The purpose of this paper is to clarify the short-term and long-term objectives of screening for various cancers, and to indicate the kinds of data that are needed to determine whether or not the objectives are met. Cancers at various sites differ with respect to their innate suitability for screening. Criteria that enhance screening suitability include the potential for serious complications and a high rate of mortality (applicable to most cancers), a prolonged preclinical phase, and an existing therapy that is simpler and more effective in reducing the mortality rate when applied to preclinical disease than to clinically evident cancer. Tests and procedures suitable for screening are simple to perform, inexpensive, acceptable to patients and physicians, safe, relatively painless, and accurate, as measured by the test's sensitivity and specificity. The actual yield of previously undiagnosed cancer arising from a screening program will depend heavily on prevalence of disease in the screened population, specificity of the screening test, and successful follow-up of screen-positive patients with diagnosis and treatment. These issues are discussed in the context of four cancers and their respective screening modalities: cervical cancer and cytologic studies, breast cancer and mammography, colon cancer and fecal occult blood tests, and lung cancer and sputum cytologic studies. The quality of data on which screening decisions have been made for each of these cancers and tests varies. The cancers vary in terms of their relevant biologic characteristics and treatment effectiveness. Similarly, each screening procedure has its own particular advantages and disadvantages. Current American Cancer Society Guidelines for early detection of three of the cancers are presented.


The relevant criteria for evaluating the suitability of a disease for screening and the adequacy of the screening test are presented. The kinds of scientific evidence that are needed and should be considered before making decisions about the value of screening are emphasized.

What is meant by the term screening? One definition is the identification of unrecognized disease by the application of tests or examinations to apparently well persons to sort out those who probably have a disease from those who probably do not. A few points should be emphasized. First, is the application of tests or procedures to apparently well persons—persons who do not have overt disease or symptoms of the disease of interest. Secondly, we are trying to sort out people who have a greater likelihood of having the disease from those who are not likely to have the disease. People who screen positively require diagnostic workup to determine whether or not they actually have the disease.

Objectives of Screening

The objectives of screening may be discussed in an operational sense or in terms of ultimate goals. An operational objective of screening is the application of a relatively simple, inexpensive test to a large number of persons to classify them as likely or unlikely to have the cancer which is the object of the screen. This objective is very similar to the definition of screening previously noted. However, in terms of the outcomes or ultimate goals, the objective of screening for a particular cancer is to reduce the complications and mortality rate from that cancer among the persons screened. This level of objective may seem remote at the day-to-day level of office practice, but the rationale and importance of this objective should become clearer in the course of this presentation.

Characteristics of Diseases Suitable for Screening

The disease should have serious consequences and these should be recognized as such by potential...
screenees. Serious consequences comprise significant complications and high mortality. In addition, the disease must have a treatment that, when applied to early stage or presymptomatic disease, is more effective than treatment applied after symptoms have appeared. Conversely, there is no point in screening for a disease that can be treated just as successfully after symptoms have appeared and a diagnosis is made in the usual course of clinical care. Nor is it reasonable to screen for an untreatable disease. Moreover, the disease must have a detectable preclinical phase, and this phase of the disease should have a high prevalence among the persons screened.1 If there is not a high prevalence of the disease among the screened group, there will be too few cases detected to justify the expense of a screening program.

As a reminder on definitions, the prevalence, or prevalence ratio, is the proportion of people afflicted with a particular condition at a point in time. For example, the detected prevalence of breast cancer might be five per 1000 women undergoing physical examination and mammography as an initial screening procedure.

The natural history of disease and the duration of the preclinical phase are also important with respect to the utility of screening. This point is illustrated in Figure 1. The arrow represents an implied time frame with the hatched area, bounded by preclinical disease inception and symptom onset, representing the duration of the preclinical phase during which screening is advantageous. Clearly, the longer the preclinical phase persists, the greater the opportunity for screening.

Characteristics of Tests Suitable for Screening

The tests should be simple to perform, inexpensive, and acceptable to patients and the persons performing the procedures. They should be safe, relatively painless, and valid—i.e., provide accurate information about the presence or absence of disease. Since no screening procedure or diagnostic test is 100% accurate, error or misclassification is always inherent in a screening test. However, the accuracy (or validity) of a test can be defined by the measures of sensitivity and specificity.

Table 1 shows the information needed to determine sensitivity and specificity. The letters a, b, c and d are symbolic representations for actual numbers. The way sensitivity and specificity are usually determined is by accumulating groups of persons who truly have the disease (a + c) and a comparable group who do not (b + d). The people in the a cell of the table are those who have the disease and who are identified as such by the screening test. Thus the sensitivity of the test is defined as a/(a + c). Sensitivity represents the likelihood that the test will be positive among people who truly have the disease.

Test specificity is represented by d/(b + d). Specificity indicates the probability of a negative test result among people who truly do not have the disease. Another way of viewing the letters in Table 1 is that the b cell represents the false-positive group, and the c cell represents the false-negative group. The larger these numbers appear, the poorer is the specificity and sensitivity, respectively, of the screening test.

A screening test should ideally have both high sensitivity and high specificity. To some extent this is possible, but in general, there is a trade-off; sensitivity can be increased with a resulting decrease in specificity or specificity can be improved with lower sensitivity resulting.

Table 2 provides an estimate of sensitivity and specificity for several screening procedures. For every procedure, there is a range of sensitivity and specificity estimates recorded in the literature, numbers that represent a middle ground have been extracted. In general, the specificity of cervical cytology is not very good, resulting in a high false-positive rate.5-4 The specificity is lower if minimal levels of atypia are considered positive whereas specificity is higher when only cells consistent with cervical intraepithelial neoplasia (CIN) III are considered positive. Mammography, however, is very sensitive and specific.5-7 References can also be cited for colon cancer8-10 and lung cancer.11

There are a number of pertinent issues contributing to the large variation in reported sensitivities and specificities of screening tests. The decision as to where to put the cutoff between a positive and a negative test result

CaNeR SCREENING • Hulka

No. 8

1777
TABLE 2. Screening Tests for Selected Cancer Sites

<table>
<thead>
<tr>
<th>Site</th>
<th>Test</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uterine cervix</td>
<td>Cytology</td>
<td>67</td>
<td>60</td>
</tr>
<tr>
<td>Breast</td>
<td>Mammography</td>
<td>85</td>
<td>80</td>
</tr>
<tr>
<td>Colon</td>
<td>Fecal blood</td>
<td>50</td>
<td>?</td>
</tr>
<tr>
<td>Lung</td>
<td>Sputum cytology</td>
<td>42</td>
<td>99</td>
</tr>
</tbody>
</table>

will affect both measures, and usually in opposite directions; i.e., sensitivity can be improved at the expense of specificity, and vice versa. Furthermore, both measures will vary depending on the source and characteristics of the diseased and non-diseased population, prevalence of disease, frequency of screening and whether the screen is an initial or subsequent event.

Measures of Screening Effectiveness

Short-Term

Once a decision is made to undertake screening as a programatic activity or within an office practice, it is important to know whether or not the operational and long-term objectives of screening are being met. A variety of measures can be used to evaluate the effectiveness of screening, as noted in Table 3. These operational measures concern the numbers of persons screened, the yield of the program in terms of preclinical disease detection rates, dollar costs, and the proportion of positive screenees brought to definitive diagnosis and therapy. Lastly, the predictive value of a positive test is noted. The positive predictive value indicates the probability that a person with a positive test, on further diagnostic workup, will have the disease in question. Two factors are the main determinants of that value; one is the specificity of the test and the other is prevalence of preclinical disease in the group being screened. The higher the specificity and the higher the prevalence, the higher the positive predictive value will be. This is a very important consideration in measuring the effectiveness of screening. Whereas a high positive predictive value is very desirable, in many settings and programs only 10% or 20% of people who screen positive ultimately will be diagnosed with the disease in question.

The positive predictive value of a screening test is illustrated in Table 1. \((a + b)\) represents the number of people with positive test results as reported to the clinician or program administrator. Of that number, only the proportion represented by \(a/(a + b)\) will subsequently be diagnosed with the disease.

Long-Term

How is the long-term effectiveness of a screening activity determined? This is difficult to demonstrate in any one practice setting because of limited numbers of patients. The empiric demonstration of the benefits of screening through outcome measures requires a large practice base or population-based data. Three measures of cancer outcome should be considered.

1. Stage distribution of detected cancers. The distribution of cancers detected should shift toward earlier stage disease.
2. Case fatality rate. The case fatality rate should be lower among the screened than among the unscreened patients.
3. Site-specific cancer mortality rate. The site-specific cancer mortality rate should be reduced among the screened population, an observation which can be made only from a large population base. This is the ultimate test of screening effectiveness and the only one which is without bias.

Quality Control in Screening

Some general points relevant to screening activities in any setting deserve mention. First, know the screening test, particularly its sensitivity and specificity, which in turn determine the frequency of false-negatives and false-positives. Secondly, the issue of quality control should be considered in performing and reading the test. With the Pap smear, for example, the source of specimens and methods for taking the smear are important in obtaining representative cell samples. Continual quality control and monitoring in the laboratory that reads the smears is necessary. Thirdly, the organizational structure should exist to provide for diagnostic evaluation and treatment. Screening of itself is of no value without the necessary diagnostic workup and treatment capabilities. In physician's offices, this is usually not a problem because the physician who screens also does the diagnostic workup or has available the appropriate referral mechanisms to other health-care facilities. The need for referral can be a major deterrent to conducting screening activities at sites other than health-care facilities. The fourth general point concerns the need for collecting a
minimum data set. At the least, these data items include the numbers of persons examined and their test results, the results of the diagnostic and therapeutic procedures for positive screenees, and health status at follow-up for those with diagnosed disease.

Selected Examples of Specific Cancers and Tests

To illustrate these points, let us consider selected cancer sites for screening, in view of the criteria previously noted. These related to the burden of suffering caused by the disease, the effectiveness of treatment of preclinical disease, the adequacy of the screening test, and the advantages and disadvantages of screening.

Cervical Cancer and Cytology

Cancer of the cervix detected by cervical cytologic studies is a classical example. The Pap smear is attractive in that it is simple to perform, safe, acceptable to patients and doctors, and inexpensive. The sensitivity is adequate but the specificity is low, which produces a high false-positive rate and a low positive predictive value. The latter can be improved somewhat by requiring two consecutive abnormal smears before engaging in a diagnostic workup.

The primary feature about cervical cancer, making it suitable for screening, is its long preclinical phase—5 to 10 years or more—providing sufficient opportunities for periodic screening and detection of preinvasive disease.

Therapy of preclinical disease, (CIN I, II, and III) is relatively simple since hysterectomy or much lesser procedures are effective in removing the lesion. Thus, death from invasive cervical cancer is prevented. The mortality rate from cervical cancer has been declining for the last 30 years. Data of several types indicate that screening with cervical cytologic studies is at least partly responsible for this reduction.

Breast Cancer and Mammography

What is the role of mammography in breast cancer detection? Data from randomized controlled trials in the US7 and Sweden14 have shown a reduction in case fatality and in breast cancer mortality of about 30% among women receiving mammography. An overall reduction in breast cancer mortality among the screened group is strong evidence for effectiveness of mammography.

As a test procedure, the mammogram has high sensitivity and specificity compared with many other tests and procedures, and in general, it meets the other criteria for a screening test. One of the negative features of the procedure is the possible risk from radiation exposure. Over the years, this risk has markedly lessened as screening technology has been modified and the doses of radiation have been reduced. Currently, the mean glandular dose to each breast is less than 0.5 rad per screening examination, which consists of two different views of each breast. At North Carolina Memorial Hospital in Chapel Hill, this figure is nearer 0.1 or 0.2 rad per examination.15 Questions about mammography still remain, however, particularly with regard to women younger than age 50 years. How often should the procedure be performed, and what is the effect on breast cancer mortality? We do not have the same convincing data of a reduction in breast cancer mortality among women younger than age 50 years that we do for women older than 50 years.

Colorectal Cancer and Fecal Occult Blood

Our third condition is colorectal cancer, for which one of the recommended screening procedures is testing for fecal occult blood. A variety of these tests are commercially available, and the colon cancer detection rate in unselected populations is about one or two per 1000 procedures.16 The tests are safe, relatively simple, and inexpensive, but patient acceptability is a problem. Multiple slide tests are required on multiple stool specimens following several days of dietary proscriptions and restrictions. This inhibits acceptability for many people. A first-order priority is to develop a more acceptable and accurate screening test for colorectal cancer.

Lung Cancer and Sputum Cytology

Although the American Cancer Society does not have guidelines for lung cancer screening, randomized clinical trials of sputum cytology obtained three times annually, in addition to annual chest x-rays, have been completed.17–19 The trials have not demonstrated an improvement in lung cancer mortality in the group of subjects receiving cytologic examinations plus annual chest x-rays as compared with those receiving annual chest x-rays.
x-rays alone. These data provide evidence against a beneficial effect of sputum cytology in terms of the ultimate goal of screening—to improve lung cancer mortality rates. There are those, however, who advocate sputum cytology for very high-risk groups, such as persons who smoke and are also exposed to occupational carcinogens like asbestos. At least for monitoring purposes, cytologic studies may have a role in such settings. However, the diagnosis of lung cancer following a positive sputum cytology may be problematic if the lesion detected is too small to be evident on chest x-rays. Abnormal cells per se are not helpful in directing bronchoscopic examination which is necessary for diagnosing and locating the lung cancer. The main problem with screening for lung cancer is its biologic aggressiveness. The natural course of this disease from its inception to clinical presentation appears to be so rapid that it is not amenable to currently existing screening modalities.

**Current American Cancer Society Recommendations**

Tables 4, 5, and 6 provide the current American Cancer Society screening recommendations for three cancer sites: the breast, uterine cervix, and colon/rectum.