The New England Journal of Medicine

TEN-YEAR RISK OF FALSE POSITIVE SCREENING MAMMOGRAMS AND CLINICAL BREAST EXAMINATIONS

JOANN G. ELMORE, M.D., M.P.H., MARY B. BARTON, M.D., M.P.P., VICTORIA M. MOCERI, PH.C., SARAH POLK, B.A., PHILIP J. ARENA, M.D., AND SUZANNE W. FLETCHER, M.D.

ABSTRACT

Background   The cumulative risk of a false positive result of a breast-cancer screening test is unknown.

Methods   We performed a 10-year retrospective cohort study of breast-cancer screening and diagnostic evaluations among 2400 women who were 40 to 69 years old at study entry. Mammograms or clinical breast examinations that were interpreted as indeterminate, aroused a suspicion of cancer, or prompted recommendations for additional workup in women in whom breast cancer was not diagnosed within the next year were considered to be false positive tests.

Results   A total of 9762 screening mammograms and 10,905 screening clinical breast examinations were performed, for a median of 4 mammograms and 5 clinical breast examinations per woman over the 10-year period. Of the women who were screened, 23.8 percent had at least one false positive mammogram, 13.4 percent had at least one false positive breast examination, and 31.7 percent had at least one false positive result for either test. The estimated cumulative risk of a false positive result was 49.1 percent (95 percent confidence interval, 40.3 to 64.1 percent) after 10 mammograms and 22.3 percent (95 percent confidence interval, 19.2 to 27.5 percent) after 10 clinical breast examinations. The false positive tests led to 870 outpatient appointments, 539 diagnostic mammograms, 186 ultrasound examinations, 188 biopsies, and 1 hospitalization. We estimate that among women who do not have breast cancer, 18.6 percent (95 percent confidence interval, 9.8 to 41.2 percent) will undergo a biopsy after 10 mammograms, and 6.2 percent (95 percent confidence interval, 3.7 to 11.2 percent) after 10 clinical breast examinations. For every $100 spent for screening, an additional $33 was spent to evaluate the false positive results.

Conclusions   Over 10 years, one third of the women screened had abnormal test results requiring additional evaluation, even though no breast cancer was present. Techniques are needed to decrease false positive results while maintaining high sensitivity. Physicians should educate women about the risk of a false positive result of a screening test for breast cancer. (N Engl J Med 1998;338:1089-96.)

©1998, Massachusetts Medical Society.

MAMMOGRAPHY and clinical breast examination are the two principal means of screening for breast cancer.[1] The effectiveness of breast-cancer screening has been well documented in eight randomized, controlled trials,[2] but there has been less attention to its accuracy in community settings and to the consequences of a false positive result. A national review of community mammography facilities in the United States found that 1 of every 10 screening mammograms gave a false positive result.[3] Equivalent information for clinical breast examination is not available.

If a woman undergoes annual screening beginning at the age of 40, she will have had 60 opportunities for a false positive result by the age of 70, with 30 mammograms and 30 clinical breast examinations. The cumulative lifetime risk of her having a result from a screening test that requires further workup, even though no breast cancer is present, is not known. An estimate of 25 percent has been given for the cumulative risk of a false positive result after 10 mammograms and 10 clinical breast examinations.[4] It is important to determine the cumulative risk of false positive tests, because women are advised to have breast-cancer screening every one to two years over several decades of their lifetimes, and false positive results can provoke anxiety, increase costs, and cause morbidity.[5-13]

Using the computerized clinical records of a health maintenance organization (HMO) for a group of women over a 10-year period, we determined the

cumulative risk of a false positive result of breast-cancer screening, the number and type of subsequent diagnostic workups resulting from the false positive results, and the costs of the false positive results. The HMO we studied has long encouraged women who are 40 or older to undergo routine breast-cancer screening. By studying the medical records, we ascertained the 10-year cumulative rates of false positive results for both mammography and clinical breast examination. We then determined the number of diagnostic examinations generated by the false positive results and estimated their costs.

METHODS

Setting

This retrospective cohort study was conducted at 11 staff-model health centers of Harvard Pilgrim Health Care, a large HMO in New England. The health centers serve nearly 300,000 adults in and around Boston. Although the majority of members belong to the HMO through an employer or a spouse's employer, approximately 5 percent are enrolled through the state Medicaid program for low-income persons. This study was approved by the Human Studies Committee of Harvard Pilgrim Health Care and the institutional review board of the University of Washington School of Medicine.

Breast-cancer screening for the members of Harvard Pilgrim Health Care is encouraged by internal guidelines and a computerized reminder system that prompts health care providers to perform clinical breast examinations and order mammograms for screening. Beginning in 1984 and throughout the study period, the HMO recommended that all women from 40 to 49 years of age be screened with mammography every two years and that women 50 years of age or older be screened annually. Most of the women were referred to sites outside the HMO for mammography, including local community and academic radiology centers. All of the radiologists who read the mammograms were board certified and worked in groups that contracted with the HMO.

Study Population

All 14,382 women who were members of the HMO and who were between 40 and 69 years of age on July 1, 1983, were potentially eligible for the study. Women were excluded for the following reasons: a lapse in enrollment in the HMO between July 1, 1983, and June 30, 1995 (8816 women); health coverage from a source other than Harvard Pilgrim Health Care or from a noncomputerized HMO center during the study period (1093 women); and a history of breast cancer or a prophylactic mastectomy or breast implants before July 1, 1983, (146 women) or a prophylactic mastectomy or breast implants during the study period (8 women). From the cohort of 4319 remaining eligible subjects, a random sample was chosen, consisting of 1200 women 40 to 49 years of age, 600 women 50 to 59 years of age, and 600 women 60 to 69 years of age, for a total sample of 2400 women.

Review of Medical Records

Harvard Pilgrim Health Care keeps computerized records of patients' visits for ambulatory care services.[14,15] Data on demographic characteristics, risk factors for breast cancer, screening clinical breast examinations, screening mammography, diagnostic testing performed as a result of breast-cancer screening, and breast cancers diagnosed were abstracted from these records onto standardized forms. If information was missing or clarification was needed, the original test reports were reviewed. Household income was estimated by matching each patient's address on December 1, 1995, with census-tract data.[16]

For development and training purposes, the data on the first

581 patients were extracted independently by a research assistant and one of the authors. To ensure quality, the data on a randomly selected sample of 5 percent of the next 1443 patients were extracted by a second person who was unaware of the first research assistant's results. Inconsistencies between the reviewers were found for 175 (0.9 percent) of the 19,407 variables reviewed among the 1443 patients. Forty of these inconsistencies (0.2 percent) were considered clinically important because they concerned the specific reasons for ordering the test. In view of this low rate of discrepancy, a 5 percent sample of the remaining 376 charts was reviewed by two different persons, who were not blinded to each other's results. In this final review, 1 inconsistency was noted among 4093 variables (0.02 percent). All inconsistencies were resolved by consensus.

Information was recorded for each appointment at which screening occurred. The diagnostic impressions for both mammography and clinical breast examination were classified as normal; abnormal and probably benign; abnormal, indeterminate; or abnormal and arousing a suspicion of cancer. Recommendations for additional testing were recorded, including diagnostic mammography within the next 12 months or second-opinion review of the screening films, ultrasound examination, physical examination (by the primary care provider or a surgeon), and biopsy (including fine-needle aspiration, open, and core biopsies). Information was recorded on all diagnostic procedures and follow-up visits resulting from positive breast-cancer screening tests.

The computerized medical records, the HMO's tumor registry, and, if needed, the original paper copy of the test results were searched to identify incident cases of breast cancer. In addition, the study participants' records for two years after completion of the study (from July 1, 1993, to June 30, 1995) were searched to be certain that all breast-cancer cases were identified.

Definitions of Screening Tests and False Positive Results

Mammography or clinical breast examinations performed on asymptomatic women without previously noted abnormalities were classified as screening tests. Mammography or clinical breast examinations performed because of abnormalities previously noted by clinicians or patients were classified as diagnostic tests.

A false positive result was defined in a manner consistent with current recommendations regarding mammography audits[17-19] and reported by other investigators.[3,20-22] A test was classified as positive if the results were indeterminate or aroused a suspicion of cancer, or if there was a recommendation for nonroutine follow-up, including physical examination, diagnostic mammography within the next 12 months, ultrasound examination, or biopsy. A positive test was classified as true positive if breast cancer (invasive or ductal carcinoma in situ) was diagnosed in the patient on the basis of pathological findings within one year of the test, and as false positive otherwise. False positive results were independent of each other.

Analysis and Assessment of Costs

The data were double-entered and verified for computer analysis. Initial comparisons were made with use of the chi-square test for categorical data and Student's t-test or analysis of variance for continuous data. Tests for trend were made with use of the Mantel–Haenszel chi-square statistic with one degree of freedom. These analyses were performed with SAS software.[23]

Estimates of the cumulative risk of having a false positive test are based on a Bayesian version of a product or an estimate of the Kaplan–Meier type, in which screening events (mammography or clinical breast examinations) are used instead of time (see Appendix 1).

A current procedural and technical code[24] was assigned to all workups resulting from the breast-cancer screening. The national Medicare fee schedule[25] and the average HMO payment were used to estimate the average payment. Inpatient care was assigned a payment specific to the diagnosis-related group (see Appendix 2).

RESULTS

Characteristics of the Patients

Most of the 2400 women were white (75 percent); 11 percent were black, 3 percent were of other races, and 11 percent were of unknown race. The median household income was $47,940 (range, $13,230 to $161,710). Eighteen percent of the women had a family history of breast cancer recorded, and 28 percent used estrogen-replacement therapy at some time during the study period.

Frequency of Breast-Cancer Screening

On average, the women underwent screening mammography and a screening clinical breast examination every two years. A small number of women (88 [3.7 percent]) had no documented breast-cancer screening, either by clinical examination or by mammography, during the 10 years of the study.

Over the 10-year period, 9762 screening mammograms were obtained for 2227 women; 173 women (7.2 percent) underwent no screening mammography. For those who had at least one mammogram, the median number of mammograms obtained was four (range, one to nine). The mammograms were read by 93 radiologists at 28 radiology facilities, consisting of 17 community, 7 HMO, and 4 academic sites. Four radiologists each read more than 1000 mammograms, one radiologist read 624 mammograms, and the others each read fewer than 500 mammograms.

Also during these 10 years, 10,905 screening clinical breast examinations were performed on 2245 women; 155 women (6.5 percent) had no screening breast examination. For those who had at least 1 breast examination, the median number of examinations performed was 5 (range, 1 to 16). The breast examinations were performed by 381 health care providers; 9290 were performed by internists, 1385 by registered nurses, nurse practitioners, or physician's assistants, 160 by obstetrician-gynecologists, 50 by surgeons, and 20 by providers with unknown credentials.

Detection of Cancer

Between July 1, 1983, and June 30, 1994, breast cancer was diagnosed in 88 women. The mean age of the women was 59 years (range, 42 to 76). Local disease was present in 67 women, and regional disease in 21. Ductal carcinoma in situ was diagnosed in 15 of the 88 women. In 58 women, the breast cancer was diagnosed as a result of an abnormality first noted on a screening mammogram (50 cancers were diagnosed within 12 months after mammography and 8 after more than 12 months). In 7 women, the cancer was diagnosed as a result of a clinical breast examination (4 cancers were diagnosed within 12 months after the breast examination and 3 after 12 months). In the remaining 23 women, the cancers were diagnosed after the women themselves noted an abnormality and sought medical evaluation.

False Positive Results

False positive results occurred in 6.5 percent of the mammograms and 3.7 percent of the clinical breast examinations (Table 1). Among the women who were screened, 23.8 percent had at least one false positive mammogram and 13.4 percent had at least one false positive breast examination during the 10 years. A false positive test due to either type of screening was noted in 31.7 percent of the women. In the majority of these women, there was only one false positive result; 89 women had two or more false positive mammograms, 72 had two or more false positive breast examinations, and 96 had false positive results on both a mammogram and a breast examination.

The percentage of screening mammograms that were false positive increased from 4.2 percent in the first three years (1983 to 1986) to 6.1 percent (1986 to 1990) and 7.6 percent (1990 to 1993) (P for trend = 0.001); however, the false positive rate varied substantially according to year. The percentage of screening clinical breast examinations with false positive results decreased slightly over the same three periods, from 4.5 percent to 3.4 percent to 3.5 percent (P = 0.08).

The false positive rates were higher for younger women than for older women (Table 2). The percentage of mammograms that were false positive decreased from 7.8 percent for women 40 to 49 years of age to 4.4 percent for women 70 to 79 years of age (P = 0.001). The false positive rate for clinical breast examination was highest for women 40 to 49 years of age (6.0 percent) and decreased to 2.2 percent for women 70 to 79 years of age (P = 0.001).

Cumulative Risk of a False Positive Result

The risk of a woman's ever having a false positive result increased as she underwent more screening.

TABLE 1. FALSE POSITIVE BREAST-CANCER SCREENING TESTS OVER A 10-YEAR PERIOD.

VARIABLE	MAMMOGRAPHY	CLINICAL BREAST EXAMINATION	EITHER TEST
		number (percent)	
Screening tests	9762	10,905	20,667
False positive tests	631 (6.5)	402 (3.7)	1,033 (5.0)
Women screened	2227	2,245	2,312
Women with at least one false positive test	530 (23.8)	300 (13.4)	734 (31.7)

TABLE 2. RATE OF FALSE POSITIVE BREAST-CANCER SCREENING TESTS ACCORDING TO THE AGE OF THE WOMAN AT THE TIME OF SCREENING.

AGE (YR)	MAMMOGRAPHY*	CLINICAL BREAST EXAMINATION*
	no. of false positives/total no. of tests (%)	
40–49	156/1996 (7.8)	162/2715 (6.0)
50–59	269/3639 (7.4)	136/3848 (3.5)
60–69	148/2810 (5.3)	74/2966 (2.5)
70–79	58/1317 (4.4)	30/1376 (2.2)
Total	631/9762 (6.5)	402/10,905 (3.7)

*P for trend = 0.001.

The estimated cumulative risk of having at least one false positive result after 10 screenings was 49.1 percent (95 percent confidence interval, 40.3 to 64.1 percent) for mammograms and 22.3 percent (95 percent confidence interval, 19.2 to 27.5 percent) for clinical breast examinations (Fig. 1 and 2).

The cumulative risk of a false positive result was higher for women 40 to 49 years of age at the time of the test than for those 50 to 79 years of age, for both mammography and clinical breast examination. For example, for women 40 to 49 years of age the cumulative risk after 5 screening mammograms was 30.3 percent (95 percent confidence interval, 23.9 to 39.4 percent) as compared with 23.8 percent (95 percent confidence interval, 21.6 to 26.4 percent) among women who were 50 or older, and after 10 screening mammograms it was 56.2 percent (95 percent confidence interval, 39.5 to 75.8 percent) as compared with 47.3 percent (95 percent confidence interval, 37.8 to 63.0 percent). For these two groups of women, the cumulative risk of a false positive clinical breast examination after 5 examinations was 20.8 percent (95 percent confidence interval, 16.7 to 25.4 percent) in the women 40 to 49 years old, as compared with 11.1 percent (95 percent confidence interval, 9.4 to 13.2 percent) in the older women, and after 10 examinations it was 34.1 percent (95 percent confidence interval, 22.8 to 59.0 percent) as compared with 18.7 percent (95 percent confidence interval, 14.8 to 26.1 percent).

Diagnostic Evaluations Performed

The relevant diagnostic evaluations performed within one year of the false positive result are shown in Table 3. False positive mammograms led to more outpatient visits, diagnostic imaging examinations, and biopsies than false positive clinical breast examinations. In one patient, cellulitis requiring hospitalization for surgical débridement and intravenous antibiotic therapy developed after a biopsy prompted by a false positive mammogram. In addition to the workups shown in Table 3, we found documentation of 260 telephone calls to patients, 32 letters to patients, and 64 second opinions obtained from radiologists after false positive mammograms. These additional events were less common after false positive clinical breast examinations, which resulted in 23 telephone calls, 5 letters, and 9 second opinions from radiologists.

One hundred fourteen women (5.1 percent) had at least one documented invasive procedure (open, core, or fine-needle aspiration biopsy) within one year as a result of a false positive mammogram, after an average of four screening mammograms. False positive clinical breast examinations led to biopsy in 45 women (2.0 percent), after an average of five examinations.

A woman's estimated cumulative risk of having at least one biopsy (open, core, or fine-needle aspiration biopsy) as a result of a false positive test also increased with repeated screenings. The risk was 6.2 percent (95 percent confidence interval, 5.1 to 7.3 percent) after 5 screening mammograms and 18.6 percent (95 percent confidence interval, 9.8 to 41.2 percent) after 10 screening mammograms. For clinical breast examinations, the risk was 2.4 percent (95 percent confidence interval, 1.8 to 3.2 percent) after 5 examinations and 6.2 percent (95 percent confidence interval, 3.7 to 11.2 percent) after 10 examinations.

Cost of Diagnostic Evaluations

The payment allowances for the initial screenings (mammography and clinical breast examinations) were $993,870 according to HMO payment-allowance estimates and $1,042,311 according to Medicare estimates. The payment allowances for the diagnostic workups undertaken as a result of the false positive tests shown in Table 3 were $329,649 according to HMO estimates and $309,755 according to Medicare payment allowances. The payment allowances for workups after false positive mammograms were three times as high as those for workups after false positive clinical breast examinations.

DISCUSSION

On average, the women in this community-based cohort underwent breast-cancer screening every two years. Over a 10-year period, the result in one third of these women required additional evaluation when no breast cancer was present. The risk of a false positive test increased with the number of breast-cancer screening tests, so that by the time a woman had undergone 10 tests, the estimated cumulative risk of at least one false positive mammogram was about 50 percent and the estimated cumulative risk of at least one false positive breast examination was about 25 percent.

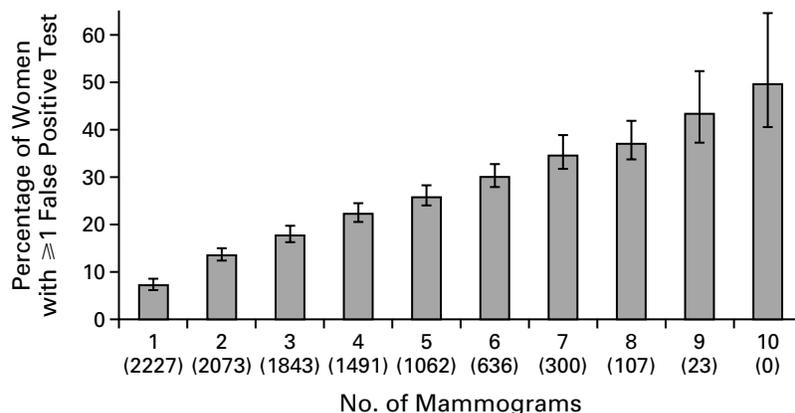The HMO setting provided complete follow-up for each woman, and the computerized medical rec-

Figure 1. A Woman's Estimated Risk of Having at Least One False Positive Screening Mammogram, According to the Total Number of Screening Mammograms Performed.

The numbers in parentheses are the numbers of women with at least that many mammograms. The I bars indicate 95 percent confidence intervals.
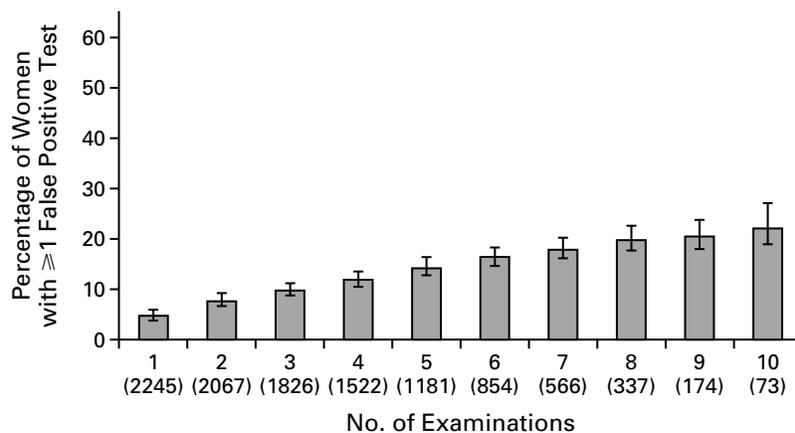

Figure 2. A Woman's Estimated Risk of Having at Least One False Positive Screening Clinical Breast Examination, According to the Total Number of Clinical Breast Examinations Performed.

The numbers in parentheses are the numbers of women with at least that many examinations. The I bars indicate 95 percent confidence intervals.

ords allowed simplified data extraction for a large, community-based cohort. Results in managed-care populations do not necessarily apply to other populations; however, our findings are likely to be applicable to other communities, since 93 different radiologists read the mammograms and 381 different health care providers performed the clinical breast examinations.

The cumulative risk of false positive results that we estimated from community-based data is twice the risk estimated by Eddy in 1989.[4] In addition, our estimates of the cumulative risk of false positive results for screening mammography may be low, because the overall percentage of abnormal screening mammograms in our study was 6.5 percent, whereas the national rate is nearly twice as high.[3]

Our definition of a false positive result is consistent with current recommendations regarding mammography audits.[17-19] We classified as positive results all instances in which screening mammograms were interpreted as indeterminate or as arousing a suspicion of cancer or in which additional workup was recommended. This definition, which has been used by others,[3,20-22,26] may be considered too broad. However, we found that even with this definition of false positive results, some follow-up diagnostic procedures were not counted. For example, five women who did not have cancer had breast biopsies prompted by screening mammograms that were interpreted by the radiologist as "abnormal–benign" but with no specific recommendations made. These cases were not included in our definition of a false positive test.

TABLE 3. DIAGNOSTIC WORKUPS PERFORMED WITHIN
ONE YEAR AFTER THE 631 FALSE POSITIVE
MAMMOGRAMS AND THE 402 FALSE POSITIVE
CLINICAL BREAST EXAMINATIONS.

TYPE OF WORKUP	WORKUPS AFTER FALSE POSITIVE MAMMOGRAMS	WORKUPS AFTER FALSE POSITIVE CLINICAL BREAST EXAMINATIONS
	number	
Outpatient visit		
To nonsurgeon (e.g., primary provider)	162	55
To surgeon	439	214
Diagnostic imaging		
Mammography	384	155
Ultrasonography	176	10
Biopsy*		
Open or core biopsy	100	25
Fine-needle aspiration	28	35
Hospitalization†	1	0

*Five additional biopsies were prompted by a mammo-
graphic reading of "abnormal–benign." Because the reading
did not fit our definition of a false positive test, these biopsies
are not included in the table.

†One patient was hospitalized for 15 days for treatment of
cellulitis that developed after a biopsy.

Some authors have calculated false positive rates by counting as positive only impressions that led to breast biopsy.[27] When this definition was applied to our data, the estimated cumulative risk of having at least one biopsy as a result of a false positive test was 19 percent after 10 mammograms and 6 percent after 10 clinical breast examinations.

We advocate a broad definition of false positive tests, because in formulating policy for the delivery of health care, it is important to determine the type and cost of all follow-up procedures, not just breast biopsies. In addition, it is increasingly clear that being told of an abnormal mammogram can cause increased anxiety in women for extended periods, regardless of whether a biopsy is performed.[5-13] In the United States, Lerman et al.[10] found that three months after they had false positive results on mammography, 47 percent of women who had highly suspicious readings reported that they had substantial anxiety related to the mammogram, 41 percent reported that they had worries about breast cancer, 26 percent reported that the worry affected their daily mood, and 17 percent reported that it affected their daily function.[10] In Norway, 18 months after screening mammography, 29 percent of women with false positive results reported anxiety about breast cancer, as compared with 13 percent of women with negative results.[7] Two studies in Britain also found that women with false positive mammograms had more anxiety than those with normal mammograms.[5,6]

The standard definition of false positive tests in breast-cancer screening uses a one-year cutoff date,[17-20,27] so that women without a diagnosis of breast cancer in the 12 months after a positive test are counted as having had false positive tests. However, a patient's actual clinical course does not always fit into the one-year period. In our study, eight breast cancers were diagnosed after a series of evaluations that took longer than one year after a positive mammogram. This situation can arise when a radiologist recommends follow-up mammography, which is then repeated for several six-month periods before the diagnosis is made. Similarly, three breast cancers were diagnosed more than 12 months after positive clinical breast examinations. If the cutoff date were changed to two years, the false positive rates we reported would change very little; the number of false positive mammograms would decrease from 631 to 624, and the number of false positive clinical breast examinations would decrease from 402 to 397.

Abnormal mammographic readings are more common in the United States than in other countries (for example, in the United States approximately 11 percent of mammograms are read as abnormal, as compared with 2 to 5 percent in Sweden), whereas the sensitivity is about the same.[2,3] The possibility that radiologists in the United States are interpreting too many mammograms as abnormal should be investigated. Little is known about the accuracy of clinical breast examinations in a community setting, despite the fact that these examinations are recommended for all women over the age of 40.[28]

The cost of working up patients with false positive results in this study was approximately one third the cost of performing the screening. The costs of evaluating women with false positive mammograms in the Stockholm randomized clinical trial were about one fourth of the costs of the initial screening.[26]

If our rates are representative, the number of breast-cancer screenings in the United States in which abnormalities are noted that require additional testing in women who do not have cancer may be substantial. For example, if 32 million American women who are 40 to 79 years old received breast-cancer screening annually for 10 years, 16 million women would have at least one false positive mammogram and 7 million would have at least one false positive clinical breast examination.

This study was a retrospective review, and therefore some missing data were inevitable. However, because we required all patients to be enrolled for at least two years beyond the study period and we searched for data from multiple sources, we are confident we did not misclassify the breast-cancer status of patients with abnormal test results. The sensitivity of breast-cancer screening cannot be calculated from these data, because eligibility for the study was confined to women enrolled continuously in the HMO for 12 years. Women who died during the study pe-

riod (some of whom may have died of advanced breast cancer) were therefore ineligible. Exclusion of these women may have led to an artificial lowering of the stage of the breast cancers noted. It is theoretically possible that women might have quit the health plan because they had false positive breast-cancer screening results; these women would also not have been eligible for the study.

Although much is known about breast-cancer screening from randomized clinical trials and academic settings, little information is available about the effect of repeated breast-cancer screening on women in a community setting. This study indicates that we need to develop ways to reduce the false positive rates of breast-cancer screening and their associated psychological and economic costs. One possibility for reducing the psychological sequelae is to use on-site radiologists to obtain immediate work-ups instead of requiring women to return for follow-up. In the meantime, women should be educated about their chances of having an abnormality noted on breast-cancer screening tests, and health care providers should be trained to deal with positive results when they occur.

APPENDIX 1. THE BAYESIAN MODEL USED TO ESTIMATE THE CUMULATIVE RISK OF A FALSE POSITIVE RESULT

The cumulative risk of a false positive result is estimated by using a Bayesian version of an estimator of the product or Kaplan–Meier type in which the number of screening events (mammography or clinical breast examinations) is used instead of time. For the ith subject in the study, we define W_i as the "time," measured by the number of screenings, until the first false positive result. That is, $W_i = j$ if her first $j-1$ tests were negative and a false positive result occurred at the jth screening. If the ith subject had a total of k_i screenings during the study, with none of them false positive, we denote this by $W_i > k_i$. Then

$$\Pr(W_i = j) = q_j \prod_{l=1}^{j-1}(1-q_l),$$

where $q_j = \Pr$ (false positive at the jth screening, giv-

en no false positive result for the first $j-1$ screenings). Also,

$$\Pr(W_i > j) = \prod_{l=1}^{j}(1-q_l).$$

We define the cumulative risk of a false positive test as a sequence of probabilities. In particular, we let p_j denote the probability of at lease one false positive test within the first j screenings. Then

$$p_j = 1 - \prod_{l=1}^{j}(1-q_l),$$

which increases in j.

For a sample of n women, the likelihood for this model takes the form

$$\prod_{l=1}^{K} q_l^{r_l - s_l}(1-q_l)^{s_l},$$

where K is the largest number of screenings observed for any subject in the study, r_l is the observed number of women with at least l screenings and no false positive result for the first $l-1$ of them, and s_l is the number of women with at least l screenings and no false positive results of these screenings. A priori, we assume that the unknown q_l values for both mammography and clinical breast examination arise independently from a beta(0.6, 9.4) distribution. This distribution reflects the prior knowledge that about 6 percent of tests are false positive, but allows a high degree of imprecision (standard deviation, 7 percent). The resulting estimator (posterior mean) of q_l is

$$\hat{q}_l = (r_l - s_l + 0.6)/(r_l + 10).$$

The resulting estimator of p_j is

$$\hat{p}_j = 1 - \prod_{l=1}^{j}(1-\hat{q}_l) = \prod_{l=1}^{j}(s_l + 9.4)/(r_l + 10).$$

The estimates of p_j can be plotted against j to display the estimated cumulative risk graphically. By using simulation, 95 percent equal-tail-interval estimates may be assigned to each p_j, providing confidence bands for the cumulative-risk curve.

APPENDIX 2. THE 1997 MEDICARE PAYMENT ALLOWANCE AND THE 1995 HMO AVERAGE PAYMENT USED TO ESTIMATE COSTS

The values used to estimate the total costs of breast-cancer screening and the evaluations performed after false positive results are shown below. Current procedural and technical codes (CPT) and, if applicable, ambulatory-surgical-center (ASC) pay-

ment levels were assigned for each procedure. The relative-value units for each CPT code were multiplied by $36. No costs were assessed for the documented telephone calls and letters noted after the false positive tests.

Type of Procedure	Medicare Payment Allowance	Average HMO Payment
	dollars	
Screening		
Screening mammogram (CPT 76092)	63.34	75.00
Screening clinical breast examination (CPT 99213 for Medicare, CPT 99397 for HMO)	38.88	24.00
Diagnostic evaluation		
Outpatient appointment (CPT 99244)	141.12	77.00
Diagnostic mammogram (CPT 76091)	79.92	82.00
Diagnostic ultrasound examination (CPT 76645)	68.40	117.68
Fine-needle aspiration biopsy (CPT 19000, CPT 88170)	131.04	381.00
Core biopsy (CPT 19100, CPT 88305, ASC group 1)	452.40	1,220.00
Open biopsy with wire guidance (CPT 19125, CPT 19290, CPT 88305, ASC group 3)	946.64	1,418.79
Open biopsy with no wire guidance (CPT 19120, CPT 88305, ASC group 3)	864.56	1,072.79
Hospitalization (DRG 263, CPT 99254 for 1 day, CPT 99262 for 10 days)*	9,855.99	9,855.99

*DRG denotes diagnosis-related group.

REFERENCES

1. Trends in cancer screening — United States, 1987 and 1992. MMWR Morb Mortal Wkly Rep 1995;45:57-61.
2. 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.
3. Brown ML, Houn F, Sickles EA, Kessler LG. Screening mammography in community practice: positive predictive value of abnormal findings and yield of follow-up diagnostic procedures. AJR Am J Roentgenol 1995;165:1373-7.
4. Eddy DM. Screening for breast cancer. Ann Intern Med 1989;111:389-99.
5. Bull AR, Campbell MJ. Assessment of the psychological impact of a breast screening programme. Br J Radiol 1991;64:510-5.
6. Ellman R, Angeli N, Christians A, Moss S, Chamberlain J, Maguire P. Psychiatric morbidity associated with screening for breast cancer. Br J Cancer 1989;60:781-4.
7. Gram IT, Lund E, Slenker SE. Quality of life following a false positive mammogram. Br J Cancer 1990;62:1018-22.
8. Gram IT, Slenker SE. Cancer anxiety and attitudes toward mammography among screening attenders, nonattenders, and women never invited. Am J Public Health 1992;82:249-51.
9. Lerman C, Rimer BK, Engstrom PF. Cancer risk notification: psychosocial and ethical implications. J Clin Oncol 1991;9:1275-82.
10. Lerman C, Trock B, Rimer BK, Boyce A, Jepson C, Engstrom PF. Psychological and behavioral implications of abnormal mammograms. Ann Intern Med 1991;114:657-61.
11. Lerman C, Trock B, Rimer BK, Jepson C, Brody D, Boyce A. Psychological side effects of breast cancer screening. Health Psychol 1991;10:259-67.
12. Paskett ED, Rimer BK. Psychosocial effects of abnormal Pap tests and mammograms: a review. J Womens Health 1995;4:73-82.
13. Sutton S, Saidi G, Bickler G, Hunter J. Does routine screening for breast cancer raise anxiety? Results from a three wave prospective study in England. J Epidemiol Community Health 1995;49:413-8.
14. Barnett GO, Justice NS, Somand ME, et al. COSTAR: a computer-based medical information system for ambulatory care. Proc I E E E 1979;67:1226-37.
15. Barnett GO. The application of computer-based medical-record systems in ambulatory practice. N Engl J Med 1984;310:1643-50.
16. MapMarker 2.0: desktop mapping software. Troy, N.Y.: MapInfo, 1995 (software).
17. Linver MN, Osuch JR, Brenner RJ, Smith RA. The mammography audit: a primer for the Mammography Quality Standards Act (MQSA). AJR Am J Roentgenol 1995;165:19-25.
18. Sickles EA. Quality assurance: how to audit your own mammography practice. Radiol Clin North Am 1992;30:265-75.
19. Bassett LW, Hendrick RE, Bassford TL, et al. Quality determinants of mammography. Clinical practice guideline no. 13. Rockville, Md.: Agency for Health Care Policy and Research, 1994. (AHCPR publication no. 95-0632.)
20. 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.
21. Bird RE. Low-cost screening mammography: report on finances and review of 21,716 consecutive cases. Radiology 1989;171:87-90.
22. Robertson CL. A private breast imaging practice: medical audit of 25,788 screenings and 1,077 diagnostic examinations. Radiology 1993;187:75-9.
23. SAS version 6.11. Cary, N.C.: SAS Institute, 1996 (software).
24. Kirschner CG, Burkett RC, Kotowicz GM, et al. Physicians' current procedural terminology: CPT '95. Chicago: American Medical Association, 1994.
25. Health Care Financing Administration. Medicare program: revisions to payment policies and five-year review of and adjustment to the relative value units under the physician fee schedule for calendar year 1997. Fed Regist 1996;61(227):59490-724.
26. Lidbrink E, Elfving J, Frisell J, Jonsson E. Neglected aspects of false positive findings of mammography in breast cancer screening: analysis of false positive cases from the Stockholm trial. BMJ 1996;312:273-6.
27. Kopans DB. The positive predictive value of mammography. AJR Am J Roentgenol 1992;158:521-6.
28. U.S. Preventive Services Task Force. Guide to clinical preventive services. 2nd ed. Baltimore: Williams & Wilkins, 1996.