Cancer Incidence and Mortality: The Priority of Screening Frequency and Population Coverage

GRAHAM A. COLDITZ,
DAVID C. HOAGLIN, and
CATHERINE S. BERKEY

Harvard University; Abt Associates Inc., Cambridge, Massachusetts

NATIONAL PROGRAMS OFTEN ARE DESIGNED TO reduce the impact of a major health problem. Yet the available research base that informs these policy decisions does not always provide the answers to some critical questions. The national health promotion and disease prevention objectives for the United States, which were devised to prevent and control cancer by the year 2000, include recommendations to increase the use of screening for breast, cervical, and colorectal cancers (U.S. Department of Health and Human Services 1990):

- increase the coverage of clinical breast exams and mammography every two years to at least 60 percent of women aged 50 and older
- increase the coverage of Pap tests every one to three years to at least 85 percent of women aged 18 and older
- increase fecal occult blood testing every one to two years to at least 50 percent of people aged 50 and older

The availability of further data or consideration of additional issues might improve these recommendations. The U.S. Preventive Services

Task Force used several criteria to evaluate screening and other preventive clinical services: the burden of disease; the efficacy of the screening test; the effectiveness of early detection; and recommendations made by others (e.g., the Canadian Task Force, the American Cancer Society, the National Cancer Institute, and professional groups).

We propose that the trade-off between the frequency of screening individuals and the population coverage (the percentage of the population that is screened), and the impact of this trade-off on the population rates of cancer, should also be carefully considered by those making policy recommendations.

Prevention strategies (e.g., some immunizations) for some diseases do not require repeated "intervention." Programs like immunization have focused solely on achieving sufficient or complete population coverage (frequency is not an issue). The discussion of cancer screening, on the other hand, has tended to focus on increasing the frequency of screening individual patients, ignoring the issues of population coverage. Unfortunately, often those at highest risk for cancer remain unscreened. Screening individuals less frequently, but covering more of the population, however, may (for some cancers) be more likely to reduce total population cancer mortality, which is one of the year 2000 goals.

We examine the criteria currently used to evaluate the effectiveness of screening for cancer and highlight issues on which the policy decision-making process requires additional data. We focus on cancers of the cervix, breast, and colon, as there is sufficient information about them to permit us to discuss seriously the trade-off between screening frequency and population coverage. We extend the approach of the U.S. Preventive Services Task Force by also considering the trade-off between (a) expending resources to achieve more complete coverage with a program and (b) more frequent screening that, in all likelihood, entails reduced coverage. Thus we ask, "Working within a constrained health care budget, how do we achieve the greatest reduction in cancer mortality?" The text of the Preventive Services Task Force (1996) indicates that extending coverage to people who have not been screened for cervical cancer is important: "The effectiveness of cervical cancer screening is more likely to be improved by extending testing to women who are not currently being screened and by improving the accuracy of the Pap smears than by efforts to increase the frequency of screening."

This contrasts with their *final recommendation* on screening for cervical cancer:

Regular Papanicolaou (Pap) testing is recommended for all women who are or have been sexually active and have a cervix. Testing should begin at the age when the woman first engages in sexual intercourse. There is little evidence that annual screening achieves better outcomes than screening every 3 years. Pap tests should be performed at least every 3 years. The interval for each patient should be recommended by the physician based on risk factors. (U.S. Preventive Services Task Force 1996)

Thus, the Preventive Services Task Force offers no population-level guidance on the trade-off between frequency of screening and coverage, although they do identify it as an important consideration. From the national public health perspective, quantifying this trade-off is fundamental to the efficient allocation of resources to achieve national goals of reduced cancer burden and mortality. In fact, emphasis on the frequency of screening and on other technical aspects should not detract from the importance of the coverage rate of a program. In principle, improving coverage can increase the benefit more than shortening the screening interval (Tomatis 1990, 267–8).

Background

Screening for cancer is one approach to prevention, and it is often called "secondary prevention" because its aim is early detection and improved treatment outcomes rather than primary prevention of disease. In contrast with primary prevention, where one intervention (such as cessation from smoking [Kawachi et al. 1993] or avoidance of obesity [Manson et al. 1995]) may reduce the incidence of multiple cancers, heart disease, and diabetes, secondary prevention strategies are specific to the individual cancers. Thus, secondary prevention offers a fragmented, but still valuable, approach to reducing the societal burden of cancer, when cancers detected at an early stage respond to treatment, hence ensuring that the apparent gain in survival exceeds what would be obtained solely through earlier diagnosis and detection of less aggressive tumors (Cole and Morrison 1980; Black and Welch 1993). Furthermore, some cancers (i.e., breast cancer) currently have no established primary prevention strategies, so we must rely solely upon screening for early detection.

Table 1 summarizes some of the measures that influence the effectiveness of screening programs for cancer. Information on these mea-

TABLE 1
Leading Determinants of the Impact of Cancer Screening Programs

	Cervix	Breast	Colon
Incidence[a]			
Age 35	14.6	66.1	3.4
Age 65	15.6	392.3	133.5
Object of screen	Squamous cells (precursor)	Cancer (early dx)	Polyp (precursor)
Test	Pap smear	Mammogram	Flex sig
Sensitivity	35%	60% to 70%[b]	85% or more
Specificity	90%		
Age effect on sensitivity	No	Yes	No
Age effect on sojourn[c]	No	Yes	?
Duration of protection	High for 3 yr	Nil after 3 yr	High for at least 5 yr

[a]Incidence rates per 100,000 per year for U.S. white women (SEER). Incidence varies by age and ethnic group.
[b]Sensitivity for women over age 50.
[c]Sojourn is defined as the time from the first detectable abnormality on screening to clinical diagnosis, also referred to as the "detectable preclinical phase."
Abbreviations: dx, diagnosis; flex sig, flexible sigmoidoscopy, 60 cm; SEER, Surveillance, Epidemiology and End Results.

sures is needed to estimate the potential impact of a screening program on cancer incidence and mortality rates.

Cervical Cancer

Cervical cancer accounts for about 7,000 deaths per year among women in the United States. The incidence is higher among African-American women than among white women. The incidence (based on the Surveillance, Epidemiology and End Results [SEER] tumor registries) for African-American women doubles from age 35 to age 65 (from 20 to 43 cases per 100,000 women). In contrast, the rates among white women are relatively stable over the same age range, perhaps reflecting in part the impact of screening programs. Cervical cancer appears to be the disease in which secondary prevention through screening has been most extensively studied. Numerous studies have evaluated the efficacy of Pap-smear and screening programs (Hakama, Miller, and Day 1986; International Agency for Research on Cancer 1986). In addition, time trends in cervical cancer incidence and mortality—particularly in Nordic countries, where implementation of programs has varied—show a strong correlation between the extent of the organized screening program and changes in cervical cancer rates (Hakama and Louhivouri 1988).

The Pap smear, now the standard screening test for cervical cancer, detects exfoliative cytology, a defined precursor of squamous cell carcinoma of the cervix. A meta-analysis suggests that a typical sensitivity (i.e., the probability that a diseased patient tests positive) is approximately 35 percent when specificity (i.e., the probability that a nondiseased person tests negative) is greater than 90 percent (Fahey, Irwig, and Macaskill 1995). A screening program, of course, requires more than a screening test; it must include referral and effective treatment for the precursor lesion identified by the test. The treatment typically involves cryosurgery to remove the premalignant lesion from the cervix and so to prevent progression to malignant disease.

For cervical cancer, the protective effect (i.e., the reduced probability of a diagnosis after a negative screening test) is high for three years after the last negative screen (Pap test), and then it declines by 4 percent per year. Thus, even six to nine years after a negative screen, substantial protection remains for screened, compared with unscreened, women.

Because the protective effect is reported by time since last negative screen, one can estimate the cumulative rate of disease that would be expected from different schedules of screening (that is, assuming different intervals between screens).

The International Agency for Research on Cancer (IARC) has coordinated a large number of studies on cervical cancer screening and has reported its work extensively. IARC estimates the percentage reduction in cumulative incidence of cervical cancer from age 35 to age 65 for a series of screening intervals (table 2). With data like these, it is possible to estimate the marginal gain, from the population perspective, when screening is intensified from every three years, say, to every two years. As seen in table 2, this increase produces a reduction in incidence of cervical cancer of 1.7 percent and an increase in the number of Pap smears from 10 to 15 per woman over the 30-year interval from age 35 to age 65. The marginal benefit (1.7 percent reduction in incidence) is low relative to the marginal cost (five additional Pap smears) per woman.

In the United States in 1992, 65 percent of women had had a Pap smear within the past three years. Furthermore, some 49 percent of women had had a Pap smear within the past year (Anderson and May 1995). The frequency of screening was related to age. Women 18 to 39 years of age were more likely to have had a Pap smear (72 percent) than those 65 to 74 years of age (54 percent). These survey data suggest that approximately half of all women in the United States have an annual Pap smear.

TABLE 2

Efficacy of Pap Screening and Reduction in Cumulative Incidence
of Cervical Cancer According to the Frequency of Screening
after Age 35 in a Population That Is Uniformly Screened at Age 35[a]

Interval between screens (yr)	Reduction in cumulative incidence (%)	Tests from age 35 to age 65 (N)
1	93.5	30
2	92.5	15
3	90.8	10
5	83.6	6
10	64.1	3

[a]Data from the International Agency for Research on Cancer (1986).

Trade-Offs in Screening Frequency and Cancer Rate Reductions

If we assume that the National Health Interview Survey data reflect current U. S. population screening patterns, we can estimate the population reduction in cervical cancer according to the frequency of screening. For this purpose, we shall assume optimistically that the 35 percent of women who have not been screened in the past 3 years receive a Pap test every 10 years, although evidence suggests that through the 1980s some 12 percent of women had never had a Pap screen. Applying the estimates of reduction reported by the IARC working group (see table 2), we calculate that, for women aged 20 to 65, based on the distribution of screening currently observed in the United States, we have an overall reduction of cervical cancer incidence by 82.8 percent from the rate that would be observed in the complete absence of screening (see table 3). These estimates assume full follow-up and treatment of lesions among all participants in the screening program.

From table 3 we see that moving the 49 percent of women who now undergo annual screening to a schedule of screening every three years would reduce drastically the number of tests per woman; from age 20 to age 65 these women would require 30 fewer tests each. The protection against cervical cancer among these women would decrease only slightly, from 93.5 to 90.8 percent. By allocating the saved resources to ensure that the 35 percent of the population that is currently screened approximately every 10 years shifts to a three-year interval between screens, we increase their number of tests from 5 to 15. These women would observe a substantial gain in protection against cervical cancer (from a 64 percent reduction to a 90.8 percent reduction in cancer incidence). At the population level, this translates into both fewer cases of cervical cancer (an 8 percent reduction) and roughly half the current number of Pap smears (from 26 to 15 tests per woman). Accounting for the false positive results of screening tests that require additional workup, the cost savings with this lower number of screening tests would be substantially more than just the cost of tests that were not performed.

This example demonstrates that a program aimed at increasing the coverage of women screened for cervical cancer, which would result in the 35 percent of women not being screened regularly now obtaining Pap smears every three years while the frequency of screening is reduced in other groups, may be a desirable strategy. It results in fewer tests to

TABLE 3

U.S. Population Distribution of Frequency of Pap Test[a] and Protection against Cervical Cancer

Screening patterns	Tests per woman from age 20 to 65 (N)	Reduction in cervical cancer, cumulative incidence (%)
Current scenario		
49% annual	45	93.5
16% every 3 yr	15	90.8
35% every 10 yr[b]	5	64.1
Current population average	26.2	82.8
Scenario of increased population coverage		
Whole population every 3 yr	15	90.8
Change from current scenario	~11 tests per woman	8% reduction in cervical cancer

[a]Based on 1992 National Health Interview Survey data.
[b]This assumption may be optimistic, as the 1987 Health Interview Survey indicates that up to 12% of the population has never been screened.

the total population, but offers greater aggregate protection against cervical cancer. Thus, this approach would be superior to the current clinical practice of testing every one to three years at the discretion of the clinician and would speed us toward attaining the national goal of reduced cervical cancer burden.

Applying a Markov modeling approach to the evaluation of screening frequency, the Office of Technology Assessment reported that a one-time screen would save women entering the Medicare program (i.e., age 65) 14,400 life years per million women screened and would cost the health care system $1,666 per year of life saved. The incremental costs per year of life saved is least for five-year screening ($1,453) and is progressively greater as screening frequency increases (see table 4). Further, on defining high-risk groups as those who have never had a prior Pap screen, the investigators observed cost savings for screening every five years (Office of Technology Assessment 1990).

A cost-effectiveness analysis conducted in Holland addressed the optimal number of screens for a woman in that country and concluded that seven to ten Pap smears over a lifetime produced the maximum cost-effectiveness. Just seven screens in a lifetime could produce the same benefit as standard Dutch clinical practice (spontaneous screening by physicians similar to current U.S. practices), but at half the cost. In large part this is due to excessive screening applied to younger women and insufficient screening of older women (Koopmanschap et al. 1990).

It is noteworthy that Finland, with a national program of Pap screens every five years, has achieved a 60 to 70 percent reduction in incidence of cervical cancer. In contrast, Knox and Woodman modeled screening behavior in the United Kingdom and observed that the distribution of Pap smears was far from optimal. They estimated that only 8 percent of cervical cancer deaths were avoided by Pap smears through 1975. When optimizing the age distribution of Pap smears and removing social stratification of use, they estimate that seven screens per woman over her life would prevent 52.4 percent of deaths from cervical cancer (Knox and Woodman 1988).

Conclusion. Screening programs for cervical cancer are so effective that annual and biennial screens are unnecessary for average-risk women, in part because the test screens for precancerous cells and the progression from the detectable, precancerous state to cancer is slow. Reallocating resources away from annual Pap smears should be an explicit public health policy, as should the emphasis on expanding coverage in

TABLE 4
Cost-Effectiveness of Cervical Cancer Screening beginning at Age 65[a]

Screen schedule	Discounted life-years[b] (in 1,000s)		Cost (millions)		Cost-effectiveness ratio (added cost/life year gained)
	Cohort	Added	Total	Added	
No screen	11,383	—	217.8[c]	—	—
One time	11,397	14.4	241.8	24.0	$ 1,666
Five-year	11,419	21.8	273.5	31.7	1,453
Three-year	11,426	7.0	315.2	41.7	5,956
Annual	11,433	6.8	585.1	269.9	39,693

[a]Five percent discount rate.
[b]Per one million women beginning at age 65.
[c]Costs include diagnosis and treatment of women presenting with cervical cancer in the absence of screening. This primarily reflects treatment of advanced disease.
Source: Office of Technology Assessment (1990).

the unscreened population. Implementing policies to expand population coverage is currently a priority at the Centers for Disease Control and Prevention.

Breast Cancer

Breast cancer is the second leading cause of death among U.S. women, but the leading cause of cancer incidence. Among women over age 50, incidence is higher in white women than in African-American women, perhaps reflecting childbearing patterns. Yet among African-American women diagnosed with breast cancer from 1986 to 1991, the cancer was diagnosed later, at a more advanced stage, and survival rates are significantly lower at each stage of diagnosis than among white women (Kosary et al. 1995), despite equivalent mammography use (Anderson and May 1995).

Much debate continues over both the age to begin mammographic screening and the frequency of screening (Fletcher et al. 1993; Chalmers 1993; Sickles and Kopans 1993). In particular, the level of coverage prescribed for health plans remains unclear: should screening be annual? every two years, as in many European countries? or perhaps every three years among older women? Few data address this issue of screening frequency and the relative benefits in terms of breast cancer mortality.

The issue of screening frequency is complicated by physiological factors. Screening is less sensitive among premenopausal women (i.e., those under 50), whose breasts are denser. In one clinical trial, sensitivity was 60 percent in younger women, compared with 86 percent in older women (Tabar et al. 1992). The rate of growth of tumors may also be faster among premenopausal women, partly because of their higher levels of circulating hormones. Consistent with these physiological conditions, studies of mammography among women under age 50 indicate that the sojourn time (i.e., the interval from the least detectable abnormality on screening to clinical diagnosis) is perhaps only one year for women aged 40 to 50, in contrast with two years among those over 50 years of age (Day and Chamberlain 1988). Modeling U.S. incidence over the past decade also indicates that lead times of two years for women aged 40 to 49 and 50 to 59, four years for ages 60 to 69, and five years for ages greater than 70 provide a good fit to the data (Feuer and Wun 1992). The combination of lower test sensitivity and faster-growing

tumors, together with lower incidence rates at younger ages, leads to inconclusive findings from the published trials regarding the benefit of screening younger women (aged 40 to 49) (Elwood, Cox, and Richardson 1993).

The actual magnitude of a program required to obtain a reduction in breast cancer mortality is exemplified by data from the Swedish trial of mammography (screening every two years). For women aged 50 to 69 at entry to the trial, one breast cancer death was prevented per 4,000 woman-years, per 1,460 mammographic examinations, per 13.5 biopsies, and per 7.4 breast cancers detected (Tabar et al. 1989). These results, based on nine years of follow-up, reflect 58 fewer deaths among the 47,000 women in the screened group. At the same time, the mammography-screened group had 20 percent more cases of breast cancer diagnosed than the unscreened group.

Trade-Offs in Screening Frequency and Cancer Rate Reductions

This challenging example of the trade-off between decreasing screening frequency and expanding population coverage has been limited by the sparse data on the reduction in risk of breast cancer according to time since the last mammogram. A report from an ongoing screening program in northern California, which includes 8,547 women who have had a first mammogram interpreted as normal and who have later undergone a second screening examination, provides insufficient information to document changes in risk with the interval since the last negative screening test; only 16 cancers were diagnosed among these women during follow-up screens (Kerlikowske et al. 1993).

Data from Europe, including a national screening program in the United Kingdom, indicate that the rate of breast cancer is reduced during the first year after a screening mammogram, rises during the second year, and, during the third year, is close to that among women who have never had a screening mammogram (table 5). Cancer may be detected after a negative screen for several reasons. The cancer may arise as a result of failure to detect it by screening mammography, or it may be a new cancer. Review of data from several trials indicates that at least half of the cancers diagnosed in the first 24 months after screening mammography are true new cancers. Thus screening every two years is

TABLE 5

Distribution of Time Interval since Mammogram for Breast Cancers in Screening Program, United Kingdom, North West Regional Health Authority

Mammo screen (yr)	Women screened (N)	Time (months) after a negative mammogram to diagnosis							
		<12		12–23		24–36	
		Cases	Rate[a]	Cases	Rate	Cases	Rate
1988–89	13,359	5	3.7	14	10.5	24	18.0
1989–90	24,390	18	7.4	23	9.4	33	13.5
1990–91	40,892	36	6.4	38	9.3	—	—
1991–92	58,781	30	5.1	—	—	—	—

[a]Rate of breast cancer per 10,000 women.
Source: Woodman et al. 1995.

viewed as the most efficient frequency for a screening program (Wood-
man et al. 1995). These British data are summarized in table 5, where
the follow-up of entry cohorts is summarized for the first 12 months,
then for months 12 to 23 and, finally, for months 24 to 36. These data
are comparable to those from other trials in Europe (Peeters et al. 1989)
and begin to define the issues around screening frequency. However,
whether optimal screening frequency is constant across age remains to
be shown, particularly as lead time varies with age.

Cost-Effectiveness of Mammography

Using an elaborate model of screening applied to the entire Dutch
population of women aged 50 to 70, van der Maas and colleagues (1989)
estimate that, with a 70 percent attendance at screening, a program of
mammography every two years can produce a 12 percent reduction in
breast cancer mortality. The cost-effectiveness of mammography is in-
fluenced mainly by the reduction in the need to treat advanced disease.
Further analysis of data for the national Dutch screening program sug-
gests that screening every two years will reduce mortality from breast
cancer in the total population by 16 percent, whereas screening every
three years reduces it by 10 percent (de Koning et al. 1991). Adjust-
ment for quality of life has little impact on these estimates. Addressing
issues of screening frequency and population coverage, de Koning esti-
mates that expanding coverage to include women aged 70 to 75 results
in a marginal cost-effectiveness of U.S. $8,000 per additional life-year
gained, which is substantially more favorable than the expansion of
screening to women under age 50.

Using data on incidence in the United States and charges from South-
ern California Medicare, Katlove and colleagues have developed a basic
package of services representing benefits that should be provided to all
citizens (Katlove et al. 1995). Based on a charge of $88.50 for mam-
mography, they estimate that, for women aged 50 to 59, charges would
be $8,280 per extra year of life and for women aged 60 to 69, $9,890.
To save one potential life at ten years of follow-up, 690 women aged 50
to 59 would need to follow regular screening; among women aged 60 to
69, the number would be 553.

Conclusion.   Mammography is moderately effective among women
aged 50 and older, where it is associated with decreased mortality from

breast cancer. The frequency of screening is not uniform across countries where programs have been implemented. European guidelines favor screening every two years for women over age 50. Because the benefit from having a mammogram substantially dissipates beyond two years, we do not propose any decrease in screening frequency for women already screened in order to increase population coverage. Expanded population coverage is needed for the detection of breast cancer, as only 48 percent of women have had a mammogram in the last three years (Anderson and May 1995).

Barriers to screening include lack of physician advice to women (Ackermann and Cheal 1994), and, among older women, a lower knowledge of the usefulness and benefits of mammography, particularly in the absence of symptoms (Costanza 1994). Further, there is evidence that, despite Medicare coverage for mammography, older women in low-income, inner-city populations still experience financial barriers to screening (Kiefe et al. 1994). However, systematic programs, like that implemented by the Group Health Cooperative of Puget Sound, have produced substantial benefits, such as a 32 percent reduction in late stage breast cancer between 1989 and 1990 (Thompson et al. 1995).

Colon Cancer

The incidence of colon cancer rises rapidly with age and is similar among women and men. It is the second most common cancer in the United States and the second leading cause of death from cancer (behind lung cancer). Screening for colon cancer currently takes one of two general approaches: either *testing for occult blood* or *flexible sigmoidoscopy*. Both aim to detect cancers at an early stage or, more important, to detect and then remove colon polyps, precursor lesions that over time will progress to malignant colon cancers if untreated.

Screening typically involves testing for fecal occult blood; the main benefit of this screening is the gain from additional tests performed on those who screen positive. Several trials of fecal occult blood testing show high rates of screening sigmoidoscopy, no doubt as a result of positive fecal occult blood tests.

Flexible sigmoidoscopy is being used more widely since the publication of results of two observational studies, one from Kaiser Permanente and one from an HMO in Wisconsin. These studies show protection

against mortality from colorectal cancer of approximately 50 percent over ten years after a screening test (sigmoidoscopic examination) (Selby et al. 1992; Newcomb et al. 1992). Although these and other recent reports suggest that screening sigmoidoscopy may reduce mortality from colon cancer, both major studies have used relatively small numbers of subjects and have had limited ability to evaluate the potential confounding effect of healthful lifestyle and likelihood of screening.

The efficacy of this screening procedure may come in part from excision of precursor lesions, or polyps, that are detected (Winewar et al. 1993). It appears that many providers proceed to more aggressive surveillance and testing among those who have at least one polyp. To date there is no rational recommendation either for the frequency of screening or for the age at which screening should begin. Despite sparse data, the American Cancer Society (ACS) recommends sigmoidoscopy every three to five years, starting at age 50 (American Cancer Society 1980).

The Preventive Services Task Force concluded in 1996 that evidence was sufficient to recommend annual fecal occult blood testing or sigmoidoscopy, or both (U.S. Preventive Services Task Force 1996, 89–103). They note that evidence is insufficient to recommend for or against routine screening with digital rectal examination, barium enema, or colonoscopy. Currently, sigmoidoscopy appears to be spreading as a routine screening procedure, with large HMOs providing this procedure as a screening service. Although most screening recommendations are made for the total population, they nevertheless focus on high-risk individuals, usually defined according to age. This practice, in part, reflects the distribution of the burden of disease. For colorectal cancer, it may also be possible to define risk according to lifestyle factors and to clarify the current ACS recommendation, which does not mention that after several negative sigmoidoscopic screens one gains little from additional screening.

The National Health Promotion and Disease Prevention Objectives specify that fecal occult blood testing every one to two years should be extended to achieve coverage of at least 50 percent of people aged 50 and older. This recommendation has been expanded, in the report *Health, United States, 1992*, to include proctosigmoidoscopy in addition to fecal occult blood testing (National Center for Health Statistics 1993, 328–9). Estimates of the prevalence of this screening procedure in the United States indicate that in 1987 only 25 percent of adults reported having ever had a proctosigmoidoscopic examination.

*Trade-Offs in Screening Frequency
and Rate Reductions*

Because clinical practice is rapidly changing as new studies are published, we review the trade-off between frequency of sigmoidoscopic examinations and reduction in cancer rates, as estimated by Eddy (1990). Using Markov models to obtain estimates, he presents numerous scenarios, but for illustration we use his estimates for screening among 50-year-old women of average risk who are screened from age 50 to age 75 (table 6).

These data suggest that little marginal gain results from increasing the frequency of sigmoidoscopy screening from every five years to every three years. For three additional tests (between ages 50 and 75) performed per person, there is an estimated reduction of 13 cases of invasive colorectal cancer per 10,000 people and 30,000 screening tests, and a reduction of four deaths per 10,000 people. Unfortunately, Eddy does not present data that allow estimation of population cancer rates if the frequency of screening were reduced to, say, every 10 years. The biology of colon polyps justifies such an interval, and published studies of screening effectiveness show decreased mortality for ten years after a sigmoidoscopy. Because the frequency of screening in the U.S. population is not currently available, we do not present estimates of the trade-off. But, as for the Pap smear, it is likely that less frequent screening, perhaps every ten years starting at age 50, but covering more of the population, would be an efficient screening strategy that would result in fewer cases and fewer deaths from colon cancer.

Although not presented in detail here, it is worth noting that Eddy's analysis shows that a barium enema every five years would prevent twice as many cancers per 10,000 women (301) as a sigmoidoscopy every three years (153), and with half as many perforations (14 vs. 37). The barium enema strategy is also cheaper. However, medical practice is not pursuing this approach to screening, but rather is exploring approaches, such as training in sigmoidoscopy for nurse practitioners, who are equally effective and have higher rates of return for follow-up examinations than gastroenterologists (Maule 1994).

Conclusion.   Flexible sigmoidoscopy, which may prevent nearly half of all colorectal cancer deaths, appears promising as a screening procedure for colon cancer. Questions of screening frequency remain important, including the need for information on longer screening intervals

TABLE 6

Alternative Scenarios for Colon Cancer Screening at 50 to 75 Years of Age for 10,000 Asymptomatic Women

Benefits and risk	Annual FOBT and flex sig every 5 yr	Annual FOBT and flex sig every 3 yr	Annual FOBT and colonoscopy every 5 yr	Annual FOBT and colonoscopy every 3 yr
Decrease in invasive colorectal cancer per 10,000 people	140	153	340	349
Decrease in death from colorectal cancer	102	106	172	173
Increase in life expectancy (days)	42.6	43.9	73.8	75.2
Number of perforations	24	37	115	56

Source: Eddy (1990).
Abbreviations: FOBT, fecal occult blood test; flex sig, flexible sigmoidoscopy, to 60 cm.

(ten years) and the relative merits of sigmoidoscopy, barium enema, and colonoscopy. It is likely that less frequent screening, but covering more of the population, will result in fewer cases and deaths. These issues must be addressed to help inform policy and social strategies to implement broader screening programs.

Discussion

Screening for precursor lesions is more useful than early detection of cancers, in part because there is a longer interval from a negative screening test to a subsequent new cancer. Less frequent screening is preferable when the target is a precursor lesion, thus resulting in more cost-effective programs to reduce cancer mortality. This approach, as exemplified by cervical and colon cancer screening, also offers the advantage of preventing cancer; treating a precursor lesion is cheaper than treating a diagnosed cancer and also less traumatic for the patient. However, for cancers (like breast cancer) in which the target of screening is early detection of malignancies, few data have been available until recently regarding the time sequence from a negative test to subsequent cancer, resulting in poorly informed screening recommendations and policy. For cancers that are not detectable in premalignant stages, research leading to better screening tests is more important in the long run, although it will not provide information of value for current screening programs.

Although others have mentioned that population-level coverage is important (Williams and Vessey 1990), perhaps more so than frequency of screening, no reports that we identified show how these measures trade off in benefits in health care strategies that assume finite resources. From an economic perspective, however, Torgerson and Donaldson note that attendance for screening may be a misleading indicator of the benefits of a screening program if this requires screening those at lower risk (Torgerson and Donaldson 1994). For example, if women not currently screened for cervical cancer were at lowest risk, then the benefit of increased coverage would be relatively small. Data from the United States suggest that just the opposite applies for cervical cancer screening, so that the benefits of expanding coverage would be even greater than we estimated. However, this point should be considered when setting goals for screening programs. Additional data and examples may aid the debate over health care reform. The allocation of resources among

screening tests, programs to increase population coverage, and savings in resources from having to treat fewer cancers diagnosed at a more advanced stage may all be considered.

When goals are set for screening programs in the face of limited resources, reduction in cancer mortality should remain the primary objective. Although comparative data documenting the population-level or public health benefits of alternative screening programs are not available, one additional measure that may help in the choice among policy options is the years of life lost due to each form of cancer. For the cancers we consider in this presentation, this measure ranges from 845,000 person-years of life lost for breast cancer to 758,000 person-years for colon cancer, and 116,000 for cervical cancer (although this figure reflects the impact of current Pap screening).

An alternative to aiming at broad population coverage with a screening program is to target a subpopulation that is at high risk for disease. Such high-risk groups may be better served by more intense screening frequencies. The literature addresses the problems of defining high-risk groups, although it usually considers only those with a positive family history of the cancer in question (for example, see Eddy [1990] on screening for colon cancer). But for breast and colon cancer, only 15 percent of cancers occur among those with a family history of disease. For cervical cancer, Hakama and colleagues show that the use of risk factors to define a high-risk group small enough to reduce the costs of a screening program was not effective (Hakama, Pukkala, and Saastamoinen 1979). They defined a high-risk group using epidemiological risk factors, but this group included only 39 percent of all cases. Similar results have been found for breast cancer.

Most policies and recommendations by national organizations and government agencies routinely use age and sex as markers to define high risk. As cancer rates increase markedly with age, the relative effectiveness of screening increases also. That is, fewer women are screened at older ages per cancer detected, but we also note that the savings in life-years decrease as life expectancy diminishes with increasing age. The effect of age on the sensitivity and specificity of the screening test may also be important, especially for breast cancer, where mammography has a higher sensitivity in older women. This has particular importance for those who test positive and then require additional diagnostic tests. In addition to the diagnostic test procedures, for all cancers the period of uncertainty

regarding the diagnosis is clearly stressful. In general, for all cancers (except lung cancer), high-risk groups defined by risk factors other than age contribute only a small fraction to the total disease burden and hence do not offer a useful approach to prevention or screening.

Emerging data do suggest that, as an alternative, low-risk groups may be defined by a number of epidemiological markers of risk. For example, for ovarian cancer we may consider defining a low-risk group that may include women who have used oral contraceptives for five or more years (which yields a 50 percent reduction in risk) (Hankinson et al. 1992), who do not have a family history of breast or ovarian cancer, and who have had a tubal ligation (which further reduces risk by approximately 50 percent) (Hankinson et al. 1993). This may be a rare example among cancers, although colon cancer may also offer the potential to define risk groups according to lifestyle factors and to recommend screening strategies separately for these groups (e.g., high-risk—defined by intake of red meat, moderate alcohol intake, and low folate intake; low-risk—defined by low red meat, low alcohol, and high folate intake) (Giovannucci et al. 1993).

Public health policy emerges from interactions among forces in society. For health promotion and disease prevention, we require a sound scientific knowledge base, social strategies, and political will. These factors must be in balance (Richmond and Kotelchuck 1984). Social strategies that facilitate broad participation in screening are essential. Barriers to compliance are a necessary consideration in determining the extent of coverage set as a goal in a screening program. These barriers go beyond the cost of the screening test and may include health beliefs, fear of unpleasant tests (e.g., sigmoidoscopy), access to services, and competing time constraints. In addition, factors like overdiagnosis, complications of the screening tests, false positive test results, the complications of treatment for the cancer in question and their impact on quality of life may also represent barriers in some populations.

There is not enough information about sigmoidoscopy and other colon-screening modalities that have been available for a considerable time to permit us to choose among screening intervals. Hence we must rely upon statistical models, such as the Markov model used by Eddy (1990), to inform decision making. If we are to maximize the use of secondary prevention to reduce cancer mortality, then obtaining these data is an urgent priority.

We note that comprehensive efforts to increase population coverage have been undertaken in the United States in recent years. The CDC received 64 million dollars in July 1992 to develop comprehensive multistate programs (summarized in table 7) for the early detection of breast and cervical cancers (Centers for Disease Control and Prevention 1993). Additional funds awarded to CDC in 1993 allowed for expansion of the program. To date, CDC notes a substantial increase in the number of screening sites and in implementation of public and professional education programs. Some 556,000 screening tests have been provided to women who are medically underserved (Henson, Wyatt, and Lee 1996).

Although the U.S. Public Health Service Year 2000 goals recommend greater screening to reduce cancer mortality, an optimal screening policy for the United States at this time should aim to achieve a maximum reduction in cancer burden for the lowest expenditure. An alternative to such an optimal approach would be to fund only screening tests shown to significantly reduce cancer mortality without regard for the dollar costs of the screening program. Under this alternative, no insurance coverage would currently be justified for PSA screening for prostate cancer, ovarian cancer screening, or mammography before age 50 because mortality reductions have not been demonstrated. If the focus is on optimal approaches, then, with sufficient funds available and evidence that screening reduced mortality, we would need to rank screening tests according to a criterion such as cost per life saved or cost per year of life saved. We would likely begin by focusing on those

TABLE 7
CDC Goals for Comprehensive Programs for Early Detection of Cancer

- Establish, expand, or improve screening services in communities with women at risk of breast and cervical cancer
- Provide appropriate referrals for medical treatment of women screened through this program and ensure the appropriate follow-up services
- Develop and implement a public education program about the importance of screening for breast and cervical cancer
- Develop and implement a professional education program
- Improve quality assurance of screening tests
- Establish a surveillance and evaluation system to monitor the program
- Establish and maintain a state-based cancer control plan

cancers responsible for the greatest life-years lost. Such an approach would rank the cancers in decreasing order of life-years lost as follows: breast, colon, prostate, and ovary. The very low estimate of life-years lost from cervical cancer already takes into account an effective screening program. To achieve optimal efficiency of national screening programs, we would focus on the total costs of the screening programs (screening, diagnosis, treatment of detected disease, and costs for overcoming population barriers to screening). In this situation, the interval or spacing between screening tests has a major impact on the relative cost-effectiveness of screening (less frequent tests are far more cost-effective, assuming that diagnosis and treatment costs are similar). An alternative to this approach, which relies heavily on life-years saved, could consider the most cost-effective screening strategy for individuals of a given age. This approach deals instead with the issue of how we can best provide health care to our population of 60-year-old women, leading us to consider whether screening dollars for women of this age are better spent by screening for colon cancer, breast cancer, or cervical cancer. The optimal screening strategy will be different for 40-year-old women.

Conclusion

In contrast to the usual measures of screening test performance (sensitivity, specificity), this review highlights the importance of the frequency of screening tests in considering a screening program. Data to inform recommendations for the timing of screening tests are sparse and are urgently needed if we are to implement programs to achieve maximum efficiency in reducing cancer mortality. To achieve national goals with limited resources, screening tests that will contribute most to reducing mortality in a cost-effective manner must be given priority. Sometimes this task can be accomplished by increasing the population coverage rather than the frequency of screening tests. The example of cervical cancer screening highlights the important trade-offs between frequency of screening and population coverage and the relative impact of these components on population cancer rates as well as on costs. To achieve the National Cancer Institute's goals for cancer mortality reduction in the United States, these trade-offs must be reviewed for each

cancer, and prevention strategies must be developed that emphasize population coverage.

References

Ackermann, S.P., and N. Cheal. 1994. Factors Affecting Physician Adherence to Breast Cancer Screening Guidelines. *Journal of Cancer Education* 9:96–100.

American Cancer Society. 1980. Guidelines for the Cancer-Related Checkup: Guidelines and Rationale. *Cancer* 30:194–240.

Anderson, L.M., and D.S. May. 1995. Has the Use of Cervical, Breast, and Colorectal Cancer Screening Increased in the United States? *American Journal of Public Health* 85:840–2.

Black, W.C., and H.G. Welch. 1993. Advances in Diagnostic Imaging and Overestimations of Disease Prevalence and the Benefits of Therapy. *New England Journal of Medicine* 328:1237–43.

Centers for Disease Control and Prevention. 1993. Update: National Breast and Cervical Cancer Early Detection Program, 1992–1993. *Morbidity and Mortality Weekly Report* 42:747–9.

Chalmers, T.C. 1993. Screening for Breast Cancer: What Should National Health Policy Be? *Journal of the National Cancer Institute* 85:1619–21.

Cole, P., and A.S. Morrison. 1980. Basic Issues in Population Screening for Cancer. *Journal of the National Cancer Institute* 64:1263–72.

Costanza, M.E. 1994. The Extent of Breast Cancer Screening in Older Women. *Cancer* 74(7 suppl.):2046–50.

Day, N.E., and J. Chamberlain. 1988. Screening for Breast Cancer: Workshop Report. *European Journal of Cancer and Clinical Oncology* 24:55–9.

de Koning, H.J., B.M. van Ineveld, G.J. Oortmarssen, et al. 1991. Breast Cancer Screening and Cost-Effectiveness: Policy Alternatives, Quality of Life Considerations and the Possible Impact of Uncertain Factors. *International Journal of Cancer* 49:531–7.

Eddy, D.M. 1990. Screening for Colorectal Cancer. *Annals of Internal Medicine* 113:373–84.

Elwood, J.M., B. Cox, and A.K. Richardson. 1993. The Effectiveness of Breast Cancer Screening by Mammography in Younger Women. *Online Journal of Current Clinical Trials* [serial online] (doc. no. 34 and 32).

Fahey, M.T., L. Irwig, and P. Macaskill. 1995. Meta-Analysis of Pap Test Accuracy. *American Journal of Epidemiology* 141:680–9.

Feuer, E.J., and L-M. Wun. 1992. How Much of the Recent Rise in Breast Cancer Incidence Can Be Explained by Increases in Mammography Utilization? *American Journal of Epidemiology* 136:1423–36.

Fletcher, S.W., W. Black, R. Harris, B.K. Rimer, and S. Shapiro. 1993. Report of the International Workshop on Screening for Breast Cancer. *Journal of the National Cancer Institute* 85:1644–56.

Giovannucci, E., M.J. Stampfer, G.A. Colditz, et al. 1993. Folate, Methionine and Alcohol Intake and Risk of Colorectal Adenoma. *Journal of the National Cancer Institute* 85:875–84.

Hakama, M., and K. Louhivouri. 1988. A Screening Programme for Cervical Cancer That Worked. *Cancer Survey* 7(3):404–16.

Hakama, M., A.B. Miller, and N.E. Day. Eds. 1986. *Screening for Cancer of the Uterine Cervix.* Lyon: International Agency for Research on Cancer.

Hakama, M., E. Pukkala, and P. Saastamoinen. 1979. Selective Screening: Theory and Practice on High Risk Groups of Cervical Cancer. *Journal of Epidemiology and Community Health* 33:257–61.

Hankinson, S.E., G.A. Colditz, D.J. Hunter, T.L. Spencer, B. Rosner, and M.J. Stampfer. 1992. A Quantitative Assessment of Oral Contraceptive Use and Risk of Ovarian Cancer. *Obstetrics and Gynecology* 80:708–14.

Hankinson, S.E., D.J. Hunter, G.A. Colditz, et al. 1993. Tubal Ligation, Hysterectomy, and Risk of Ovarian Cancer: A Prospective Study. *Journal of the American Medical Association* 270:2813–18.

Henson, R.M., S.W. Wyatt, and N.C. Lee. 1996. The National Breast and Cervical Cancer Early Detection Program: A Comprehensive Public Health Response to Two Major Health Issues for Women. *Journal of Public Health Management Practice* 2(2):36–47.

International Agency for Research on Cancer, Working Group on Evaluation of Cervical Cancer Screening Programmes. 1986. Screening for Squamous Cervical Cancer: Duration of Low Risk after Negative Results of Cervical Cytology and Its Implication for Screening Policies. *British Medical Journal* 293:659–64.

Katlove, H., A. Liberati, E. Keeler, and R.H. Brook. 1995. Benefits and Costs of Screening and Treatment for Early Breast Cancer. Development of a Basic Benefits Package. *Journal of the American Medical Association* 273:142–8.

Kawachi, I., G.A. Colditz, M.J. Stampfer, et al. 1993. Smoking Cessation in Relation to Total Mortality Rates in Women. A Prospective Cohort Study. *Annals of Internal Medicine* 119:992–1000.

Kerlikowske, K., D. Grady, J. Barclay, E.A. Sickles, A. Eaton, and V. Ernster. 1993. Positive Predictive Value of Screening Mammogra-

phy by Age and Family History of Breast Cancer. *Journal of the American Medical Association* 270:2444–50.

Kiefe, C.I., S.V. McKay, A. Halevy, and A. Brody. 1994. Is Cost a Barrier to Screening Mammography for Low-Income Women Receiving Medicare Benefits? A Randomized Trial. *Archives of Internal Medicine* 154:1217–24.

Knox, E.G., and C.B.J. Woodman. 1988. Effectiveness of a Cancer Control Program. *Cancer Surveys* 7(3):379–401.

Koopmanschap, M.A., K.T.N. Lubbe, G.J. van Oortmarssen, H.M.A. van Agt, M. van Ballegooijen, and J.D.F. Habbema. 1990. Economic Aspects of Cervical Cancer Screening. *Social Science and Medicine* 30:1081–7.

Kosary, C.L., L.A.G. Ries, B.A. Miller, B.F. Hankey, A. Harras, and B.K. Edwards. Eds. 1995. *SEER Cancer Statistics Review, 1973–1992: Tables and Graphs.* NIH pub. no. 96-2789. Bethesda, Md.: National Cancer Institute.

Manson, J.E., W.C. Willett, M.J. Stampfer, et al. 1995. Body Weight and Mortality Among Women. *New England Journal of Medicine* 331:677–85.

Maule, W.F. 1994. Screening for Colorectal Cancer by Nurse Practitioners. *New England Journal of Medicine* 330:183–7.

National Center for Health Statistics. 1993. *Healthy People 2000 Review. Health, United States, 1992.* Hyattsville, Md.: U.S. Public Health Service.

Newcomb, P.A., R.G. Norfleet, B.E. Storerr, et al. 1992. Screening Sigmoidoscopy and Colorectal Cancer Mortality. *Journal of the National Cancer Institute* 84:1572–5.

Office of Technology Assessment. 1990. *The Costs and Effectiveness of Screening for Cervical Cancer in Elderly Women—Background Paper.* OTA-BP-H-65. Washington, D.C.: U.S. Congress.

Peeters, P.H.M., A.L.M. Verbeek, J.H.C.L. Hendriks, R. Holland, M. Mravunac, and G.P. Vooijs. 1989. The Occurrence of Interval Cancers in the Nijmegen Screening Programme. *British Journal of Cancer* 55:929–32.

Richmond, J.B., and M. Kotelchuck. 1984. Co-ordination and Development of Strategies and Policy—The United States Example. In: *Oxford Textbook of Public Health,* ed. W. Holland. Oxford: Oxford University Press.

Selby, J.V., G.D. Friedman, C.P. Quesenberry, and N.S. Weiss. 1992. A Case-Control Study of Screening Sigmoidoscopy And Mortality from Colorectal Cancer. *New England Journal of Medicine* 326:653–7.

Sickles, E.A., and D.B. Kopans. 1993. Deficiencies in the Analysis of Breast Cancer Screening Data. *Journal of the National Cancer Institute* 85:1621–4.

Tabar, L., G. Fagerberg, S.W. Duffy, et al. 1992. Update of the Swedish Two-County Program of Mammographic Screening for Breast Cancer. *Radiology Clinics of North America* 30:187–210.

Tabar, L., G. Fagerberg, S.W. Duffy, and N.E. Day. 1989. The Swedish Two County Trial of Mammographic Screening for Breast Cancer: Recent Results and Calculation of Benefit. *Journal of Epidemiology and Community Health* 43:107–114.

Thompson, R.S., S.H. Taplin, T.A. McAfee, M.T. Mandelson, and A.E. Smith. 1995. Primary and Secondary Prevention Services in Clinical Practice. *Journal of the American Medical Association* 273:1130–5.

Tomatis, L. Ed. 1990. *Cancer: Causes, Occurrence and Control*. Lyon: International Agency for Research on Cancer.

Torgerson, D.J., and C. Donaldson. 1994. An Economic View of High Compliance as a Screening Objective. *British Medical Journal* 308:117–19.

U.S. Department of Health and Human Services. 1990. *Healthy People 2000. National Health Promotion and Disease Prevention Objectives*. DHHS pub. no. (PHS) 91-50213. Hyattsville, Md.: U.S. Public Health Service.

U.S. Preventive Services Task Force. 1996. Guide to Clinical Preventive Services: Screening for Colorectal Cancer. Baltimore: Williams & Wilkins.

van der Maas, P.J., H.J. de Koning, B.M. van Ineveld, et al. 1989. The Cost-Effectiveness of Breast Cancer Screening. *International Journal of Cancer* 43:1055–60.

Williams, E.M.L, and M.P. Vessey. 1990. Compliance with Breast Screening Achieved by the Aylesbury Mobile Service (1984–1985). *Journal of Public Health Medicine* 12:51–5.

Winewar, S.J., A.G. Zauber, M.N. Ho, et al. 1993. Prevention of Colorectal Cancer by Colonoscopic Polypectomy. *New England Journal of Medicine* 329:1977–81.

Woodman, C.B.J., A.G. Threlfall, C.R.M. Boggis, and P. Prior. 1995. Is the Three Year Breast Screening Interval Too Long? Observance of Interval Cancers in the NHS Breast Screening Programme's North West Region. *British Medical Journal* 310:224–6.