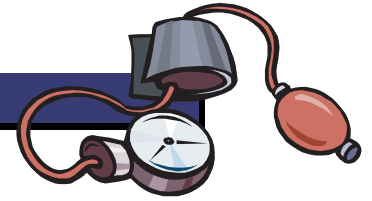


Screening in Public Health Practice

Transcript



Screening in Public Health Practice

Welcome to “Screening in Public Health Practice.” I’m Victoria Holt, your narrator for this module. As a nurse, I’ve worked in a variety of hospital and clinical practice settings, including public health clinics in East Tennessee and North Carolina. More recently, as an epidemiologist, I’m a faculty member at the Northwest Center for Public Health Practice at the School of Public Health and Community Medicine at the University of Washington in Seattle.

Since 1993, I have also been a faculty member in the Department of Epidemiology at the University of Washington, where I currently teach courses in epidemiologic methods.

About This Module

I’d like to mention a few points that may help make your learning experience more enjoyable.

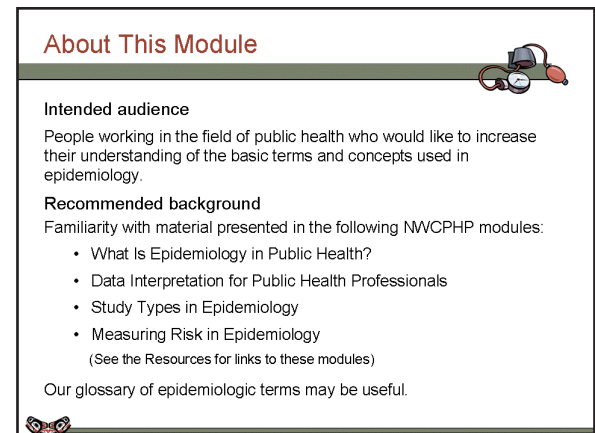
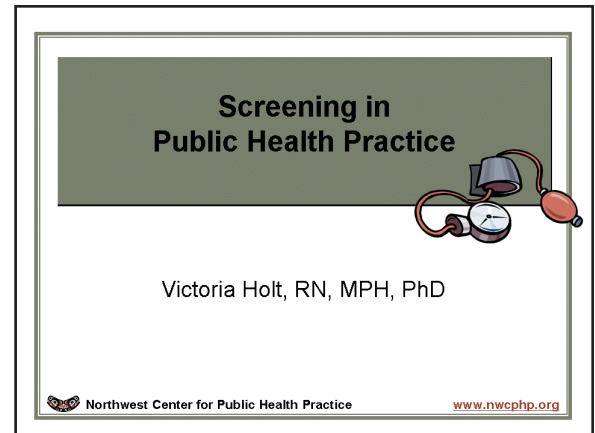
This module and others in the epidemiology series from the Northwest Center for Public Health Practice are intended for people working in the field of public health who are not epidemiologists but who would like to increase their familiarity with and understanding of the basic terms and concepts used in epidemiology. Before you go on with this module we recommend that you become familiar, if you haven’t already, with the background material presented in the following modules, which you can find on the Center’s Web site or under the Resources link on the attachments tab: What Is Epidemiology in Public Health? Data Interpretation for Public Health Professionals; Study Types in Epidemiology; Measuring Risk in Epidemiology.

We introduce a number of new terms in this module. If you want to review their definitions at any time, the glossary in the attachments link at the top of the screen may be useful.

Module Objectives

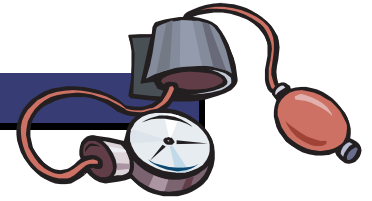
This course offers an overview of the concepts and use of screening for disease.

By the end of this module you should be able to define screening and describe its role in the work of public health practitioners. You will also be able to use the



Screening in Public Health Practice

Transcript



definition of screening to determine diseases for which it is appropriate to screen and to determine which tests are appropriate to use for screening, based on the various characteristics of a good screening test. You will learn how to list and describe important factors to consider when designing and evaluating screening programs. Finally, you should be able to help clients interpret the results of their screening tests through knowledge of the concept of predictive value.

Let's start this module by defining the word screening as it's used in public health.

What Is Screening?

Screening essentially is using tests to detect the likely presence of diseases or conditions before they cause illness or symptoms. According to the U.S. Commission on Chronic Illness, "screening is the presumptive identification of unrecognized disease or defect by the application of tests, examinations, or other procedures that can be applied rapidly." One key aspect of this definition is the word "presumptive." Screening is not the same as diagnosis, which is the definite identification of disease.

As shown in this diagram, screening is done on an asymptomatic population, that is, people who show no evidence yet of the disease for which they are being screened. The results of the screening test separate them into two categories—those who are likely to have the disease, that is, they screened positive, and those who screened negative and thus are unlikely to have the disease. Those who screen as likely to have disease then undergo the definitive diagnostic tests that come after and are separate from the screening process to determine whether they in fact have the disease for which they have screened positive. Those who are determined by screening to be unlikely to have the disease are not examined further as part of the screening process.

Examples of Screening

Among many examples of screening currently used in public health settings are the following: screening for breast cancer using mammographic radiologic exams; screening newborns for phenylketonuria, or PKU, a genetic abnormality that results in an enzyme deficiency, using a heel stick to obtain a small sample of blood; and using a single blood pressure measurement to screen for hypertension.

Module Objectives

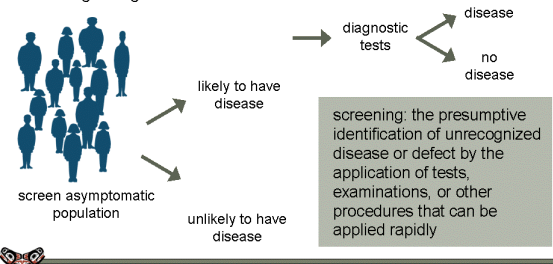
By the end of this module you should be able to:

- Define screening and describe its role in public health work
- Determine diseases for which it is appropriate to screen
- Determine which tests are appropriate to use
- Describe important factors in the design and evaluation of screening programs
- Help clients interpret results of screening tests

What Is Screening?

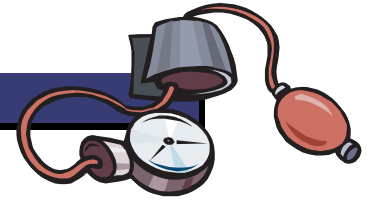
Using tests to detect the likely presence of diseases or conditions before illness or symptoms.

Screening ≠ diagnosis

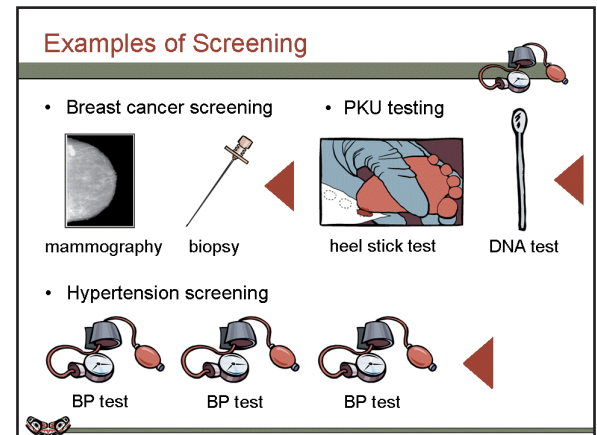


Screening in Public Health Practice

Transcript



A positive test by mammography screening typically would be followed up by more invasive diagnostic tests for breast cancer, such as a biopsy of the breast. A positive screening test for PKU may be followed by confirmatory DNA testing for the disease. And an individual who screens positive for hypertension, that is, has a high blood pressure reading, may be followed up clinically with repeat measurements of their blood pressure. So you can see that the diagnostic process following a positive screening test can have quite a range—from repeating the screening test for confirmation to instituting different, and often more elaborate or invasive testing procedures. And here also you see that screening, though most often performed in relation to a disease such as cancer or PKU, can also be done for a risk factor for disease, such as hypertension, a known risk factor for stroke.



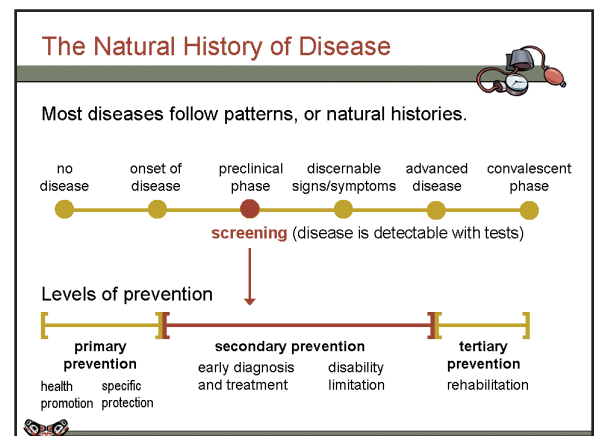
The Natural History of Disease

To understand how screening operates and why we do it, it may be helpful to think about the idea of the natural history of any disease. Most diseases follow certain patterns, or natural histories, moving from the onset of the disease process through a preclinical phase, and eventually on to the appearance of discernable signs and symptoms and advanced disease progression.

As we noted on the previous slide, screening is an activity that occurs when there are no symptoms of the disease; asymptomatic people are the target of screening programs. Screened individuals in whom the disease process had begun would be in the preclinical phase of that disease.

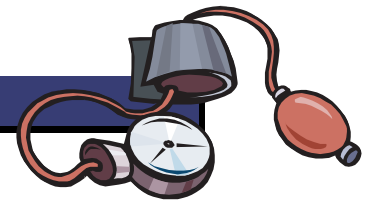
A basic assumption of screening is that in these individuals, while symptoms of the disease do not yet exist, the presence of disease is detectable, through radiologic or laboratory or other tests.

In public health we talk about three levels of disease prevention—primary, secondary, and tertiary. In primary prevention, disease is prevented, through general health promotion activities such as dietary advice, or through activities that protect against specific diseases, such as immunization for measles. In the category of secondary prevention, the disease process is not prevented, but the disease is discovered and addressed early after the process begins. This is where screening operates. The importance of screening in public health is that it provides an opportunity to intervene in the disease process at an early time—when such intervention should be more successful, thus limiting illness and death. Just to complete the



Screening in Public Health Practice

Transcript



picture here, the final category of prevention is tertiary prevention, which occurs after a disease has become clinically evident—it is defined as activities undertaken to help one recover from the effects of the disease, such as rehabilitation after a stroke.

Using Screening in Public Health Practice

Screening for disease is important in public health practice for several reasons. Prevention is an essential and traditional activity of the public health field, and screening is a cornerstone of prevention. By discovering and treating disease early, as noted on the previous slide, we undertake secondary prevention, minimizing the burden of disease in the population and using our health care resources more cost-effectively.

Public health practitioners often screen for disease directly—giving specific screening tests to individual clients. Later in this module we'll talk about how to interpret screening test results.

Finally, public health practitioners may design and provide prevention activities on a community-wide basis. On this level, understanding screening tests and the screening process can help practitioners determine diseases for inclusion in mass screening or selective screening programs, and also select the best and most appropriate screening tools to use to implement these screening programs.

Using Screening in Public Health Practice

Screening is important in public health:

- As cornerstone of prevention orientation of public health
 - Minimize burden of disease in the population
 - Use health care resources cost-effectively
- In screening for disease directly
 - Giving screening tests to individuals
 - Interpretation of test results to screenees
- In community-wide design and provision of prevention activities
 - Determine diseases for selection in mass screening or selective screening programs
 - Select best screening tools

What Makes a Disease Appropriate for Screening?

How do public health practitioners go about determining which diseases are appropriate to screen for in their communities?

There are several established guidelines. First, for a disease to be a suitable target for a screening program, it must be considered an important public health problem. This is subjective and is usually based on the magnitude or the seriousness of the disease or both. For instance, diabetes is considered to be a disease suitable for screening because diabetes is common in our communities, and this disease has a serious long-term impact on health, including premature mortality. On the other hand, one of the diseases for which we perform newborn screening, phenylketonuria, is quite uncommon. Nonetheless, this disease is a target of screening because of the likelihood of rapid mortality if it is not diagnosed and treated extremely early in life.

What Makes a Disease Appropriate for Screening?

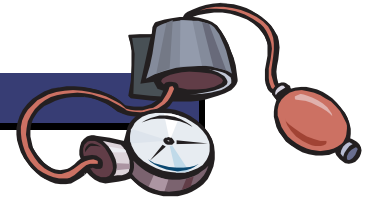
Important public health problem	Detectable preclinical phase	Preclinical disease progresses to clinical disease	
magnitude (diabetes) seriousness (phenylketonuria)	Otherwise screening program won't find many undiagnosed cases	Otherwise may increase burden of disease if treatment has adverse effects	

Factors Making Diseases Appropriate for Screening



Screening in Public Health Practice

Transcript



As we would predict from looking at the slide of the natural history of disease just shown, to be suitable for a screening program, the disease must have a detectable preclinical phase. Otherwise, all individuals in whom the disease process has begun would be symptomatic and would seek diagnosis, and our screening program would be unlikely to find many undiagnosed cases of the disease.

Additionally, it's necessary that, when untreated, the disease progresses to a clinical disease phase rather than remaining indefinitely without signs, symptoms, or effects. Screening for and ultimately detecting disease that would never have progressed to the clinical phase is likely to be counterproductive—diagnosing and treating those individuals would not decrease the burden of disease in the population, in the sense of decreased health—but may very well increase the burden of disease if the diagnostic process or subsequent treatment have adverse effects.

What Makes a Disease Appropriate for Screening? (cont)

To be an appropriate target of a screening program, it is essential that the disease, whether chronic, infectious, or genetic, is treatable. If nothing can be done to alter the course of the disease, there's no point in discovering and diagnosing disease early in the disease process. Early diagnosis is likely to just increase distress among those diagnosed earlier and their families.


A related concept is the final one listed here. For a disease to be a suitable candidate for screening, early detection and treatment of that disease must offer improved survival likelihood or fewer long-term problems than late treatment does. If these last two criteria are not met, the burden of disease will not be decreased as a result of the screening program—in fact, just finding disease early in the disease process, in the absence of effective early treatment, is likely to increase the prevalence of disease in your community. Let's see how that would happen.

What If a Disease Isn't Treatable?

Here's a situation in which we screen for a disease for which there is no effective treatment—so the outcome of disease isn't improved because the disease was discovered by screening. I'll explain how early recognition of the disease can appear erroneously to improve survival from that disease, an effect called lead time bias.


The disease process begins at time B. A person with this disease is usually diagnosed at a point somewhat later in

What Makes a Disease Appropriate for Screening? (cont)




Important public health problem magnitude (diabetes) seriousness (phenylketonuria)	Detectable preclinical phase Otherwise screening program won't find many undiagnosed cases	Preclinical disease progresses to clinical disease Otherwise may increase burden of disease if treatment has adverse effects	Treatable Early diagnosis of untreatable disease may increase distress among patients and families	Early recognition & treatment has better results than late treatment Otherwise likely to increase prevalence of disease
---	--	--	--	---

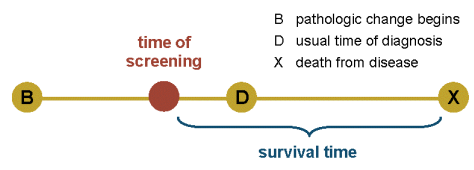
Factors Making Diseases Appropriate for Screening



What If a Disease Isn't Treatable?



Early recognition can appear erroneously to improve survival.




B pathologic change begins
 D usual time of diagnosis
 X death from disease

Diagnosis made earlier in disease process can make survival time appear longer.

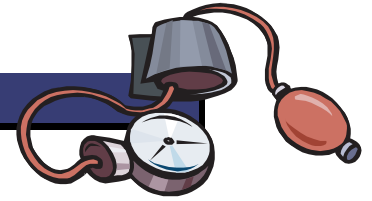
Prevalence of disease in community may increase

Incorrect conclusion that screening has a good impact on survival




Screening in Public Health Practice

Transcript



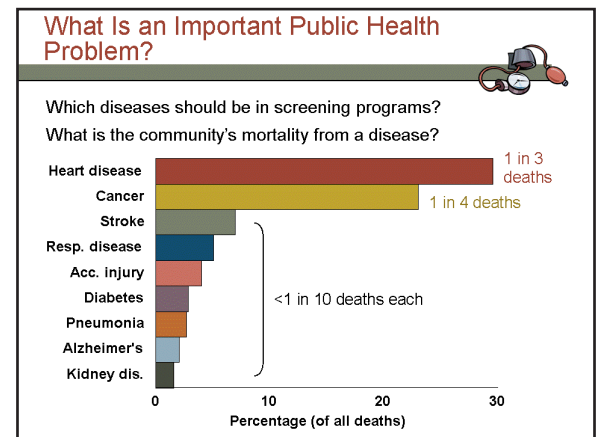
time—time D, after symptoms appear. The time between D and X is the length of time this person survives after being diagnosed with the disease if he or she is diagnosed as a result of seeking medical care after becoming symptomatic.

But if this person is screened for this disease, the discovery of disease and a diagnosis will come earlier, before symptoms appear. Even if there is no effective treatment as a result of this early screening and diagnosis, it will appear that this person survives longer with the disease. In reality, the amount of time from B, the beginning of the disease process, to X, death from the disease hasn't actually changed. It is just because the diagnosis is made earlier in the disease process and he or she lives more time after the diagnosis is made that survival mistakenly appears longer.

This concept is important for two reasons. First, as noted previously, screening for diseases for which there is no effective treatment may lead to increased prevalence of those diseases, because diseased individuals have a longer time from diagnosis to death and therefore are more likely to be in the population when the prevalence is measured. Second, as we'll discuss later on, this perceived shift in survival time could lead you to incorrectly conclude that your screening program has a beneficial impact on survival.

What Is an Important Public Health Problem?

Deciding which diseases are important problems and therefore should be included in screening programs is the first step in program development. There are many ways to determine which diseases are important problems, and one that is commonly used is how much of the community's mortality is from that disease. Let's consider the major causes of death in the United States, shown here. Coronary heart disease is responsible for the single largest fraction of mortality in our country, and it accounts for almost one in three deaths overall. The second biggest contributor is cancer of all types combined, accounting for nearly one in four deaths. The other major players are much less common, accounting for fewer than 1 in 10 deaths each. Some of these causes of mortality are associated with substantial long-term morbidity, however, and may deserve consideration for screening programs on that basis.



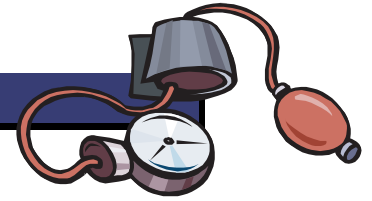
Is the Disease an Important Problem?

Let's talk a bit more about cancer, the second leading cause of death in the US, definitely an important public health problem. Consider the example of breast cancer to explain why we screen women for this disease. We know about breast cancer incidence from the National Cancer Institutes' SEER (Surveillance, Epidemiology, and



Screening in Public Health Practice

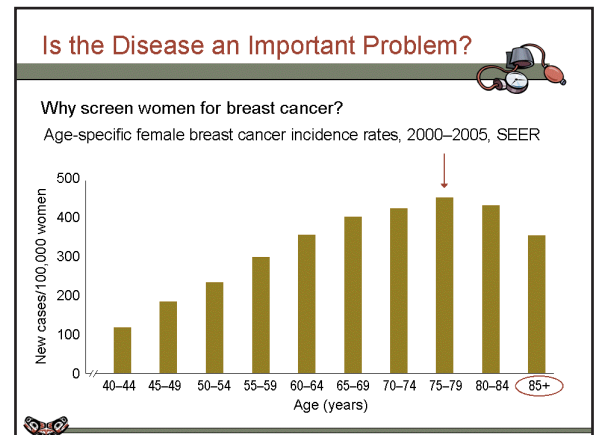
Transcript



End Results program) population-based registries in several locations around the US, covering over one-quarter of the US population.

Here you see the incidence of invasive breast cancer in women. Invasive means the cancer has spread from where it started into surrounding, healthy tissue in the breast (and sometimes into other parts of the body). The incidence is low early in a woman's life, and peaks at age 75–79, when around 450 out of 100,000 women develop breast cancer each year. Breast cancer incidence then drops off a bit—to around 350 per 100,000 women per year from age 85 on.

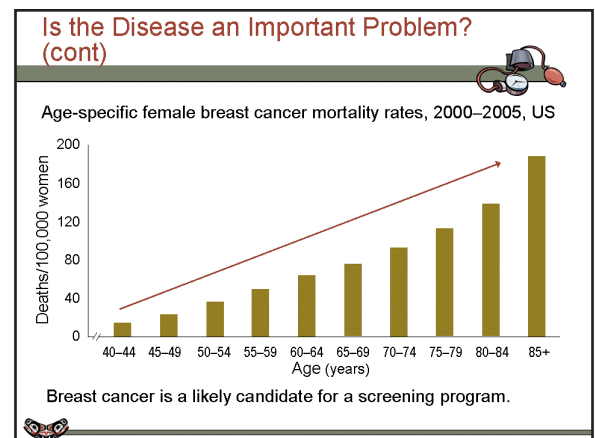
While 350 or 450 women out of 100,000 (or around 1 in 200) may not seem common, this is the most common cancer in women, and there are nearly 200,000 new cases in the US each year. Most people would consider a serious disease occurring this frequently to be an important public health problem, particularly if it causes substantial mortality. Let's see what we know about the mortality of invasive breast cancer.



Is the Disease an Important Problem? (cont)

We know about mortality from breast cancer among women through statistics gathered from US death certificates. We see here that the rate of death from breast cancer increases steadily with increasing age, with the highest rate, nearly 200/100,000 in the oldest women, those 85 years of age and above. There are approximately 40,000 deaths from breast cancer among American women each year, making this the second most common cause of cancer death for females, after lung cancer.

Because breast cancer is the most commonly occurring cancer among US women, and the second most common cause of cancer death, breast cancer would be a likely candidate for a public health screening program, if there is a lengthy preclinical phase and effective treatment is available, and if treatment works better when started earlier in the course of the disease. Let's talk a bit more about those issues.



Is There a Preclinical Phase?

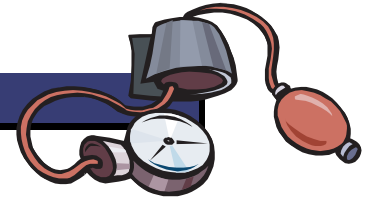
Is there a preclinical phase in female breast cancer, that is, a time during which the disease process is already present but the disease is not yet apparent clinically?

Yes, there is. Breast cancer begins with a single cancerous cell, and then



Screening in Public Health Practice

Transcript



progresses gradually. The growth potential of a breast cancer varies widely among patients, and it's estimated that the preclinical phase lasts from two to eight years. During this time, the disease can be detected with a good screening test for breast cancer.

This preclinical phase can include a time during which the cancer is in situ. In situ breast cancer is noninvasive; it's confined to the milk ducts or lobules in the breast and has not spread to the surrounding tissues in the breast or other parts of the body.

Better to Detect Breast Cancer Early?

As we've already noted, to be a suitable disease for screening, discovering and treating the disease early in the disease process must be better than late treatment. Better is defined as improved survival or decreased long-term effects of the disease.

Is it better to detect breast cancer early? Yes, it appears to be. One hundred percent of women with in situ breast cancer survive at least five years after diagnosis. Survival is lower for invasive breast cancer, and it depends on the extent of spread of the disease at the time of diagnosis. Women diagnosed while the cancer is still confined to the primary site (which we call localized) also have nearly 100% five-year survival, while women diagnosed after the cancer has spread to the regional lymph nodes or directly beyond the primary site (that is, those with regional cancer) have somewhat lower survival—around 84%. For the low number of women diagnosed after the cancer has already metastasized (that is, those with cancer also at a distant site) only 27% survive for five or more years.

So we see that early discovery of breast cancer, while it remains noninvasive, localized, or regional, is better than discovery late in the disease process. This information tells us that breast cancer screening and early intervention are potentially useful in decreasing mortality from this disease.

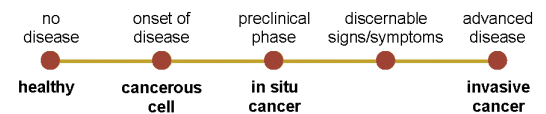
Should We Screen for Breast Cancer?

In terms of whether we should screen for breast cancer, the US Preventive Services Task Force recommends yes. This body recommends that we screen for breast cancer with mammography, either with or without a clinical breast examination every one or two years for women aged 40 or older.

They make this recommendation because it's been shown that having a mammo-

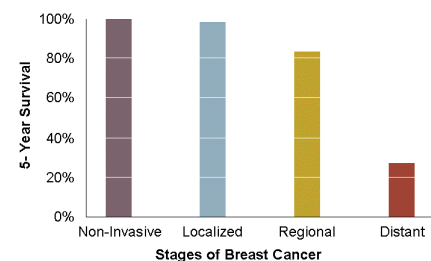
Is There a Preclinical Phase?

preclinical phase: time when disease is present, but not apparent



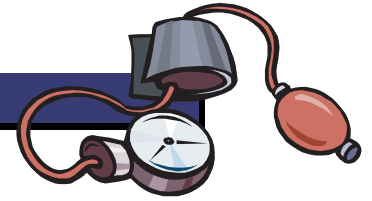
Better to Detect Breast Cancer Early?

Discovery and treatment early in the disease must be better than late treatment.



Screening in Public Health Practice

Transcript



gram every 12–33 months significantly reduces the risk of death from breast cancer. This health benefit is particularly evident in women between the ages of 50 and 69.

This recommendation has been made on the basis of the importance of the disease that might be prevented with screening, and also on the availability of a suitable and effective screening test—in this case mammography. Information about this and other screenable diseases and screening test recommendations can be found on the US Preventive Services Task Force Web site. After the upcoming exercise we'll talk about what makes a good screening test.

Now we'll pause for the first of several interactive exercises about the material you have just covered. Please note that the exercises sometimes take several seconds to load.

Exercise 1

What Makes a Test Suitable for Screening?

For a screening test to be considered effective, it must be reliable, applicable, acceptable, and valid, which includes both sensitivity and specificity. We'll talk about each of these characteristics in this next section moving forward.

Reliability

Reliability measures precision. Reliability is the ability of the test to give consistent results across repeated trials. For the breast cancer example, we would say that mammography is a reliable screening test for breast cancer if there are the same results (negative or positive) on the same woman each time she gets a mammogram within a short time period. This is an important aspect of a good screening test.

There are several different types of variations that can affect a screening measure's level of reliability. A screening test may not appear reliable if there is intra-individual variation in the characteristic being measured, that is, an individual's value on that characteristic varies. For example, blood pressure is subject to intra-individual variation—a person's blood pressure naturally varies somewhat over even small periods of time. This doesn't necessarily mean that a screening test

Should We Screen for Breast Cancer?

- The US Preventive Services Task Force recommends screening mammography:
 - With or without clinical breast examination
 - Every 1–2 years for women 40 and older
- Mammograms reduce risk of mortality from breast cancer—evidence strongest in women ages 50–69.
- Recommendation based on:
 - Importance of disease
 - Availability of suitable and effective test

US Preventive Services Task Force www.ahrq.gov/clinic/uspstfix.htm

What Makes a Test Suitable for Screening?

- Reliable
- Applicable
- Acceptable
- Valid
 - Sensitivity
 - Specificity

Reliability

Measures precision—ability to give consistent results repeatedly

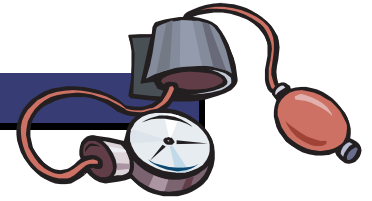
- Intra-individual variation
- Inter-observer variation
- Intra-observer variation

Screening test can be reliable yet inaccurate



Screening in Public Health Practice

Transcript



is unreliable. The screening test gets different results because the characteristic being measured differs from one time to the next, not because the test is wrong.

The next two issues have a greater effect on a test's reliability. Inter-observer variation occurs when two people interpreting the results of a screening test come to different conclusions. For example, if one radiologist reads a mammogram and calls it positive, while another reading the same mammogram calls it negative, the reliability of a mammogram as a screening test is questionable. Intra-observer variation occurs when the same person reads that same mammogram differently on two different readings.

Even if these variations are kept to a minimum and the screening measure has a high level of reliability, that does not necessarily mean it's accurate. Accuracy is a different characteristic of any screening measure, and we'll talk about that when we talk about validity in a minute.

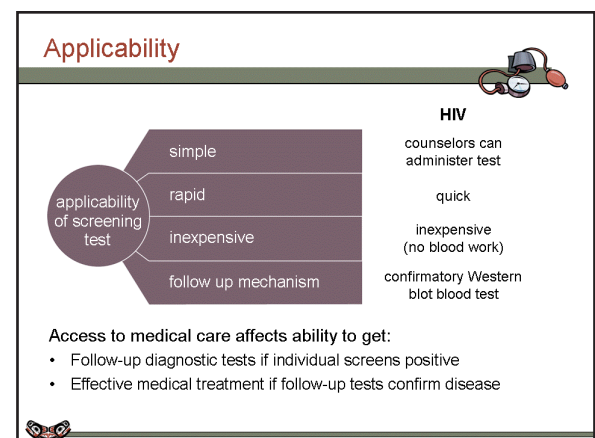
Applicability

For a screening measure to be useful in public health, it must also be applicable to the population. A screening test should be simple enough to be administered by health care personnel other than physicians. Ideally, it would provide rapid results and be inexpensive enough to use on a broad scale.

One example of an applicable test is the current method for HIV screening with rapid HIV antibody tests administered by counselors in clinical and non-clinical settings. In some settings antibody tests first are done on oral fluid samples. These initial screening tests are simple, quick, inexpensive and do not require blood to be drawn.

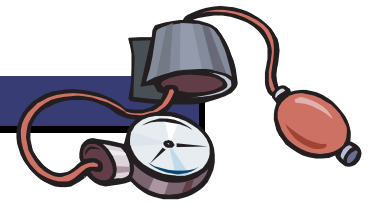
Another consideration is whether there's a mechanism for the follow-up of positive test results. Access to medical care affects the ability to obtain follow-up diagnostic tests if individuals screen positive, and to obtain effective medical treatment if follow-up tests confirm they have the disease. In situations in which follow-up testing and care are not available to members of a certain population group, screening tests are usually seen as not applicable for that group—as these individuals would not themselves benefit from merely being screened.

For example, while rapid HIV tests usually are followed by confirmatory Western blot blood tests as part of the screening procedure, it's possible that some individuals who are confirmed positive do not have health care coverage and therefore would not be able to obtain medical treatment to decrease their likelihood of future morbidity and mortality from AIDS.



Screening in Public Health Practice

Transcript



Acceptability

The third characteristic that defines the usefulness of a screening measure in public health is how acceptable the measure is to the population being screened.

The population being screened must always understand exactly what the measure is and what it's designed to detect, and in most cases they should provide consent before testing is done. A current exception is HIV screening, which the CDC now recommends be considered a routine part of medical care, to be performed unless the patient declines, without separate written consent for testing.

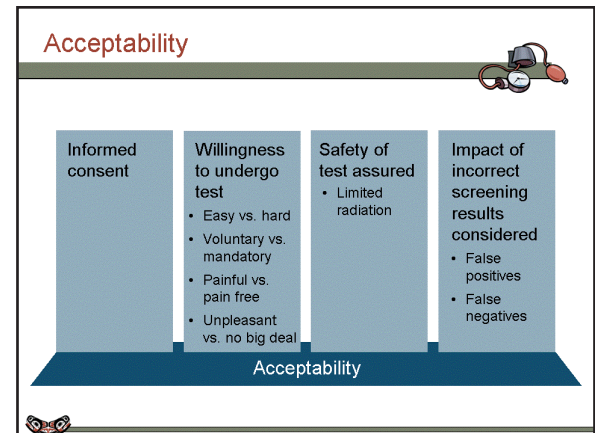
Once they understand the test, individuals must be willing to undergo the required screening test, which can sometimes be fairly unpleasant. For instance, oral rapid HIV tests are appealing because they do not require a finger prick or blood draw, and they have increased willingness of individuals to be screened.

Heel sticks are required to obtain blood for new born screening. Some parents don't want their newborn infants to undergo what might be their first painful procedure since birth, but since this test is mandated by law, the choice is not theirs to make. Some women, however, may choose not to undergo screening mammography because of the pain associated with the procedure of compressing breast tissue. And it's certain that some people avoid screening colonoscopy because of the unpleasantness of the procedure itself, or the bowel-cleansing preparation that is required before the test. These are important issues to consider when contemplating initiating mass screening programs.

To be an acceptable screening procedure the test also must be safe for the population being screened, and perceived as such. This means that there should be no adverse health effects from the screening test. Unsafe screening tests would include those that expose the screenee to substantial radiation, for example.

Finally, screenees must clearly understand and accept the potential impact not only of getting a positive diagnosis from the screening, but also the chances and the possible implications of getting incorrect screening results. We'll talk more about both false positives and false negatives later in the module. These risks must be clearly communicated for a screening test to be acceptable.

Let's pause now while you answer some questions on what you have just learned.



Exercise 2



Screening in Public Health Practice

Transcript



Validity

The final characteristic of an effective screening test is that it is valid. This speaks to the accuracy of the test, and helps you determine whether this test is a good one to use in screening for a particular disease.

The accuracy of a screening test is determined by studies designed just for this purpose, and also by evaluations of the effect of large long term screening programs using the test.

Validity is composed of two separate measures, sensitivity and specificity. Sensitivity is, in practical terms, the likelihood that, among individuals who have the disease, the test will pick it up—that is, in what proportion will the test come back positive? Ideally, a screening test you select would be 100% sensitive—it would miss no individuals who are truly diseased. What do we use sensitivity for? To help us answer the question of whether this is the best screening test to use for a certain disease.

Specificity is the other side of that equation; the likelihood that, among individuals who don't have the disease, the test will come back negative. Ideally, a screening test would also be 100% specific—it would return a positive result for no individuals who are not truly diseased. The level of specificity is another part of how we determine whether a test is good at screening for a certain disease.

No screening test is both 100% sensitive and 100% specific. We'll look at some typical ranges in a moment.

Validity

Validity: accuracy of test

- Determined by validity studies
- Evaluations of long-term screening programs
- Indicates whether the test is the best one to use in screening for a certain disease

No screening test is 100% sensitive and 100% specific.

Sensitivity

Before we calculate the level of sensitivity of a screening measure, let's talk a bit about the terminology around test results. If we're doing a study to determine the accuracy of a screening test, we would start by screening all the study subjects, categorizing them as test positive or test negative, and then determine their true disease status using a diagnostic test, which we sometimes call the "gold standard." Then we can place them into the four boxes shown here, based on their screening test results and their true disease status.

If the disease is present and the screening test picks it up, that's known as a true positive. The other correct result is if the disease is not present and the test result comes back negative, that's known as a true negative. If the disease is not present but the test comes back as positive, that's classified as

Sensitivity

To determine accuracy of a test, screen and categorize all study subjects.

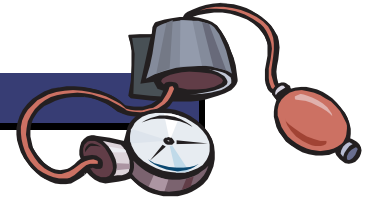
		True disease status	
		Present	Absent
Test result	Positive	↑ True positive	False positive
	Negative	↓ False negative	True negative

$$\text{Sensitivity} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$$



Screening in Public Health Practice

Transcript



a false positive. And if the disease is present and the test doesn't register that, it's known as a false negative.

To calculate sensitivity, we take the number of true positives and divide it by the number of true positives added to the false negatives. A sensitive test is going to have a high level of true positives and a low level of false negatives, that is, most of the time if there is disease the screening test picks it up.

This module makes use of several formulas in both the content and in the quizzes. You may want to access the formulas document available under the attachments tab of this module for your ongoing reference.

Example: Screening for Colon Cancer

Let's practice calculating sensitivity. Here's a hypothetical example of a study designed to test the validity of fecal occult blood tests—stool guaiacs—to determine the presence of cancer of the colon.

Fecal occult blood test smears were taken from 120 persons who had proven cancer of the colon and from 580 who did not have colon cancer. The smears were then read by people with no knowledge of the cancer status of the subjects. They reported the smears as either positive (cancer) or negative (no cancer).

Of the smears taken, 200 were reported positive, only 90 of which came from the proven colon cancer cases.

In this example, what is the sensitivity of the colon cancer screening test?

Example: Screening for Colon Cancer

Test the validity of fecal occult blood test (FOBT):

- 120 persons with colon cancer
- 580 persons who did not have colon cancer
- Tests read without knowledge of cancer status
- Reported as either positive (cancer) or negative (no cancer)

Of the samples taken, 200 were reported positive, 90 of which came from the proven cancer cases.

Example: Sensitivity of FOBT

Let's go back and calculate the sensitivity by filling in the table we introduced previously with the number of true positives, true negatives, false positives, and false negatives. We know that we have 120 subjects with colon cancer—this is the combined total of the true positives and the false negatives. We also know that 90 of the positive tests were from subjects with colon cancer—these are the true positives. So by subtraction we have 30 people who are false negatives—they truly have colon cancer, but they had negative occult blood tests. Now we can fill in the rest of the table. We know that 200 subjects had positive occult blood tests—this is the combined total of the true positives and the false positives. We already know that 90 people are true positives, so that leaves 110 as false positives. And for the final cell of the table, we just subtract the

Example: Sensitivity of FOBT

Test result	True colon cancer status		Total
	Present	Absent	
Positive	TP = 90	FP = 110	200
Negative	FN = 30	TN = 470	
Total	120	580	

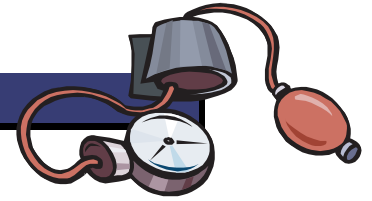
Sensitivity = $\frac{TP}{TP + FN} = \frac{90}{90 + 30} = \frac{90}{120} = 75\%$

FOBT picks up the cancer 75% of the time.



Screening in Public Health Practice

Transcript



110 false positives from the total of 580 subjects without colon cancer—leaving us with 470 true negatives.

We know that sensitivity is defined as the number of true positives divided by the number of true positives plus false negatives. For the fecal occult blood test example shown here, sensitivity is calculated as the 90 true positives divided by the 90 true positives plus the 30 false negatives—or 90 divided by 120—which equals 75%. This means that, for people who have colon cancer, 75% will have positive fecal occult blood test results when screened. The screening test picks up the cancer 75% of the time. The higher the sensitivity, the less likely it is that there will be false negative tests.

Let's pause now while you answer some questions on what you have just learned.

Exercise 3

Specificity

Now let's take a look at the specificity of the screening test. While sensitivity is looking for the number of true positives, specificity focuses on the number of true negatives. Specificity is calculated by dividing the number of true negatives by the number of true negatives added to the number of false positives. The result of the calculation is the likelihood that a screening test will be negative if the person truly does not have the disease.

What do we use specificity for? Similar to sensitivity, we use it to help us determine whether this is a good test—should we use this test to screen for a certain disease in our population or might another test be more accurate?

Let's make this a little more practical by returning to our colon cancer and fecal occult blood test example.

Example: Specificity of FOBT

Looking at this table of the test results that we filled in earlier, let's calculate the specificity. There were 470 true negatives, so we take that number and divide it by the combination of true negatives and false positives. This leaves us with 470 divided by 580, or 81%. Our interpretation of this is that for people who do not have colon cancer, 81% will have negative fecal occult blood test results when screened. The higher the specificity, the less likely it is that there will be false positive tests. High

Specificity

		True disease status		
		Present	Absent	
Test result	Positive	True positive	False positive	Sensitivity
	Negative	False negative	True negative	Specificity

$$\text{Specificity} = \frac{\text{true negatives}}{\text{true negatives} + \text{false positives}}$$

Used to determine whether test is the most accurate

Example: Specificity of FOBT

		True colon cancer status	
		Present	Absent
Test result	Positive	TP = 90	FP = 110
	Negative	FN = 30	TN = 470
Total		120	580

$$\text{Specificity} = \frac{470}{470 + 110} = \frac{470}{580} = 81\%$$

High specificity important when screening for disease with costly, invasive, or risky follow-up tests.



Screening in Public Health Practice

Transcript



specificity might be particularly important when screening for a disease that requires costly, invasive, or risky follow-up tests of individuals who screen positive—such as colonoscopies to follow up for colon cancer. In that situation you might not want to subject a large number of people without the disease (the false positives) to the follow-up tests. Let’s pause now while you answer some questions on what you’ve just learned.

Exercise 4

Predictive Value of a Positive Test

The final two characteristics of screening tests that we’ll consider involve the test’s predictive value. The predictive value of a test depends on how good the test is (the sensitivity and specificity) and how common the disease being screened for is in the population being screened.

Let’s start by talking about the predictive value of a positive test. This answers the question—among all people who test positive on the screening test, what proportion really have the disease? Or, on an individual level: If I test positive on the screening test, what’s the likelihood that I really have this disease? To calculate the positive predictive value of a test, we divide the number of true positives by the number of true positives combined with the number of false positives.

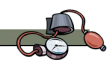
This concept is an important one for public health professionals designing screening programs, as we’ll discuss in a bit. It is helpful for deciding which subset of the population should be screened, and knowing the predictive value in the population that has been screened helps us interpret test results for screenees. For this is the question that people who test positive on a screening test usually ask; for instance, after being told of a positive Pap smear, a woman may ask “How likely is it that I have cervical cancer?” The predictive value of a positive test answers that question. Let’s calculate the answer, using our previous example.

Example: Positive Predictive Value

Returning again to our example of fecal occult blood tests, how would we calculate the predictive value of this test?

In the population in this example, there are 90 true positives—those who were correctly identified from the screening test as having cancer, and 110 false positives—those who were incorrectly identified from the screening test as having cancer. The equation for the positive

Predictive Value of a Positive Test



Depends on how good the test is and how common the disease is in the population.

Among all the people who test positive on the screening test, what proportion really have the disease?

If I test positive on the screening test, what’s the likelihood that I really have the disease?

$$\text{Predictive Value Positive (PV+)} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}$$

Helps in deciding which subset of population should be screened.

Helps clinicians interpret test results.



Example: Positive Predictive Value



What is the predictive value of a positive FOBT for colon cancer in this population?

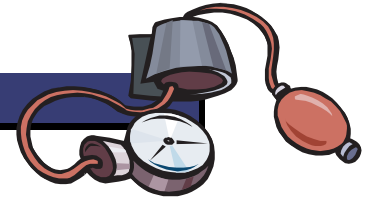
Test result	True colon cancer status		Total
	Present	Absent	
Positive	TP = 90	FP = 110	200
Negative	FN = 30	TN = 470	500

$$PV + = \frac{90}{90 + 110} = \frac{90}{200} = 45\%$$



Screening in Public Health Practice

Transcript



predictive value is 90 true positives divided by all those who tested positive—here 90 plus 110, or 200. Ninety divided by 200 is 45%. We would say, then, if someone from this population tests positive, there is a 45% likelihood that they do have colon cancer. This is quite high, for reasons we'll cover in a moment. The positive predictive value of a screening test is usually lower than this.

Let's pause for a second while you practice calculating the positive predictive value.

Exercise 5

Predictive Value of a Negative Test

Finally, predictive value can also refer to the predictive value of a negative test. This answers the question—among all the people who test negative on the screening test, what proportion really don't have the disease we're screening for? And on the individual level "If I test negative on the screening test, what's the likelihood that I really don't have the disease?"

The predictive value of a negative test is calculated by dividing the number of true negatives by the number of true negatives added to the number of false negatives.

In contrast to those who screen positive, individuals who screen negative don't usually ask upon hearing their test results "How likely is it that I really don't have the disease you just screened me for?" We should remember that a screening test is not the same as a diagnostic test. In screening for disease, we're only separating out individuals with a high likelihood of disease from those with a low likelihood of disease. Unless the predictive value of a negative test is 100%, some individuals who screen negative truly are diseased. A negative screening test result is not equivalent to the absence of disease unless the predictive value negative is 100%. While we don't want to alarm screenees, it is important that they understand this distinction and are aware that there is a slim chance that they may still have the disease.

Example: Negative Predictive Value

Returning to our fecal occult blood test and colon cancer example again, let's calculate the predictive value of a negative test. In this population tested, there were 470 true negatives and 30 false negatives. The predictive value is the 470 true negatives divided by all 500 negatives, or 94%. This means that 94% of those identified as negative by the screening test in this population were in fact free

Predictive Value of a Negative Test



Among all the people who test negative on the screening test, what proportion really don't have the disease?

If I test negative on the screening test, what's the likelihood that I really don't have the disease?

$$\text{Predictive Value Negative (PV-)} = \frac{\text{true negatives}}{\text{true negatives} + \text{false negatives}}$$

A screening test is not the same as a diagnostic test.

Unless predictive value negative is 100%, some people who screen negative have the disease.



Example: Negative Predictive Value



Test result	True colon cancer status		Total
	Present	Absent	
Positive	TP = 90	FP = 110	200
Negative	FN = 30	TN = 470	500

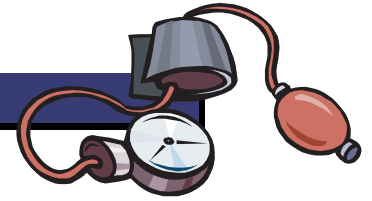
$$PV- = \frac{470}{470 + 30} = \frac{470}{500} = 94\%$$

An individual who screens negative has a 94% chance that they do not have colon cancer.



Screening in Public Health Practice

Transcript



of colon cancer. For an individual who screens negative then, there is a 94% chance that they do not have colon cancer. Let's pause for a minute while you practice calculating the negative predictive value.

Exercise 6

Populations with High Disease Prevalence

Often, disease screening is done in populations with a relatively high prevalence of the disease for which the test screens. In the recent exercises we've seen this example. This population has a fairly high prevalence of this unnamed disease—1650 out of every 100,000 people in the population have the disease. You might see disease prevalence this high, for instance, in a population of “high risk” individuals who are patients in a specialty clinic.

We have already calculated the sensitivity, specificity, and predictive value positive in this population—they are shown here.

Populations with High Disease Prevalence

Screening is often done in populations with high prevalence of disease

Test	Disease		Total
	Present	Absent	
Positive	1,240	3,994	5,174
Negative	410	94,416	94,826
Total	1,650	98,350	100,000

Sensitivity = $\frac{1,240}{1,650} = 75\%$ PV+ = $\frac{5,174}{5,174 + 94,826} = 24\%$

Specificity = $\frac{94,416}{94,416 + 3,994} = 96\%$

Populations with Low Disease Prevalence

This example is of the same screening test and the same disease as the previous slide, but the number of people in the population with the disease is lower. This may represent a population of “low risk” patients at a primary care clinic.

Let's see what happens to the sensitivity, specificity, and predictive value of the screening test when the prevalence of disease in the population is much lower—which is a more likely depiction of the general population prevalence of most diseases.

We see that the sensitivity of this test is still 75%—defined again as the number of diseased people who test positive divided by all those who are diseased, or here 124/165. Sensitivity is an intrinsic aspect of the test—measuring how good it is at picking up the disease when it's there—and it shouldn't vary from population to population.

We can calculate the specificity also, the same way we did on the previous slide. It's the number of disease-free people who test negative divided by all those who are disease-free. And here that's 95,842 divided by 99,835, or 96%. As on the previous slide we would say that among individuals without the disease 96% will screen negative. Specificity is also an intrinsic aspect of the test that doesn't vary from population to population.

Populations with Low Disease Prevalence

Test	Disease		Total
	Present	Absent	
Positive	124	3,993	4,117
Negative	41	95,842	95,883
Total	165	99,835	100,000

Sensitivity = $\frac{124}{165} = 75\%$ PV+ = $\frac{4,117}{4,117 + 95,883} = 3.0\%$

Specificity = $\frac{95,842}{95,842 + 3,993} = 96\%$

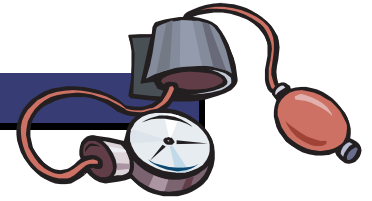
Sensitivity and specificity is about population to population.

If disease is uncommon, the likelihood that an individual who screens positive really has the disease is lower than for a screen-positive individual where the disease is common.



Screening in Public Health Practice

Transcript



In contrast, we see that the predictive value of a positive test is much lower here than it was on the previous slide. Remember, we calculate the predictive value positive as the number of diseased people who test positive divided by all people who test positive, or $124/4,117$ in this example, or 3%. This illustrates an important fact. If a disease is uncommon in a population, the likelihood that an individual from this population who screens positive really has the disease is lower than it would be for a screen-positive individual from a population in which the disease is common. This is why clinicians who see mostly patients at high risk of disease often take more seriously a positive screening test result for that disease—in their experience a positive result is more likely to indicate disease than it is for the clinician who tests primarily low risk patients.

This is also an important concept in designing screening programs. The amount of the disease that is present in the population that you screen has a big impact on the predictive value of the screening test. Because of this, you might want to focus on high risk population subgroups to maximize the efficiency of your program.


Designing Screening Programs

How do we actually use the screening test concepts we've just discussed in public health practice settings?

Knowledge of the concepts of sensitivity and specificity helps us to determine which screening tests are more accurate, therefore helps us to choose the best tests to use in public health screening programs. As we've discussed, sensitivity and specificity are attributes of a given test—that is, they do not vary depending on the population that is tested. Several commonly used screening tests have sensitivities in the range of 70–95%, that is, if the disease being screened for is present, it will be picked up 70–95% of the time. Specificities are usually a bit higher, in the range of 90–95%. This means that, if the disease being screened for is absent, the screening test will be negative 90–95% of the time.

Some screening tests do not give just a positive or negative result, but rather a numeric value—such as blood pressure. In these circumstances it is up to the individuals providing the screening test to decide what constitutes a positive test for their screening program. This decision is usually made after weighing the advantages of capturing all the “true positives” against the disadvantages of including more “false positives” in what you would consider to be a positive screening test result. For example, if you're screening for hypertension, you might decide that you would call everyone with a blood pressure of 160/110 or higher positive and refer them to medical care for further diagnosis and possible treatment. While it's likely that most individuals with a one-time blood pressure measurement that high indeed are

Designing Screening Programs









Choose the test

- Sensitivity: commonly ranges from 70–95%
- Specificity: commonly ranges from 90–95%

Determine the threshold

- Screeners decide what constitutes a positive test

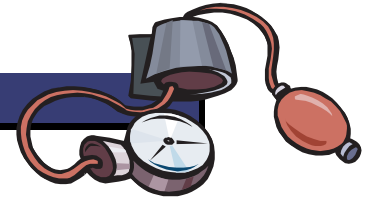
 general population	 >160/110	 general population	 >130/80
	 <160/110		 <130/80

● actual hypertensive ● actual non hypertensive



Screening in Public Health Practice

Transcript



hypertensive, it's also likely that you would miss some hypertensives using a threshold that high.

Alternatively, you might decide that in the interests of not missing any true hypertensives your definition of a positive test—with subsequent referral for diagnosis and treatment—might be 130/80. You would be more certain not to miss any true hypertensives with this cut-off point, but there would likely be some individuals referred for diagnosis and treatment who in fact did not have hypertension. Consideration of the financial, emotional, and physical costs of setting a low threshold for a positive screen, with consequent follow-up of a larger population, must be made.

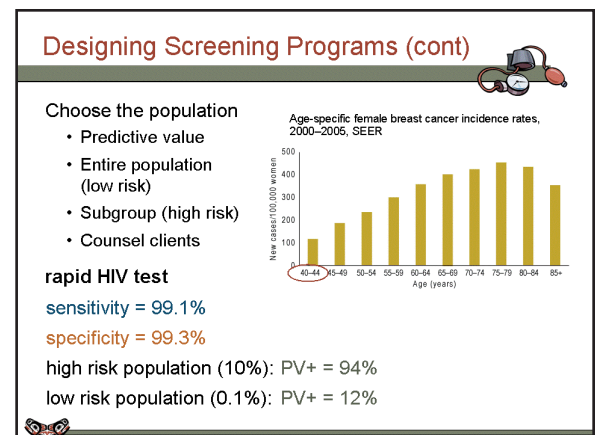
Designing Screening Programs (cont)

Public health practitioners also need to choose the population to screen. Calculating the predictive value of a positive test in your population will help determine whether screening the entire population, usually considered a low risk population, or certain subgroups—or high risk populations—is warranted. We have already seen that screening mammography is recommended by the Preventive Services Task Force for women only starting at age 40. This is because the occurrence of breast cancer is low in women before that age, therefore the predictive value of a positive mammogram will be lower than it would be among a population of older women—a higher risk population.

The concept of predictive value is also used by public health practitioners in helping counsel individual clients about the meaning of positive or negative screening tests. As we noted previously, individuals who screen positive usually want to know how likely it is that they have the disease for which they were just screened—and the predictive value positive answers this question. Because the predictive value is a function of the accuracy of the test as well as how common the disease is in the population screened, both of these statistics must be known or estimated. For example, the sensitivity of the rapid HIV test we discussed earlier is 99.1%, and the specificity is 99.3%. In a high risk population with an HIV+ prevalence of 10%, the predictive value of a positive test would be 94%, while in a general population of lower risk with an HIV prevalence of 0.1% the positive predictive value would be only 12%.

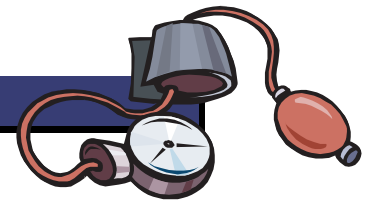
Evaluating Screening Programs

A key component of any screening program is evaluation. An initial part of the evaluation of a screening program might be process oriented—how many screening tests were performed—as an indication of the need for the program. For example, the US National Breast and Cervical Cancer Early Detection Program reports that between



Screening in Public Health Practice

Transcript



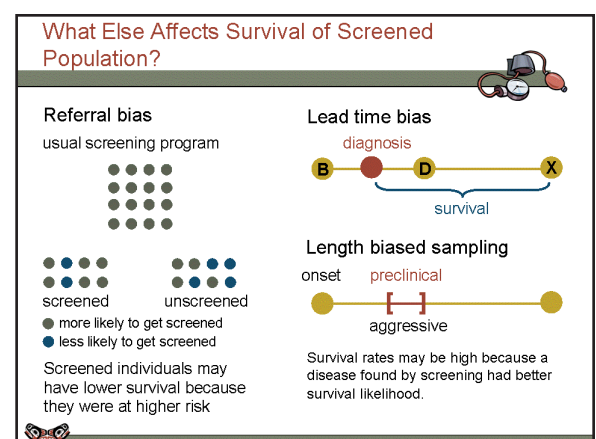
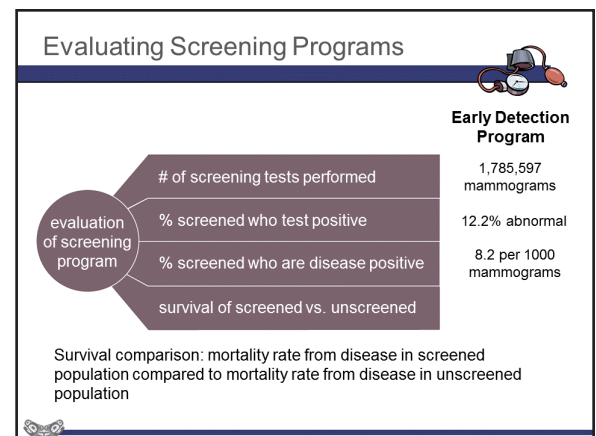
July 2002 and June 2007 1,785,597 mammograms were provided to women who would have been unlikely to obtain them otherwise.

Additional information about the utility of a screening program can come from calculation of the percent of screenees who test positive and the percent who are diagnosed with the disease for which they were screened. In the Early Detection Program just noted, for the 2002–2007 period, the screening mammograms of 217,887 women were abnormal (12.2% of all mammograms), and breast cancers were detected in 14,682 screened women, for a rate of 8.2/1000 mammograms. These statistics may be sufficient to indirectly evaluate the effect of screening programs on diseases for which the screening test has previously been shown to capture disease early, if early treatment has been determined to improve disease outcome.

The most accurate and powerful measure of a screening program's impact, however, is to directly compare the survival of screened and unscreened populations. And the most accurate survival comparison is to determine the mortality rate from the disease in the entire intended screened population and compare it with the mortality rate from the disease in the entire unscreened population. Unfortunately, this calculation is not possible when looking only at a screened population in a screening program. Let's take a look at an ideal situation and see how we would conduct evaluations in situations where information is available on both screened and unscreened populations.

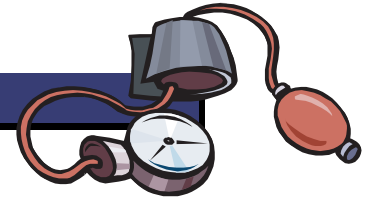
What Else Affects Survival of Screened Population?

While the ideal situation in which to evaluate screening programs and compare the mortality rates of screened and unscreened populations is a randomized trial, this is not the usual design used in screening programs. In a randomized trial, the only difference between the screened and unscreened populations is likely to be the screening, and any differences in survival can be assumed to be due to that screening. However, usually in screening programs participants self-select for screening, and their risk of the disease for which they are screened (and therefore mortality from the disease) may be unusual. It may be that a woman decides to enter a breast cancer screening program because of a strong family history of breast cancer.



Screening in Public Health Practice

Transcript



That would be called referral bias. In that circumstance screenees may have lower survival, not because of the screening, but because they were at higher disease risk to begin with.

We've already talked about another issue which can affect survival: lead time bias. Screenees may appear to have a longer survival time from the disease, but it may just be that their disease was found and diagnosed earlier in the disease process than it was in non-screened individuals, and there was no real beneficial effect.

Finally, if a study compares the survival rates of individuals with disease found through screening to survival rates of unscreened individuals, length-biased sampling may occur because of differences in the length of the preclinical phase of the disease in the two groups. Slow-growing, less aggressive disease is more likely to be found by screening because those with more aggressive disease will have a shorter preclinical phase. As a result, survival rates of those with disease found through screening may be high in comparison to those of unscreened individuals—not because screening improved survival but only because the disease found by screening was less aggressive and so had inherently better survival likelihood.

Keeping in mind these potential errors may help you evaluate studies of screening program effectiveness, and help you decide which screening programs to implement.

Summary

These completed course objectives should now help you take screening into your day-to-day job. Depending on your job role, you may be determining which diseases for which it is appropriate to screen, determining appropriate screening tests, and designing and conducting screening programs that can be evaluated appropriately.

Knowledge of the burden of specific diseases in your population—usually measured by mortality rates, but also measured for chronic diseases by prevalence and likelihood of causing substantial morbidity—is the starting point for decisions about which diseases are appropriate for screening in your population. A detectable preclinical phase, the existence of progression from preclinical to clinical disease, and effective treatment that is more effective when begun earlier in the disease process are also important attributes of screenable diseases.

Consideration of which screening tests to use for important screenable diseases in your population relies on knowledge of the reliability and validity of the available screening tests for these diseases, as well as their acceptability and applicability to your specific population.

The design and conduct of effective screening programs starts with the choice of appropriate diseases, screening tests, and populations to screen. The evaluation of

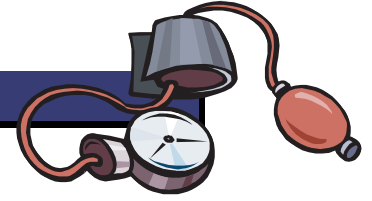
Summary

- **Determine appropriate diseases to screen**
 - Burden of specific diseases in population
 - Detectable preclinical phase progressing to clinical disease
 - Treatment more effective when begun earlier in disease process
- **Determine which tests are appropriate to use**
 - Reliability and validity of available tests
 - Acceptability and applicability to specific population
- **Design, conduct, and evaluate screening programs**
 - Appropriate diseases, tests, and populations
- **Help clients interpret results of screening tests**
 - PV + and PV -



Screening in Public Health Practice

Transcript



these programs may include process measures as well as outcome measures such as disease-specific mortality rates in screened and unscreened populations.

And finally, you may be called upon to help your clients interpret the results of screening tests. Predictive value of positive and negative tests are the concepts to consider in the interpretation of individuals' test results.

Resources

If you would like to learn more about the concepts in this module, you might want to explore some of the resources listed here. Also included is the [US Preventive Services Task Force](#) website address. This site includes recommendations for appropriate targets for screening and summaries of sensitivity and specificity of those recommended tests. Now, if you're ready, please go on to the final assessment.

Resources



Related online modules from the Northwest Center for Public Health Practice

- [What Is Epidemiology in Public Health?](#)
- [Data Interpretation for Public Health Professionals](#)
- [Study Types in Epidemiology](#)
- [Measuring Risk in Epidemiology](#)

Other useful resources for further study

Epidemiology for Public Health Practice, 3rd edition. RH Friis and TA Sellers. Jones and Bartlett Publishers, Sudbury, MA. 2004.

Essentials of Epidemiology in Public Health. A Ashengrau and GR Seage III. Jones and Bartlett Publishers, Sudbury, MA. 2003.

US Preventive Services Task Force www.ahrq.gov/clinic/uspstfix.htm

