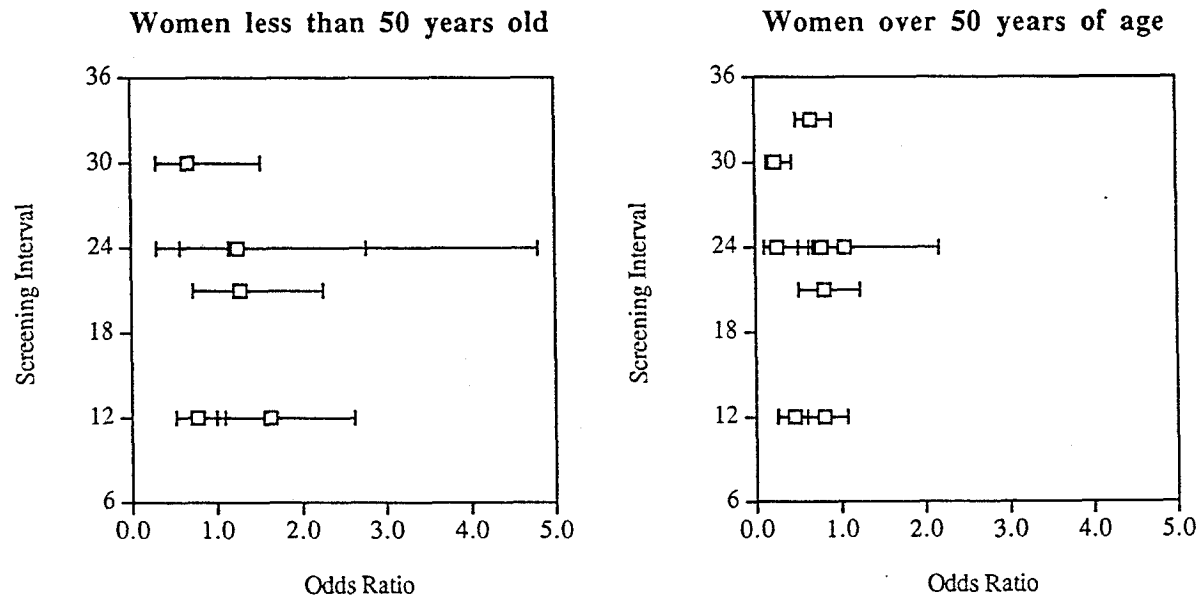


FIGURE 3 Relation between odds ratio (odds ratio and 95% confidence interval) and screening interval broken down by age group.

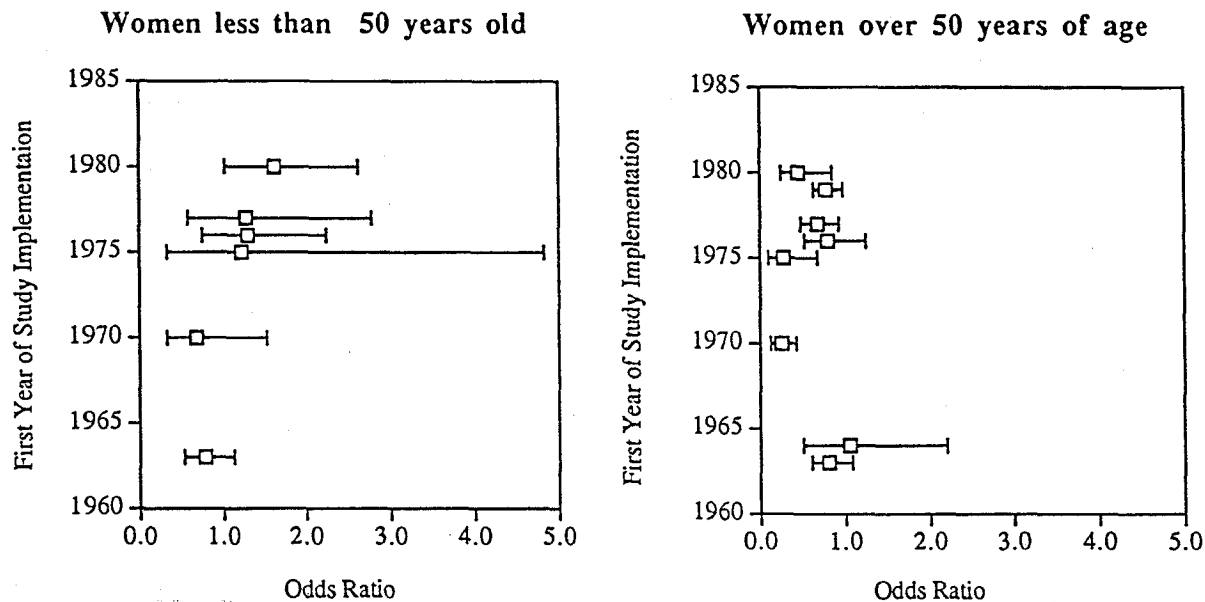


FIGURE 4 Relation between odds ratio (odds ratio and 95% confidence interval) and first year of study implementation broken down by age group.

assigned to obtain mammograms actually did not receive them. Figure 5 breaks down the odds ratios by compliance rates. As the figure shows, compliance rates were not systematically related to outcome. Indeed, the study with the lowest compliance rate (i.e., the Florence study) tended to show the greatest benefit of mammography.

LIMITATIONS OF CURRENT DATA

Overall, our review found little evidence supporting screening mammography for women under the age of 50 years. However, several issues should be taken into account. First, it has been argued that the failure to find a significant effect of screening results from low statistical power. It is possible that larger studies would show statistically significant benefits for screening (Kopans, Halpern, & Hulka, 1994). However, the results suggest that, at best, the effect is very small (summary odds ratio of 1.09). Thus, the sample size required to detect this small statistical effect would be enormous.

Second, the results of the HIP study suggest that the benefits of mammographic screening in women younger than 50 are apparent when longer follow-up data become available. In the HIP study, among women younger than 50 years at entry in the study, the relative risk was of .93 (95% confidence interval: .50; 1.75) at the 5-year follow-up and of .77 (95% confidence interval: .53; 1.11) at the 18-year follow-up. This positive relation between length of the follow-up and effect size is presumably associated with the fact that cancer incidence increases with age. Because a larger number of women develop cancer as they grow older, larger follow-ups increase statistical power. However, increasing statistical power through larger follow-ups leads to ambivalence about the relative effect of early mammography.

Women with longer follow-ups may have a greater reduction in deaths associated with breast cancer because they are older at the time of follow-up assessments. These women may have been screened closer to age 50 or after age 50, and the observed benefit may be that commonly reported for women of postmenopausal age (Kerlikowske et al., 1995). De Koning, Boer, Warmerdam, Beemsterboer, and van der Maas (1995) recently published a study in which a computer simulation program was used to analyze the Swedish breast cancer screening trials. De Koning et al. concluded that as much as 70% of the reduction in breast cancer mortality for women 40-49 years at trial entry might be attributable to a reduction due to screening these women after they reach age 50 (de Koning et al., 1995). Certainly, this issue deserves further evaluation. In particular, we need more data on women screened during their 30s and evaluated 10 to 20 years after their initial screening.

Third, an important issue is the quality of the mammography. Problems in equipment as well as human errors in reading the mammograms might result in

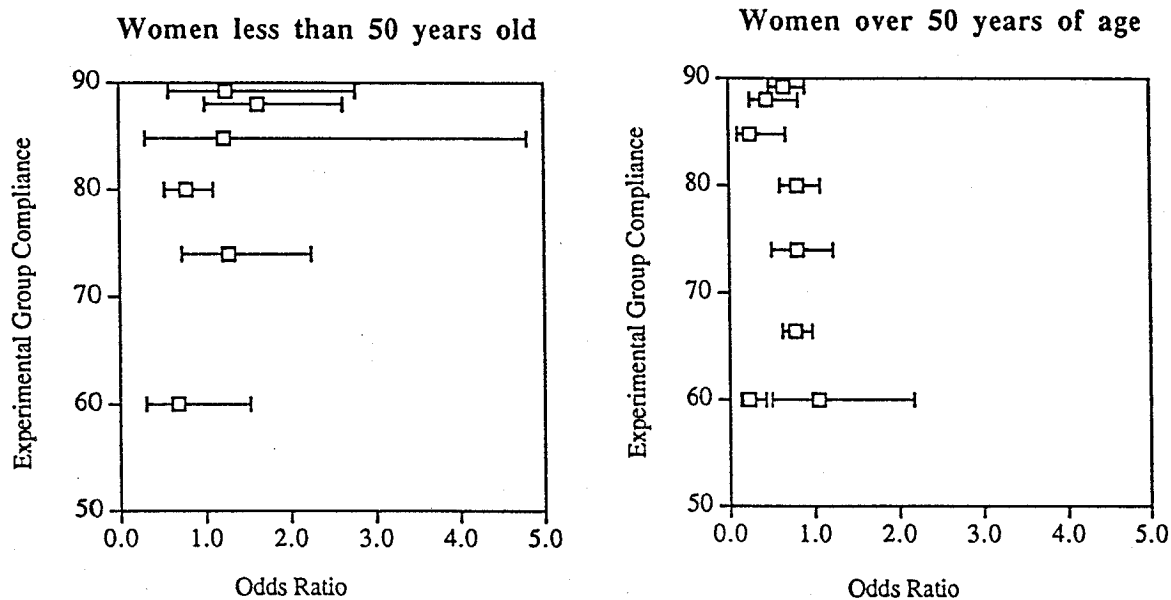


FIGURE 5 Relation between odds ratio (odds ratio and 95% confidence interval) and compliance in the experimental group broken down by age group.

both false positives and false negatives diagnostic decisions. The reviewed published studies identified in the literature do not provide complete information regarding the quality of mammographic screening. The HIP study reported that independent radiology readings of each set of films were made by two of three radiologists (Shapiro et al., 1988b). In addition, a 10% random sample of films were reviewed for quality of technique and interpretation and improvements in technique were introduced as suggested by these reviews (Shapiro et al., 1988b). Andersson et al. (1988) indicated that improved equipment was used as it became available in the Malmö study. Sensitivity of mammography was estimated to be 80% and the positive predictive value was 35% in the Nijmegen study (Verbeek et al., 1988).

The Canadian study was initially designed to provide appropriate statistical power to assess the effect of mammographic screening in women between 40 and 50 years old. However, concerns regarding the poor quality of mammographic screening in the Canadian study have often been used to question the validity of its results. The results of the Canadian study do not suggest that mammographic screening is effective in preventing deaths from breast cancer in premenopausal women. It is likely that these concerns have become apparent because the Canadian trial included the most extensive study of the quality of mammographic screening to date. The results have been reported in various published articles (Baines et al., 1988; Baines, McFarlane, et al., 1990; Baines, Miller, et al., 1990). Unfortunately, data on quality of mammography as complete as the data collected in the Canadian study are not available for the other reviewed studies. Thus, Baines et al. (1988) reported sensitivity of .70 and specificity of .94 for the first screening in the 15 centers participating in the Canadian trial. Additional data were collected in which a single reference radiologist blindly reviewed two-view mammographic screening of 5,200 women not known to have breast cancer, 575 screening-detected breast cancer cases, and 102 interval breast cancer cases. Agreement of the reference radiologist was better for breast cancer cases than for those not known to have breast cancer. Moreover, it was estimated that observer error and technical problems lead to delayed detection in 22% of the screening-detected breast cancer and 35% of the interval breast cancers (Baines, McFarlane, et al., 1990). Further data based on a review of randomly selected films by three external experts suggest that the quality of the mammograms improved over the years of the Canadian trial (Baines, Miller, et al., 1990). However, no differences were found between the quality of the mammograms in different age groups, 40-49 versus 50-59 years (Baines, Miller, et al., 1990). Additional data would be needed to clarify if indeed poor quality of mammographic screening can partially explain the lack of effectiveness of mammographic screening preventing deaths in women younger than 50 years of age but not in women 50 years old or older.

Fourth, in our review we examined compliance with mammographic screening in the experimental groups. However, complete information was not available as to the extent to which women in the experimental groups received mammographic

screening on a regular basis at the intervals offered in the respective studies. Moreover, the percentage of women in the respective control groups who received mammographic screening was not available in most of the reviewed studies. The data do not suggest that there were differences in compliance between pre- and postmenopausal women. However, we cannot rule out the possibility that women younger than 50 years of age get regular cancer screening tests and preventive medical checkups more frequently than women 50 years old or older. If that were the case, the comparison group for women younger than 50 could include a higher proportion of women getting mammographic screening than the comparison group for women over 50 years. Consequently, if similar rates of screening would be achieved in both age groups, the effect of the intervention group in the younger group of women would be smaller than in women over 50.

MAMMOGRAPHY IN CONTEXT

Our review suggests that mortality from breast cancer is not statistically significantly reduced by the use of screening mammography for women 40–49 years of age. There are clearly limitations with the studies, and we cannot say with certainty that regular mammography screening in premenopausal women produces no benefit. Several other reviews of the same literature have been published (Eddy, Hasselblad, McGivney, & Hendee, 1988; Fletcher et al., 1993; Kerlikowske et al., 1995; Miller, Chamberlain, Day, Hakama, & Prorok, 1990; Nystrom et al., 1993; Rodgers, 1990; Shapiro, 1994). These reviews favor the same conclusions. There are some differences in the studies included in the different reviews and the length of the follow-up for the different studies available at the time the respective reviews were conducted.

Table 2 presents a summary of the studies included in the different reviews as well as the length of the follow-up for the studies included in the reviews. Given the differences in the studies included as well as the statistical techniques utilized to combine the results, the effect sizes are not identical in all the reviews. However, the conclusions of the different reviews are consistent in that no statistically significant decrease in mortality of breast cancer has been found in women younger than 40 years old that can be attributed to screening mammography. We agree that future studies might result in evidence supporting the use of regular mammographic screening in premenopausal women. However, current scientific evidence tends not to support the use of screening mammography for women less than age 50 with no identified risk factors for breast cancer.

Various countries around the world have examined the evidence. Virtually all countries, except Sweden, have recommended that screening mammography begin at age 50 (Carter et al., 1993). Recently, the ACS and the NCI were split in their opinions. The NCI, after reviewing the evidence, suggested that screening begin at

TABLE 2
Primary Studies and Respective Length (in Years) of Follow-Up Included in Published Reviews of Mammography Screening Studies for Women Younger Than 40 Years

<i>Study</i>	<i>Eddy et al. (1988)</i>	<i>Fletcher et al. (1993)</i>	<i>Kerlikowske et al. (1995)</i>	<i>deKoning et al. (1995)</i>	<i>Nystrom et al. (1993)</i>	<i>Rodgers (1990)</i>	<i>Shapiro (1994)</i>
Health Insurance Plan	18	10	10	—	—	10 18	12
Malmö	—	12	12	11.8	12	9	12
Swedish Two-County Study	6	12 ^a	12 ^a	10.2 9.7 ^a	12	8	12 ^a
Stockholm	—	8	8	8.3	8	—	8
Gothenburg	—	7	7	6.2	7	—	7
U.K. Trial	—	10 ^b	7 10 ^c	—	—	7 ^c	10 ^b
National Breast Screening Study (NBSS)	—	7	7	—	—	—	7
Nijmegen	7	—	8	—	—	6	—
Florence	7	—	10	—	—	7	—

^aData for the Swedish Two-County Study are included in the review separately for each of the counties (i.e., Kopparberg and Östergötland). ^bOnly the results of the randomized trial in Edinburgh from the more comprehensive U.K. Trial are included in the review. ^cThe results from the U.K. Trial in Guilford are reported separately from the results in Edinburgh.

age 50 (Fletcher et al., 1993). The ACS still recommends that screening begin at age 40 (Hamblin & Connor, 1995, Jenks, 1993). All guidelines suggest early screening for women with a family history of breast cancer or with other identified risk factors for breast cancer. Further, diagnostic mammography is recommended for women with any sign or symptom of breast disease.

After reviewing the evidence, the U.S. Preventive services Task force, the American College of Physicians, and the HEDIS group all recommend that routine screening mammography be started at age 50 (Eastman, 1996, Mckennett, 1993). This is consistent with national policies in countries that have formed formal policies on screening such as the United Kingdom, The Netherlands, Australia, and Canada (Kaplan, 1993). In the United States, the inclusion of screening mammography has been made a major issue in the development of basic benefit packages. The initial proposal by the Clinton Administration suggested that mammography screening begin at age 50. The Administration's position stimulated significant protest by members of congress. Within weeks of the first announcement of the plan, Representative Nita M. Lowey, a member of the Congressional Women's Caucus, challenged the plan because there would be a copayment for women between ages 40 and 49. Lowey insisted on adherence to the ACS guidelines, suggesting that waiting until age 50 would place millions of women at risk. As a result, the Administration modified their proposal to include screening for women between ages 40 and 49. This recommendation is inconsistent with reviews by the NCI, the American College of Physicians, and virtually every international group that has considered the currently available data (Fletcher et al., 1993).

OPPORTUNITY COSTS IN TARGETED MAMMOGRAPHY

It is commonly argued that, despite the lack of evidence or efficacy of screening mammography for younger women, this procedure provides essentially no harm. Because mammography is probably not dangerous, why not include it in any basic benefit package? Analyses by Eddy (1989) do raise significant questions about the regular use of mammography. For example, a woman between ages 35 and 50 who obtains yearly screening mammography has either little or no probability of benefiting from the screening. However, in about one third of these women, findings will emerge that will require an additional workup, including biopsy. These workups are not without consequence because they cause significant anxiety. For instance, Lerman et al. (1991) found that women with suspicious abnormal mammograms tend to worry more about breast cancer and that these worries affect their moods and daily functioning. Moreover, these concerns persisted even after these women had received additional tests ruling out cancer (Lerman et al., 1991).

Some of the implications of screening policies have recently been evaluated. Eddy (1994) used data from the Kaiser-Permanente Medical Group of Southern

California. Currently, this HMO performs about 300,000 mammograms each year. About half of these mammograms are completed on women between the ages 50 and 75 years, and about 45% are done for women less than 50 years. The remaining 5% are done for women who are older than 75 years. Among the population of women that Kaiser serves, mammograms are given to about 22% of the women between the ages of 30 and 40 years, 60% between 40 and 50 years, and 69% between 50 to 75 years. In addition, Kaiser screens about 57% of the women between 75 and 85 years.

Using computer simulation, Eddy estimated that the current policy will prevent approximately 909 women from dying of breast cancer by the year 2010 at a cost of \$707 million. There are alternative uses of the mammogram budget. One policy might be to strongly discourage the use of mammography for women less than 50 years and greater than 75 years. Instead, the policy might aggressively recruit women for mammography between ages 50 and 75 years (Eddy, 1994). In the 1990 National Health Interview Survey, less than 40% of women over 50 report screening mammography in the last year. The most important reasons for not having a mammogram was lack of knowledge and lack of physician's recommendation. Previous studies have shown that women who do not have screening mammograms are more likely to believe that the procedure is unnecessary in the absence of symptoms than women who are screened (Breen & Kessler, 1994). An aggressive education program might significantly increase use of mammography in this group. Eddy (1994) estimated that if this program were successful in attracting 95% of the women in the 50-to-75-year age group, the number of breast cancer deaths prevented would increase to 1,206 from 909 (a net increase in 297 lives). Further, the program would cost \$210 million less than the current program. In other words, a cost-saving maneuver might result in about a 33% reduction in breast cancer deaths.

The meta-analysis may be challenged because it includes studies with a range of results. It might be argued that this important issue should be decided on the most optimistic results rather than the average results reported in the literature. To date, the most positive results have been reported by Tabár, Fagerberg, Day, and Duffy (1992) from the Swedish Two-County Program. Focusing exclusively on the Tabár results, Kattlove et al. (Kattlove, Liberati, Keeler, & Brook, 1995) performed a cost-effectiveness analysis for screening and treatment of early breast cancer. They considered the effect of screening on disease-free survival and health-related quality of life. In addition, they evaluated the costs of screening in relation to other alternatives. Even using the most optimistic assumption about the benefits of screening for breast cancer, the Kattlove analysis also concludes that basic benefit packages should not include screening of premenopausal women.

Cost-utility analyses consider the cost to produce the equivalent of a year of life. Based on the overall analysis of all studies, these equations will almost always suggest that screening mammography for women less than age 50 should not be part of a basic benefit package because there is essentially no benefit. With the

denominator of the ratio as zero, the cost-effectiveness ratio will be large to infinite. Using the more optimistic estimates of the cost-effectiveness of screening it is possible to get a cost-effectiveness ratio. The Kattlove (Kattlove, Liberati, Keeler, & Brook, 1995) analysis shows that the cost to save one potential life within 10 years by screening premenopausal women is \$1,480,000. The equivalent cost to save a life for women greater than 50 years is \$183,000. To place this in perspective, the cost to save the equivalent of one life with hypertension screening is about \$20,000.

ALTERNATIVE OPPORTUNITIES FOR MAMMOGRAPHY FUNDS

The real consequences of screening all women may accrue to the pool of women for whom health care services are not available. An estimated 17% of the U.S. population does not have health insurance. Today, public programs cannot afford to support basic services for large numbers of women. In part, this results because public funds have been used to support basic services for other groups. We suggest that more prudent use of expensive services will free resources that could be used by other women who are seriously in need of basic health care.

There are many potential alternative uses of the funds. For example, in a program like Kaiser of Southern California, restriction of mammography to women between the ages of 50 to 74 would save about \$300 million (Eddy, 1994). What could be done with the savings? It is important to emphasize that many programs are not currently available within systems such as Kaiser. For example, Eddy's analysis estimates that antismoking education programs for pregnant women may add 3,700 years of life that would have been lost to tobacco-related diseases. Although there has been an overall decrease in tobacco use, smoking prevalence rates among pregnant women have declined very little since 1969. Dissemination of tested smoking cessation methods may benefit between 12,900 and 155,000 pregnant smokers each year (Windsor et al., 1993). Systematic clinical trials have shown that a low-cost self-help smoking cessation program can significantly reduce smoking during pregnancy (Ershoff, Mullen, & Quinn, 1989). Smoking during pregnancy is a major cause of premature birth, and reductions in smoking during pregnancy produces significant health benefits as well as cost savings.

Other areas in which prevention programs could improve health status and prevent premature death include prenatal care and programs to reduce risk factors for cardiovascular diseases that kill over 300 per 100,000 American women each year and remain the most common cause death for both men and women in the United States (U.S. Department of Health and Human Services, 1991). Moreover, additional programs to improve the rate of screening mammography among women 50 years old or older could be implemented. These programs could be funded from

the savings that would accrue from more effective mammography screening. The issue is not to save money, only to use it more wisely.

WHY EVALUATE MAMMOGRAPHY?

This article focuses on mammography as a case study. Mammography screening was chosen as a case example because it has attracted such intense public interest. As noted earlier, there is a legitimate scientific controversy as to whether or not mammography for premenopausal women produces health benefits. In addition to the scientific debate, mammography screening has attracted attention as a political and discrimination issue. We do not want to give the impression that the issues of opportunity costs are unique to mammography. In fact, many other screening tests could be evaluated in the same light. For example, there is significant controversy concerning screening for prostate cancer using prostate-specific antigen. Over the last decade the incidence of prostate cancer has increased dramatically, whereas the mortality rate has remained about the same (Merrill, 1996). It has been recommended that all men 50 years of age or older be screened for this malignancy. However, cost analysis suggests that the screening would cost between \$12 and \$25 billion per year and could consume 5% of the entire health care budget. Because health care budgets are contracting rather than expanding, following the recommendation for universal prostate cancer screening would mean that other health services would need to be cut. Current analyses demonstrate that prostate cancer screening for older men does not increase life expectancy (Waterbor & Bueschen, 1995). Further, identification of indolent prostate cancer may initiate a series of treatment decisions that ultimately do not extend life expectancy and may reduce quality of life. Results from some analyses show a net reduction in life expectancy when quality of life adjustments are made (Krahn et al., 1994). As a result, we have joined others who have argued against prostate cancer screening for men (Kaplan, 1996).

SUMMARY AND CONCLUSIONS

Problems in health care are complex. Consumers have assumed that the more health care they receive, the more benefit they achieve. However, the increases in health care costs threaten consumers' willingness to pay and huge costs may have significant negative impacts on the economic viability of the United States. As a result, it is important to enact policies that produce the most health for the most people.

Mammography screening clearly provides benefits for women over age 50 but the benefits for women less than age 50 are difficult to demonstrate. A society can

choose to enact any policy. The option to have mammography screening for all women is one alternative. However, there are consequences. Enacting this policy will either increase the health care costs that will be paid by consumers or taxpayers, or require that we disregard the opportunities to offer other services. Screening programs targeted by age have the potential to produce more health at a lower cost. The savings may be used to support other effective services for women or for other citizens.

ACKNOWLEDGMENT

This article was supported by Grant U01-52948 from the National Cancer Institute, National Institutes of Health.

REFERENCES¹

- American Cancer Society. (1994). *American Cancer Society position-Clinton Health Reform Plan*. New York: Author.
- *Andersson, I., Aspegren, K., Janzon, L., Landberg, T., Lindholm, K., Linell, F., Ljungberg, T., Ranstam, J., & Sigfússon, B. (1988). Mammographic screening and mortality from breast cancer: The Malmö mammographic screening trial. *British Medical Journal*, 297, 943-948.
- *Aron, J. L., & Prorok, P. C. (1986). An analysis of the mortality effect in a breast cancer screening study. *International Journal of Epidemiology*, 15, 36-43.
- *Baines, C. J., McFarlane, D. V., & Miller, A. B. (1988). Sensitivity and specificity of first screen mammography in 15 NBSS centres. *Canadian Association of Radiologists Journal*, 39, 273-276.
- *Baines, C. J., McFarlane, D. V., & Miller, A. B. (1990). The role of the reference radiologist: Estimates of interobserver agreement and potential delay in cancer detection in the National Breast Screening Study. *Investigative Radiology*, 25, 971-976.
- *Baines, C. J., Miller, A. B., Kopans, D. B., Moskowitz, M., Sanders, D. E., Sickles, E. A., To, T., & Wall, C. (1990). Canadian National Breast Screening Study: Assessment of technical quality by external review. *American Journal of Roentgenology*, 155, 743-747.
- Bashshur, R. L., Homan, R. K., & Smith, D. G. (1994). Beyond the uninsured: Problems in access to care. *Medical Care*, 32, 409-419.
- Black, W. C., & Welch, H. G. (1993). Advances in diagnostic imaging and overestimations of disease prevalence and the benefits of therapy. *New England Journal of Medicine*, 328, 1237-1243.
- Breen, N., & Kessler, L. (1994). Changes in the use of screening mammography: Evidence from the 1987 and 1990 National Health Interview Surveys. *American Journal of Public Health*, 84, 62-67.
- Breslow, N. E., & Day, N. E. (1980). *Statistical methods in cancer research*. Lyon, France: International Agency for Research on Cancer (IARC) Scientific Publications.
- Carter, R., Glasziou, P., van Oortmarssen, G., de Koning, H. J., Stevenson, C., Salkeld, G., & Boer, R. (1993). Cost-effectiveness of mammographic screening in Australia. *Australian Journal of Public Health*, 17, 42-50.

¹References marked with an asterisk indicate studies included in the meta-analysis.

- Collins, K. S., Romero, V. L., Drummond, C., & Shannon, I. (1993). Financing women's health care. *Women's Health Issues*, 3, 79-85.
- *Dales, L. G., Friedman, G. D., & Collen, M. F. (1979). Evaluating periodic multiphasic health checkups: A controlled trial. *Journal of Chronic Diseases*, 32, 385-404.
- de Koning, H. J., Boer, R., Warmerdam, P. G., Beemsterboer, P. M. M., & van der Maas, P. J. (1995). Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials. *Journal of the National Cancer Institute*, 87, 1217-1223.
- Duerksen, S. (1994, April 27). Are they worth it? *The San Diego Union*, pp. E1, E6.
- Eastman, P. (1996). Task force issues new screening guidelines. *Journal of the National Cancer Institute*, 88, 74-76.
- Eddy, D. M. (1989). Screening for breast cancer. *Annals of Internal Medicine*, 111, 389-399.
- Eddy, D. M. (1994). Rationing resources while improving quality. How to get more for less. *Journal of the American Medical Association*, 272, 817-824.
- Eddy, D. M., Hasselblad, V., McGivney, W., & Hendee, W. (1988). The value of mammography screening in women under age 50 years. *Journal of the American Medical Association*, 259, 1512-1519.
- Ershoff, D. H., Mullen, P. D., & Quinn, V. P. (1989). A randomized trial of a serialized self-help smoking cessation program for pregnant women in an HMO. *American Journal of Public Health*, 79, 182-187.
- *Fagerberg, C. J. G., & Tabár, L. (1988). The results of periodic one-view mammography screening in a randomized, controlled trial in Sweden. In N. E. Day & A. B. Miller (Eds.), *Screening for breast cancer* (pp. 33-44). Toronto, Canada: Hans Huber.
- Fleiss, J. L. (1981). *Statistical methods for rates and proportions*. New York: Wiley.
- Fletcher, S. W., Black, W., Harris, R., Rimer, B. K., & Shapiro, S. (1993). Report of the International Workshop of Screening for Breast Cancer. *Journal of the National Cancer Institute*, 85, 1644-1656.
- Hamblin, J., & Connor, P. D. (1995). An overview of the American Cancer Society screening guidelines. *Journal of the Tennessee Medical Association*, 88, 10-16.
- Health Care Financing Administration. (1994). *Office of the Actuary Report*. Washington, DC: U.S. Government Printing Office.
- Jenks, S. (1993). ACS keeps mammography guidelines for women under 50 [news]. *Journal of the National Cancer Institute*, 85, 348-349.
- Kaplan, R. M. (1993). *Hippocratic Predicament: Affordability, access, and accountability in health care*. San Diego, CA: Academic.
- Kaplan, R. M. (1996). Measuring health outcome for resource allocation. In R. L. Gluckauf, R. Frank, G. R. Bond, & J. H. McGrew (Eds.), *Psychological practice in a changing health care system: Issues and directions* (pp. 101-133). New York: Springer.
- Kattlove, H., Liberati, A., Keeler, E., & Brook, R. (1995). Benefits and costs of screening and treatment for early breast cancer: Development of a basic benefit package. *Journal of the American Medical Association*, 273, 142-148.
- Kerlikowske, K., Grady, D., Rubin, S. M., Sandrock, C., & Ernster, V. L. (1995). Efficacy of screening mammography. A meta-analysis. *Journal of the American Medical Association*, 273, 149-154.
- Kopans, D. B., Halpern, E., & Hulka, C. A. (1994). Statistical power in breast cancer screening trials and mortality reduction among women 40-49 years of age with particular emphasis on the National Breast Screening Study of Canada. *Cancer*, 74, 1196-1203.
- Krahn, M. D., Mahoney, J. E., Eckman, M. H., Trachtenberg, J., Pauker, S. G., & Detsky, A. S. (1994). Screening for prostate cancer. A decision analytic view. *Journal of the American Medical Association*, 272, 773-780.
- Lerman, C., Trock, B., Rimer, B. K., Jepson, C., Brody, D., & Boyce, A. (1991). Psychological side effects of breast cancer screening. *Health Psychology*, 10, 259-267.

- Levit, K. R., Olin, G. L., & Letsch, S. W. (1992). Americans' health insurance coverage, 1980-1991. *Health Care Financing Review*, 14, 31-57.
- McKennett, M. (1993). Changing guidelines for screening mammography. *The Western Journal of Medicine*, 158, 520-521.
- Merrill, R. M. (1996). Trends in prostate cancer incidence and surgery. *Journal of the National Cancer Institute*, 88, 403.
- Merritt, R. E., & Demkovich, L. (1991). Exhausting alternatives: There are disadvantages to relying on the states, but many feel forced to take aggressive steps to cover the uninsured. *Health Management Quarterly*, 13, 11-15.
- Mettlin, C., & Smart, C. R. (1994). Breast cancer detection guidelines for women aged 40 to 49 years: Rationale for the American Cancer Society reaffirmation of recommendations. *CA: A Cancer Journal for Clinicians*, 44, 248-255.
- *Miller, A. B. (1988). The Canadian National Breast Screening Study. In N. E. Day & A. B. Miller (Eds.), *Screening for breast cancer* (pp. 51-58). Toronto, Canada: Hans Huber.
- *Miller, A. B., Baines, C. J., To, T., & Wall, C. (1992a). Canadian National Breast Screening Study: Vol. 1. Breast cancer detection and death rates among women aged 40 to 49 years. *Canadian Medical Association Journal*, 147, 1459-1476.
- *Miller, A. B., Baines, C. J., To, T., & Wall, C. (1992b). Canadian National Breast Screening Study: Vol. 2. Breast cancer detection and death rates among women aged 50 to 59 years. *Canadian Medical Association Journal*, 147, 1477-1488.
- Miller, A. B., Chamberlain, J., Day, N. E., Hakama, M., & Prorok, P. C. (1990). Report on workshop of the UICC project evaluation of screening for cancer. *International Journal of Cancer*, 46, 761-769.
- National Center for Health Statistics. (1994). *Health, United States, 1993*. Hyattsville, MD: Public Health Service.
- Nystrom, L., Rutqvist, L. E., Wall, S., Lindgren, A., Lindqvist, M., Ryden, S., Andersson, I., Bjurstam, N., Fagerberg, G., Frisell, J., Tabár, L., & Larsson, L. (1993). Breast cancer screening with mammography: Overview of Swedish randomized trials. *Lancet*, 341, 973-978.
- *Palli, D., Buiatti, M., Buiatti, E., Carli, S., Ciatto, S., Toscani, L., & Maltoni, G. (1986). A case-control study of the efficacy of a nonrandomized breast cancer screening program in Florence (Italy). *International Journal of Cancer*, 38, 501-504.
- Piani, A., & Schoenborn, C. (1993). *Health promotion and disease prevention: United States, 1990*. Washington, DC: National Center for Health Statistics.
- *Roberts, M. M., Alexander, F. E., Anderson, T. J., Chetty, U., Donnan, P. T., Forrest, P., Hepburn, W., Huggings, A., Kirkpatrick, A. E., Lamb, J., Muir, B. B., & Prescott, R. J. (1990). Edinburgh trial of screening for breast cancer: Mortality at seven years. *Lancet*, 335, 241-246.
- Rodgers, A. (1990). The U.K. breast cancer screening programme: An expensive mistake. *Journal of Public Health Medicine*, 12, 197-204.
- Shapiro, S. (1994). Screening: Assessment of current studies. *Cancer*, 74, 231-238.
- *Shapiro, S., Venet, W., Strax, P., & Venet, L. (1988a). Current results of the breast cancer screening randomized trial: The Health Insurance Plan (HIP) of Greater New York study. In N. E. Day & A. B. Miller (Eds.), *Screening for breast cancer* (pp. 3-15). Toronto, Canada: Hans Huber.
- *Shapiro, S., Venet, W., Strax, P., & Venet, L. (1988b). *Periodic screening for breast cancer: The Health Insurance Plan project and its sequelae, 1963-1986*. Baltimore: Johns Hopkins University Press.
- *Shapiro, S., Venet, W., Strax, P., Venet, L., & Roesser, R. (1982). Ten- to fourteen-year effect of screening on breast cancer mortality. *Journal of the National Cancer Institute*, 69, 349-355.
- Sofaer, S. & Abel, E. (1990). Older women's health and financial vulnerability: Implications of the Medicare benefit structure. *Women and Health*, 16, 47-67.

- Tabár, L., Fagerberg, G., Day, N. E., & Duffy, S. W. (1992). The Swedish two-county trial of mammographic screening for breast cancer: Recent results on mortality and tumor characteristics. *Pathological Biology*, 39, 846.
- *Tabár, L., Fagerberg, G., Day, N. E., & Holmberg, L. (1987). What is the optimum interval between mammographic screening examinations? An analysis based on the latest results of the Swedish two-county breast cancer screening trial. *The British Journal of Cancer*, 55, 547-551.
- *Tabár, L., Fagerberg, G., Gad, A., Baldetorp, L., Holmberg, L. H., Gröntoft, O., Ljungquist, U., Lundström, B., & Månson, J. C. (1985). Reduction in mortality from breast cancer after mass screening with mammography. *Lancet*, 1, 829-832.
- *U.K. Trial of Early Detection of Breast Cancer Group. (1988a). First results on mortality reduction in the U.K. trial of early detection of breast cancer. *Lancet*, 2, 411-416.
- *U.K. Trial of Early Detection of Breast Cancer Group. (1988b). Progress report of the U.K. trial of early detection of breast cancer. In N. E. Day & A. B. Miller (Eds.), *Screening for breast cancer* (pp. 45-49). Toronto, Canada: Hans Huber.
- U.S. Department of Health and Human Services. (1991). *Healthy people 2000: National health promotion and disease prevention objectives*. Washington, DC: Author.
- *Verbeek, A. L. M., Hendriks, J. H. C. L., Holland, R., Mravunac, M., & Sturmans, F. (1985). Mammographic screening and breast cancer mortality: Age-specific effects in Nijmegen project, 1975-1982. *Lancet*, 1, 865-866.
- *Verbeek, A. L. M., Hendriks, J. H. C. L., Holland, R., Mravunac, M., Sturmans, F., & Day, N. E. (1984). Reduction of breast cancer mortality through mass screening with modern mammography. *Lancet*, 1, 1222-1224.
- *Verbeek, A. L. M., Straatman, H., & Hendriks, J. H. C. L. (1988). Sensitivity of mammography in Nijmegen women under age 50: Some trials with the Eddy model. In N. E. Day & A. B. Miller (Eds.), *Screening for breast cancer* (pp. 29-32). Toronto, Canada: Hans Huber.
- Waterbor, J. W., & Bucschen, A. J. (1995). Prostate cancer screening (United States). *Cancer Causes Control*, 6, 267-274.
- White, C. C., Tolsma, D. D., Haynes, S. G., & McGee, D. (1987). Cardiovascular disease. In R. W. Amler & H. B. Dull (Eds.), *Closing the gap: The burden of unnecessary illness* (pp. 43-54). New York: Oxford University Press.
- Windsor, R. A., Li, C. Q., Lowe, J. B., Perkins, L. L., Ershoff, D. H., & Glynn, T. (1993). The dissemination of smoking cessation methods for pregnant women: Achieving the year 2000 objectives. *American Journal of Public Health*, 83, 173-178.