Introduction #2:
DIAGNOSTIC MEDICINE
The prostate is part of a man’s reproductive system. It is a gland surrounding the neck of the bladder.

A healthy prostate is about the size of a walnut and is shaped like a donut. The urethra (the tube through which urine flows) passes through the hole in the middle of that “donut”. Because of that, if the prostate grows too large, it squeezes the urethra causing a variety of urinary problems.
PROSTATE CANCER

- Cancer begins in cells, building blocks of tissues.
- When normal process goes wrong, new cells form unnecessarily and old cells do not die when they should. Extra mass of cells called a tumor; and malignant tumors are cancers.
- No one knows the exact causes of prostate cancer ... yet, but age is a significant factor. Most men with prostate cancer are over 65; if they live long enough a large proportion of men would eventually have prostate cancer.
PROSTATE CANCER SCREENING

- There are risk factors (age, family history) and symptoms (inability to urinate, frequent urination at night, etc…)

- Common screening is a blood test to measure prostate-specific antigen (PSA).

- However, a high level could be caused by benign prostatic hyperplasia (BPH – growth of benign cells); so the test is not specific.
Sometimes screening is more complicated than just making “one measurement”. For example, having one high measurement of PSA might not mean anything because – as mentioned - a high level could be caused by benign prostatic hyperplasia. There should be “some pattern” of measurements.
A PROSTATE CANCER MODEL

- Serum PSA in patients diagnosed with prostate cancer follows an exponential growth curve.

- A retrospective study of banked serum samples (Carter et al., *Cancer Research* 52, 1992) showed that the exponential growth begins 7-9 years before the tumor is detected clinically.
EXPONENTIAL GROWTH MODEL

\[ PSA_t = PSA_0 \exp(\beta_1 t) \]

\[ Y_t = \ln PSA_t \]

\[ = \beta_0 + \beta_1 t \]

This is a “A Simple Linear Regression Model” with PSA used on log scale.
PSA DOUBLING TIME

PSA-DT, the prostate specific antigen doubling time, has been used to predict clinical outcomes such as time to progression and mortality.

\[
\text{PSA}_t = \text{PSA}_0 \exp(\beta_1 t) \\
Y_t = \ln \text{PSA}_t \\
= \beta_0 + \beta_1 t \\
Y_2 - Y_1 = \beta_1 (t_2 - t_1) \\
\text{DT} = \frac{\ln 2}{\beta_1}
\]
AIDS

- **Acquired Immunodeficiency syndrome** (AIDS) is a severe manifestation of infection with the Human Immunodeficiency Virus (HIV, identified in 1983).
- The virus destroys the immune system leading to opportunistic infections of the lungs, brain, eyes, and other organs; Consequences include debilitating weight loss, diarrhea, and several forms of cancer.
- Currently, 40 millions living with AIDS; about 5 millions newly infected and 3 millions deaths in 2004 – most affected region is Sub-Sahara Africa. Diagnosed by blood tests (for example, using CD4⁺ T-cell count, or CD8⁺ T-cell count).
ULCERS

- An ulcer is a break in the lining of the stomach or in the duodenum (first part of small intestine); Gastric/peptic ulcers cost 3.2 billions in 1975 dollars.

- Most ulcers are caused by H. Pylori (identified in 1982), a bacterium living in the pylorus (the passage connecting the stomach and the duodenum).

- The two Australian physicians who discovered this bacteria won Nobel prize in 2005.

- Diagnosed by a blood test, even a breath test.
INFECTIONS

- There are many bacterial and viral diseases such as AIDS and ulcers.

- Bacterial and viral diseases are often easier, and more accurately, to be detected because: (i) the disease is “better defined”, and (ii) they are all implicated by a common marker: some form of agent-specific antibodies (provided that we know/have the right assay!)
THE DIAGNOSIS PROCESS

- It starts with an idea, it could be accidental or the result of a long search. The idea then goes through a two-stage process

- **Stage I: Developmental Stage**
  The question here is: Does the idea work?

- **Stage II: Applicational Stage**
  The question here is: Does it work for “me”? (i.e. the user; or when does it work? Or to whom does it work?)
In the Developmental Stage, the basic question is: Does the idea work? It’s the investigator’s (or producer’s) burden to prove.

Approach: Trying the test’s idea on a “pilot population” where one compares the test results versus truth; the “true diagnosis” may be based on a more refined/accurate method or evidence emerged after the passage of time (Note: here we have data).
Some tests yield dichotomous results, such as the presence or absence of a bacterium or some specific DNA sequence; but many may involve assessments measured on ordinal scale (e.g., 5-point scale for mammograms) or continuous scale (e.g., blood glucose). Let start with the simple binary case.
Let “D” and “T” denote the true diagnosis and the test result, respectively.

The key parameters are two conditional probabilities:
- Sensitivity, $S^+ = \Pr(T=+|D=+)$
- Specificity, $S^- = \Pr(T=-|D=-)$

Sensitivity is the probability to correctly identify a diseased individual and Specificity the probability of correctly identify a healthy individual.
The idea, in the developmental stage, was to classify people as “diseased” (condition present) or “healthy” (condition absent) based on certain measurement (from blood or urinary components). The basic question is “How high is high?” or “How low is low?”.
Separator X is normally distributed with the same variance, but different means; no matter where you “cut”, both errors result!
MISCLASSIFICATION

<table>
<thead>
<tr>
<th></th>
<th>Test=Positive</th>
<th>Test=Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diseased</strong></td>
<td>True Positive</td>
<td>False Negative</td>
</tr>
<tr>
<td><strong>Healthy</strong></td>
<td>False Positive</td>
<td>True Negative</td>
</tr>
</tbody>
</table>
Cervical cancer was probably the most frequent malignancy in Western Europe in the middle of the 19th century. Its natural history makes it detectable in its preclinical phase and increases the chances of cure and of mortality reduction. Before World War II there had only been pilot program of cancer detection. After World War II, the development of a cervical cancer cytological test (mainly the Pap smear) created a public health paradox previously unknown in the field of cancer: incidence increased and mortality decreased, mainly because many of the newly detected tumors were in situ and had therefore an excellent prognosis after surgery.
An early and maybe the earliest cancer detection center was established in 1937 by Dr. Elise L’Esperance in New York City and offered comprehensive examinations to asymptomatic adults for the purpose of early cancer diagnosis. The women received a cervical Papanicolaou smear with a confirming surgical biopsy on finding a suspicious lesion. If a precancerous lesion was detected, patients were referred to a surgeon for excision of the lesion or were carefully observed in the clinic. In 1937, there were only 71 applicants; in 1946, the number went up to 1356 & there were 3016 return visits.
**Example:** Some Historical Data

**CERVICAL CANCER**

This “Pap” Test is highly specific (Specificity = 98.5%) but not very sensitive (Sensitivity = 40.6%). If a healthy person is tested, the result is almost sure negative; but if a woman with cancer is tested the chance is 59.4% that the disease is undetected.

<table>
<thead>
<tr>
<th>True</th>
<th>Test</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>-</td>
<td>23,362</td>
<td>362</td>
</tr>
<tr>
<td>+</td>
<td>225</td>
<td>154</td>
</tr>
</tbody>
</table>

Sensitivity = \(\frac{154}{379} = 0.406\) or 40.6%

Specificity = \(\frac{23,362}{23,724} = 0.985\) or 98.5%
A cancer that forms in the tissues of the organ connecting the uterus and vagina

A slow-growing often asymptomatic cancer

11,070 new cases in the U.S. in 2008 (NCI)

3,870 deaths in the U.S. in 2008 (NCI)

Worldwide, it’s No. 2 “most common cancer” for women.
THE PAP TESTS

The “Pap” test or Pap Smear test is an important part of women’s health care. The smeared cells or cell suspension is placed on a glass slide, stained with a special dye (Pap stain), and viewed under a microscope.

Pap test look for abnormal cells in the lining of the cervix before they have the chance to become pre-cancer or cervical cancer. It may be not sensitive due to cases in early stage.

FDA approved in 1990s a liquid-based technique, called “Thin-layer Pap” with improved sensitivity and specificity – plus 2 DNA tests.
Certain types of the human papillomavirus (HPV) infection is a primary cause; for most women who have HPV, the virus will go away on its own. If not go away, abnormal cell can develop in the lining of the cervix; if not found early & treated, precancer and then cervical cancer can develop.

An HPV vaccine, called Gardasil; approved by the FDA in 2006. However, even with the vaccine, screening is still highly recommended by experts.
In the Applicational Stage - when the product is on the market, the basic question is: Does it work for “me”? (the user/testee; or when does it work?); It’s the user’s (or consumer’s) concern.

Problem: One can’t resolve the concern, like comparing the test result versus the truth, because if one knows the truth one would not need the test; and waiting for evidence to emerge after some passage of time may be “too late”.

Need to know if you could “trust” before trying!
Again, let “D” and “T” denote the true diagnosis and the test result, “D” is unknown in this stage.

The key parameters are two “other” conditional probabilities:

- Positive Predictive Value, $P^+ = \Pr(D=+|T=+)$
- Negative Predictive Value, $P^- = \Pr(D=-|T=-)$

Positive predictive value, or positive predictivity, is the probability have an accurate positive result and negative predictive value is the probability have an accurate negative result; Perhaps, users are more often concerned about $P^+$ than $P^-$. 
Simple Illustration: Should We Conduct “RANDOM TESTING” For Diseases, Such As AIDS?

- Those against the practice often cite concerns about errors, privacy and confidentiality, and “unwanted consequences” (such as job’s loss).

- Those promoting the practice, eg. policy makers, often want to know “the magnitude of the problem” in order to justify spending on research as well as interventions.
Assumptions: Let assume

- Complete privacy
- Complete confidentiality
- There is a good/reliable screening procedure (say, 98% sensitive and 97% specific)
- Consider 2 examples, a low-risk sub-population (Example A, prevalence is .1%) and a high-risk sub-population (Example B, prevalence is 20%) – then see what we can learn.
**Example A:**

<table>
<thead>
<tr>
<th></th>
<th>Infection=Yes</th>
<th>Infection=No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test=Positive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test=Negative</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>99900</td>
<td>100000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Infection=Yes</th>
<th>Infection=No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test=Positive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test=Negative</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>99900</td>
<td>100000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Infection=Yes</th>
<th>Infection=No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test=Positive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test=Negative</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>99900</td>
<td>100000</td>
</tr>
</tbody>
</table>
**RESULTS**

- True prevalence: 100/100,000 = .1%
- Estimated prevalence: 3,095/100,000 = 3.1%
  - Not good for policy makers: Over estimate more than 30 times;

- $P^+ = \frac{98}{3,095} = 3.2\%$; Not good for users: very low Positive Predictive Value ($P^+$)

<table>
<thead>
<tr>
<th></th>
<th>Infection=Yes</th>
<th>Infection=No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test=Positive</td>
<td>98</td>
<td>2997</td>
<td>3095</td>
</tr>
<tr>
<td>Test=Negative</td>
<td>2</td>
<td>96903</td>
<td>90905</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>99900</td>
<td>100000</td>
</tr>
</tbody>
</table>
### Example B:

<table>
<thead>
<tr>
<th></th>
<th>Infection=Yes</th>
<th>Infection=No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test=Positive</strong></td>
<td>19600</td>
<td>2400</td>
<td>22000</td>
</tr>
<tr>
<td><strong>Test=Negative</strong></td>
<td>400</td>
<td>77600</td>
<td>78000</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>20000</td>
<td>80000</td>
<td>100000</td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th></th>
<th>Infection=Yes</th>
<th>Infection=No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test=Positive</strong></td>
<td>19600</td>
<td>2400</td>
<td>22000</td>
</tr>
<tr>
<td><strong>Test=Negative</strong></td>
<td>400</td>
<td>77600</td>
<td>78000</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>20000</td>
<td>80000</td>
<td>100000</td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th></th>
<th>Infection=Yes</th>
<th>Infection=No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test=Positive</strong></td>
<td>19600</td>
<td>2400</td>
<td>22000</td>
</tr>
<tr>
<td><strong>Test=Negative</strong></td>
<td>400</td>
<td>77600</td>
<td>78000</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>20000</td>
<td>80000</td>
<td>100000</td>
</tr>
</tbody>
</table>

---
RESULTS

- True prevalence: \(20,000/100,000 = 20\%\)  
  Estimated prevalence: \(22,000/100,000 = 22\%\)  
  Good for policy makers: 22\% versus 20\%

- \(P^+ = 19600/22,000 = 89.1\%\); Good for users that they can “trust” the results (before using): reasonable

Predictive Value (P+)

<table>
<thead>
<tr>
<th></th>
<th>Infection=Yes</th>
<th>Infection=No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test=Positive</td>
<td>19600</td>
<td>2400</td>
<td>22000</td>
</tr>
<tr>
<td>Test=Negative</td>
<td>400</td>
<td>77600</td>
<td>78000</td>
</tr>
<tr>
<td>Total</td>
<td>20000</td>
<td>80000</td>
<td>100000</td>
</tr>
</tbody>
</table>
**BAYES’ RULE**

\[
Pr(B \mid A) = \frac{Pr(B \text{ and } A)/Pr(A)}{Pr(A \text{ and } B) + Pr(A \text{ and } \text{Not } B)}
\]

\[
Pr(B \mid A) = \frac{Pr(A \mid B)Pr(B)}{Pr(A \mid B)Pr(B) + Pr(A \mid \text{Not } B)Pr(\text{Not } B)}
\]
Let $A = (T=+)$ and $B = (D=+)$, we have:

$$
\text{Note: } \text{“not } B\text{”} = (D=-)
$$
Both predictive values are functions of disease prevalence, $\pi = \Pr(D = +)$:

\[
P^+ = \frac{S^+ \pi}{S^+ \pi + (1 - S^-)(1 - \pi)}
\]

\[
P^- = \frac{S^- (1 - \pi)}{S^- (1 - \pi) + (1 - S^+) \pi}
\]
EXAMPLES

Example A: \( S^+ = .977, S^- = .926, \) and \( \pi = .003: \)

\[
P^+ = \frac{(0.977)(0.003)}{(0.977)(0.003) + (0.074)(0.997)} = 0.038 \text{ or } 3.8\%
\]

Example B: \( S^+ = .977, S^- = .926, \) and \( \pi = .20: \)

\[
P^+ = \frac{(0.977)(0.20)}{(0.977)(0.20) + (0.074)(0.80)} = 0.767 \text{ or } 76.7\%
\]

(Note: Current Estimate for USA’s AIDS: .3% as above and \( S^+ \) and \( S^- \) are for ELISA in Weiss, 1985).