Unit 12
Screening in Epidemiology

Screening and evaluation of diagnostic tools

Dr. A. Sanchez-Anguiano
Epidemiology 6000

"Understand purposes of screening and why screening is important"  
"What part of the natural history of disease does screening affect?"  
"Requirements, "prices", and consequences of screening on morbidity"  
"Desirable characteristics of a screening test"  
"Validity: sensitivity and specificity and their inter-relationship"  
"Consequences of setting the cut point too low or too high"  
"What are receiver-operating characteristic curves and why are they used"  
"Reliability of screening tests"  
"Measuring the performance (yield) of screening tests: PVP and PVN"  
"Relationship between yield and prevalence of the disease in the population"  
"Sources of bias in evaluating screening programs: volunteer bias, lead-time bias, length bias"

Screening

Definition: The application of a test to people who are as yet asymptomatic (search for subclinical disease; detect disease). Systematic testing of asymptomatic individuals for preclinical disease.

Purpose: To classify individuals with respect to their likelihood of having a particular disease. To prevent or delay development of advanced disease in a subset of patients with preclinical disease through early detection and treatment.
**Screening**
- Screening procedure itself does **NOT** diagnose illness.
  - **Screening**: PH interventions among populations (tool for detection)
  - **Diagnosis**: Clinical intervention applied to individual

- Usually used for diseases in which
  - some treatment or prevention can be applied,
  - to find cases for prevalence and incidence studies.

---

**Screening for Disease Control**
- Examination of asymptomatic people
  - (screen +) likely
  - Classification as
    - (screen -) unlikely
    - ..... to have a disease

---

**Screening for Disease Control**

```
| "Unlikely"        | referred to next screening cycle |
| "Likely"          | further testing for diagnosis   |
```

```
Diagnosed - Treatment
Not Diagnosed - Referred to next screening cycle
```
Screening for Disease Control

Screening objective:
- Operationally: application of a relatively simple, inexpensive test to a large number of persons in order to classify them as likely or unlikely to have the disease that is the object of the screen.
- In terms of outcome is to lower morbidity and mortality of disease in a screened population, at a reasonable cost.
- Screening provides access to the medical care system which is not an actual goal of screening, but is a benefit.

Screening is important because:
1) While etiologic (risk factor), diagnostic and therapeutic advances take many years and progress may be slow, screening may provide a “direct solution” to modify the history of a disease in a population;
2) Screening provides a model for studying disease mechanisms and the natural history of a disease (from sub-clinical through treatment).

Assumptions behind a Screening Program
- Implicit assumption is that early detection, before development of symptoms, will lead to a more favorable prognosis
- Assumption: treatment begun before the disease becomes clinically manifest will be more effective than later treatment.
- Other conditions to be considered:
  - Risks
  - Costs
**Primary requirements for screening:**

1) Early detection of screened disease should lead to a **more favorable prognosis due to early treatment** (compared to delayed treatment).

2) The disease should be **serious** (relates to cost effectiveness, ethics, and prognosis).

3) Prevalence of pre-clinical disease should be **relatively high among those screened** (might be sub-population at high risk).

**Diseases for which screening has been recommended**

- Cervical cancer
- Breast cancer
- Prostate cancer
- Colon cancer
- Obesity
- Diabetes
- Hypertension
- Hypercholesterolemia
- Down's syndrome
- Sickle-cell anemia
- Huntington's disease
- Neonatal hypothyroidism
- Phenylketonuria
- Infect: TB, HIV, STDs
- etc...

**Prices** of screening:

1) **Financial** - may be very costly if screening is spread out over an entire population or if screening method is expensive;

2) **Anxiety** - Individuals may have to be screened more often if found to be at high risk but not meeting screening criteria for more complete workup;

3) Creation of "lead time" morbidity.
Lead Time Definition

- Interval between detection of a disease at screening and when it would be detected due to the development of symptoms.
- Amount of time by which the diagnosis has been advanced as a result of screening.
- Depends on how diseases progress from pre to clinical and also how soon the screening is performed after the preclinical conditions becomes detectable.

<table>
<thead>
<tr>
<th>Time</th>
<th>Biological onset of disease</th>
<th>Disease detectable by screening</th>
<th>Screening takes place</th>
<th>Disease clinically detectable</th>
<th>Outcome: recovery/disability/death</th>
</tr>
</thead>
</table>

Lead Time Bias

Diagram shows: with screening, time of diagnosis is advanced by lead time provided by positive test result. If earlier diagnosis has no effect on time of death from disease, then survival with testing is equal to survival without testing plus lead time.

Natural History of Disease (ex: Cancer)

Age of Individual

20 30 40 45 50 55 60

Birth Exposure Neoplasia Cells exfoliate Screening Symptom test diagnosis Death

5 years additional “duration” of disease
**Natural History of Disease**

*Age of Individual*

- Birth
- Exposure
- Cells exfoliate
- Neoplasia
- Screening test
- Symptom diagnosis
- Detectable Pre-Clinical Phase (DPCP)
- Total Pre-Clinical Phase (TPCP)

**TPCP:** Begins at the true initiation of disease pathology; ends when the disease is clinically manifested (30 years in this example).

---

**Impact of Screening on Epidemiologic Measures of Morbidity: Prevalence**

- Prevalence of disease (found by either symptoms or screening)

**Steady state**

- Result of screening: decrease of secular trend due to increased recognition

**TIME**

---

**Impact of Screening on Epidemiologic Measures of Morbidity: Incidence**

- Incidence of disease

**Note:** incidence rises, and then drops sharply because the "pool at risk" is temporarily depleted

**Steady state**

**Time**
Evaluating Screening Tests

Characteristics of a screening test:
- **Validity**: ability of the test to correctly classify cases and non-cases
  - High sensitivity
  - High specificity
- **Reliability**/reproducibility: ability of the test to give the same result time after time
- **Feasibility**: Low cost, minimum level of invasiveness and discomfort, acceptance
- **Performance** (Yield) - Positive predictive value

Validity of Screening Tests

<table>
<thead>
<tr>
<th>True disease status</th>
<th>Results of screening test</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

a = true positive  
c = false negative  
b = false positive  
d = true negative

Validity

How well does the test measure true disease status (gold standard, true underlying pathology)?

Of all the people with and without the disease, how well does the test classify a person with respect to his/her true disease status?

- The test will correctly classify a diseased person as likely to have the condition ("sensitivity")
- The test will classify a non-diseased person as unlikely to have the condition ("specificity")
Validity of Screening Tests

<table>
<thead>
<tr>
<th>True Disease Status</th>
<th>+</th>
<th>-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Results of screening test</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>-</td>
<td>c</td>
<td>d</td>
</tr>
</tbody>
</table>

Sensitivity: Given that the disease is truly present, what is the probability of testing positive?

Sensitivity = \( \frac{a}{a + c} \)

Specificity: Given that the disease is truly absent, what is the probability of testing negative?

Specificity = \( \frac{d}{b + d} \)

Breast Cancer

<table>
<thead>
<tr>
<th>Physical exam</th>
<th>+</th>
<th>-</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>132</td>
<td>983</td>
</tr>
<tr>
<td>-</td>
<td>45</td>
<td>63650</td>
</tr>
</tbody>
</table>

Sensitivity: \( \frac{a}{a + c} \)

Sensitivity = \( \frac{132}{132 + 45} = 74.6\% \)

Specificity: \( \frac{d}{b + d} \)

Specificity = \( \frac{63650}{983 + 63650} = 98.5\% \)
Validity of Screening Tests

Setting the criterion for positivity: simple situation: Bi-modal distribution

Number of people in population

Blood sugar values

True non-diabetics

True diabetics

Question: What is the best cutpoint?

Validity of Screening Tests

<table>
<thead>
<tr>
<th>True Diabetes</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100% sens</td>
<td>100% spec</td>
</tr>
<tr>
<td>(i)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cutpoint is perfect: The never happens!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Problem: None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test is not specific: work up a lot of people who don’t have D</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>True Diabetes</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100% sens low spec</td>
<td>100% spec</td>
</tr>
<tr>
<td>(ii)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cutpoint set too low: All true diabetics are classified positive by test (up FP)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>True Diabetes</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>low sens</td>
<td>100% spec</td>
</tr>
<tr>
<td>(iii)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cutpoint set too high: Some diabetics will not be found (up FN)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test is not sensitive: it is missing people w/ the D</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Validity of Screening Tests

Relationship between sensitivity and specificity:

1. Lowering the criterion for positivity results in increased sensitivity but at the expense of decreased specificity (situation ii)

2. Making the criterion for positivity more stringent increases the specificity, but at the expense of decreased sensitivity (situation iii)
3. The goal is to have both high sensitivity and high specificity, but this is often not possible or feasible.

4. The decision for the cutpoint involves weighing the consequences of leaving cases undetected (false negatives) against erroneously classifying healthy persons as diseased (false positives).

5. In general, specificity should be at least 98% - but depends how much work you’re willing to put up with.

6. Sensitivity should be increased:
   • when the penalty associated with missing a case is high (serious disease that has treatment available: PKU)
   • when the disease can be spread (eg., HIV, STDs)
   • when you are screening to estimate prevalence/incidence of a condition in a population (FN will result in underestimation of rates).
   • when more diagnostic evaluations are associated with minimum cost or risk (hypertension).

7. Specificity should be increased when:
   - the costs or risks associated with further diagnostic techniques are substantial (e.g., positive screen requires that a biopsy be performed: breast cancer for which the definitive diagnostic evaluation is a biopsy.) (Sens=80% young women)

In this circumstance it must be made clear to those screened negative that there is not a guarantee of being disease-free but rather that the probability of having disease is low.
1. Construct a series of 2x2 tables; each 2x2 table uses the next possible cut-point – numbers in cells change each time (ths w/ D and w/o D remain constant).

2. For each 2x2 table, calculate sensitivity and specificity.

3. Plot sensitivity (y) by 1-specificity (x).

4. Choose point of curve that is closest to 100% sensitivity and 100% specificity.

Receiver-Operator Curves (ROC) - used to determine the best cut-point for a test

Receiver-Operator Curves (ROC)

<table>
<thead>
<tr>
<th>T4 value</th>
<th>Hypothyroid</th>
<th>Euthyroid</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 or less</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>5.1 - 7</td>
<td>7</td>
<td>17</td>
</tr>
<tr>
<td>7.1 - 9</td>
<td>6</td>
<td>36</td>
</tr>
<tr>
<td>9 or more</td>
<td>3</td>
<td>39</td>
</tr>
<tr>
<td>Totals</td>
<td>32</td>
<td>93</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cutpoint</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.56</td>
<td>0.99</td>
</tr>
<tr>
<td>7</td>
<td>0.78</td>
<td>0.81</td>
</tr>
<tr>
<td>9</td>
<td>0.91</td>
<td>0.62</td>
</tr>
</tbody>
</table>