12

Principles of Screening for Cancer

Russell Harris and Linda S. Kinsinger

Screening is defined as testing for a condition when the person has no recognized signs or symptoms of that condition. The purpose of screening is not to merely detect a condition, but rather to help people live better or longer. This is an important distinction: the detection of earlier disease by itself is insufficient to justify a screening program. The program must additionally demonstrate that people live longer or better because of the earlier detection.

A positive screening test result does not indicate that a person has the condition, but rather that he or she has a higher probability of having the condition. People with positive screening tests usually undergo diagnostic testing to determine whether the condition is present. For example, a woman with a positive mammogram result does not necessarily have breast cancer, but she may undergo a needle localization biopsy to determine whether she has breast cancer.

Screening is not a single test, but rather a cascade of events that can lead to either benefits or harms (see The Cascade of Screening, later in this chapter). Potential benefits include living better or living longer and are usually experienced some years after screening. Potential harms include the effects of false-positive or false-negative screening tests and problems that result from overdiagnosis and overtreatment. Harms are usually suffered soon after screening.

Because screening programs may lead to either net benefit or net harm, decision makers must carefully evaluate proposed programs. Eight criteria distinguish effective programs:

1. Disease: The disease should cause a sufficient burden of suffering to warrant attention and should have a detectable preclinical phase of sufficient length to allow early detection (see The Critical Point in Cancer Treatment, later in this chapter).
2. Test: The screening test should be sufficiently sensitive to detect those cancers that could benefit from earlier treatment. Note that the test does not need to be maximally sensitive but rather "sensitive enough" to detect those cancers that it is important to detect. Cancers that are important to detect are those which are treatable when detected by screening but not when detected clinically.
3. Test: There are usually many more false-positive test results than true-positive results. The screening test should be specific enough to minimize the number of false-positive test results so as to minimize their negative consequences.
4. Availability and acceptability: The screening test, workup, and resultant treatment should be available to all and acceptable both to clinicians and to the people being screened.
5. Treatment: There must be a treatment for the disease that is more effective when applied to screening-detected cancers than clinically detected cancers. By "more effective," we mean that people will live longer or better as a result of this earlier treatment.
6. Harms of overtreatment: Often earlier detection includes detection of people with intermediate lesions that would never progress to invasive cancer. Screening may lead many people with these lesions to be subjected to treatment that they do not need and which causes harm. To minimize harms, people with lesions that will not progress to clinically important disease should rarely be subjected to potentially harmful and unnecessary treatment.
7. Benefits and harms: Overall benefits (in terms of people living longer or better) must outweigh overall harms (including harms from the screening test, harms from the workup, the adverse effects of earlier treatment and overtreatment, the psychologic effects of labeling, and the downstream effects of surveillance).
8. Costs: The net health benefits must come at a reasonable cost.

Screening for cancer is a popular idea, but this popularity may be based more on intuition than on understanding. Studies show that the great majority of Americans are convinced that being screened is part of being a responsible citizen.[1] What is less certain is how well the public comprehends the magnitude of the potential benefits of cancer screening; even less certain is whether the public appreciates the magnitude of the potential harms. Further, one could wonder whether the public has a reasonable grasp of the gaps in our knowledge of the effects of screening.

One might argue that whether the public understands these issues is irrelevant. The fact is that the public has decided that cancer screening is a good that it desires. We suggest that there are several important reasons for the public to better understand screening. The first is that screening consumes resources, such as money and the time of medical personnel. In a system strapped for resources to appropriately care for all our people, expending resources on services that offer little benefit and risk greater harm reduces the contribution medical care can make to the health of the public. The second reason is that if widespread screening fails to reduce the rate at which people die of cancer, ultimately the public will ask why it does not. If the medical care system has not

informed the public of the limitations of cancer screening as a strategy to reduce cancer mortality, its credibility will be damaged. Finally, if the public pins all its hopes for cancer control on screening, this attitude may inhibit creative new ideas and research that could develop alternative strategies for cancer control.

The public needs to have a clear idea of both sides of the cancer screening coin: benefits and harms. Clinicians must play a large role in this educational effort. This chapter attempts to help clinicians better understand these issues so that they can appropriately advise the public.

The Idea of Cancer and the Idea of Screening

The public's understanding of how cancer works is central to its understanding of how screening works and thus to its strong interest in screening. Especially relevant is the public's perception of the development and progression of cancer and of the degree of homogeneity of cancers with the same name (e.g., breast cancer) in their malignant potential.

Although the process of cancer development is not completely understood, it is clear that a normal cell does not become cancer suddenly, all at once. Rather, cells undergo a number of assaults over time, with various results.[2] Some of these assaulted cells develop various abnormal forms, or "intermediate lesions," such as cervical intraepithelial neoplasia (CIN), colonic polyps, or ductal carcinoma in situ of the breast (DCIS). Although not cancer themselves, these intermediate lesions do at times develop into cancer.

As screening frequently detects intermediate lesions, their natural history is important. If nearly all intermediate lesions progress to malignant cancer, then early detection and treatment would appear to be an effective strategy for cancer control. The detection of intermediate lesions would be a triumph. By interdicting the developing cancer at this early point (i.e., even before it can be called a *cancer*), treatment could eradicate a lesion that would have caused major health problems in the years to come.

With many intermediate lesions, however, the majority (most often, the great majority) never progress to invasive cancer. Thus, screening often results in detecting and treating intermediate lesions that do not need to be detected or treated. If there are any harms to this early detection and treatment, the magnitude of these harms must be counted against the magnitude of the benefit. It is doubtful if many people understand this result of screening, or at least the frequency with which it occurs.

After cancer develops, a critical issue is the extent to which it uniformly progresses in a linear and inevitable manner to cause symptoms and death. If cancer is always an inexorably progressive condition, it is intuitively appealing to think that early detection is an effective strategy for cancer control. The experiences of people who have cancer with the same name (e.g., breast cancer) would vary little; all would be destined for a difficult death because the cancer had progressed too far for effective treatment. Again, the facts are otherwise.

Cancers, even cancers with the same name (e.g., breast cancer), vary widely in their growth rate and malignant potential. Studies have found that cancers that vary with respect to certain cell markers have different prognoses.[3] Gene expression profiling using DNA microarrays[4,5] has shown the genetic heterogeneity of individual breast cancers. There is not one type of breast (or colorectal or prostate) cancer, but a number of types, each with a different natural history. Together with the probable but largely unknown ways in which individual susceptibility varies, these cancer types produce great variation in the ways a particular cancer is expressed. Some cancers in certain individuals grow rapidly and are lethal within a short time, regardless of our best treatments. Screening is unlikely to make a difference for people with such cancers, which may metastasize from the first cell.

Other cancers with the same name grow more slowly, or not at all. People with some of these latter cancers may be greatly helped by early detection and treatment; others have cancers that do not need to be detected and treated at all. In some cases, lesions that clearly meet histologic criteria for cancer do not cause important clinical problems. Experts have termed this last group *pseudodisease*, lesions that appear to be cancer but do not progress to clinically important disease. It is the existence of this type of cancer, less malignant and less requiring of treatment, that gives pause to the push for screening. Here are cancers that do not need to be found early; some of them do not need to be found at all.

Much of the public has another conception of how cancer works. The word *cancer* usually means a condition that universally and inevitably progresses, a condition that is potentially fatal in every case. The fact that some people have long-term survival after cancer diagnosis is attributed to some exceptional characteristic of the individual or to effective treatment. Intermediate lesions are called *premalignant*; the popular conception is that they too inexorably progress to cause major clinical problems. This incorrect view of the nature of cancer is an important underlying reason for the popularity of cancer screening. As people have commented to the authors, cancer screening "simply makes sense." Given this view of cancer, one can understand their thinking.

The Critical Importance of Treatment Effectiveness in Determining the Benefits of Screening

The purpose of screening is not simply to detect disease earlier, but rather to help people live better or longer (i.e., improve health outcomes) because of early detection. Thus, the question we need to ask ourselves in considering a screening situation is not how many early cancers we find but how many people avoid poor health outcomes.

Although many people and their clinicians view the potential benefits of screening as primarily a function of the accuracy (especially the sensitivity) of a screening test, in fact, the factor that most commonly limits the benefit from screening is the treatment. For a screening program to improve health outcomes, it must include a treatment that is not only effective but which is more effective if applied earlier than if applied later. That is, the critical issue with screening is the *timing* of treatment. If the treatment is not effective at any time, obviously screening is not useful. If treatment is excellent and just as effective for clinically detected cancer as screening-detected cancer, then again early detection by screening is not helpful. Screening is only useful in improving health outcomes when the treatment is effec-

tive for screening-detected cancer but not clinically detected cancer.

This treatment criterion for a screening program is often misunderstood. The important question is this: where in the natural history of this cancer is the critical point (see following discussion) at which a particular treatment becomes ineffective? Theoretically, at least, many treatments may be effective when a potentially fatal cancer is only a few cells in size. As this cancer grows, however, there comes a point at which treatment is no longer effective in altering its natural history and helping the person to live better or longer. It is the location of this *critical point*, and especially its relationship with the point of detection by the screening test, that determines the effectiveness of the screening-and-early-treatment program. If the critical point is earlier than the point at which the cancer can be detected by screening, then screening cannot be helpful. If the critical point is during the "lead time" produced by the screening test, then screening will be helpful. If the treatment is very effective and the critical point is after the point at which regular, competent medical care would detect the cancer, then screening is not useful because treatment after usual clinical detection is as effective as treatment after screening detection.

When treatment is particularly effective, the critical point for some potentially fatal cancers may be at a far-advanced stage. Even very effective treatments may become ineffective for far-advanced stage cancers. Far advanced stage cancers at diagnosis may occur in several situations: in people who neglect their health; in people without access to regular, competent medical care; or in people without understanding that early signs or symptoms should be evaluated. In the past, for example, some women presented with breast tumors that were the size of a lemon or even an orange. It would be difficult to deny that many such cancers could have been treated more successfully had they been evaluated at an earlier stage. Few women present with such advanced tumors now, at least partly, because most women in this country recognize that breast lumps of any size should be examined by a physician.

The treatment requirement for a screening program is that the treatment must be more effective after detection by screening than after usual clinical detection. It does not require that the treatment be effective for far-advanced stage cancers. One does not need to implement a screening program to prevent the development of far-advanced cancers by helping people understand that new symptoms and signs should be reported to one's physician. This educational effort is different from screening.

The issue of the effectiveness of treatment at different points in the natural history of cancer is made more complex by the marked variation in cancers and individuals, as just discussed. It is not surprising, for example, that early detection and treatment rarely reduce mortality by 100%. For example, in the overviews of the randomized controlled trials of breast cancer screening, mortality is reduced by less than 20%.[6] This finding would imply that about 20% of women destined to die of breast cancer have a type of cancer that is better treated earlier than later. The other 80% of women destined to die of breast cancer have a type of cancer for which earlier treatment is not useful. These women may have a particularly malignant form of cancer in which metastasis occurs at an early stage, too early to be detected by screening.

Many other women are detected by breast cancer screening, of course, but these may be women not destined to die of breast cancer. They may have either a less-malignant form of the disease for which later treatment is as effective as earlier treatment or a benign form of cancer that would never have caused major adverse health outcomes even without treatment.

The Critical Point in Cancer Treatment

As shown in Figure 12.1, cancer begins as a small number of cells. If it were possible to detect every cancer at this point, and accurately distinguish the potentially fatal ones from the nonfatal, then our treatments would have a high rate of success. As the cancer grows, however (moving to the right in the figure), the potentially fatal cancers reach a point at which they are less effectively treated. This critical point varies between cancers and within cancers with the same name. It also varies between treatments. An important advance in treatment may mean that cancers can be effectively treated at a later stage in their development (i.e., farther toward the right of the figure).

The relationship of the critical point to the point at which a screening test can detect a cancer helps determine the potential benefits of screening. If the critical point is between points A and B in the figure (i.e., before the screening test can detect the cancer), then screening with the present test will not reduce the burden of suffering of the cancer. If the critical point is between points B and C (i.e., within the detectable preclinical phase of the cancer), then screening may well be helpful in reducing mortality and/or morbidity. If the critical point is to the right of point C in the figure, then the treatment is effective for even advanced cancers, and earlier detection is not needed.

Sensitivity of the Screening Test: A Less Important Criterion

In contrast to effective treatment, the sensitivity of the screening test, that is, its ability to detect early cancer, may or may not be an important factor in determining the benefit from a screening program. If a screening test is made more sensitive (for example, by reducing the cut-point for defining "abnormal"), it is likely that the test will detect more cancers. However, if these additional cancers are either more

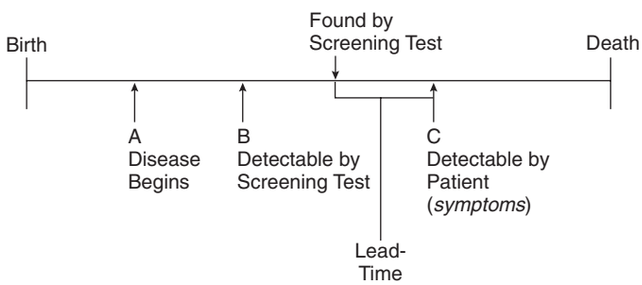

FIGURE 12.1. Natural history of cancer. *Critical Point*, point at which treatment becomes less effective; could be between *A* and *B*, between *B* and *C*, or after *C*.

benign (and would never cause problems) or more malignant (and thus have already metastasized), then the extra sensitivity would not have made a contribution to improving health outcomes. The operative question is not how many more cancers are found by a more sensitive test but rather whether screening has moved detection for at least some potentially fatal cancers back to a more treatable stage. If this has not occurred, then the more sensitive test has not been a useful addition to the screening program. This rationale includes such strategies as screening more frequently (i.e., reducing the screening interval), which may increase sensitivity but may or may not improve health outcomes.

For this reason, the sensitivity of a screening test may not be related to its ability to improve health outcomes. For example, screening for cervical cancer with the Pap smear probably has a fairly low sensitivity,[7,8] yet screening every 3 years apparently reduces cervical cancer mortality by more than 80%.[9] Developments in the technology of screening tests that seek to improve screening programs by increasing the sensitivity of the screening test may increase sensitivity without improving health outcomes. Such approaches may increase the cost of screening without providing additional health benefit.

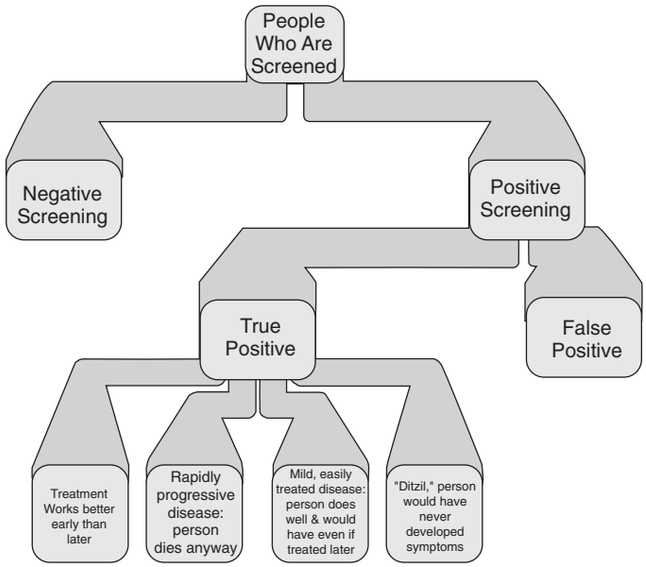

FIGURE 12.2. Cascade of screening.

Potential Harms of Screening: False-Positive Results and Overtreatment

It is difficult to understand how finding cancer earlier could cause harm. In the popular paradigm of cancer being an inexorably progressive disease, the idea of harms from screening makes little sense. It is not difficult to understand why people report having little concern about being harmed by a screening test. But in a real world in which not every cancer is an enemy, most of the intermediate lesions never progress to invasive cancer, workups and diagnostic tests have side effects, and cancer treatments can cause suffering of their own, the possibility of doing harm with screening is easier to understand.

The Cascade of Screening

Screening is not a single test, but rather a cascade of events that can result in either benefit or harm (Figure 12.2). The first step is the screening test itself. Although some diagnostic tests can have useful intermediate results, with a screening test the result is either positive or negative. If a recommendation for anything other than continued routine screening comes from the screening test, it is a positive test. The patient is notified that all is not well and that further evaluation of some kind is needed.

Typically, with cancer screening, many more people have a negative screening test than a positive. After a positive test, further workup is required to determine whether the screening test is a true positive or a false positive. The workup may vary depending on the degree of positivity of the screening test or other circumstances. Some people with false-positive tests have ongoing anxiety related to the experience of screening whereas others do not.

People with a true-positive test do not all benefit from earlier detection of their cancer. These people fall into four categories. Category 1 includes people with fast-growing, malignant disease for which treatment is ineffective. These people do not benefit from earlier detection. Category 2 includes people whose cancer is easily treated regardless of when it is detected. These people also do not benefit from earlier detection. Category 3 includes people whose cancer would never have caused important clinical problems and does not need to be treated. These people have "pseudodisease" and do not benefit from earlier detection. Category 4 includes people whose cancer is more effectively treated earlier, after screening detection, than later, after clinical detection. These are the people who benefit from earlier detection. Treatment of people from category 3 is overtreatment; any adverse effects from treatment of this group must be counted as among the harms of screening. This situation happens frequently because category 3 cancers cannot always be accurately distinguished from other categories.

We next consider two categories of potential harms from screening: false-positive screening tests and overtreatment. Although false-negative tests could also theoretically cause harm by false reassurance, we know of little evidence to substantiate this potential harm.

False-Positive Test Results

False-positive screening test results can cause harm both psychologically and by adverse effects from unnecessary workups. Although a positive screening test does not mean that one has the disease in question, it does mean a person has been placed into a higher risk group than previously. That is, the risk of having breast cancer is higher among women with a positive mammogram than among women who have not yet had a mammogram. The individual may experience the uncertainty of not knowing whether she has breast cancer. This occurrence usually causes stress to the individual involved; any delay in the diagnostic workup adds to the person's concern.

The psychologic trauma from a false-positive screening test can be increased by incomplete resolution of the situa-

tion. For example, women with an abnormal mammogram are sometimes asked to return for follow-up mammograms every 6 months (rather than annually). Similarly, some men with a high prostate-specific antigen (PSA) value and a negative prostate biopsy are asked to return in 3 to 6 months for a second set of biopsies because the cancer may have been missed in the first set. People who have had benign colonic polyps removed are sometimes asked to have repeat colonoscopy at more frequent intervals. Some of these people may suffer psychologic stress as a result of prolonging the experience of uncertainty.

Some people who have had seemingly complete resolution of the false-positive screening test (e.g., a woman with an abnormal mammogram who had a negative biopsy and was told she does not have cancer) still have lasting concerns. A study of women after having a false-positive mammogram found that many still had lingering doubts that interfered with sleep or function 6 months after a negative biopsy.[10] A recent study found a similar result among men with a high PSA screening test and a negative biopsy for prostate cancer.[11]

Because false-positive tests lead to workups without clinical benefit, any complication from the workup of a false-positive screening test (e.g., colonic perforation from a false-positive fecal occult blood test) is also a harm from screening. Most workups for positive screening tests will be negative.

Although these psychologic effects and complications from workups may seem of little consequence when compared with the potential for extending life by screening, the weighing of these effects on a population level must take into account that the actual number of false-positive screening tests is far larger than the number of true-positive tests, and larger still then the number of true-positive tests that lead to extended life.

The rate at which a screening test yields false positives is determined by its specificity. Many screening tests have specificities above 90%. Although this sounds very high, specificity in the 90% range guarantees a large number of false-positive tests. This is because specificity is the percentage of all people without the cancer who are classified correctly as having a negative test; 1-specificity is the percentage of people without disease who are incorrectly classified as having cancer (i.e., false positive). In a screening program, however, the number of people without cancer is very large; thus, even 10% (or even 5%) of a large number is still a large number.

In most cases, the number of false-positive screening tests outnumbers true-positive tests by a factor of from 4:1 (e.g., prostate cancer) to 10:1 (e.g., breast cancer) or higher. If we consider the proportion of people who have at least one false-positive screening test over a period of years of repeated screening, the ratio of false-positive to true-positive tests is even larger. In one study, nearly 50% of women had at least one abnormal mammogram over 10 years of annual screening.[12]

Because the prevalence of cancer in a screening population is low, the number of true-positive tests is usually low. If, as noted previously (see Figure 12.2), only a fraction of the true-positive tests lead to extended life, then the number of people who could, over a period of years, potentially suffer the harms of a false positive screening test so far outnumbers the people who may reap the benefits that weighing benefit and harm overall is not straightforward.

As noted previously, improving the sensitivity of a screening test may or may not lead to increased benefits from screening. However, improving the specificity of a screening test often leads to less harm because there are fewer false-positive tests. A smaller number of false-positive tests gives less opportunity for adverse psychologic effects of screening and for adverse effects of negative workups. Thus, improving the specificity of screening tests should often be a priority.

For most tests, whether screening or diagnostic, sensitivity and specificity are inversely related. Thus, increasing the specificity of a screening test may well reduce the sensitivity. The optimal screening test, then, may be neither the most sensitive nor the most specific test, but rather the test (or test cut-point) that gives the optimal trade-off between benefits and harms.

Overdiagnosis and Overtreatment

In addition to false-positive tests, harms may also follow from true-positive tests. Not all people with true-positive screening tests benefit from the earlier detection of cancer. One can think of people with true-positive tests as having cancers in one of four categories.

Category 1: People with an aggressive, malignant cancer may not benefit from screening because the cancer has already metastasized before it can be detected. We are learning, in fact, that some cancers may metastasize within the first few cell divisions, too early to be the target of screening.

Category 2: Other people with slower-growing cancers may be highly treatable even after clinical detection. Testicular cancer may be such a tumor; our treatments are highly effective without the need of early detection. People with such cancers do not benefit from screening.

Category 3: Some people may have pseudodisease, cancers that do not need treatment at all. These people either have intermediate lesions that would not progress but are still considered positive tests (e.g., small colonic adenomas) or have cancer that would not cause clinically important problems for the person in his/her lifetime. These are lesions that appear to be cancer but do not act as we think cancer usually acts. These people cannot benefit from early detection of their "cancer."

Category 4: These are people who can benefit from earlier detection. These people have cancers that are potentially lethal but which can be treated more effectively because they were found earlier. In this case, the criterion is met that the treatment must be more effective if applied after screening detection than later, after clinical detection. Usually, this group of true-positive cancers is a minority of all true positives. The randomized controlled trials of breast cancer screening, for example, tell us that less than 20% of potentially lethal breast cancers (categories 1 and 4) belong to group 4.

A problem with this formulation, however, is that many cancers can only be placed in their proper category retrospectively. That is, the people in category 3, who do not need to be detected or treated, are often initially difficult to distinguish from the other groups. Thus, people in this category are still treated. An example is men with prostate cancer

detected by screening. The majority of men with screening-detected prostate cancer have tumors that are moderately differentiated. Some cancers of this type are potentially lethal whereas others will never cause clinical problems. Because it is impossible to distinguish these cancers with high confidence at the time of diagnosis, virtually all men with this type of cancer are treated. This constitutes overdiagnosis, as we are diagnosing some men with cancer who do not need to be diagnosed, and overtreatment, as we are treating some men who do not need treatment.

The fact of overtreatment is undeniable and likely occurs with many cancers. The most important question is how often it occurs. Determining the number of people in category 3 (the primary group that is affected by overtreatment), however, is not simple. One can consider the issue in either of two ways: pathologically or epidemiologically. These different approaches explain much of the debate about the frequency of "clinically important" prostate cancers.

The pathologic approach to determining the frequency of cancers that do not need treatment uses grade and other cellular prognostic characteristics to determine prognosis at the time of diagnosis. People who are at risk of overtreatment have cancers with more benign characteristics. The problem with this approach is that none of the known prognostic characteristics is able to separate benign from malignant cancers with a high degree of accuracy. For example, one population-based study found that from 40% to 70% of men (depending upon age) with localized Gleason score 7 prostate cancer died of prostate cancer within 15 years of diagnosis.[13] This finding also means that 30% to 60% of men with this type of cancer did not die of prostate cancer in that time. As these men were diagnosed before widespread PSA screening, it is likely that these survival figures would be higher today, independent of any changes in the effectiveness of treatment. Thus, the Gleason score and extent of tumor only give partial information about prognosis, and we are uncertain about whether an individual man will die of prostate cancer.

Another approach is based on the epidemiology of the cancer. This approach examines such issues as the difference between incidence and mortality; trends over time; changes in the effectiveness of treatment; and the lead time produced by the screening test. Using these assumptions with statistical modeling, investigators can calculate an approximation of the proportion of cancers that would not have caused problems during the person's lifetime. The problem with this approach is that it is based on a number of assumptions, at least some of which may be incorrect.

The best way to calculate the percentage of cancers that would never become clinically apparent is an analysis of results from a randomized controlled trial (RCT) of screening, comparing invited and control groups. If the trial screens people in the invited group for several years and then stops screening, the initial increase in incidence usually seen in the invited group compared with the control group should gradually decrease after the end of screening, as the cancers in the control group are detected at a later time. If the cumulative incidence of cancer in the control group never catches up with the invited group, this is evidence of detection by screening (in the invited group) of cancers that would never become clinically apparent. This approach may theoretically underestimate the true frequency of overdiagnosis, however, as it does not count cancers that produce only minimal symptoms (but symptoms sufficient to be diagnosed) in the overdiagnosis category. Although such cancers do cause some symptoms, they may grow so slowly that they would never progress to important clinical problems within the lifetime of the individual. The extent to which such cancers exist is unknown, but they do not need to be diagnosed early.

Overtreatment causes harm in a number of ways. First, the individual has been labeled as a "cancer patient," with likely important consequences for the person's life. Second, most cancer treatments have some side effects, some of which may be long lasting. Thus, in attempting to gain additional life in the future, people must undergo immediate harm from treatment. Finally, the large number of people being treated leads to an exaggerated view by professionals and the public of the true frequency of the cancer and the effectiveness of treatment.[14] Further, *5-year survival* statistics, which improve as more benign cancers are detected and treated, provide an erroneous overestimate of the efficacy of treatment (15) (see following discussion), and many "cancer survivors" are actually people who had either benign-type cancers (category 2) or pseudodisease (category 3) (see Cancer Survivors, later in this chapter).

The Fallacy of 5-Year Survival in Indicating the Effectiveness of a Screening Program

The 5-year survival rate is frequently cited as evidence for the effectiveness of screening in reducing cancer mortality. Nearly every cancer has a longer 5-year survival for early-stage disease than late-stage disease. It should then follow that finding the cancer at an earlier stage leads to improved outcomes and lower mortality.

Factors other than the effectiveness of screening, however, are important determinants of the 5-year survival rate.[15] As survival is defined as the time from diagnosis to death, it is heavily influenced by early detection, even if death is not postponed. Thus, improved 5-year survival for early-stage cancers could simply reflect the stage at which the cancer is found, with no effect of screening on the natural history of the cancer.

A second problem with the 5-year survival rate as a measure of the effectiveness of screening is related to the heterogeneity of cancers with the same name. Cancers diagnosed at an early stage may be pathologically different from cancers diagnosed at a later stage. Screening may have little to do with the higher 5-year survival rate for early-stage cancers: they would have lower malignant potential regardless of how (or when) they were detected.

Biases in Cancer Screening

Several biases may lead one to believe that screening is effective even in situations where it is not. The first of these is "lead time bias" (Figure 12.3). As shown in the figure, lead time is the time by which earlier detection advances diagnosis. If treatment is ineffective, however ("situation 2" in Figure 12.3), then the patient will die at the same time he/she would have without earlier detection ("situation 1"). The patient's "survival," measured from diagnosis, has been prolonged but the patient has not benefited. Thus, studies that compare survival between people whose cancers were

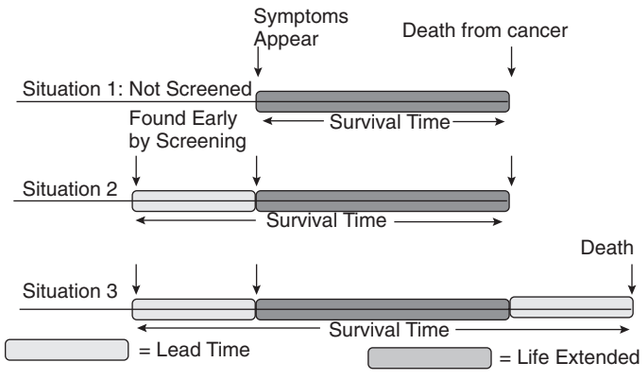

FIGURE 12.3. Lead time bias.

detected by screening with those whose cancer was detected clinically are flawed. The important issue with screening is whether there is a "situation 3," in which people's lives have actually been extended. The best study design to avoid this bias is the randomized controlled trial.

A second bias also may cause people to conclude that screening is useful when it is not. This bias, termed "length-biased sampling" (or "length-time bias"), is associated with the heterogeneity of cancer growth rates and malignant potential (Figure 12.4). Patients 1 and 4 in the figure have rapidly progressive tumors that spend relatively little time in the "detectable preclinical phase" area. As a result, these cancers are often missed by screening tests. Patients 2 and 3, however, have slower-growing, less-malignant cancers that are less likely to be fatal. These cancers spend a longer time in the detectable preclinical phase area and thus are more likely to be detected by screening. Thus, length-biased sampling makes us believe that screening is effective because people with screening-detected cancers do better than people with clinically detected cancers. Slower-growing and less-malignant cancers are preferentially detected by screening programs.

Interestingly, patient 4, whose cancer was detected at a later age, does not die of his or her cancer, even though the cancer is faster growing and malignant, because of competing risks: he or she is more likely to die of another cause at this older age. Patient 4 is not helped by screening.

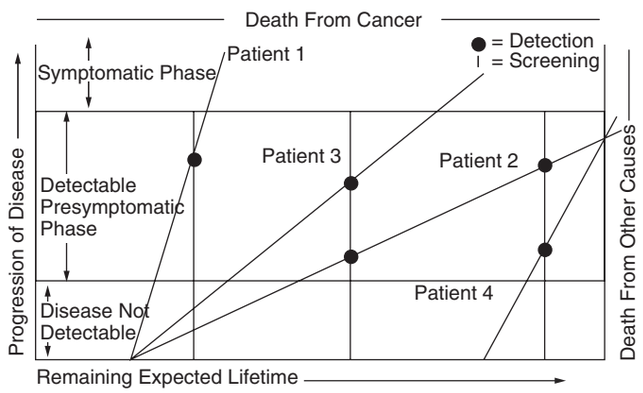

FIGURE 12.4. Length-biased sampling.

Cancer Survivors

The prominence of cancer survivors is a factor in the public's strong interest in cancer screening. A number of people who appear to have been cured of cancer attribute their well-being to detection of their cancer by screening and resultant early treatment. Their testimony to the power of screening contributes to the public's perception that cancer screening is a responsibility. However, at least some of these people likely had either an easily treated benign-type cancer (category 2) or pseudodisease (category 3). Neither of these cancers requires earlier detection. Easily treated cancers are slower growing and can be treated as well after clinical detection as after screening detection. Pseudodisease cancers would never have caused important problems and thus do not need treatment at all. Because it is often difficult to distinguish these cancers at diagnosis from more malignant forms, people who do well after treatment tend to attribute their well-being to screening and early detection. Because many of these people choose to have high public visibility, this creates a bias in favor of the public's view that screening is highly effective in reducing mortality from cancer.

Weighing Benefits and Harms: Decision Making About Screening

Weighing the benefits and harms of screening programs is difficult. First, one must determine the presence and magnitude of the benefits and harms. Benefits seem intuitive, but closer inspection shows that they are easy to overestimate. Evidence about the accuracy or yield of screening tests by itself is inadequate, as is evidence about the effectiveness of treatment in people detected clinically. Etiologic data about trends in mortality over time are open to multiple interpretations. To avoid the strong biases involved, the ideal study design is a randomized controlled trial (RCT) of screening. Even this design, however, is open to criticism, as has been shown by the controversy over the breast cancer screening trials.[16,17] In the end, the determination of the benefits of screening depends on examining evidence from many sources with different designs, and then considering the relevance of the studies to the community setting.[18] One should consider not only whether there are benefits, but how many people benefit and by how much. Judgment is involved in this complex process.

The same process is involved in determining the harms of screening. This part of the equation is often forgotten: a screening program is justified if, and only if, the benefits outweigh the harms of the program. Even RCTs of screening often do not report on the adverse effects caused by screening.

After determining the magnitude of benefits and harms, decision makers need to consider whether one outweighs the other. A problem is that the benefits and harms are usually in different metrics. Benefits should be stated in terms of the estimated number of lives extended in 1,000 people screened over a given time period. Some have claimed reassurance from negative screening tests as a benefit, but careful examination shows that such reassurance is based on questionable assumptions. The actual reduction in the probability of cancer after a negative cancer screening test is very small. Harms should include the number of people in 1,000 screened with false-positive tests, people who are referred for workups

and may suffer psychologic distress. Ideally, one should also estimate the number of people who were overtreated and the consequences of overtreatment.

Weighing the (usually) small number of lives extended against the larger number of people with various types of harms requires a value judgment. What value is placed on the experience of the people suffering harms and what value is placed on the people whose lives are extended? Currently, our culture seems to have decided that even a small number of lives extended by cancer screening outweighs a larger number of people having problems with false positives and overtreatment. It is not clear whether this concept will change as the public becomes more aware of the magnitude of the benefits and harms from cancer screening.

For many cancer screening decisions, some have suggested that the people undergoing the screening should be informed of the potential harms as well as the potential benefits and should be involved in the decision.[19] This approach is called shared decision making (SDM). Research is currently exploring ways of making SDM more feasible and effective in the clinical setting, using decision aids and other decision support resources.

The benefits and harms of cancer screening usually vary by age. The incidence and mortality of most cancers usually associated with screening increase with age. If screening leads to a constant relative reduction in mortality risk, then the absolute reduction in risk (i.e., mortality benefit) increases with age. Thus, younger people with a low probability of having cancer may receive few benefits and expose themselves to important harms by having screening. For example, women in their forties probably receive some benefit from mammography screening for breast cancer, but the magnitude of benefit is small. These women do, however, have a higher probability of having a false-positive mammogram. Given this information, some women will choose to have mammography in their forties and some will not. The age at which the benefits of screening overcome the harms depends on personal values. In general, however, screening is not offered to people at a younger age who have a very low probability of having cancer.

In considering an upper age limit for screening, it is important to remember that any life extension from screening does not occur immediately after screening but rather some years in the future. Thus, to realize the benefit from screening one must live a certain period into the future. As people age, however, the risk of dying of a cause other than cancer increases. Some older people, then, do not live long enough to benefit from screening. Clearly, people with a limited life expectancy have little to gain from screening.

The harms of screening may or may not vary by age. In some cases (e.g., mammography), the screening test may yield fewer false positives in older people, thus decreasing the harms of this finding. In other cases (e.g., PSA for prostate cancer), the screening test may yield more false positives and thus potentially increase the probability of harms. Harms should be carefully considered in every screening decision; the weight they are given by the patient may vary by age. Interventions to reduce psychologic harms by educating people about false positives before screening may be useful but need more research.

An important decision for people who have decided to be screened is the frequency of repeat screening. It is unusual to have RCT evidence about the relative benefits and harms of various screening intervals. More commonly, we reason about this issue with indirect evidence, including our understanding of the natural history of the cancer and measurements of cancer incidence in people who have waited different times to be rescreened.

In general, cancer detection is greatest with the first screening round, the so-called prevalence screen, because there are more asymptomatic cancers to be detected initially than on later screening rounds. Cancers detected in a previous round of screening have been removed from the pool of remaining asymptomatic cancers. Thus, a short screening interval will likely find fewer cancers than a longer interval. The longer the interval, the closer cancer detection will revert to the initial round.

The balance of benefits and harms from various screening intervals is more complex. A shorter screening interval usually means increased sensitivity, but specificity may be reduced. In deciding the most appropriate screening interval, one must consider the trade-offs between finding all appropriate cancers (i.e., increased sensitivity) and increasing the rate of harms related to false positives.

The value of cancer screening after previously negative screening tests is uncertain and needs further study. It is possible that various results from previous screening tests (e.g., men with very low PSA values or women with benign types of parenchymal findings on a mammogram) may be markers of people at decreased risk of developing cancer. These markers may help us define a group of people who do not need further screening, thus allowing us to target screening to people who have the greatest possibility of benefit. If by defining a low-risk population (rather than a high-risk population) we could reduce the number of people requiring screening (thus reducing costs and potential false positives and overtreatment), the balance between the benefits and harms of screening could be improved.

The Costs and Cost-Effectiveness of Screening

Even after gathering the evidence about the benefits and harms of screening, policy makers must still ask the question of whether the net benefits are worth the costs and resource utilization. It may be, for example, that annual abdominal computed tomography (CT) scans detect unsuspected cancers of several types and even sometimes extend a life. But the costs of such a strategy (leaving aside the likely harms for the large number of false positives) may be prohibitive. In other words, the benefits may not be worth the cost.

Opportunity costs are also important. That is, if clinicians spend large amounts of time discussing cancer screening that has little probability of benefit, this may take away from time that could be used, for example, to counsel people about stopping tobacco use, or spending more time on SDM for potentially beneficial cancer screening.

Cancer Screening Examples

To illustrate these principles, we have included examples of screening from four different cancers (cervical, prostate, breast, and colorectal) among those most commonly consid-

ered for screening. In each case, there is the clear potential for benefits and harms. In each case, the benefits are not large in an absolute sense, whereas the harms are not inconsequential. Rational people may decide to have or not to have screening for these cancers based on the same understanding of the evidence. It is important for the public to come to a better understanding of the potential benefits and harms of cancer screening.

Example of Cervical Cancer Screening

Incidence and Mortality

In 2004, an estimated 10,520 new cases of and 3,900 deaths from invasive cervical cancer were expected.[20] In 2000, the age-adjusted incidence rate in nine Surveillance, Epidemiology, and End Results Program (SEER) registries was 8 per 100,000 women; the age-adjusted mortality rate was 3 per 100,000.[21] From 1950 to 1970, the incidence and mortality rates of invasive cervical cancer fell impressively by more than 70%. From 1970 to 2000, the rates continued to decrease by more than 40%.[22] This trend has been attributed largely to screening with the Papanicolaou (Pap) test.

Screening Tests

The Pap test, the standard screening test for cervical cancer, has never been studied in an RCT. A large body of consistent observational data, however, supports its effectiveness in reducing mortality from cervical cancer. Both incidence and mortality from cervical cancer have sharply decreased in a number of large populations following the introduction of well-run screening programs.[23–26] Reductions in cervical cancer incidence and mortality were proportional to the intensity of screening.[22,27]

Case-control studies have found that the risk of developing invasive cervical cancer is 3 to 10 times greater in women who have not been screened.[28–32] Risk also increases with longer duration following the last normal Pap test, or similarly, with decreasing frequency of screening.[33,34] Screening every 2 to 3 years, however, has not been found to increase significantly the risk of finding invasive cervical cancer above the risk expected with annual screening.[34,35]

The precise sensitivity and specificity of Pap tests has been difficult to determine because of the methodological limitations of studies.[36] Studies that compare the Pap test with repeat Pap testing have found that the sensitivity of any abnormality on a single test for detecting high-grade lesions is 55% to 80%.[7,8] Because of the usual slow-growing nature of cervical cancer, the sensitivity of a program of regular Pap testing is likely higher.

Specificity of the Pap test is probably above 90%; it seldom categorizes a woman without any degree of cervical intraepithelial neoplasia (CIN) as having anything more than a mild cytologic abnormality. Specificity is lower, however, for women with mild, clinically unimportant degrees of dysplasia, who are often categorized as having cytologic abnormalities that require further testing and even treatment. Women with such cytologic findings as atypical squamous cells of undetermined significance (ASCUS) are often shown on further evaluation to have neither severe dysplasia nor invasive cancer. If these women are counted as false positives, then specificity will be calculated as lower.[7]

Newer techniques that employ liquid-based cytology (e.g., ThinPrep) have been developed to improve the sensitivity of screening. As with the Pap test, the optimal studies to determine the sensitivity and specificity of these technologies have not been done. Some less than optimal studies show that sensitivity is modestly higher for detecting any degree of CIN, with modestly lower specificity.[37,38] One careful study, however, showed that conventional Pap testing was slightly more sensitive and specific than liquid-based cytology.[39]

The evidence is also mixed about whether liquid-based techniques improve rates of test adequacy.[37,38] One advantage of liquid-based cytology is that human papillomavirus (HPV) testing can be done on the same preparation; one disadvantage is that liquid-based approaches are more expensive than conventional Pap testing. No study has examined whether liquid-based cytology actually reduces the number of women dying of cervical cancer compared with conventional Pap testing.

Rationale for Screening

Invasive squamous carcinoma of the cervix results from the progression of preinvasive precursor lesions called cervical intraepithelial neoplasia (CIN), or dysplasia. Not all these lesions progress to invasive cancer; many mild and moderate lesions regress. The rate at which invasive cancer develops from CIN is usually slow, measured in years and perhaps decades.[40] This long natural history provides the opportunity for screening to effectively detect this process during the preinvasive phase, thus allowing early treatment and cure. Because many of these preinvasive lesions (especially low-grade lesions) would have never progressed to invasive cancer,[41–43] screening also runs the risk of leading to treatment of women who do not need to be treated. This approach leads to harms of screening by overtreatment.

The leading etiologic factor in the development of preinvasive and invasive cervical cancer is infection with specific types of HPV transmitted by sexual contact. Thus, women who are not sexually active rarely develop cervical cancer, whereas sexual activity at an early age with multiple sexual partners is a strong risk factor. About 95% of women with invasive cervical cancer have evidence of HPV infection.[44–47] Many women with HPV infection, however, never develop cervical cancer; thus, this infection is necessary but not sufficient for the development of cancer.[48]

Harms of Screening

The major potential harm of screening for cervical cancer lies in the detection of many lesions [such as most cases of low-grade squamous intraepithelial lesions (LSIL)] that would never progress to cervical cancer. Women with abnormal LSIL or high-grade squamous intraepithelial lesions (HSIL) on Pap testing are usually referred for colposcopy and treated with cryotherapy or loop electrosurgical excision procedure (LEEP), which permanently alters the cervix and has unknown fertility and pregnancy consequences. As younger women have the highest incidence of acquisition of HPV and LSIL, they are disproportionately at risk of receiving intervention for a condition that often spontaneously resolves.

The cost of newer screening methods is also problematic. A cost-effectiveness analysis found little effect on life expectancy with the new technologies when used for annual screening.[49] They may be more cost-effective when used on a less frequent (e.g., every 3 years) basis.

Balance of Benefits and Harms

Based on an analysis of screening records from nearly 350,000 women in Bristol, England, investigators projected that 1,000 women would need to be screened for cervical cancer for 35 years to prevent 1 death from the disease.[50] For each death prevented, the authors estimated that more than 150 women have an abnormal result, more than 80 are referred for investigation, and more than 50 receive treatment.

Annually in the United States, 50 million women undergo screening; about 3.5 million (7%) will be referred for further evaluation. Of these, more than 2 million will be referred for further evaluation of atypical squamous cells of undetermined significance (ASCUS).[51] Fewer than 11,000 cases of invasive cervical cancer were expected in 2004. Thus, Pap test screening results in a large number of colposcopies for benign conditions. Strategies to improve the specificity of the cervical cytopathology test are being evaluated by the ASCUS/LSIL Triage Study (ALTS).[52] Improved specificity, even at the cost of sensitivity, will likely improve the balance between benefits and harms, given the large burden of false positives, abnormalities that do not represent risk for invasive cancer or death.

Improved cervical cancer screening practices may also favor a more positive benefit-to-harm balance. Such practices include not screening women who have had hysterectomies (with removal of the cervix) for benign disease. More than one-third of U.S. women have had hysterectomies by age 65, more than 90% of which are done for noncancer indications.[53] These women rarely have important abnormalities on Pap testing.[54,55] In addition, continued Pap test screening for women over age 65 who previously have had regular cervical cancer screening with normal test results provides little benefit. The risk of cervical cancer and yield of screening decline steadily through middle age.[56] The majority of older women who are found to have invasive cervical cancer have not been screened recently, if at all.[53] Thus, the focus of cervical cancer screening practices should be on finding and screening women at increased risk because of inadequate past screening rather than continuing to screen women at low risk.

Example of Prostate Cancer Screening

Incidence and Mortality

The American Cancer Society estimated that, in 2004, 230,110 men would be diagnosed with prostate cancer; 29,900 men would die of this disease.[20] The age-adjusted prostate cancer incidence in nine Surveillance, Epidemiology, and End Results (SEER) registries between 1996 and 2000 was about 173 per 100,000 men.[21] The mortality during that period was about 33 per 100,000 men. The probability at birth of being diagnosed with prostate cancer by age 80 is about 14%; the probability at birth of dying of this disease by age 80 is about 1.26%.[22] The difference between prostate cancer incidence and mortality is one of largest for any cancer; this difference increased greatly after PSA screening became widespread. This is a strong indication that at least some prostate cancers now detected by screening would never become clinically important.

The incidence of prostate cancer increased dramatically after the beginning of PSA screening in the late 1980s and then stabilized in the later 1990s. Mortality from prostate cancer decreased after about 1992, a total reduction of about 20% by 2000. Screening is one of several possible interpretations of this reduction in mortality.[57]

Screening Tests

The two most common screening tests for prostate cancer are prostate-specific antigen (PSA) and digital rectal exam (DRE). No well-conducted RCT of prostate cancer screening has been completed; two large studies are under way.

Because of the uncertainty about which prostate cancers are clinically important, the sensitivity and specificity of screening is difficult to determine. DRE detects fewer cancers than PSA. Various approaches have been suggested for increasing the sensitivity and specificity of screening, but whether these approaches improve detection of clinically important cancers and reduce detection of unimportant cancers is unknown.[58]

Rationale for Screening

Because of the absence of clear evidence that screening reduces mortality from prostate cancer, the rationale for screening is not established. However, many men are still being screened.[59] Many believe that the ecologic evidence (showing a reduction in mortality after the start of PSA screening) justifies screening; others find that screening has a strong intuitive appeal.

A single well-conducted RCT compared radical prostatectomy and watchful waiting in men with clinically detected prostate cancer.[60] After 8 years, fewer men in the prostatectomy group had died of prostate cancer [13.6% versus 7.1%; absolute difference, 6.6% (2.1%–11.1%)]. The groups did not differ in all-cause mortality. As the cancers in this study were more advanced than those usually detected by screening PSA, this study does not provide adequate evidence about the effectiveness of screening.

Harms of Screening

Two major sources of the harms of prostate cancer screening are false positives and overdiagnosis and overtreatment. False-positive tests are common. On the initial screening round, from 5% to 27% of men (depending on age) have a PSA greater than 4.0 (the traditional cut-point); about 30% of these men will have prostate cancer diagnosed by biopsy.[57] A problem for men with a negative biopsy is that biopsies often miss some prostate cancers; thus, even a negative biopsy does not assure a man that he does not have cancer,[61] and this uncertainty could increase anxiety.

The frequency of overdiagnosis and overtreatment of prostate cancer caused by screening is uncertain. Surveillance data show a large increase in the number of new cases of prostate cancer, with only a small absolute reduction in mortality, after the introduction of PSA screening in the late 1980s.

If, as seems likely, most of the new cases detected would never have been fatal, then more than half of screening-detected prostate cancers do not require major treatment.

Although a small percentage of prostate cancers have histologic characteristics that reliably predict either a very small or a very large malignant potential, most prostate cancers have intermediate histology, leaving us uncertain about the likely prognosis. Because of the inability to determine prognosis from clinical and histologic data, most men under age 70 years receive aggressive treatment: either radical prostatectomy or radiation therapy. These treatments have important adverse effects, including impotence and incontinence, for some 50% of men being treated.[57] Thus, if there are a substantial number of men who do not need treatment but receive it, many of them will be harmed unnecessarily. The exact magnitude of this problem is uncertain, but it may be quite large.

Balance of Benefits and Harms

For screening for prostate cancer, the benefits are not clear whereas the harms are very clear. Thus, the net balance between the two is currently impossible to determine. Given this information, some men will choose to have screening while others will choose not to be screened. Several professional associations and expert groups recommend shared decision making (SDM), informing men of the pros and cons of screening and encouraging them to participate in the decision about whether to be screened.[62–67]

Example of Breast Cancer Screening

Incidence and Mortality

The American Cancer Society estimated that, in 2004, 215,990 women would be newly diagnosed with breast cancer; 40,110 would die of this disease.[20] About 59,390 women will be diagnosed with carcinoma in situ of the breast, primarily by mammography.[20] From 1996 to 2000, the age-adjusted incidence in nine Surveillance, Epidemiology, and End Results (SEER) registries was about 137 per 100,000 women; the age-adjusted mortality during this period was about 28 per 100,000.[21] The probability at birth of being diagnosed with breast cancer in 80 years of life is about 11%; the probability at birth of dying of breast cancer by age 80 is about 2%.[21] Breast cancer incidence for all women increased from 1980 to 2000, although the increase slowed considerably in the late 1990s. Between 1990 and 2000, breast cancer mortality for all women decreased by about 2.3% per year.[22,68] The reasons for this decrease are not clear and may be due to a combination of screening and improved treatment.[69]

Screening Tests

Three primary tests are currently in use for breast cancer screening. Although still controversial, the overall evidence shows that mammography results in a reduction in breast cancer mortality by less than 20%.[6] Indirect evidence suggests that clinical breast examination (CBE), when well conducted, may also lead to a small reduction in mortality, but uncertainty about this remains.[70,71] Breast self-examination (BSE) has been shown in a large RCT to be ineffective in reducing mortality.[72]

The accuracy of mammography depends on a number of factors. One large prospective cohort study of 329,495 women of ages 40 to 89 years from seven population-based mammography registries found sensitivity ranged from 62.9% in women with dense breasts to 87% in women with fatty breasts. Specificity ranged from 89.1% in women with dense breasts to 96.9% in women with fatty breasts.[73]

The accuracy of mammography varies among radiologists and among countries.[74–78] In general, North American radiologists tend to interpret a higher percentage of mammograms as positive than radiologists in other countries, without evident additional benefit. In one study of community radiologists in New England, false-positive rates ranged from 2.6% to 15.9%.[76] The accuracy of CBE also varies widely among clinicians.[79]

Newer approaches to breast cancer screening are being studied, targeted especially to increasing sensitivity.[80] Interestingly, although earlier mammography from the 1970s and 1980s was certainly less sensitive than present-day mammography, the Health Insurance Project (HIP) study from this era found a similar reduction in breast cancer mortality as more recent studies.[81,82] Thus, it is not clear whether increasing sensitivity will provide additional reduction in breast cancer mortality.

Rationale for Screening

The primary rationale for screening comes from the RCTs of screening that have been conducted over the past 30 years.[17,83] Although the overall evidence suggests that breast cancer mortality is reduced by mammographic screening, the reduction is less than 20%[6]; this means that 80% of the women who have potentially fatal cancers are not helped by screening and earlier treatment. Clearly, some breast cancers are aggressive and metastasize before they can be detected by mammography. Some, however, respond better to earlier treatment than to later, thus reducing mortality.

Given the relatively low reduction in mortality from breast cancer from screening, the absolute number of women whose lives would be extended is small. From one to two women in their forties and from two to six women in their fifties and sixties would have their lives extended by screening annually for 10 years.[84]

Harms of Screening

The two major potential harms of screening for breast cancer are false-positive tests and overtreatment of ductal carcinoma in situ (DCIS). One study estimated that 49% of women would have at least one false-positive mammogram after 10 rounds of screening; almost 19% would undergo a biopsy as a result of the false positive.[12] False-positive mammograms sometimes lead to a recommendation of a short-interval follow-up (e.g., 6 months rather than a year), despite the evidence that such a policy rarely leads to increased cancer detection.[85] False-positive mammograms do lead to increased anxiety, both in the short run and after 6 to 12 months, for some women.[10,86]

Ductal carcinoma in situ (DCIS) is a heterogeneous intermediate lesion with an uncertain prognosis. This lesion was rare before screening mammography but has increased dramatically as the number of women undergoing mammograms has increased. About 1 in 1300 mammograms detects DCIS;

from 16% to 28% of all breast "cancers" are DCIS.[87] Probably less than 50% of untreated women with DCIS ever develop invasive breast cancer.[88–90] Treatment is often surgical; some women have mastectomy whereas others have breast conservation surgery. Few women treated for DCIS eventually die of breast cancer.[91]

Because DCIS is so common (an estimated 59,000 cases in 2004)[1] and because its prognosis is so uncertain, many women undergo unnecessary surgery because of its diagnosis. This is an important area of overtreatment. One modeling study found that detection of DCIS plays a minor role in the reduction in breast cancer mortality from screening.[92]

Balance of Benefits and Harms

Screening for breast cancer is an important example of the trade-offs involved in the decision to be screened. On the one hand, screening likely does extend some women's lives. On the other hand, screening also leads to many women having workups for false-positive screening tests, and other women having treatment for DCIS, a lesion that would never develop into invasive breast cancer for many women. It is important for women to understand these trade-offs; women should be offered the opportunity to participate in the decision about screening.

Improving screening programs should seek not only to improve sensitivity. Improved sensitivity may or may not further reduce mortality. Improved specificity should also be a priority. If the number of women with false-positive tests can be reduced, potential harms could be decreased, thus improving the balance between benefits and harms.

Another way of improving breast cancer screening programs would involve finding ways of determining which women with DCIS are truly at risk of invasive cancer, allowing some women to avoid unnecessary surgery.

Example of Colorectal Cancer Screening

Incidence and Mortality

In 2004, an estimated 146,940 new cases of and 56,730 deaths from colon and rectal cancers were expected.[20] Colorectal cancer (CRC) is the third leading cause of new cancer cases (11% of all new cases) and cancer deaths (10% of all cancer deaths) in both men and women.[20] In 2000, the age-adjusted incidence rate in nine Surveillance, Epidemiology, and End Results Program (SEER) registries was 55 per 100,000; the age-adjusted mortality rate was 21 per 100,000.[21] The lifetime risk from birth of being diagnosed with CRC is about 6%; the lifetime risk of dying from CRC is about 2%. Thus, about 1 in 3 people who develop CRC die of this disease. Between 1992 and 2001, mortality from CRC declined by 1.8% per year[93] and incidence declined by 0.8% annually in the United States.[94] The early detection and removal of precancerous colorectal polyps may have contributed to the decline in CRC incidence and mortality.[95]

Screening Tests

The major screening tests currently available for CRC screening are the fecal occult blood test, sigmoidoscopy, colonoscopy, and double-contrast barium enema. These tests are used to identify precancerous or cancerous lesions in the colon and rectum. No one screening strategy has been shown to be superior to the others, although they differ in regard to accuracy, effectiveness, and potential harms.

Fecal occult blood testing (FOBT) has been examined in three RCTs involving more than 250,000 people followed for up to 18 years.[96,97] All three trials found a reduction in CRC mortality from 15% to 33%, with an absolute risk reduction for CRC deaths ranging from 0.8 per 1000 with biennial screening in the United Kingdom over 8 years of follow-up[98] to 4.6 per 1000 with annual screening in Minnesota during 18 years of follow-up.[99] The Minnesota study also noted a 17% to 20% decrease in incidence of CRC.[99] The sensitivity of a single test is approximately 30% to 50%, with a specificity of 90% to 98%, depending on how the test is done. Fecal occult blood tests find about 25% to 50% of patients with colorectal cancer, but only 2% of patients with a positive test had cancer in the Minnesota trial.

The effectiveness of sigmoidoscopy to reduce CRC deaths has been examined in three well-designed case-control studies.[100–102] These studies showed a mortality reduction of 60% to 80%.[97] In a small, randomized trial of sigmoidoscopy, in which persons with polyps were followed up with colonoscopy, the incidence of colorectal cancer was decreased by 80% but no decrease in mortality was found.[103] Using full examination of the colon as the "gold standard," sigmoidoscopy has been found to identify 70% to 80% of patients with advanced adenomas or cancer.[104,105] The sensitivity and specificity of sigmoidoscopy are difficult to determine, because all visible polyps are typically removed, many of which may have little to no malignant potential.

The ability of screening colonoscopy to reduce colorectal cancer morbidity or mortality has not been directly studied to date. Data from studies of other modalities have been extrapolated to support the effectiveness of colonoscopy. Because it is often used as the gold standard, determining its sensitivity and specificity has been difficult. A recent study by Pickhardt et al., comparing optical colonoscopy with CT virtual colonoscopy,[106] in which 1,233 patients underwent both procedures, found the sensitivity of optical colonoscopy for adenomatous polyps to be 88% to 92%, depending on the size of the polyps. As with sigmoidoscopy, the natural history of many polyps found on colonoscopic examination is not known; thus, the potential for identifying false positives must be considered.

No screening studies of double-contrast barium enema with a mortality outcome have been published; thus, the accuracy and effectiveness of this procedure are unknown.[96] Its sensitivity is likely lower than that of endoscopic procedures, but if it misses primarily polyps that are small and not likely to progress to invasive cancer, its effectiveness for screening may be adequate.

Rationale for Screening

A variety of different types of polyps occur in the colon and rectum. Hyperplastic polyps are the most common of those that have little potential for becoming malignant. They cannot be distinguished visually from adenomatous polyps, so biopsy is required for diagnosis. Whether the presence of distal hyperplastic polyps increases the risk of proximal neoplastic polyps is uncertain.[107] A systematic review of 18 studies[108] estimated a 21% to 25% risk for any proximal neoplasia in patients with a distal hyperplastic polyp, includ-

ing a 4% to 5% risk of an advanced neoplasm (cancer or polyp with severe dysplasia or villous histology). In 4 of the studies in which colonoscopy was performed regardless of distal findings, however, the relative risk of finding any proximal neoplasia was 1.3 (95% confidence interval, 0.9–1.8).

Two-thirds of all colonic polyps are adenomatous, which are defined as dysplastic and thus have malignant potential. Most colorectal cancers arise from adenomatous polyps. Some proportion of these grow from small (less than 5 mm) to large (greater than 1.0 cm) to cancer, generally over a period of 10 years or longer. The proportion that makes this transition is thought to be small; adenomatous polyps occur in 30% to 40% of adults over the age of 50, but the risk of developing colorectal cancer is only about 6%.[107] However, removal of adenomatous polyps is associated with a reduced risk of colorectal cancer incidence and mortality.

Harms of Screening

The harms of screening for colorectal cancer include the risk of the screening tests themselves, the risks of the subsequent workup from positive screening tests, the potential for false-negative screening results, and the potential for overdiagnosis and treatment of lesions that would not have become malignant over the person's lifetime. No direct adverse effects of FOBT exist (other than the inconvenience and some patients' distaste for performing the test). Both sigmoidoscopy and colonoscopy are associated with low risks for major complications, including bleeding and perforation of the colon during the examination. A large population-based study of Medicare beneficiaries aged 65 and older found perforation rates of nearly 1 per 1,000 for sigmoidoscopy and 2 per 1,000 for colonoscopy.[109] The risk of death following colonic perforation was 52 to 65 per 1,000 perforations.

Fecal occult blood tests may miss small adenomas, as these lesions frequently do not bleed. Whether that represents a true negative or a false negative is uncertain, as these small adenomas may not be likely to develop into neoplastic lesions. Even though some consider colonoscopy to be the optimal examination of the colon and rectum for detection of precancerous and cancerous lesions, studies have shown that significant lesions (i.e., those larger than 1 cm) may be missed. The Pickhardt study comparing virtual with optical colonoscopy found that virtual colonoscopy missed 5 of 59 advanced neoplasms (defined as adenomatous polyps 10 mm or more in diameter or demonstrating high-grade dysplasia, villous changes, or cancer) and optical colonoscopy missed 7 of the 59 lesions.[106]

The risk of overdiagnosis and treatment of lesions that do not have long-term malignant potential (false-positive lesions) is more difficult to quantify. Most adenomas (60%–75%) are smaller than 1 cm on endoscopic examination.[107] The risk for high-grade dysplasia increases from 1% in small adenomas (less than 5 mm) to 6% for medium-sized adenomas (5–10 mm) to 21% for large adenomas (greater than 1 cm).[107]

Balance of Benefits and Harms

In a recent study of a screening colonoscopy program at a work site,[110] the authors created a clinical index to stratify risk for advanced proximal neoplasia (defined as an adenoma 1 cm or larger or one with villous histology, severe dysplasia, or cancer) and to identify a subgroup at low risk for whom screening sigmoidoscopy alone might be sufficient. Scores were based on age, sex, and distal findings. In the validation arm of the study, the 47% of the cohort determined to be in the low-risk subgroup had a risk for advanced proximal neoplasia of 0.4%. Use of the index in this population identified 92% of persons with advanced proximal neoplasia. The number needed to screen (NNS) to detect advanced proximal neoplasia among patients with any distal polyp was 16 and, among everyone, the NNS was 36. The NNS to extend one life from colorectal cancer mortality was not calculated and would be higher.

Colorectal cancer screening reduces death from colorectal cancer and decreases the incidence of invasive cancer by finding and removing adenomatous polyps. These benefits of screening, however, are tempered somewhat by the effort involved, the harms of the screening procedures themselves, and the possibility of overdiagnosis and overtreatment of small lesions with low malignant potential. As tests with greater sensitivity are developed, the risk of overdiagnosis increases.

References

1. Schwartz LA, Woloshin S, Fowler FJ, Welch HG. Enthusiasm for cancer screening in the United States. JAMA 2004;291: 71–78.
2. Hahn WC, Weinberg RA. Rules for making human tumor cells. N Engl J Med 2002;347:1593–1603.
3. Keyomarsi K, Tucker SL, Buchholz TA, et al. Cyclin E and survival in patients with breast cancer. N Engl J Med 2002; 347:1566–1575.
4. van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 2002;347(25):1999–2009.
5. Ramaswamy S, Perou CM. DNA microarrays in breast cancer: the promise of personalised medicine. Lancet 2003;361:1576–1577.
6. Nystrom L, Andersson I, Bjurstam N, et al. Long-term effects of mammography screening: updated overview of the Swedish randomized trials. Lancet 2002;359:909–919.
7. Soost HJ, Lange HJ, Lehmacher W, et al. The validation of cervical cytology. Sensitivity, specificity, and predictive values. Acta Cytol 1991;35:8–14.
8. Benoit AG, Krepart GV, Lotocki RJ. Results of prior cytologic screening in patients with a diagnosis of Stage I carcinoma of the cervix. Am J Obstet Gynecol 1984;148:690–694.
9. Peto J, Gilham C, Fletcher O, Matthews FE. The cervical cancer epidemic that screening has prevented in the UK. Lancet 2004; 364:249–356.
10. Lerman C, Trock B, Rimer B, et al. Psychological and behavioral implications of abnormal mammograms. Ann Intern Med 1991; 114:657–661.
11. McNaughton-Collins M, Fowler FJ, Caubert J-F, et al. Psychological effects of a suspicious prostate cancer screening test followed by a benign biopsy result. Am J Med 2004;117:719–725.
12. Elmore J, Barton M, Moceri V, et al. Ten-year risk of false positive screening mammograms and clinical breast examinations. N Engl J Med 1998;338:1089–1096.
13. Albertsen PC, Hanley JA, Gleason DF, Barry MJ. Competing risk analysis of men aged 55 to 74 years at diagnosis managed conservatively for clinically localized prostate cancer. JAMA 1998;280:975–980.
14. Black WC, Welch HG. Advances in diagnostic imaging and overestimations of disease prevalence and the benefits of therapy. N Engl J Med 1993;328:1237–1243.

15. Welch HG, Schwartz LM, Woloshin S. Are increasing 5-year survival rates evidence of success against cancer? JAMA 2000; 283:2975–2978.
16. Olsen O, Gotzsche PC. Cochrane review on screening for breast cancer with mammography. Lancet 2001;358:1340–1342.
17. Fletcher SW, Elmore JG. Mammographic screening for breast cancer. N Engl J Med 2003;348:1672–1680.
18. Harris RP, Helfand M, Woolf SH, et al. Current methods of the U.S. Preventive Services Task Force: a review of the process. Am J Prev Med 2001;20(3S):21–35.
19. Sheridan SL, Harris RP, Woolf SH, et al. Shared decision making about screening and chemoprevention. Am J Prev Med 2004;26: 56–66.
20. Jemal A, Tiwani RC, Murray T, et al. Cancer statistics, 2004. CA Cancer J Clin 2004;54:8–29.
21. Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER* Stat Database: Incidence: SEER 9 Regs Public-Use, Nov 2002 Sub (1973–2000). National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 2003, based on the November 2002 submission.
22. Ries LAG, Eisner MP, Kosary CL, et al. (eds). SEER Cancer Statistics Review, 1975–2000. National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2000, 2003.
23. Laara E, Day NE, Hakama M. Trends in mortality from cervical cancer in the Nordic countries: association with organised screening programmes. Lancet 1987;1(8544):1247–1249.
24. Christopherson WM, Lundin FE Jr, Mendez WM, et al. Cervical cancer control: a study of morbidity and mortality trends over a twenty-one-year period. Cancer 1976;38:1357–1366.
25. Miller AB, Lindsay J, Hill GB. Mortality from cancer of the uterus in Canada and its relationship to screening for cancer of the cervix. Int J Cancer 1976;17:602–612.
26. Johannesson G, Geirsson G, Day N. The effect of mass screening in Iceland, 1965–1974, on the incidence and mortality of cervical carcinoma. Int J Cancer 1978;21:418–425.
27. Sigurdsson K. Effect of organized screening on the risk of cervical cancer. Evaluation of screening activity in Iceland, 1964–1991. Int J Cancer 1993;54:563–570.
28. Benedet JL, Anderson GH, Matisic JP. A comprehensive program for cervical cancer detection and management. Am J Obstet Gynecol 1992;166:1254–1259.
29. Aristizabal N, Cuello C, Correa P, et al. The impact of vaginal cytology on cervical cancer risks in Cali, Colombia. Int J Cancer 1984;34:5–9.
30. Clarke EA, Anderson TW. Does screening by "Pap" smears help prevent cervical cancer? A case-control study. Lancet 1979; 2(8132):1–4.
31. La Vecchia C, Franceschi S, Decarli A, et al. "Pap" smear and the risk of cervical neoplasia: quantitative estimates from a case-control study. Lancet 1984;2(8406):779–782.
32. Herrero R, Brinton LA, Reeves WC, et al. Screening for cervical cancer in Latin America: a case-control study. Int J Epidemiol 1992;21:1050–1056.
33. Celentano DD, Klassen AC, Weisman CS, et al. Duration of relative protection of screening for cervical cancer. Prev Med 1989;18:411–422.
34. Screening for squamous cervical cancer: duration of low risk after negative results of cervical cytology and its implication for screening policies. IARC Working Group on evaluation of cervical cancer screening programmes. Br Med J (Clin Res Ed) 1986;293(6548):659–664.
35. Kleinman JC, Kopstein A. Who is being screened for cervical cancer? Am J Public Health 1981;71:73–76.
36. Nanda K, McCrory DC, Myers ER, et al. Accuracy of the Papanicolaou test in screening for and follow-up of cervical cytologic abnormalities: a systemic review. Ann Intern Med 2000;132: 810–819.
37. Hartmann KE, Hall SA, Nanda K, et al. Screening for Cervical Cancer. Rockville, MD: Agency for Health Research and Quality, 2002.
38. McCrory DC, Matchar DB, Bastian L, et al. Evaluation of Cervical Cytology. Evidence Report/Technology Assessment No. 5. AHCPR Publication No. 99-E010. Rockville, MD: Agency for Health Research and Quality, 1999.
39. Coste J, Cochand-Priollet B, de Cremoux P, et al. Cross sectional study of conventional cervical smear, monolayer cytology, and human papillomavirus DNA testing for cervical cancer screening. BMJ 2003;326(7392):733.
40. Holowaty P, Miller AB, Rohan T, et al. Natural history of dysplasia of the uterine cervix. J Natl Cancer Inst 1999;91:252–258.
41. Nasiell K, Roger V, Nasiell M. Behavior of mild cervical dysplasia during long-term follow-up. Obstet Gynecol 1986;69: 665–669.
42. Nash JD, Burke TW, Hoskins WJ. Biologic course of cervical human papillomavirus infection. Obstet Gynecol 1987;69:160–162.
43. Melnikow J, Nuovo J, Willan AR, et al. Natural history of cervical squamous intraepithelial lesions: a meta-analysis. Obstet Gynecol 1998;92(4 pt 2):727–735.
44. Bosch FX, Manos, MM, Muñoz N, et al. Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. International biological study on cervical cancer (IBSCC) Study Group. J Natl Cancer Inst 1995;87:796–802.
45. Wallin KL, Wiklund F, Angström T, et al. Type-specific persistence of human papillomavirus DNA before the development of invasive cervical cancer. N Engl J Med 1999;341:1633–1638.
46. Alani RM, Münger K. Human papillomaviruses and associated malignancies. J Clin Oncol 1998;16:330–337.
47. Walboomers JM, Jacobs MV, Manos MM, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J Pathol 1999;189:12–19.
48. Ho GY, Bierman R, Beardsley L, et al. Natural history of cervicovaginal papillomavirus infection in young women. N Engl J Med 1998;338:423–428.
49. Fox J, Remington P, Layde P, Klein F. The effect of hysterectomy on the risk of an abnormal screening Papanicolaou test result. Am J Obstet Gynecol 1999;180:1104–1109.
50. Raffle AE, Alden B, Quinn M, et al. Outcomes of screening to prevent cancer: analysis of cumulative incidence of cervical abnormality and modeling of cases and deaths prevented. BMJ 2003;326:901–905.
51. Solomon D, Schiffman M, Tarone R, et al. Comparison of three management strategies for patients with atypical squamous cells of undetermined significance: baseline results from a randomized trial. J Natl Cancer Inst 2001;93:293–299.
52. Schiffman M, Adrianza MF. ASCUC-LSIL Triage Study. Design, methods, and characteristics of trial participants. Acta Cytol 2000;44:726–742.
53. Sawaya GF. Should routine screening Papanicolaou smears be done for women older than 65 years? Arch Intern Med 2004; 164:243–245.
54. Fox J, Remington P, Layde P, et al. The effect of hysterectomy on the risk of an abnormal screening Papanicolaou test result. Am J Obstet Gynecol 1999;180:1104–1109.
55. Pearce KF, Haefner HK, Sarwar SF, et al. Cytopathological findings on vaginal Papanicolaou smears after hysterectomy for benign gynecologic disease. N Engl J Med 1996;335:1559–1562.
56. US Preventive Services Task Force. Screening for Cervical Cancer: Recommendations and Rationale. Accessed at http://www.ahrq.gov/clinic/uspstf/uspscerv.htm on Feb. 29, 2004.
57. Harris R, Lohr KN. Screening for prostate cancer: an update of the evidence for the US Preventive Services Task Force. Ann Intern Med 2002;137:917–929.

58. Barry MJ. Prostate-specific-antigen testing for early diagnosis of prostate cancer. N Engl J Med 2001;344:1373–1377.
59. Sirovich BE, Schwartz LM, Woloshin S. Screening men for prostate and colorectal cancer in the United States. JAMA 2003; 289:1414–1420.
60. Holmberg L, Bill-Axelson A, Helgesen F, et al. A randomized trial comparing radical prostatectomy with watchful waiting in early prostate cancer. N Engl J Med 2002;347:781–789.
61. Djavan B, Ravery V, Zlotta A, et al. Prospective evaluation of prostate cancer detected on biopsies 1, 2, 3 and 4: when should we stop? J Urol 2001;166:1679–1683.
62. Prostate-specific antigen (PSA) best practice policy. American Urological Association (AUA). Oncology 2002;14:267–272.
63. American Cancer Society. ACS Cancer Detection Guidelines: Cancer-Related Checkup. Accessed at www.cancer.org on November 23, 2004.
64. American Medical Association. Policy H-425.980 Screening and Early Detection of Prostate Cancer. Accessed at www.ama-assn.org on November 23, 2004.
65. US Preventive Services Task Force. Screening for prostate cancer: recommendations and rationale. Ann Intern Med 2002; 137:915–916.
66. Periodic Health Examinations. Revision 5.6, August 2004. American Academy of Family Physicians. Accessed at www.aafp.org on November 23, 2004.
67. American College of Physicians. Screening for prostate cancer. Ann Intern Med 1997;126:480–484.
68. Ghafoor A, Jemal A, Ward E, Cokkinides V, Smith R, Thun M. Trends in breast cancer by race and ethnicity. CA Cancer J Clin 2003;53:342–355.
69. Jatoi I, Miller AB. Why is breast cancer mortality declining? Lancet Oncol 2003;4:251–254.
70. Barton MB, Harris R, Fletcher SW. Does this patient have breast cancer? JAMA 1999;282:1270–1280.
71. Harris R, Kinsinger LS. Routinely teaching breast self-examination is dead. What does this mean? J Natl Cancer Inst 2002;94:1420–1421.
72. Thomas DB, Gao DL, Ray RM, et al. Randomized trial of breast self-examination in Shanghai: final results. J Natl Cancer Inst 2002;94:1445–1457.
73. Carney PA, Miglioretti DL, Yankaskas BC, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. Ann Intern Med 2003;138:168–175.
74. Smith-Bindman R, Chu PW, Miglioretti DL, et al. Comparison of screening mammography in the United States and the United Kingdom. JAMA 2003;290:2129–2137.
75. Elmore JG, Nakano CY, Koepsell TD, Desnick LM, D'Orsi CJ, Ransohoff DF. International variation in screening mammography interpretations in community-based programs. J Natl Cancer Inst 2003;95:1384–3193.
76. Elmore JG, Miglioretti DL, Reisch LM, et al. Screening mammograms by community radiologists: variability in false-positive rates. J Natl Cancer Inst 2002;94:1373–1380.
77. Beam CA, Conant EF, Sickles EA. Association of volume and volume-independent factors with accuracy in screening mammogram interpretation. J Natl Cancer Inst 2003;95:282–290.
78. Esserman L, Cowley H, Eberle C, et al. Improving the accuracy of mammography: volume and outcome relationships. J Natl Cancer Inst 2002;94:369–375.
79. Fletcher SW, O'Malley MS, Bunce LA. Physicians' abilities to detect lumps in silicone breast models. JAMA 1985;253: 2224–2228.
80. Liberman L. Breast cancer screening with MRI: what are the data for patients at high risk? N Engl J Med 2004;351:497–500.
81. Shapiro S, Venet W, Strax P, Venet L. Periodic Screening for Breast Cancer: The Health Insurance Plan Project and Its Sequelae, 1963–1986. Baltimore: Johns Hopkins University Press, 1988.
82. Fletcher SW, Black W, Harris R, Rimer BK, Shapiro S. Report of the International Workshop on Screening for Breast Cancer. J Natl Cancer Inst 1993;85:1644–1656.
83. Humphrey LL, Helfand M, Chan BK, Woolf SH. Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med 2002;137:347–360.
84. Harris R, Leininger L. Clinical strategies for breast cancer screening: weighing and using the evidence. Ann Intern Med 1995;122:539–547.
85. Yasmeen S, Romano PS, Pettinger M, et al. Frequency and predictive value of a mammographic recommendation for short-interval follow-up. J Natl Cancer Inst 2003;95:429–436.
86. Yasmeen S, Romano PS, Pettinger M, et al. Re: Short-interval follow-up mammography: are we doing the right thing? J Natl Cancer Inst 2003;95:1175–1176.
87. Barton MB, Moore S, Polk S, Shtatland E, Elmore JG, Fletcher SW. Increased patient concern after false-positive mammograms: clinician documentation and subsequent ambulatory visits. J Gen Intern Med 2001;16:150–156.
88. Ernster VL, Ballard-Barbash R, Barlow WE, et al. Detection of ductal carcinoma in situ in women undergoing screening mammography. J Natl Cancer Inst 2002;94:1546–1554.
89. Page DL, Dupont WD, Rogers LW, Landenberger M. Intraductal carcinoma of the breast: follow-up after biopsy only. Cancer (Phila) 1982;49:751–758.
90. Page DL, Jensen RA. Ductal carcinoma in situ of the breast: understanding the misunderstood stepchild. JAMA 1996;275: 948–949.
91. Welch HG, Black WC. Using autopsy series to estimate the disease "reservoir" for ductal carcinoma in situ of the breast: how much more breast cancer can we find? Ann Intern Med 1997;127:1023–1028.
92. Ernster VL, Barclay J, Kerlikowske K, Wilkie H, Ballard-Barbash R. Mortality among women with ductal carcinoma in situ of the breast in the population-based surveillance, epidemiology and end results program. Arch Intern Med 2000;160:953–958.
93. Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER* Stat Database: Mortality: All COD, Public-Use With State, Total U.S. (1969–2001). National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 2004. Underlying mortality data provided by NCHS (www.cdc.gov/nchs).
94. Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER* Stat Database: Incidence: SEER 9 Regs Public-Use, Nov 2003 Sub (1973–2001). National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 2004, based on the November 2003 submission.
95. Weir HK, Thun MJ, Hankey BF, et al. Annual report to the nation on the status of cancer, 1975–2000, featuring the uses of surveillance data for cancer prevention and control. J Natl Cancer Inst 2003;95:1276–1299.
96. Pignone M, Rich M, Teustch SM, Berg AO, Lohr KN. Screening for colorectal cancer in adults at average risk: a summary of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med 2002;137:132–141.
97. Walsh JME, Terdiman JP. Colorectal cancer screening. Scientific review. JAMA 2003;289:1288–1296.
98. Hardcastle JD, Chamberlain JO, Robinson MH, et al. Randomized controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet 1996;348:1472–1477.
99. Mandel JS, Bond JH, Church TR, et al. Reducing the mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Screening Study. N Engl J Med 1993; 328:1365–1371.
100. Selby JV, Friedman GD, Quesenberry CP Jr, Weiss NS. A case-control study of screening sigmoidoscopy and mortality from colorectal cancer. N Engl J Med 1992;326:653–657.

101. Newcomb PA, Norfleet RG, Storer BE, Surawicz TS, Marcus PM. Screening sigmoidoscopy and colorectal cancer mortality. J Natl Cancer Inst 1992;84:1572–1575.
102. Muller AD, Sonnerberg A. Prevention of colorectal cancer by flexible endoscopy and polypectomy: a case-control study among veterans. Arch Intern Med 1995;123:904–910.
103. Thiis-Everson E, Hoff GS, Sauar J, Langmark F, Majak BM, Vatn MH. Population-based surveillance by colonoscopy: effect on the incidence of colorectal cancer. Telemark Polyp Study I. Scand J Gastroenterol 1999;34:414–420.
104. Lieberman DA, Weiss DG, Bond JH, Ahnen DJ, Garewal H, Chejfec G. Use of colonoscopy to screen asymptomatic adults for colorectal cancer. Veterans Affairs Cooperative Study Group 380. N Engl J Med 2000;343:162–168.
105. Imperiale TF, Wagner DR, Lin CY, Larkin GN, Rogge JD, Ransohoff DF. Risk of advanced proximal neoplasms in asymptomatic adults according to the distal colorectal findings. N Engl J Med 2000;343:169–174.
106. Pickhardt PJ, Choi JR, Hwang I, et al. Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. N Engl J Med 2003;349:2191–2200.
107. Lawrence SP, Ahnen DJ. Approach to the patient with colonic polyps. In: Rose BD (ed) UpToDate. Wellesley: UpToDate, 2004.
108. Dave S, Hui S, Kroenke K, Imperiale TF. Is the distal hyperplastic polyp a marker for proximal neoplasia? A systematic review. J Gen Intern Med 2003;18:128–137.
109. Gatto NM, Frucht H, Sundararajan V, Jacobson JS, Grann VR, Neugut AI. Risk of perforation after colonoscopy and sigmoidoscopy: a population-based study. J Natl Cancer Inst 2003; 95:230–236.
110. Imperiale TF, Wagner DR, Lin CY, Larkin GN, Rogge JD, Ransohoff DF. Using risk for advanced proximal colonic neoplasia to tailor endoscopic screening for colorectal cancer. Ann Intern Med 2003;139:959–965.

http://www.springer.com/978-0-387-24291-0

Oncology
An Evidence-Based Approach
Chang, A.E.; Ganz, P.A.; Hayes, D.F.; Kinsella, T.; Pass, H.I.; Schiller, J.H.; Stone, R.M.; Strecher, V. (Eds.)
2006, LXXXII, 2012 p. In 2 volumes, not available separately., Hardcover
ISBN: 978-0-387-24291-0