Ten- to Fourteen-Year Effect of Screening on Breast Cancer Mortality [1, 2]

Sam Shapiro, *B.S.*,[3] Wanda Venet, *B.S., R.N.*,[4] Philip Strax, *M.D.*,[5] Louis Venet, *M.D.*,[6] *and* Ruth Roeser, *M.A.*[4, 7]

ABSTRACT—Results from the randomized trial conducted by the Health Insurance Plan (HIP) to determine the efficacy of breast cancer screening with mammography and palpation are reported for longer periods than previously available. By the end of 10 years after entry, the study group's mortality due to breast cancer was about 30% below the control group's. Arithmetic gains due to screening were maintained through year 14; relative gains declined. With increases in the period of follow-up, cumulative survival rates among cases detected by mammography alone (palpation negative during screening), decreased more rapidly than rates among other subgroups, but survival rates for mammography cases remained relatively high. Study women aged 40–49 years at entry began to show lower breast cancer mortality than those in the control group as duration of follow-up increased. Reservations are advanced about the acceptance of this finding as evidence of the efficacy of screening under age 50 under the conditions of the HIP study. The reservations are based on the observation that the decrease of mortality among the study group aged 45–49 at entry is concentrated entirely among cases diagnosed after they reached 50 years of age.—JNCI 1982; 69:349–355.

The favorable impact of periodic screening with mammography and clinical examination of the breast on mortality from breast cancer shown by the HIP Study has been extensively reported for the short-term period of 5 years after the start of screening and for several subsequent years (*1*). Stimulated by these results, a number of issues have attracted attention that affect the spread of large-scale screening programs, e.g., the independent contribution of mammography (*2*), the age groups that benefit from screening, the radiogenic effect of mammography, and the risk–benefit effects of screening at various intervals (*3–5*). At the same time, it has been well recognized that the efficacy of screening will not be adequately comprehended until we know the long-term effects on mortality.

Short- to intermediate-term reduction in mortality is close to one-third, but the absence of information on the natural history of breast cancer detected through screening leaves uncertain whether this reduction represents a cure or a postponement in mortality from breast cancer. Evidence on this issue would have great significance not only for screening but also for the more general understanding of the biology of breast cancer. At this point, the data on breast cancer detected in clinical practice show that even among cases diagnosed at a localized stage, there continues to be excess mortality for long periods after diagnosis (*6*).

The discussion that follows represents a start in addressing the question of long-term effects related to screening through an extension of previously reported observations into a 12- to 14-year period of follow-up.

METHODS

Population source and observation methods.—To test the efficacy of periodic screening with mammography and clinical examination, we selected two systematic random samples, each consisting of about 31,000 women aged 40–64 years, from members of HIP, a comprehensive, prepaid group practice. Study women were offered screening examinations; the 65% who appeared for an initial examination were offered three additional examinations at annual intervals unless earlier follow-up or biopsy was indicated. All but a small proportion (12%) had at least one additional annual examination. Screening consisted of a clinical examination, usually by a surgeon; mammography, in which two views were taken of the breast (cephalocaudad and lateral); and an interview to obtain relevant demographic information and a health history. Independence between the two examining modalities was strictly maintained. Initial screenings began in December 1963 and continued through June 1966; annual rescreenings ended in June 1970. Control women continued to receive their usual medical care. The extent to which women in the study and control groups had general physical examinations (including breast examinations) or engaged in breast self-examination after the completion of the screening project is not known. At HIP, mammography was utilized extensively for differential diagnostic purposes but not as part of a program for early detection among asymptomatic women.

Follow-up procedures.—All observations relate to women in the study and to control groups with no breast cancer diagnosed prior to their entry into the study. Follow-up procedures for identification of breast cancers and deaths due to breast cancer or other causes have been applied equally to the study and control groups. These include periodic communication with women (or their next of kin) on the breast cancer registries established in the study, searches of hospital insurance and death records, and surveys

Abbreviations used: CFR=case fatality rate(s); CSR=case survival rate(s); HIP=Health Insurance Plan of Greater New York.

[1] Received August 21, 1981; accepted February 26, 1982.

[2] Supported in part by Public Health Service (PHS) contract N01CP-43278 from the Division of Cancer Cause and Prevention, National Cancer Institute; and by PHS contract NIH-69-88 to The Health Insurance Plan of Greater New York from the National Cancer Institute.

[3] Health Services Research and Development Center, The Johns Hopkins University School of Hygiene and Public Health, 624 North Broadway, Baltimore, Md. 21205.

[4] The Health Insurance Plan of Greater New York, 220 West 58th St., New York, N.Y. 10019.

[5] New York Medical College, Valhalla, N.Y. 10595.

[6] Department of Surgery, Beth Israel Medical Center, 10 Nathan D. Perlman Place, New York, N.Y. 10003.

[7] We thank Dr. George B. Hutchison of the Department of Epidemiology, Harvard University School of Public Health for his valuable suggestions on reviewing the draft of this paper.

to Fourteen-Year Effect of Screening on Breast Cancer Mortality[1,2]

Sam Shapiro, *B.S.*,[3] Wanda Venet, *B.S., R.N.*,[4] Philip Strax, *M.D.*,[5] Louis Venet, *M.D.*,[6] *and* Ruth Roeser, *M.A.*[4,7]

ABSTRACT—Results from the randomized trial conducted by the Health Insurance Plan (HIP) to determine the efficacy of breast cancer screening with mammography and palpation are reported for longer periods than previously available. By the end of 10 years after entry, the study group's mortality due to breast cancer was about 30% below the control group's. Arithmetic gains due to screening were maintained through year 14; relative gains declined. With increases in the period of follow-up, cumulative survival rates among cases detected by mammography alone (palpation negative during screening), decreased more rapidly than rates among other subgroups, but survival rates for mammography cases remained relatively high. Study women aged 40–49 years at entry began to show lower breast cancer mortality than those in the control group as duration of follow-up increased. Reservations are advanced about the acceptance of this finding as evidence of the efficacy of screening under age 50 under the conditions of the HIP study. The reservations are based on the observation that the decrease of mortality among the study group aged 45–49 at entry is concentrated entirely among cases diagnosed after they reached 50 years of age.—JNCI 1982; 69:349–355.

The favorable impact of periodic screening with mammography and clinical examination of the breast on mortality from breast cancer shown by the HIP Study has been extensively reported for the short-term period of 5 years after the start of screening and for several subsequent years (*1*). Stimulated by these results, a number of issues have attracted attention that affect the spread of large-scale screening programs, e.g., the independent contribution of mammography (*2*), the age groups that benefit from screening, the radiogenic effect of mammography, and the risk–benefit effects of screening at various intervals (*3–5*). At the same time, it has been well recognized that the efficacy of screening will not be adequately comprehended until we know the long-term effects on mortality.

Short- to intermediate-term reduction in mortality is close to one-third, but the absence of information on the natural history of breast cancer detected through screening leaves uncertain whether this reduction represents a cure or a postponement in mortality from breast cancer. Evidence on this issue would have great significance not only for screening but also for the more general understanding of the biology of breast cancer. At this point, the data on breast cancer detected in clinical practice show that even among cases diagnosed at a localized stage, there continues to be excess mortality for long periods after diagnosis (*6*).

The discussion that follows represents a start in addressing the question of long-term effects related to screening through an extension of previously reported observations into a 12- to 14-year period of follow-up.

METHODS

Population source and observation methods.—To test the efficacy of periodic screening with mammography and clinical examination, we selected two systematic random samples, each consisting of about 31,000 women aged 40–64 years, from members of HIP, a comprehensive, prepaid group practice. Study women were offered screening examinations; the 65% who appeared for an initial examination were offered three additional examinations at annual intervals unless earlier follow-up or biopsy was indicated. All but a small proportion (12%) had at least one additional annual examination. Screening consisted of a clinical examination, usually by a surgeon; mammography, in which two views were taken of the breast (cephalocaudad and lateral); and an interview to obtain relevant demographic information and a health history. Independence between the two examining modalities was strictly maintained. Initial screenings began in December 1963 and continued through June 1966; annual rescreenings ended in June 1970. Control women continued to receive their usual medical care. The extent to which women in the study and control groups had general physical examinations (including breast examinations) or engaged in breast self-examination after the completion of the screening project is not known. At HIP, mammography was utilized extensively for differential diagnostic purposes but not as part of a program for early detection among asymptomatic women.

Follow-up procedures.—All observations relate to women in the study and to control groups with no breast cancer diagnosed prior to their entry into the study. Follow-up procedures for identification of breast cancers and deaths due to breast cancer or other causes have been applied equally to the study and control groups. These include periodic communication with women (or their next of kin) on the breast cancer registries established in the study, searches of hospital insurance and death records, and surveys

ABBREVIATIONS USED: CFR=case fatality rate(s); CSR=case survival rate(s); HIP=Health Insurance Plan of Greater New York.

[1] Received August 21, 1981; accepted February 26, 1982.

[2] Supported in part by Public Health Service (PHS) contract N01CP-43278 from the Division of Cancer Cause and Prevention, National Cancer Institute; and by PHS contract NIH-69-88 to The Health Insurance Plan of Greater New York from the National Cancer Institute.

[3] Health Services Research and Development Center, The Johns Hopkins University School of Hygiene and Public Health, 624 North Broadway, Baltimore, Md. 21205.

[4] The Health Insurance Plan of Greater New York, 220 West 58th St., New York, N.Y. 10019.

[5] New York Medical College, Valhalla, N.Y. 10595.

[6] Department of Surgery, Beth Israel Medical Center, 10 Nathan D. Perlman Place, New York, N.Y. 10003.

[7] We thank Dr. George B. Hutchison of the Department of Epidemiology, Harvard University School of Public Health for his valuable suggestions on reviewing the draft of this paper.

5 and 10 years after entry of all 62,000 women whose status was not already known through other sources. Other tracing methods are being applied for residual groups, and a new survey 15–17 years after entry is under way. Results thus far indicate that nearly all cases of breast cancer that occurred during the first 10 years after entry are now known; however, this may not be so for later years. Accordingly, mortality due to breast cancer, the primary measure for determining efficacy of screening, is restricted to cases detected in the 10-year interval. The end point for detection of breast cancer deaths among these cases is 14 years from entry; all study and control group women were observed for this period of time, and survival status of all cases on the breast cancer register has been ascertained.

Also presented are case survival rates among women with breast cancer histologically confirmed during the first 5 years after entry. The secondary position given to these rates in the determination of efficacy of screening results from *a*) the lead time in detection gained by the cases detected through screening and from *b*) the tendency of screening to select individuals with longer mean duration of preclinical disease. (In these instances each subsequent stage of disease may be more indolent, giving rise to a bias termed "length-biased sampling" (*7–9*), the effect of which diminishes as new cases are accrued.)

Life-table methods have been applied with the use of 6-month intervals after diagnosis. All breast cancer cases had a minimum of 9.5 years of exposure by December 31, 1980, the cut-off date for this report. The maximum period for which survival rates are given is 12 years[8] after which the numbers of cases withdrawn due to incomplete exposure become relatively large. Recalculation of survival rates between 9.5 and 12 years with the use of 1-month intervals shows that the resulting cumulative survival rates differ from the rates based on 6-month calculations by less than 1%.

Comparability of study and control groups.—High levels of comparability between the study and control groups have been demonstrated for a wide range of demographic and other characteristics (*10*) and for general mortality other than breast cancer (table 1). There are, however, large differences in general mortality between the study women who participated in the screening program and those who refused. The differential appears to be decreasing as the interval from entry increases, but the margin is still large (28%)—6–10 years from entry. This evidence of bias in the self-selection for screening has emphasized the importance of comparisons based on the total study group (screenees and refusers combined).

RESULTS

Breast Cancer Detection

Case detection information has been described elsewhere in detail (*1*); only a few of the results will be given here. By the end of 5 years from entry, which covers about 1.5 years after the last screening cycle, the gap between the breast cancer rates of the total study and control groups had become small (2.03 and 1.94/1,000 person-yr, respectively). Women who refused screening had a low rate of breast cancer (1.58/1,000) as compared with the rate among controls, indicating that women with a higher risk for breast cancer tended to self-select themselves for screening. Clinical and mammography examinations both contributed cases not detected by the other; the relative contribution of mammography (in the absence of clinical findings) was less among women under 50 years of age at diagnosis than among those 50 and over (14.4 vs. 37.6%). The proportion with no histologic evidence of axillary nodal involvement was higher among study cases than in the control group (56.4 vs. 46.3%). Breast cancers detected through screening had an especially high proportion with no nodal involvement (70.5%).

TABLE 1.—*Mortality from all causes excluding breast cancer: 10-year follow-up after entry*[a]

Population	Mortality at following intervals from entry:		
	10 yr	1–5 yr	6–10 yr
Total study	65.5	54.9	76.7
Screened	54.9	42.3	68.1
Refused screening	87.2	80.4	94.3
Control	64.8	56.5	73.4

[a] Rates per 10,000 person-yr.

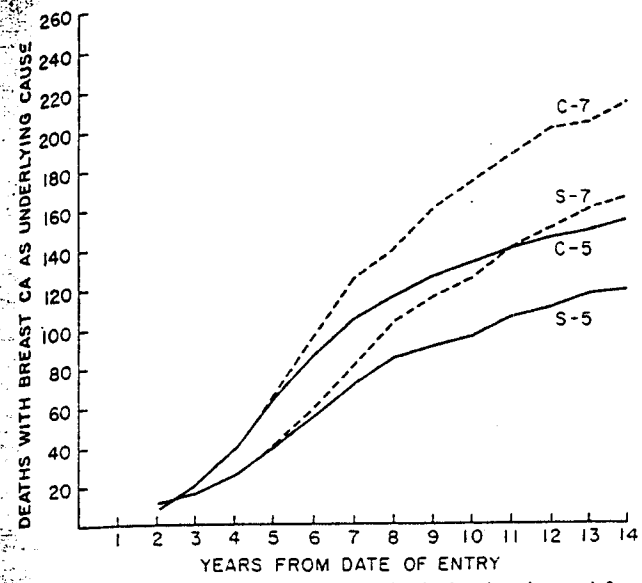
Mortality Differentials

The primary measure in the determination of the efficacy of screening is the differential between study women (those screened one time or more plus those who refuse screening) and control women in the number or rate of deaths with breast cancer as the underlying cause. Included are *a*) deaths due to breast cancer among women with histologically confirmed breast cancer and *b*) deaths among women with breast cancer as the underlying cause of death but with no histologic confirmation prior to death. Category *b*) consists of deaths among women for whom clinical or autopsy evidence indicated that breast cancer was the underlying cause of death. There were 9 and 16 such cases in the study and control groups, respectively, during the first 10 years after entry.

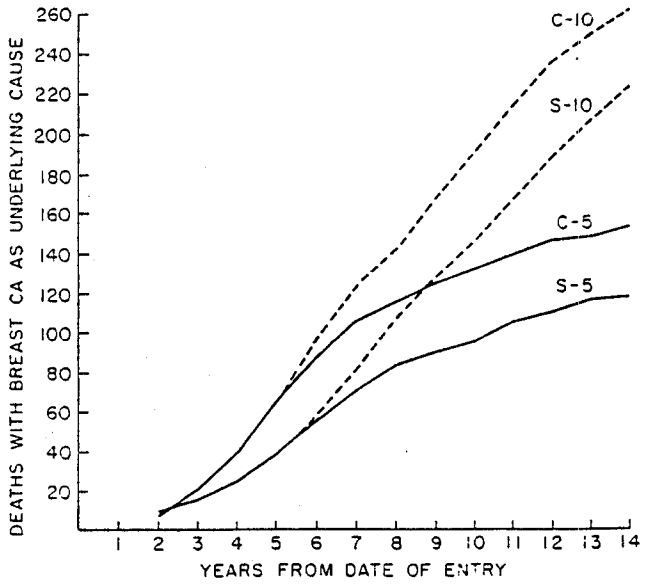
Follow-up periods of 5, 7, 10, and 14 years from date of entry are included to indicate changes in relationships as observations progress from short to intermediate to long intervals of follow-up.

Text-figure 1 displays two trend series. The curves representing cumulative breast cancer mortality from entry date among women with breast cancer diagnosed in the first 5 years have previously served as the basis for conclusions about the impact of periodic screening. Five years has been of interest because it includes an average of 3.5 years of screening and 1.5 years post screening—an interval that extends only moderately beyond the point when women would have been due for an additional examination if the program had continued. The gap between the curves for the

[8] In the total study group, there were 176 women alive at the start of follow-up interval 9.5–10 yr; over the next 1.5 yr 12 women had incomplete exposure; at the start of interval 11–11.5 yr, there were 146 women alive, 29 of whom had incomplete exposure by the end of yr 12. Corresponding figures for the control group are similar.



TEXT-FIGURE 1.—Cumulative breast cancer deaths by time interval from date of entry. S-5 and C-5 refer to breast cancer deaths within 5 yr of entry; S-7 and C-7 refer to corresponding situation for cancers diagnosed within 7 yr of entry. S=total study; C=controls.



TEXT-FIGURE 2.—Cumulative breast cancer deaths by time interval from date of entry. S-5 and C-5 refer to breast cancer deaths among cancers diagnosed within 5 yr of entry. S-10 and C-10 refer to corresponding situation for cancers diagnosed within 10 yr of entry. S=total study; C=controls.

study and control groups widens numerically until year 6 or 7 from entry and then remains fairly constant through year 14 (table 2).

Addition of breast cancer deaths to cases diagnosed through year 7 results in increased numerical differentials between study and control groups; no further increases occur with the inclusion of mortality among cases diagnosed in years 8–10 (text-fig. 2). From this, it would appear that mortality due to breast cancer among cases detected starting 3-3.5 years after screening ended became very similar for study and control groups of women.

Thus far, the discussion has been concerned with numerical differences. For purposes of estimating the influence of screening on reduction in the breast cancer mortality rate of a population, relative differences are important. Table 2 shows that the percent decrease becomes smaller with increases in duration from entry. This follows from the constancy in arithmetic differences just presented. Addition of breast cancers diagnosed through year 10 also reduces the differential between study and control groups in breast cancer mortality. The attenuation reflects the effect on breast cancer mortality in a population that participated in screening, because the cases are added long after screening ends. The magnitude of the attenuation is appreciable as suggested by the "observed" percentages in text-figure 3.

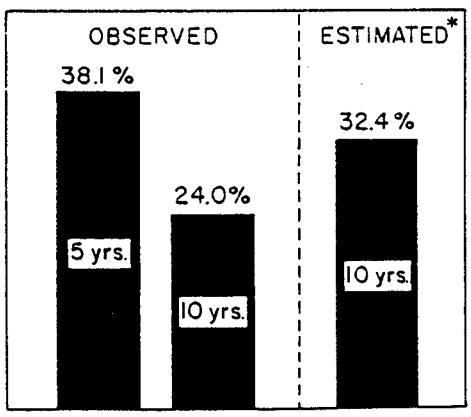
TABLE 2.—*Cumulative numbers of deaths due to breast cancer by selected time intervals from date of entry*

Interval to breast cancer diagnosis, yr	No. of breast cancers[a]	No. of deaths with breast cancer as underlying cause through following yr after entry:			
		5	7	10	14
Within 5					
Study	306	39	71	95	118
Control	300	63	106	133	153
Difference, %		(38.1)[b]	(33.0)[c]	(28.6)[b]	(22.9)[b]
Within 7					
Study	425	39	81	123	165
Control	443	63	124	174	212
Difference, %		(38.1)[b]	(34.7)[c]	(29.3)[c]	(22.2)[b]
Within 10					
Study	600	39	81	146	218
Control	604	63	124	192	262
Difference, %		(38.1)[b]	(34.7)[c]	(24.0)[b]	(16.8)[b]

[a]No. indicate breast cancers histologically confirmed within a specified interval after entry plus deaths among women with breast cancer as the underlying cause but with no histologically confirmed diagnosis prior to death.

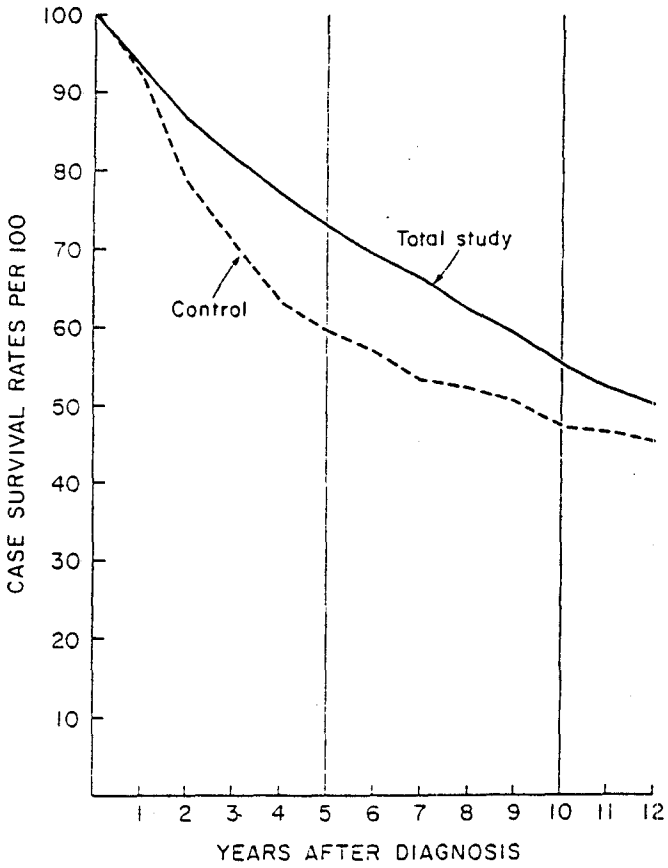
[b]$0.01<P<0.05$. Conditional 2-tailed test of the ratio between 2 Poisson parameters (study:control No. of deaths due to breast cancer) applied.

[c]$P<0.01$. Statistical test same as explained under footnote *b*.



*Estimated under assumption screening program renewed in years 6-10 after entry.

TEXT-FIGURE 3.—Difference between total study and control group deaths due to breast cancer 5 and 10 yr after entry.



TEXT-FIGURE 4.—CSR by interval from diagnosis in total study breast cancer patients and in controls.

Under the assumption that a new cycle of screening in years 6–10 among the women aged 45–64 years resulted in the same reduction in mortality as that found in years 1–5, the margin between study and control groups is 32.4%. This is to be compared with a differential of 24.0% under the condition of no continuation in screening as in the HIP study. Nevertheless, there is some decrease between the 5-year effect of screening and the effect over a 10-year period.

TABLE 3.—*Cumulative CSR (per 100) among confirmed breast cancer cases*[a]

Population	No. of cases	CSR through following yr after diagnosis:[b]		
		5	10	12
Total study	303	73.9	54.8	50.4 (2.9)
Detected through screening	132	87.1	64.4	59.7 (4.3)
Screened but not detected through screening	93	62.4	46.2	44.7 (5.2)
Refused screening	78	65.4	48.7	41.9 (5.6)
Control	294	59.5	46.3	43.1 (2.9)
Total study, adjusted[c]	303	71.3	53.8	48.9 (2.9)

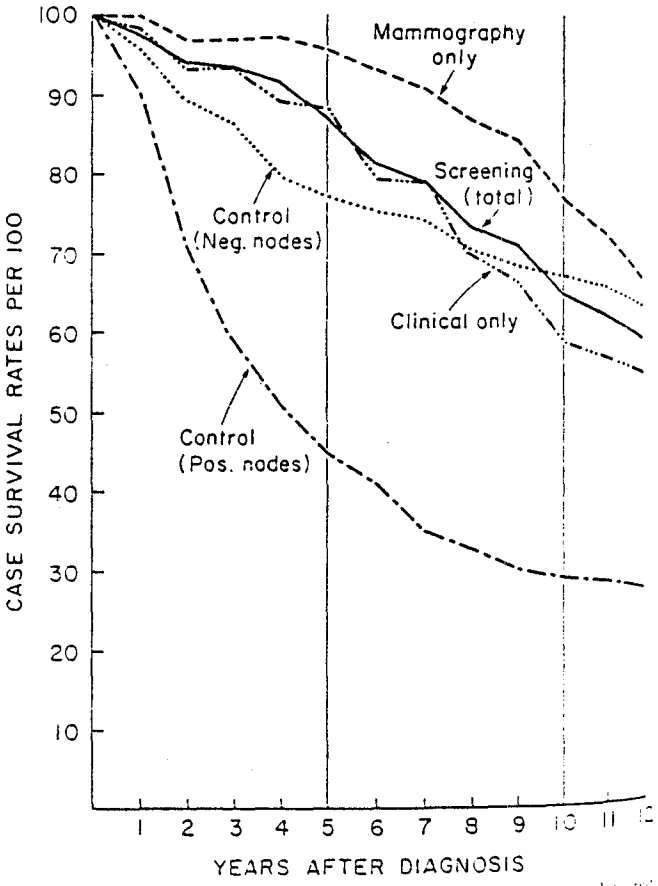
[a] Included are cases histologically confirmed within 5 yr of entry.
[b] *Numbers in parentheses* are standard errors due to sampling variability.
[c] Lead time of 1 yr is included for cases detected through screening. The adjusted CSR through yr n is derived as a weighted average of the rates for screened cases (yr n+1), screened but not detected through screening (yr n), and refused screening (yr n).

Case Survival Differentials

Past presentations on mortality from the HIP study paralleled the discussions of mortality due to breast cancer with a consideration of CFR, adjusted to take into account an estimated 1-year lead time in the detection of cases through screening (*8*). In general, the results have been highly consistent with those from the analysis of the population-based breast cancer mortality data. A reason for turning to the CFR has been that this is the only possible approach to the examination of the course of mortality related to the screening process itself, i.e., how CFR vary with the mode of detection (mammography or clinical examination) and with the round of screening examinations (initial examination vs. subsequent screenings). There are major restrictions in the interpretation of such CFR because of the uncertainty about how lead time or length-biased sampling, previously discussed, varies with the subgroup of cases detected.

The data to be discussed from this point represent a shift from CFR to CSR. All of the considerations related to CFR are applicable to CSR. The shift to CSR is useful because it places the experience in the study within the context of the usual analysis of breast cancer experience as exemplified by publications from the Cancer Surveillance, Epidemiology, and End Results (SEER) Program (*6*).

Text-figure 4 and table 3 show that screening has had a profound effect on the contour of the trend in cumulative CSR among breast cancer cases diagnosed in the total study



TEXT-FIGURE 5.—CSR by interval from diagnosis of breast cancers detected in screening and control cases.

TABLE 4.—*CSR (per 100) among breast cancer cases detected on screening by modality and control cases*[a]

Breast cancer category of cases	No. of cases	CSR through following yr after diagnosis:[b]		
		5	10	12
Detected through screening[c]				
Total	132	87.1	64.4	59.7 (4.3)
Mammography only	44	95.5	77.3	67.8 (7.1)
Clinical only	59	88.1	59.3	55.9 (6.5)
Mammography and clinical	29	72.4	55.2	55.2 (9.2)
Control				
Negative nodes[d]	136	77.2	66.9	62.5 (4.2)
Positive nodes[d]	123	44.7	29.3	27.4 (4.0)

[a] Included are cases histologically confirmed within 5 yr of entry.
[b] *Numbers in parentheses* are standard errors due to sampling variability.
[c] Initial evidence for biopsy recommendation made independently by the two modalities.
[d] Classification based on histologic evidence of axillary nodal involvement. Not shown are 35 cases with unknown nodal involvement.

group within 5 years after entry. The corresponding control cases have a marked convex curve similar to what is seen in general population series based on relative CSR. In contrast, the total study group barely begins to flatten out at year 10, and, in fact, a linear regression equation fits the time series quite well.

Clearly, within this time period there has been a fundamental alteration in the survival experience among breast cancer cases in the study group. The large bulge between the two trend lines is in the interval 4–8 years post diagnosis. At all subsequent points on the survival curve for the study group, the rate in the control group is 3–4 years earlier, which is well beyond the effect of introducing a 1-year average lead time for cases detected through screening, and provides a basis for estimation of person-years of life gained by the study group. Relative survival rates are currently being developed that take into account age distributions of breast cancer cases detected in the HIP study and age-specific general mortality in the study and control groups. Preliminary results indicate that the relationships discussed in this paper are not materially affected when based on relative survival rates.

Text-figure 5 and table 4 give further evidence of the complex changes that occur in cumulative CSR in a screening program. The solid line represents rates among all cases detected through sceening and is slightly concave. Data plotted for subgroups of these cases are based on small numbers and are subject to large sampling variability. Nevertheless, it would appear that the trend in CSR for cases detected on mammography only is decidedly concave. For "clinical only" cases, the trend is slightly concave. In contrast, the control cases with negative nodes show a modest convex trend in CSR, and by 12 years post diagnosis the cumulative CSR in this group is higher than that in the clinical only cases (≈3/4 of which were diagnosed with negative nodes). Finally, the difference between cumulative CSR for mammography only and clinical only cases appears to peak in year 10 after diagnosis. Some reductions may then occur. Only additional years of observation will determine whether this represents the start of closure or just chance variation. As expected, a wide gap persists between the rates for control cases with positive nodes and all of the other rates.

Mortality by Age

The last issue to be discussed in this paper concerns the much debated question whether the lowered mortality from breast cancer among study women is related to age at entry. Table 5 indicates that the direction of the difference is in favor of the study group at all ages. However, the firmest statements have been related to age group 50–59 years, which showed a statistically significant difference through year 10 ($P<0.05$); in year 14 the difference was not significant ($P>0.10$).

A major issue that has arisen concerns the situation in the age group 40–49 years at entry (*11, 12*) when there has been an increasing differential in breast cancer mortality in favor of the study group. At no time has the difference been statistically significant. In addition, the reason for our conclusion that the study has not demonstrated a benefit at ages under 50 is the observation that for this age group the smaller number of breast cancer deaths in the study group is almost entirely attributable to fewer deaths among cases diagnosed after women advanced to ages 50–54 years.

The rationale for examination of data by age at diagnosis in addition to age at entry is that a repetitive screening program that begins at young ages would eventually show benefits as women advanced beyond age 50, according to the HIP study. The relevant group of women were 45–49 years of age at entry; increasing proportions of these women passed their 50th birthday in successive rounds of screening.

As noted in table 6, the observed difference in numbers of deaths due to breast cancer between study and control groups 45–49 years old at entry (28 vs. 35, respectively, in 14 yr of follow-up) is attributable to the situation among cases diagnosed when the women were 50–54 years of age. The question that might be raised is whether this finding reflects relatively large numbers in the study group of breast cancer cases diagnosed under age 50 and small numbers at 50–54 years of age. However, ratios between deaths and cancer cases classified by age at diagnosis show no difference

TABLE 5.—*Breast cancer deaths by age at entry 5 and 14 years from entry*[a]

Age at entry, yr	No. of deaths through following yr after entry			
	5		14	
	Study	Control	Study	Control
Total	39	63	118	153
40–49	19	20	46	61
40–44	9	11	18	26
45–49	10	9	28	35
50–59	15	33	53	68
50–54	8	23	29	36
55–59	7	10	24	32
≥60	5	10	19	24

[a] Included are deaths with breast cancer as an underlying cause among breast cancers diagnosed during the first 5 yr after entry.

TABLE 6.—*Breast cancer deaths and cases with age at entry, at 40–49 years, by age at diagnosis, 14 years from entry*[a]

Age at diagnosis, yr	Deaths due to breast cancer		Breast cancer cases	
	Study	Control	Study	Control
	Age at entry 40–44 yr			
40–49	18	26	49	46
40–44	7	11	17	15
45–49	11	15	32	31
	Age at entry 45–49 yr			
45–54	28	35	68	68
45–49	18	12	40	30
50–54	10	23	28	38

[a]Included are deaths with breast cancer as an underlying cause among breast cancers diagnosed during the first 5 yr after entry.

between study and control groups at ages 45–49 (0.45 vs. 0.40), but they show a major difference at ages 50–54 (0.36 vs. 0.61).

Women 40–44 years old at entry remained under 50 in the interval 5 years after entry. If the deaths are aggregated by age at diagnosis, there is no difference between the study and control groups at 45–49 years (11+18, study group vs. 15+12, control group). The numbers at 40–44 years, age at diagnosis, are too small for meaningful assessment.

DISCUSSION

Follow-up in the HIP randomized trial for determination of efficacy of periodic screening for breast cancer with mammography and palpation of the breast has started to produce information beyond the short to intermediate periods after the start of screening. Results continue to be promising for a role for screening in secondary prevention of breast cancer mortality, although some changes are taking place in the magnitude of differentials that indicate effects. It seems clear that repetitive screening has led to about a 30% reduction in breast cancer mortality over a 10-year period. The impact of screening on mortality was felt rapidly, and the margin between study and control groups of women in mortality due to breast cancer reached a peak in the first 6–7 years after screening started. The arithmetic gains made by year 7 have been maintained through year 14; relative gains have declined to about 20%. Information from follow-up now under way will determine the nature of changes in differentials over the long run, thereby providing a basis for deriving estimates of person-years of life saved through screening.

Cumulative CSR have been exceptionally favorable among women with breast cancer detected through screening after allowance is made for an estimated 1-year lead time. This accounts for the entire difference between the total study group (screenees and refusers combined) and the control group in their CSR through year 12 after diagnosis of breast cancer. Nevertheless, as the interval from diagnosis increases, trends in CSR differ among various subgroups, the net effect of which is to reduce the differential between the study and control groups' rates. In the early years, cases detected on screening had relatively high CSR even when compared with the rates for control cases diagnosed with negative nodes. But with follow-up approaching 10 years, the advantage is eliminated.

CSR among "mammography only" cases were extraordinarily high the first 5 years after diagnosis and then started to decrease fairly rapidly. One possible explanation is that the mammography only cases were markedly affected by the lead time and length-biased sampling factors. Accordingly, the experience thus far indicates that caution needs to be exercised in drawing conclusions about the advantage for long-term prognosis attributable to screening with mammography from rates over a short period. Differentials in breast cancer mortality by age at entry may also be entering a new phase. From the early years of the study, screening's impact on mortality has appeared at ages 50–54; at entry ages 40–49, study–control differences show up later (although not statistically significant at any time). A complicating factor is the age at diagnosis, which, when introduced in the case of the age group 45–49, strongly suggests that whatever gain may have occurred is attributable to cases diagnosed after the women passed their 50th birthday. At this point, the prudent conclusion is that for ages under 50 years, the HIP study does not provide evidence that screening has had an effect on mortality. Small numbers may be a factor as well as the level of effectiveness of case detection under screening conditions in the 1960's.

All of the preceding observations are derived from an experiment in which mammography was in an earlier stage of development than it is now, and that screening ended after the offering of an initial examination and three additional screening examinations at annual intervals to women in the study group. Despite these limitations, there have been gains in the saving of lives from breast cancer or postponement of death among women with breast cancer. While it is important to assess the longer run impact of screening, the gains have extended over a sufficiently long period of time (12–14 yr) to be significant.

The questions that remain have been well defined (*13*). A high-priority question is whether repetitive screening with physical examination tied to instruction in breast self-examination achieves most of the effects found in screening that also utilizes mammography. Reduced costs for screening and elimination of radiation exposure, although very low under controlled conditions, are the issues at stake. At the same time, improvement in early case detection among women under age 50 and among older women through modern mammography has been demonstrated (*13*). This fact raises the possibility that under present conditions screening would have a beneficial effect under age 50. Resolution of the issues of relative value of palpation alone versus palpation and mammography and the efficacy of screening under age 50 has such great consequences for the spread of breast cancer screening to merit new investments in appropriately designed randomized trials. Fortunately, a move in this direction has already been taken by the National Cancer Institute of Canada (*14*).

REFERENCES

(1) SHAPIRO S. Evidence on screening for breast cancer from a randomized trial. Cancer 1977; 39 (suppl 6):2772–2782.

(2) Bailar JC III. Mammography—a contrary view. Ann Intern Med 1976; 84:770–784.

(3) Schwartz M. Analysis of the benefits of serial screening to detect early breast cancer. Cancer 1978; 41:1550–1564.

(4) Eddy DM. A theory of screening. In: Screening for cancer: Theory, analysis and design. Englewood Cliffs, N.J.: Prentice-Hall, 1980.

(5) Chiacchierini RP, Lundin FE, Scheidt PC. A risk/benefit analysis of life table modeling of an annual breast cancer screening program which includes X-ray mammography. In: Niebergs H, ed. Proceedings of the third international symposium on detection and prevention of cancer, New York City, April 26–May 1, 1976. New York: Marcel Dekker, 1980.

(6) Axtell L, Asire JA, Meyers MH. Cancer patient survival. Report No. 5. Washington, D.C.: U.S. Govt Print Off, 1976 [DHEW publication No. (NIH) 77-992].

(7) Zelen M. Theory of early detection of breast cancer in the general population. Heuson JC, Mattheiem WH, Rosenweig M, ed. Breast cancer: Trends in research and treatment. New York: Raven Press 1976:287–300.

(8) Shapiro S, Goldberg J, Hutchison G. Lead time in breast cancer detection and implication of periodicity of screening. Am J Epidemiol 1974; 100:357–366.

(9) Prorok PC. The theory of periodic screening. I. Lead time and proportion detected. Adv Appl Prob 1976; 8:127–143.

(10) Fink R, Shapiro S, Lewison J. The reluctant participant in a breast cancer screening program, 1968. Public Health Rep 83:479–490.

(11) Dubin N. Benefits of screening for breast cancer: Application of a probabilistic model to a breast cancer detection project. J Chron Dis 1979; 32:145–151.

(12) Prorok PC, Hankey BF, Bundy BN. Concepts and problems in the evaluation of screening programs. J Chron Dis 1981; 34:159–171.

(13) Beahrs OH, Shapiro S, Smart C, et al. Report of the working group: to review the National Cancer Institute-American Cancer Society breast cancer detection demonstration projects. I. General issues related to breast cancer screening. JNCI 1979; 62:655–662.

(14) Miller AB, Howe GR, Wall C. The national study of breast cancer screening. Clin Invest Med. In press.