Screening for Disease

“An Ounce of Prevention is Worth a Pound of Cure.”

Actually, an ounce of prevention is better than a pound of cure, but if prevention hasn’t been effective, perhaps early identification of disease by a screening test would be beneficial.
A Hypothetical Time Line for Disease

With screening, early diagnosis and treatment *may* result in longer survival, less disability, decreased recurrence, etc.
Many Types of “Screening”

You can think about these as screening tests that hopefully enable you to identify disease at an early stage:

- Self breast exam
- Clinical Breast exam
- Mammography
- Digital rectal exam
- PSA
- Dip stick test for sugar
- “Routine” blood tests
- Routine EKG or “stress test”
- Skin test for TB

Characteristics of a good screening test?
Characteristics of a Good Screening Test

- Inexpensive (dip stick for diabetes vs. MRI for brain tumor)
- Easy to administer
- Minimal discomfort
- Reliable: The test gives the same result each time.
- High Test Accuracy (Test Validity) – Test results accurately identify diseased & non-diseased people?
Sources of Variability (Error)

BP = 165/95

Observation (Test) Variability:

- **Within-an-observer**: Consistency when a single observer performs repeated measurements.

- **Between-observers**: Similarity of values when different observers perform measurements on a set of subjects.

- **Within-an-instrument**: Consistent measurements from a single instrument.

- **Between-instruments**: Do different instruments give consistent measurements?
Other Sources of Variability

Biological (Subject) Variability:

- **Within-subject**: Is the measurement the same over time?
  - Did he just walk up the stairs?
  - Did he just smoke a cigarette?
  - Is he stressed out?
  - Does he have “white coat” syndrome?
  - Is he over weight?

- **Between-subjects**: How much variability is there from subject to subject?
Even if the PSA test is:

- Precise (consistent measurements with repeated tests) &
- Accurate (close to his true PSA level)…

… how good is the PSA test with respect to determining whether he has prostate cancer or not? (Test validity)
You have prostate cancer.

Could the screening test be incorrect?
<table>
<thead>
<tr>
<th>Measurement</th>
<th>NCI Risk Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 – 2.5 ng/ml</td>
<td>Low</td>
</tr>
<tr>
<td>2.6 – 10 ng/ml</td>
<td>Slightly to moderately elevated</td>
</tr>
<tr>
<td>10.1 – 19.9 ng/ml</td>
<td>Moderately elevated</td>
</tr>
<tr>
<td>≥ 20 ng/ml</td>
<td>Significantly elevated</td>
</tr>
</tbody>
</table>

PSA = 5.6
Does the test accurately distinguish healthy & diseased people?

The ideal is to have a test that is exquisitely sensitive & specific....
The reality is that test values from diseased and non-diseased people often overlap….

What if >4 is deemed abnormal? Is the test valid?
How can we think about test validity in a structured way?

… but maybe not as much overlap as in the previous example.
Men with a Variety of PSA Test Results

These have low probability of having cancer.

These have a higher probability of having cancer.
Of those who were **diseased**, what was the probability (%) of correctly being identified by a + screening test?

Of those who were **NOT diseased**, what was the probability (%) of correctly having a (-) screening test?

If you screened (+), what was the probability that you had cancer?

If you screened (-), what was the probability that you did NOT have cancer?
Measures of Test Validity

Does the test accurately distinguish between healthy and diseased people?

Two Perspectives

Probability of correct test (screening)

- **Sensitivity**: probability that diseased people test +
- **Specificity**: probability that non-diseased people test -

Probability of disease

- **Predictive value (+)**: if someone has a + test, what is the probability that they actually *have* the disease?
- **Predictive value (−)**: if someone has a - test, what is the probability that they *don’t* have the disease?
### True disease status

<table>
<thead>
<tr>
<th></th>
<th>Diseased</th>
<th>Not Diseased</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Positive</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Negative</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Test Results

<table>
<thead>
<tr>
<th></th>
<th>Diseased</th>
<th>Not Diseased</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Positive</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Negative</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Among those who really had disease, the probability that the test correctly identified them as positive.

\[
\text{Sensitivity} = \frac{a}{a+c}
\]

- Test Positive
- Test Negative
- Total diseased = a+c
Sensitivity:

Among those who really had disease, the probability that the test correctly identified them as positive.

<table>
<thead>
<tr>
<th></th>
<th>Diseased</th>
<th>Not Diseased</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Positive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>True Positive</td>
<td>132</td>
<td>983</td>
</tr>
<tr>
<td>Test Negative</td>
<td>45</td>
<td>63,650</td>
</tr>
</tbody>
</table>

Sensitivity = \( \frac{132}{177} = 74.6\% \)
Among those who really did *NOT* have disease, the probability that the test correctly identified them as (-).
The ideal is to have a test that is exquisitely sensitive & highly specific, but this frequently isn’t the case.

If PSA values for men with prostate cancer overlap those of men without cancer, what do you use as the criterion for “abnormal”? 

The Trade-Off: Sensitivity Versus Specificity
<table>
<thead>
<tr>
<th>Test Positive (PSA &gt;4)</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>818</td>
<td>1,132</td>
</tr>
<tr>
<td></td>
<td>112</td>
<td>558</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test Negative</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>930</td>
</tr>
<tr>
<td></td>
<td>1,690</td>
</tr>
<tr>
<td></td>
<td>2,620</td>
</tr>
</tbody>
</table>

Criterion of positivity = 4 ng/ml

Sensitivity = 818/930 = 86%
Specificity = 558/1,690 = 33%

We’ve been looking at the performance of the test (its validity) by asking:

Among people who truly have the disease, what is the probability that the test will identify them as diseased?

and

Among people who don’t have the disease, what is the probability that the test will categorize them as non-diseased?
Another Perspective

If I have a test done and it is **positive** (abnormal), what is the probability that I really **have** the disease?

If I have a test done and it is **negative** (normal), what is the probability that I really **don’t have** the disease?
1) If the test was *positive*, how likely is it that he really *has* the disease? [How *worried* should he be?]

2) If the test was *negative*, how likely is it that he really *does NOT* have it? [How *reassured* should he be?]
Among men with positive test results, what is the probability that they have disease?

<table>
<thead>
<tr>
<th>Diseased</th>
<th>Not Diseased</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Positive</td>
<td>132</td>
</tr>
<tr>
<td>Test Negative</td>
<td>45</td>
</tr>
</tbody>
</table>

Predictive value (+) = \( \frac{a}{a+b} = \frac{132}{1,115} = 11.8\% \)
Predictive value (−) = \[ \frac{d}{c+d} = \frac{63,650}{63,695} = 99.9\% \]
"Methods: PSA testing results were compared with a reference standard of prostate biopsy. Subjects were 2,620 men 40 years and older undergoing (PSA) testing and biopsy from 1/1/95 through 12/31/98 in the Albuquerque, New Mexico metropolitan area. Diagnostic measures included the area under the receiver-operating characteristic curve, sensitivity, specificity, and likelihood ratios. 

Results: Cancer was detected in 930 subjects (35%). The area under the ROC curve was 0.67 and the PSA cutpoint of 4 ng/ml had a sensitivity of 86% and a specificity of 33%.

Question: What was the positive predictive value in this study?

Hint: You have to use the information provided to piece together the complete 2x2 table; then compute the PPV. See if you can do this before looking at the answer.
<table>
<thead>
<tr>
<th>Biopsy-Proven Prostate Cancer</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Positive (PSA &gt;4)</td>
<td>800</td>
<td>1,132</td>
</tr>
<tr>
<td>Test Negative</td>
<td>130</td>
<td>558</td>
</tr>
</tbody>
</table>

- **Sensitivity** = 800/930 = 86%
- **Specificity** = 558/1690 = 33%
- **Positive PV** = 800/1932 = 41%
- **Negative PV** = 558/688 = 81%

The table presents the prevalence of HIV in this study population. The numbers in the table represent the counts of individuals in each category:

- **Positive Test**
  - HIV+: 10
  - HIV-: 510
- **Negative Test**
  - HIV+: 0
  - HIV-: 99,480

The total population is 100,000, with 10 positive cases of HIV and 99,990 negative cases. The prevalence of HIV in the study population can be calculated as follows:

Prevalence = (Number of HIV+ / Total Population) * 100

Substituting the numbers:

Prevalence = (10 / 100,000) * 100 = 0.1%

Therefore, the prevalence of HIV in this study population is 0.1%.
Prevalence of HIV in this study population?

<table>
<thead>
<tr>
<th>True Disease Status</th>
<th>HIV+</th>
<th>HIV -</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>10</td>
<td>510</td>
</tr>
<tr>
<td>Negative</td>
<td>0</td>
<td>99,480</td>
</tr>
</tbody>
</table>

Prevalence = 10/100,000 = 0.0001 = 0.01%
True Disease Status

<table>
<thead>
<tr>
<th>HIV+</th>
<th>HIV -</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>Test</td>
</tr>
<tr>
<td>Negative</td>
<td>0</td>
</tr>
</tbody>
</table>

Sensitivity = 100%
Specificity = 99.5%

Prevalence in Female Blood Donors = 0.01%

Predictive value (+) = \( \frac{a}{a+b} = \frac{10}{520} = 1.9\% \)
True Disease Status

<table>
<thead>
<tr>
<th></th>
<th>HIV+</th>
<th>HIV -</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Positive</strong></td>
<td>4,000</td>
<td>480</td>
</tr>
<tr>
<td><strong>Negative</strong></td>
<td>0</td>
<td>95,520</td>
</tr>
<tr>
<td><strong>Test</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>4,000</strong></td>
<td>96,000</td>
<td>100,000</td>
</tr>
</tbody>
</table>

Sensitivity = 100%
Specificity = 99.5%

Predictive value (+) = \( \frac{a}{a+b} = \frac{4,000}{4,480} = 89\% \)

Prevalence in Males at an STD Clinic = 4%
True Disease Status

<table>
<thead>
<tr>
<th>HIV+</th>
<th>HIV -</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>20,000</td>
</tr>
<tr>
<td>Negative</td>
<td>79,600</td>
</tr>
</tbody>
</table>

Predictive value (+) = \( \frac{a}{a+b} = \frac{20,000}{20,400} = 98\% \)

Sensitivity = 100%
Specificity = 99.5%

Prevalence in IV Drug Users = 20%

NOTE: Predictive value is GREATLY influenced by the prevalence of disease in the population being screened.
What About Cells “b” and “c”???

<table>
<thead>
<tr>
<th></th>
<th>Diseased</th>
<th>Not Diseased</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive Test</td>
<td>132 True Positives</td>
<td>983 False Positives</td>
</tr>
<tr>
<td>Negative Test</td>
<td>45 False Negatives</td>
<td>63,650 True Negatives</td>
</tr>
</tbody>
</table>

- **a**: 132 True Positives
- **b**: 983 False Positives
- **c**: 45 False Negatives
- **d**: 63,650 True Negatives

Total:
- Diseased: 1,115
- Not Diseased: 63,695
- Positive Test: 177
- Negative Test: 64,633
- Total: 64,810
False Positives: worry, cost, complications, risk, of diagnostic tests

False Negatives: false reassurance
What Should We Screen For?

• Periodic Pap smears for cancer of cervix?

• Annual chest x-ray for lung cancer?

• Annual ultrasound for gallstones?

• Should everyone be tested for HIV?

• Fecal occult blood testing to screen for colorectal cancer?
  ➢ Annual colonoscopy for all adults?
What Diseases are Appropriate for Screening?

1) **Serious disease.** (Cervical cancer vs. gallstones)

2) **When treatment before symptoms is better than treatment after symptoms appear.** (e.g. cervical cancer.)
   - Need a detectable pre-clinical phase (DPCP)
   - Presumes existence of an effective test

3) **High prevalence of undiagnosed disease in the DPCP.**
   (Prevalence of HIV in couples applying for a marriage certificate is *extremely* low.)

---

High blood pressure leads to kidney damage, atherosclerosis, & stroke. It has a *long* DPCP and can be effectively treated with diet and medication. Undiagnosed HBP is very common.
How Do We Assess Screening Programs?  
(Feasibility and Effectiveness)

- Follow-up of those who test positive (positive predictive value)
- Assessment of cost per case detected
- Compare outcome measures to assess effectiveness in screened vs. unscreened groups (RCT is best):
  - Overall mortality rates
  - Disease-specific mortality rates
If one compares survival time in cancer patients identified by screening to those identified clinically, biases can occur.

Those identified by screening may appear to have longer survival times because of:

- Self-selection bias
- Lead time bias
- Length time bias
People who choose to participate in screening programs:

- tend to be healthier and have lower mortality rates
- tend to adhere to therapy better

- but, may also represent the “worried well”, people who are asymptomatic, but at higher risk (e.g. breast cancer)
Lead time is a good thing! But, it exaggerates the survival time.

How do the survival times compare (from diagnosis to death)?
By how much was life extended?

Lead Time Bias

Survival time with screening
Death with screening

Survival time without screening
Death without screening

DPCP

Disease detectable by screen

Biologic onset of disease

• How do the survival times compare (from diagnosis to death)?
• By how much was life extended?
Death with screening

Survival time with screening

Survival time without screening

Actual increase in survival

Death without screening

DPCP

Disease detectable by screen

Biologic onset of disease

Lead Time Bias
Some cancers are biologically aggressive and have short DPCPs.

Note that the length of the DPCP varies from person to person.

What does a short DPCP mean?

... others are slower growing and have longer DPCPs.
Screening has a better chance of detecting those with a long detectable pre-clinical phase, e.g. less aggressive tumors with more favorable prognosis.

Mean DPCP of screen +: 8 yrs
Mean survival of screen +: 6 yrs

Mean DPCP of unscreened: 6 yrs
Mean survival: 4 yrs
Evaluating a Screening Program

- Correlational studies
- Case-control studies
- Cohort studies
- Randomized clinical trials*