Chapter 7 Screening and Diagnostic tests Research and Evaluation

Lina Wang

Department of Epidemiology and Biostatistics
School of Public Health, Southeast University
Phone: 025-83272569 (o)
E-mail: lnwang@seu.edu.cn

Patient Profile

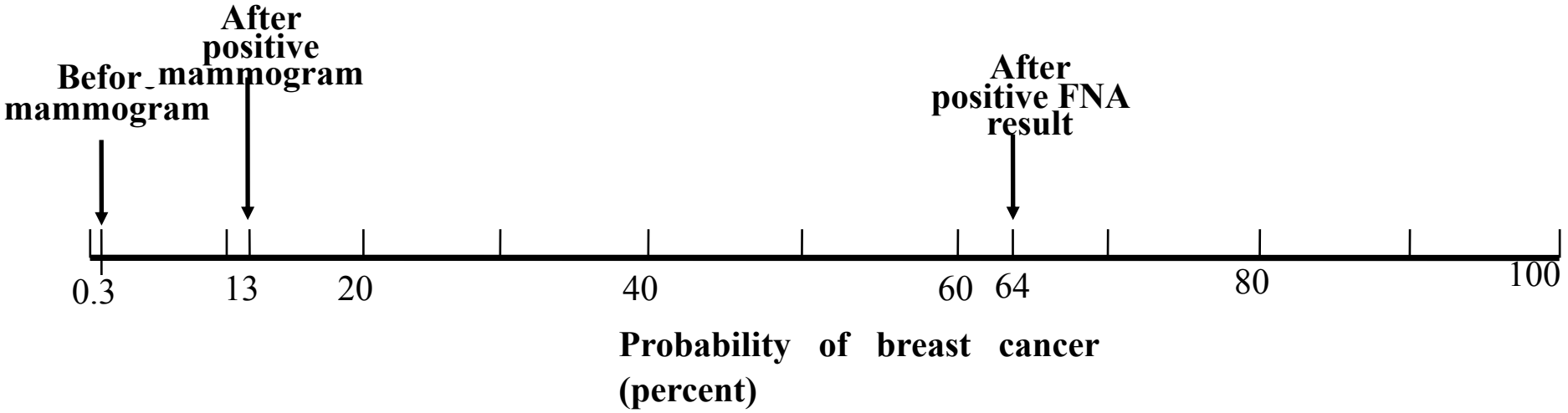
- A 54-year-old high school teacher visited her family practitioner for an annual checkup. She reported no illnesses during the preceding year, felt well, and had no complaints.

- The physician performed a physical examination, comprising breast, pelvic (including a Papanicolaou smear), rectal examinations and mammogram.

- **The results of the mammogram were not normal, so she was suggested to perform a breast biopsy.**

- **Evaluation of the fine needle aspiration (FNA) specimen by a pathologist revealed cancer cells, and the patient was scheduled for further surgery the following week.**

Schematic diagram of the estimated probabilities of breast cancer in a 54-year-old woman without a palpable breast mass

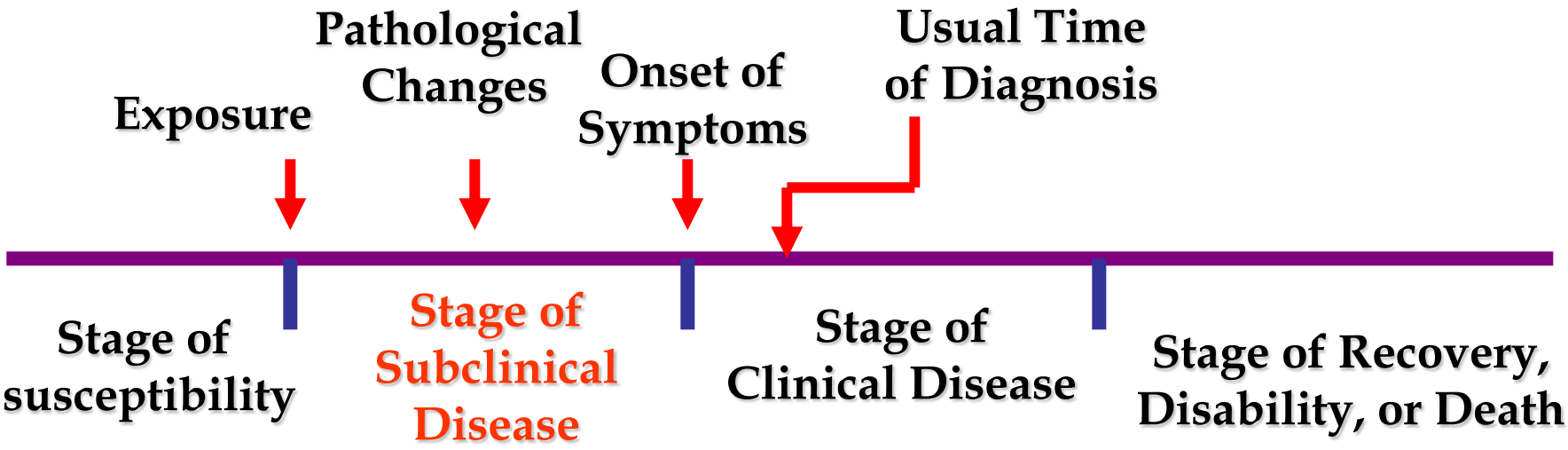


The clinical decision-making process is based on probability.

The purpose of a diagnostic test

- **is to move the estimated probability of the presence of a disease toward either end of the probability scale(0 or 100%), thereby providing information that will alter subsequent diagnostic or treatment plans.**

Natural History of Disease



Types of early diagnosis of diseases

1. Screening
2. Periodic health examination
3. Routine examination
4. Case finding
5. Self-examination

Diagnostic test

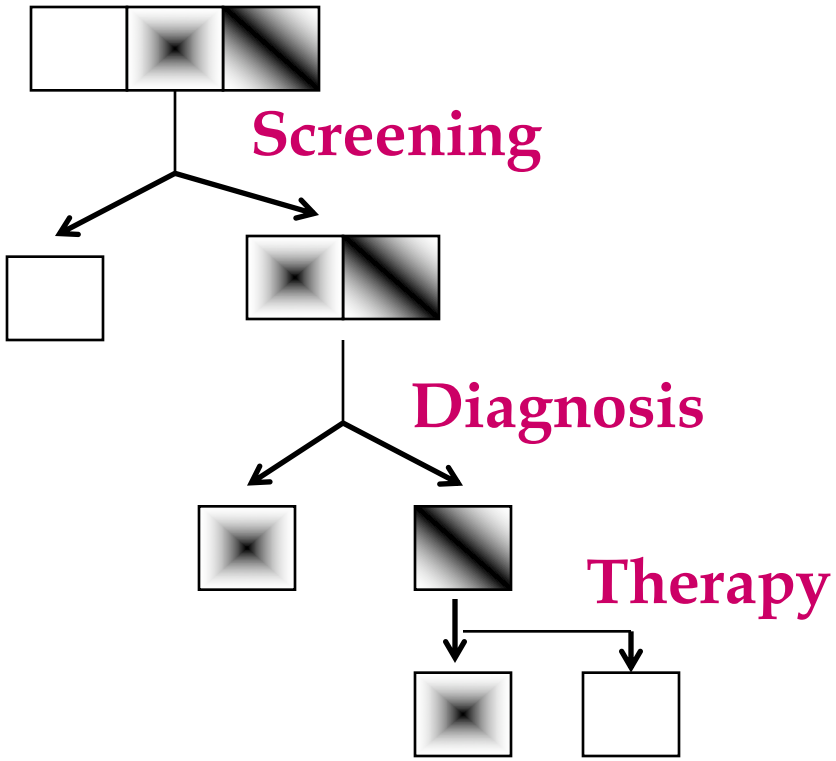
Some questions :

- What is the diagnosis?
- How would you go about establishing the diagnosis?
- What is the differential diagnosis?
- What is a case?
- What diagnostic criteria should be applied?
- How confident are you in the diagnostic test(s)?

The situation

- Patient presents with symptoms, and is suspected of having some disease. Patient either **has the disease** or **does not have the disease**.
- Physician performs a **diagnostic test** to assist in making a diagnosis.
- Test result is either **positive** (diseased) or **negative** (healthy).

Screening and diagnosis



True free of disease

Probable disease

True disease

Diagnostic test

- Diagnostic testing is used to:
  - Diagnose disease
  - Assess severity of disease
  - Predict disease outcome
  - Monitor treatment
- A diagnostic testing evaluates the probability of an outcome

Evaluation of diagnostic and screening test

- **Accuracy**
- **Reliability**

Accuracy

- How close is a test result to the truth
- Proportion of all tests, both positive and negative, that are correct

What is Truth?

- **Gold Standard**
  - The test that is used to determine if a disease is truly present or not
  - Other tests are compared to it to determine their accuracy

- Diagnostic test results maybe:
  - True positive and false positive
  - True negative and false negative
- The validity of a diagnostic testing can be evaluated in a contingency table using a "***gold standard", ---the most definitive diagnostic method***.

Format for comparison of the results of a diagnostic test with the "true" status of a disease.

Test result	"True" or "Gold standard"		Total
	Disease	No disease	
Positive	*a* True-positive	*b* False-positive	a+b
Negative	*c* False-negative	*d* True-negative	c+d
Total	a+c	b+d	a+b+c+d

Evaluation of diagnostic test

- For the FNA test, a surgical excisional biopsy procedure (ie, histopathologic examination of the tissue in question), is the gold standard, and is considered to represent the true status of the disease.

Comparison of FNA test results with findings from surgical excisional biopsies in women without palpable breast masses.

FNA result	Surgical biopsy		Total
	Cancer	No cancer	
Positive	14	8	22
Negative	1	91	92
Total	15	99	114

Sensitivity

- **The sensitivity of a test is defined as the percentage of persons with the disease of interest who have positive test results.**

$$\text{Sensitivity} = \frac{\text{True-positives}}{\text{True-positives+False-negatives}} \times 100$$

$$= \frac{a}{a+c} \times 100$$

- **So, the sensitivity of the FNA test is**

$$\text{Sensitivity} = \frac{\text{True-positives}}{\text{True-positives+False-negatives}} \times 100$$

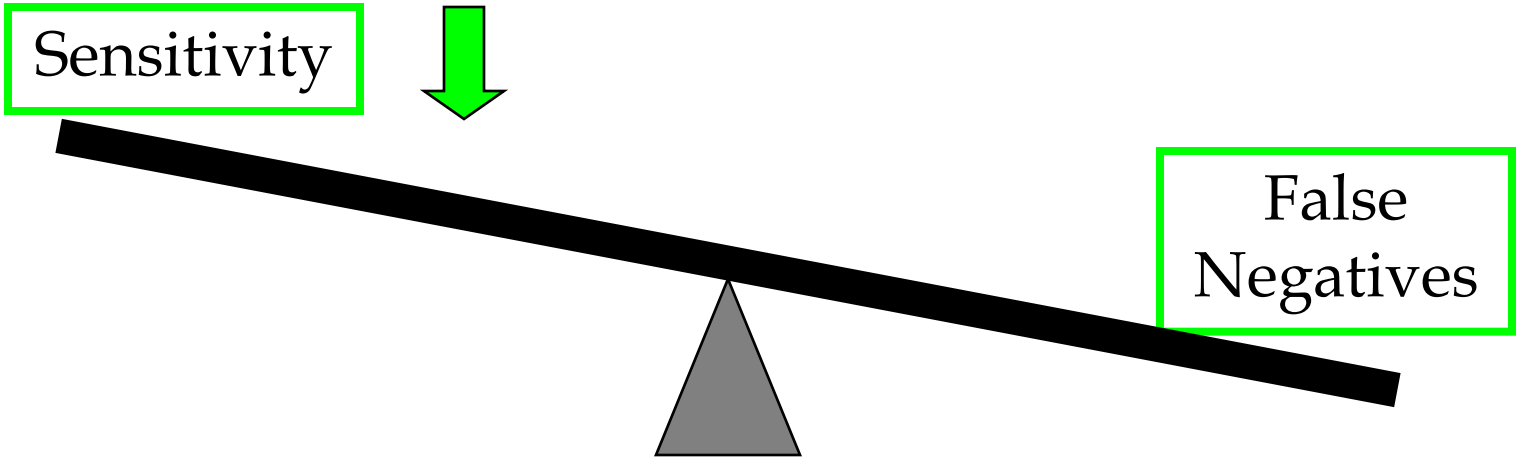
$$= \frac{14}{14+1} \times 100 = 93\%$$

The meanings of sensitivity

- The greater the sensitivity of a test, the more likely the test will detect persons with the disease of interest.
- For the FNA test, 93% of all the patients with breast cancer had positive test results.
- Tests with great sensitivity are useful clinically to rule out the presence of a disease.

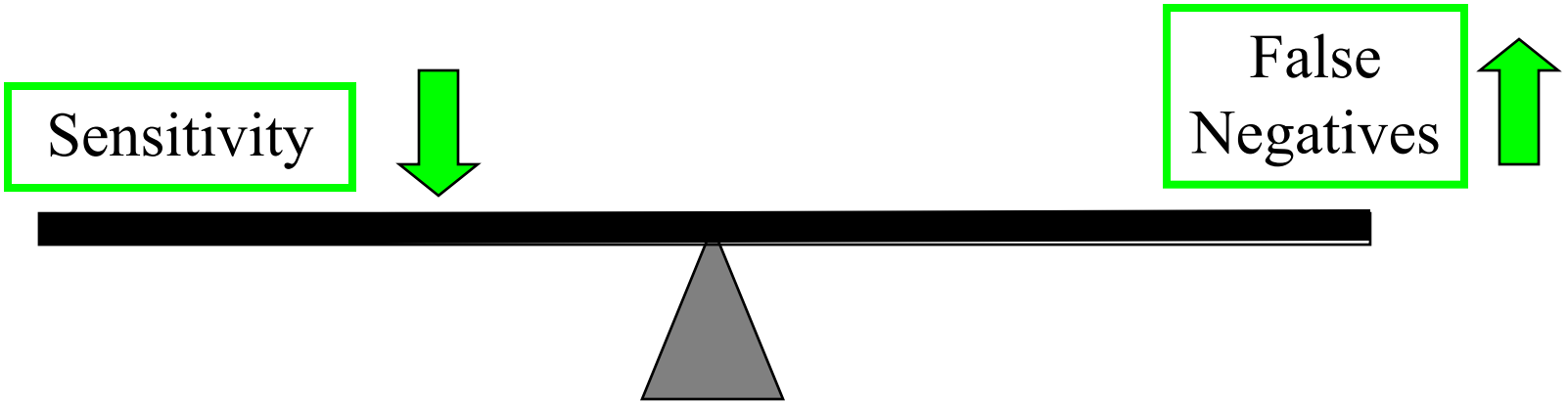
False Negative Rate = $\frac{c}{a+c} \times 100\%$

Likelihood of a negative result when patient actually has disease



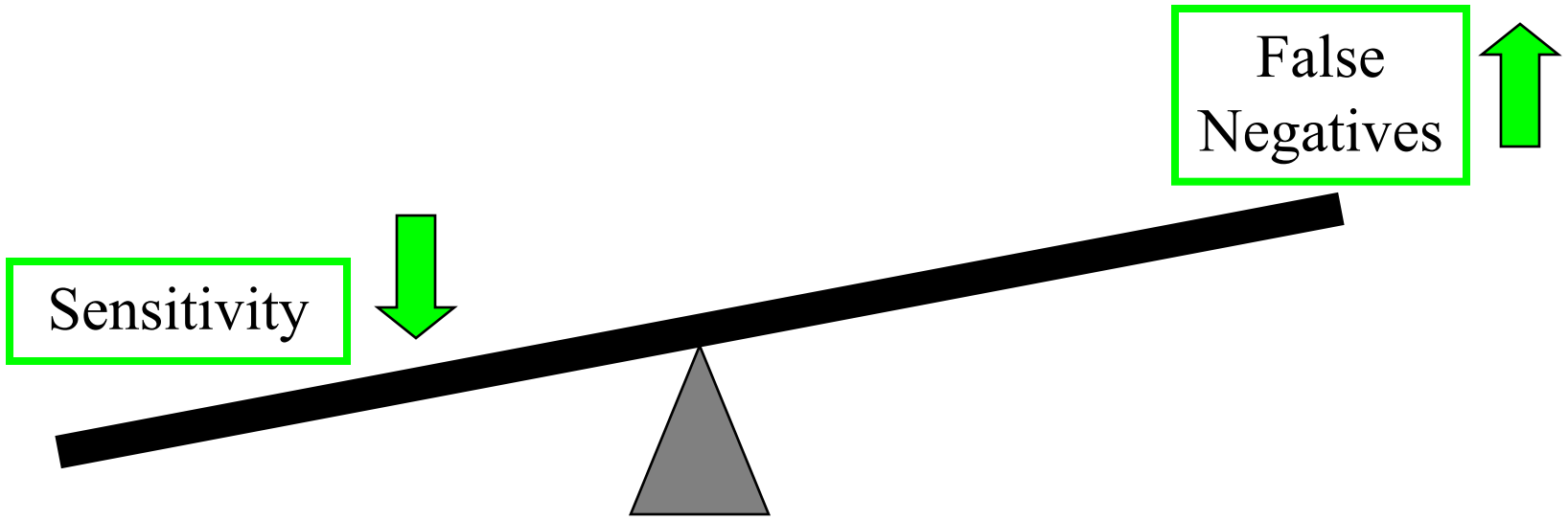
False Negative Rate

Likelihood of a negative result when patient actually has disease



False Negative Rate

Likelihood of a negative result when patient actually has disease (missed diagnostic)



False negative rate increases with decreased sensitivity

Specificity

- **Specificity of a test is defined as the percentage of persons without the disease of interest who have negative test results.**

$$\text{Specificity} = \frac{\text{True-negatives}}{\text{True-negatives+False-positives}} \times 100$$

$$= \frac{d}{d+b} \times 100$$

The ability of the test to detect freedom from disease

- **So the specificity of the FNA test is**

$$\text{Specificity} = \frac{\text{True-negatives}}{\text{True-negatives+False-positives}} \times 100$$

$$= \frac{91}{91+8} \times 100 = 92\%$$

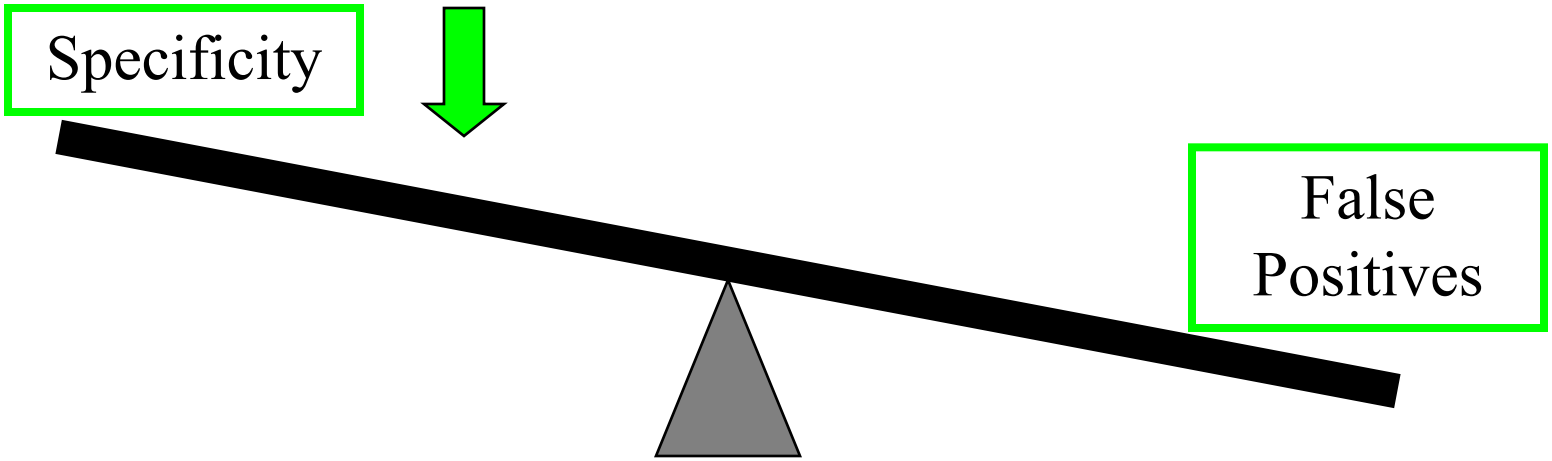
The meanings of specificity

- The greater the specificity of a test, the more likely it is that persons without the disease of interest will be excluded from consideration of having the disease.

- Very specific tests often are used to confirm the presence of a disease.

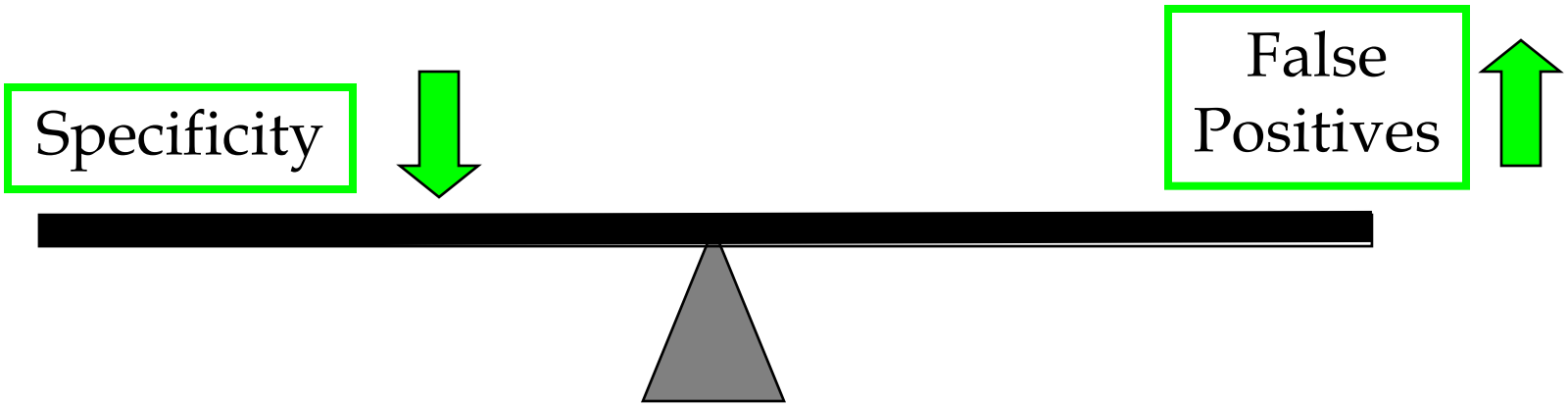
False Positive Rate = $\frac{b}{b+d} \times 100\%$

Likelihood of a positive result when patient does not have the disease (misdiagnostic)



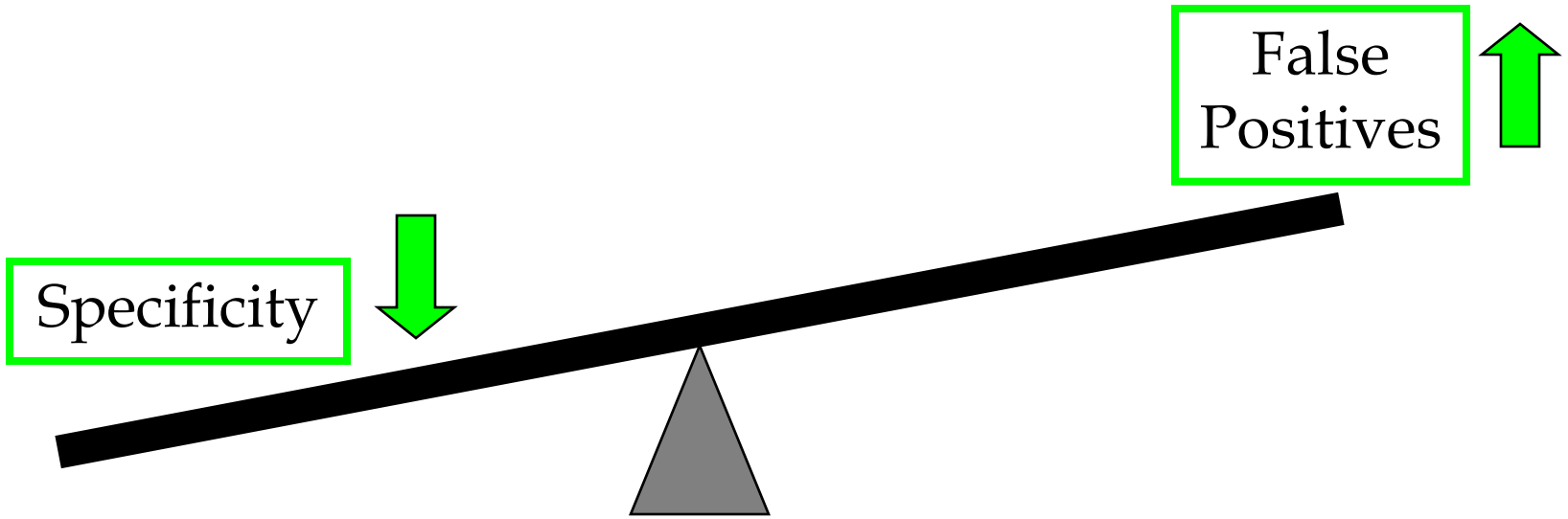
False Positive Rate

Likelihood of a positive result when patient does not have the disease



False Positive Rate

Likelihood of a positive result when patient does not have the disease



False positive rate increases with decreased specificity

Positive & Negative Predictive Value

- **Positive predictive value (*PV*+)**
- **Negative predictive value (*PV*-)**

Positive Predictive Value (*PV*+)

- **The PV+ is defined as the percentage of persons with positive test results who actually have the disease of interest.**

$$PV+ = \frac{\text{True-positives}}{\text{True-positives+False-positives}} \times 100$$

$$= \frac{a}{a+b} \times 100$$

- **The PV+ is the percentage of persons with positive test results who have the disease.**
- **The calculation of the PV+ for the FNA test described as:**

$$PV+ = \frac{\text{True-positives}}{\text{True-positives+False-positives}} \times 100$$

$$= \frac{14}{14+8} \times 100 = 64\%$$

The probability of breast cancer for a woman with a positive test result increased to 64%.

Negative predictive value (*PV-*)

- **The PV－is defined as the percentage of persons with negative test results who do not have the disease of interest.**

$$PV- = \frac{\text{True-negatives}}{\text{True-negatives+False-negatives}} \times 100$$

$$= \frac{d}{d+c} \times 100$$

- For the FNA test data, the ***PV-*** is

$$PV- = \frac{\text{True-negatives}}{\text{True-negatives+False-negatives}} \times 100$$

$$= \frac{91}{91+1} \times 100 = 99\%$$

After a negative FNA result, the probability of not having breast cancer increased to 99%.

The usefulness of the FNA test

- **A positive test result increased the probability of breast cancer from 13%(15/114) to 64%(14/22) ,a further workup such as surgical test biopsy is indicated.**

- **A negative test result, however, would reduce the probability that breast cancer is present to 1% (100%- PV^- = 1%) , accepting a 1 in 100 risk of mistakenly delaying treatment of an existing cancer.**

Influences of the Predictive value

- The pretest probability of the presence of disease in an individual.
- The prevalence of disease in a population.

Comparison of FNA test results with findings from surgical excisional biopsies in women with or without palpable breast masses.

Without palpable masses

FNA result	Surgical biopsy		Total
	Cancer	No cancer	
Positive	14	8	22
Negative	1	91	92
Total	15	99	114

Prevalence =13%
Sensitivity =93%
Specificity =92%
PV^+ =64%
PV^- =99%

With palpable masses

FNA result	Surgical biopsy		Total
	Cancer	No cancer	
Positive	113	15	128
Negative	8	181	189
Total	121	196	317

Prevalence =38%
Sensitivity =93%
Specificity =92%
PV^+ =88%
PV^- =96%

Relationship between the PV and the prevalence rate of the disease

- The PV+ rose with an increase in the pretest probability of disease.

- The PV- decreased as the pretest probability of the presence of disease increased, it becomes easier to exclude a disease as the pretest prevalence of disease decreases.

Effect of prevalence on positive predictive value when sensitivity and specificity of a test equal 95%

Prevalence (%)	Positive Predictive Value (%)
50.0	95.0
5.0	50.0
2.0	27.9
1.0	16.1
0.1	1.9

The difference in predictive values and the clinical implications

- Among women with nonpalpable lesions, a negative FNA test result (PV- = 99%) could reduce the probability of the presence of disease to 1%, and therefore obviate the need for a surgical biopsy.

- In the group with palpable masses, the probability of breast cancer still would be 4%(1-96%), which might warrant further testing, for example, a surgical biopsy.

Cutoff Points

- In reality, many test results occurs along a **continuum**.
- Extreme results reflecting lab-instrument error or abnormal variability of measure (abnormal patient response)
- **Changing cutoff point affect test performance**

- **cutoff point,**

  which is the point at which a test result is considered to change from negative to positive.

For example: blood pressure (BP)

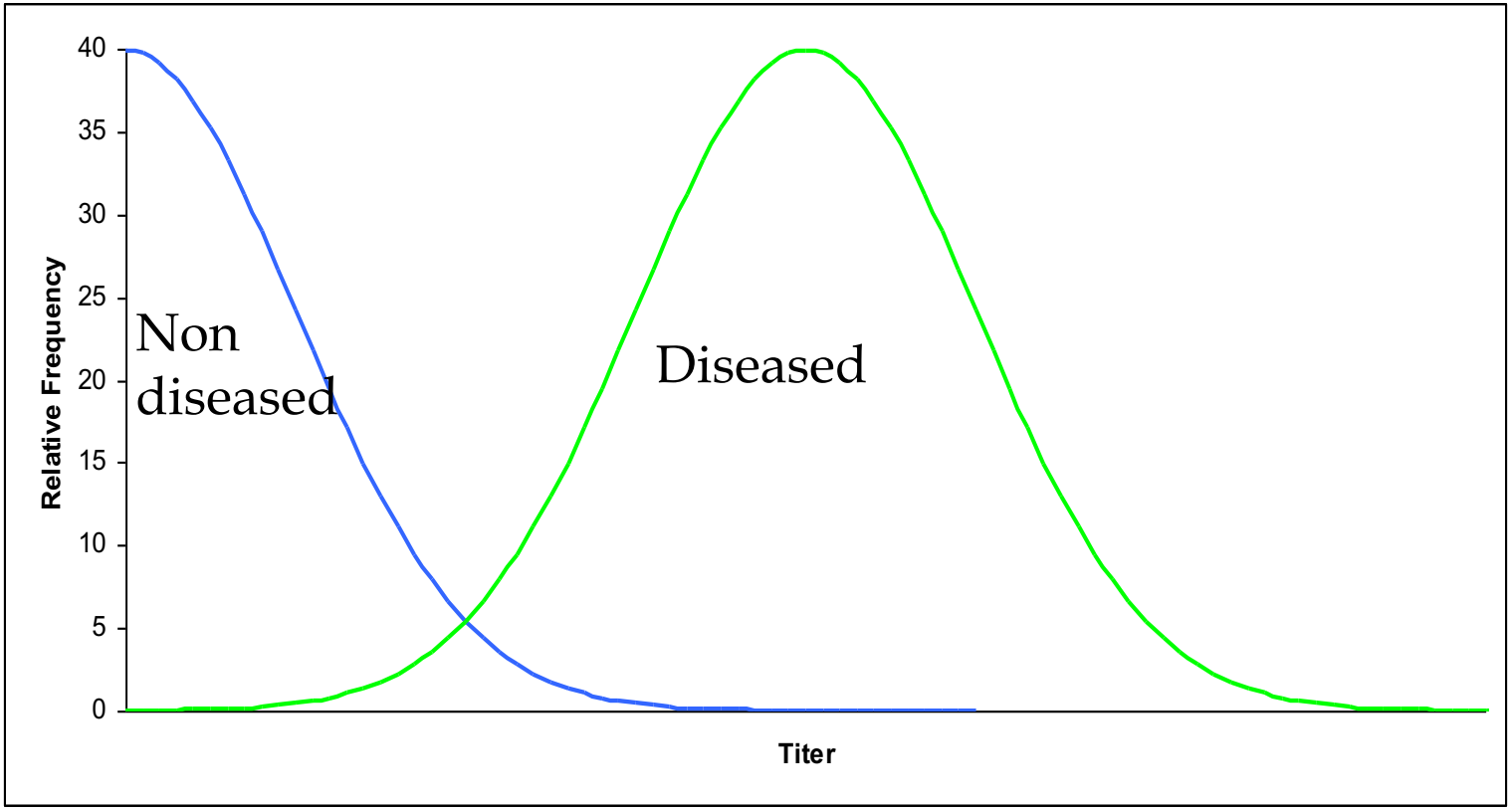
cutoff point: 140/90mmHg

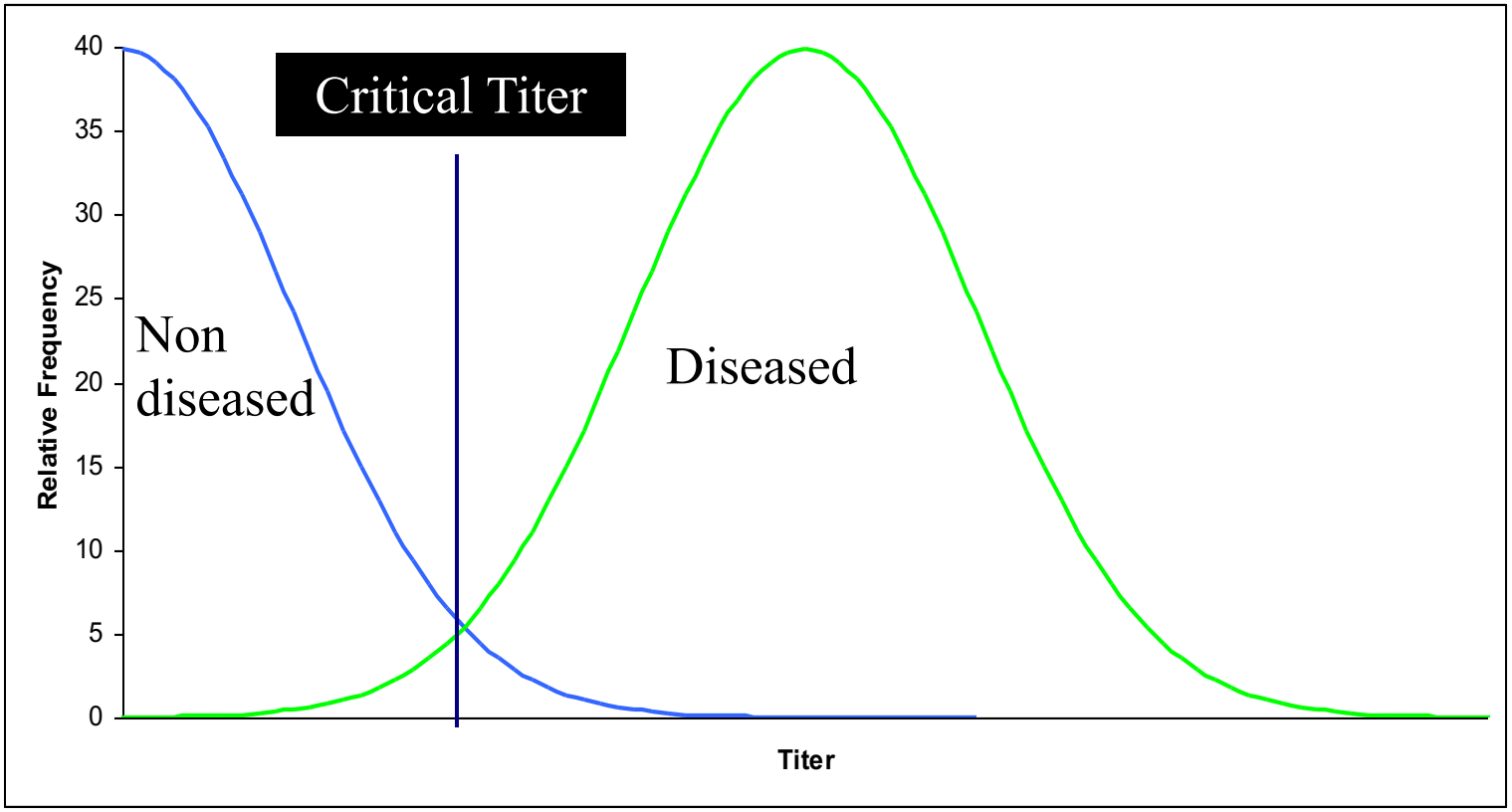
90/60mmHg≤BP<140/90mmHg: Normal

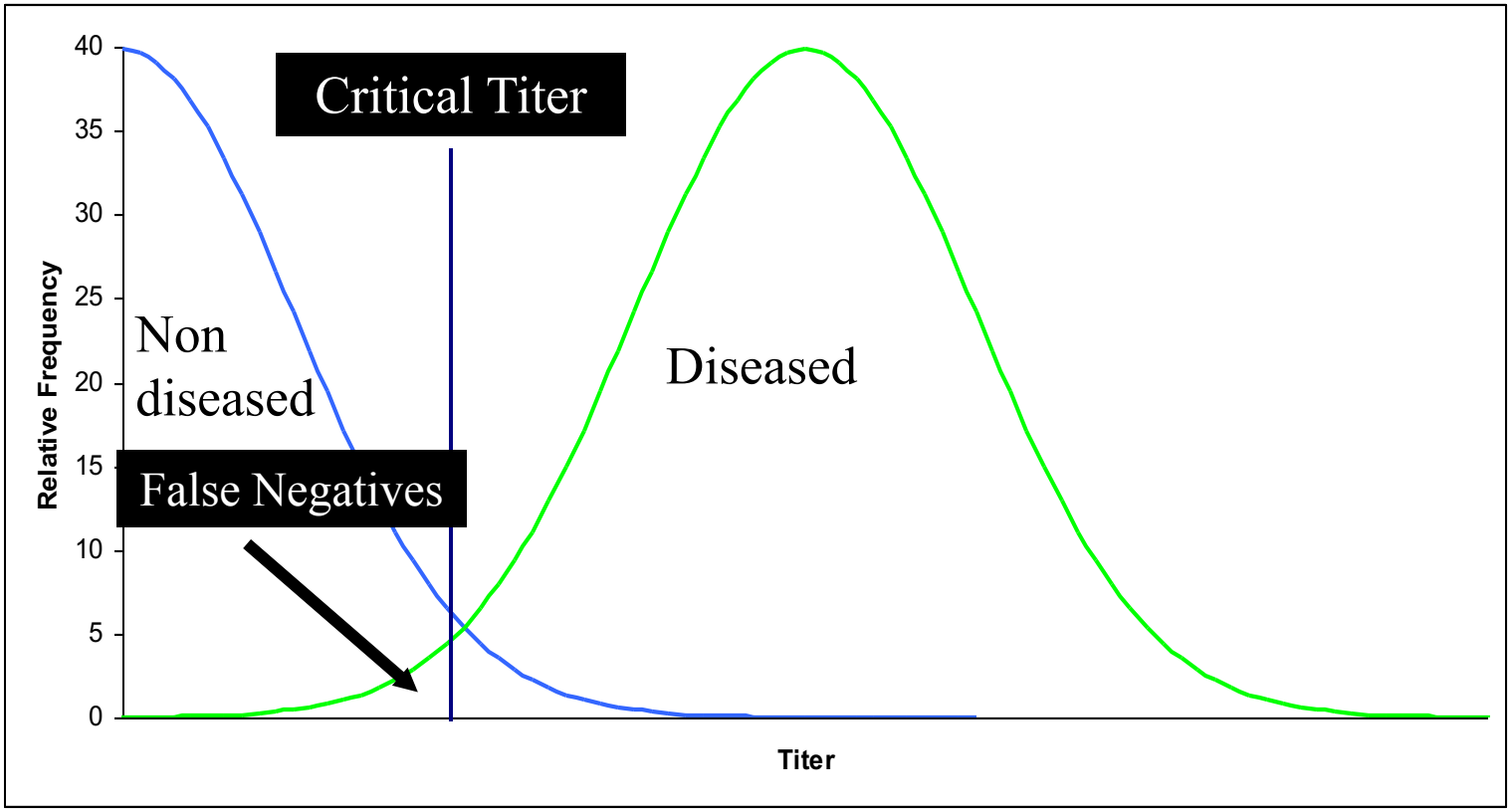
BP< 90/60mmHg: Hypotension

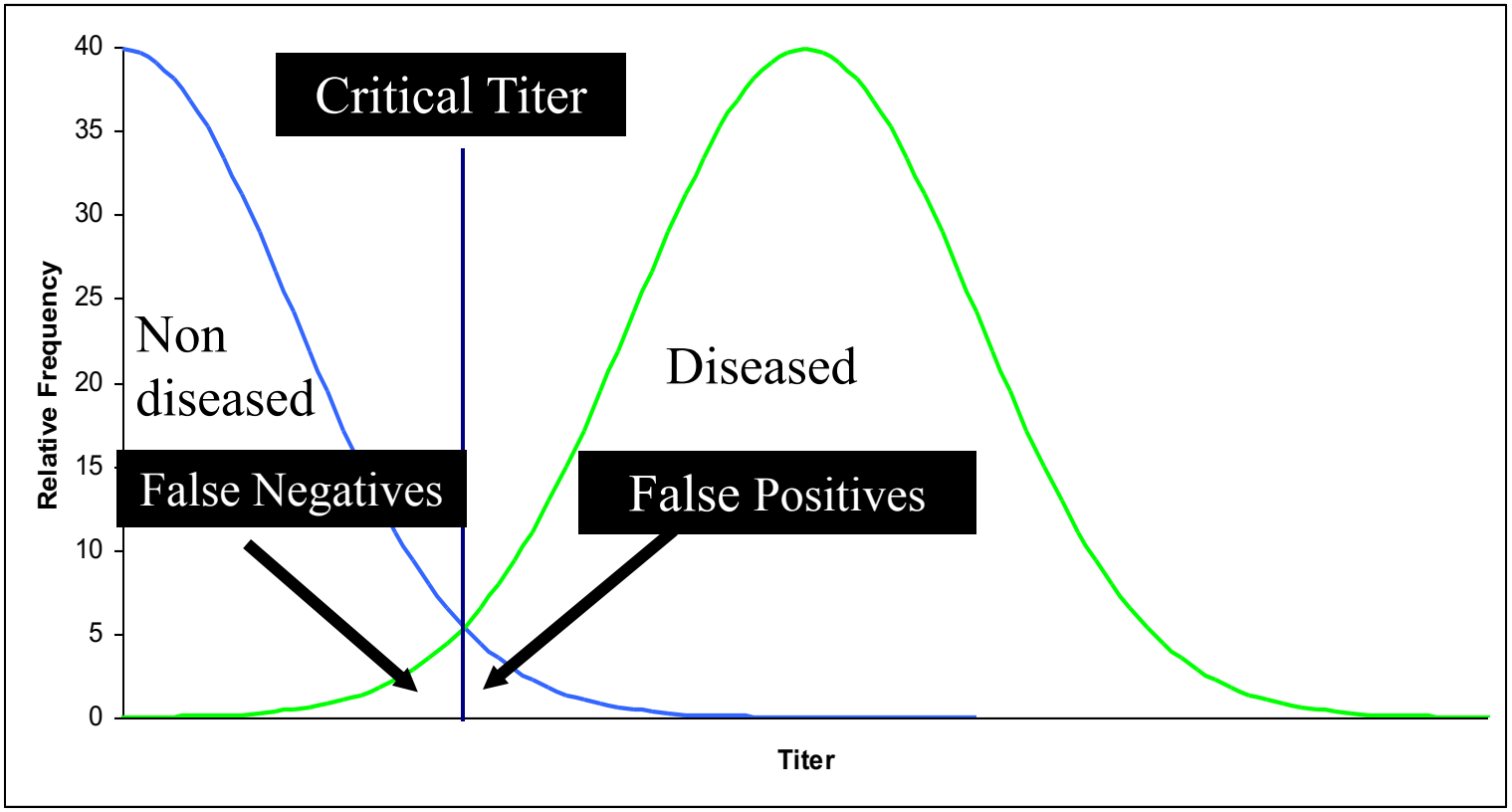
BP≥140/90mmHg: Hypertension

Relationship of Sensitivity and Specificity

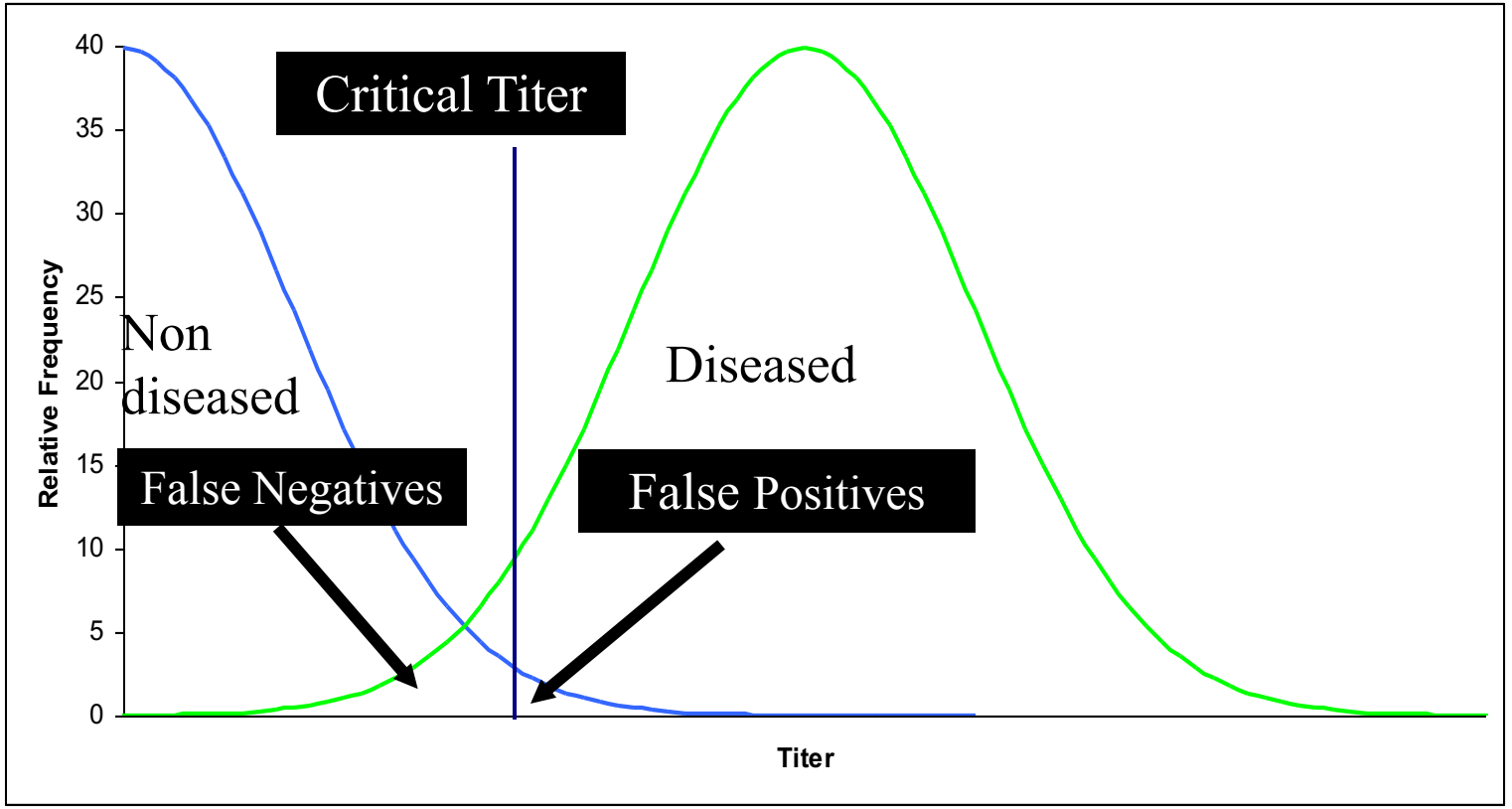



Critical Titer
Non diseased
Diseased
Relative Frequency
Titer
40
35
30
25
20
15
10
5
0

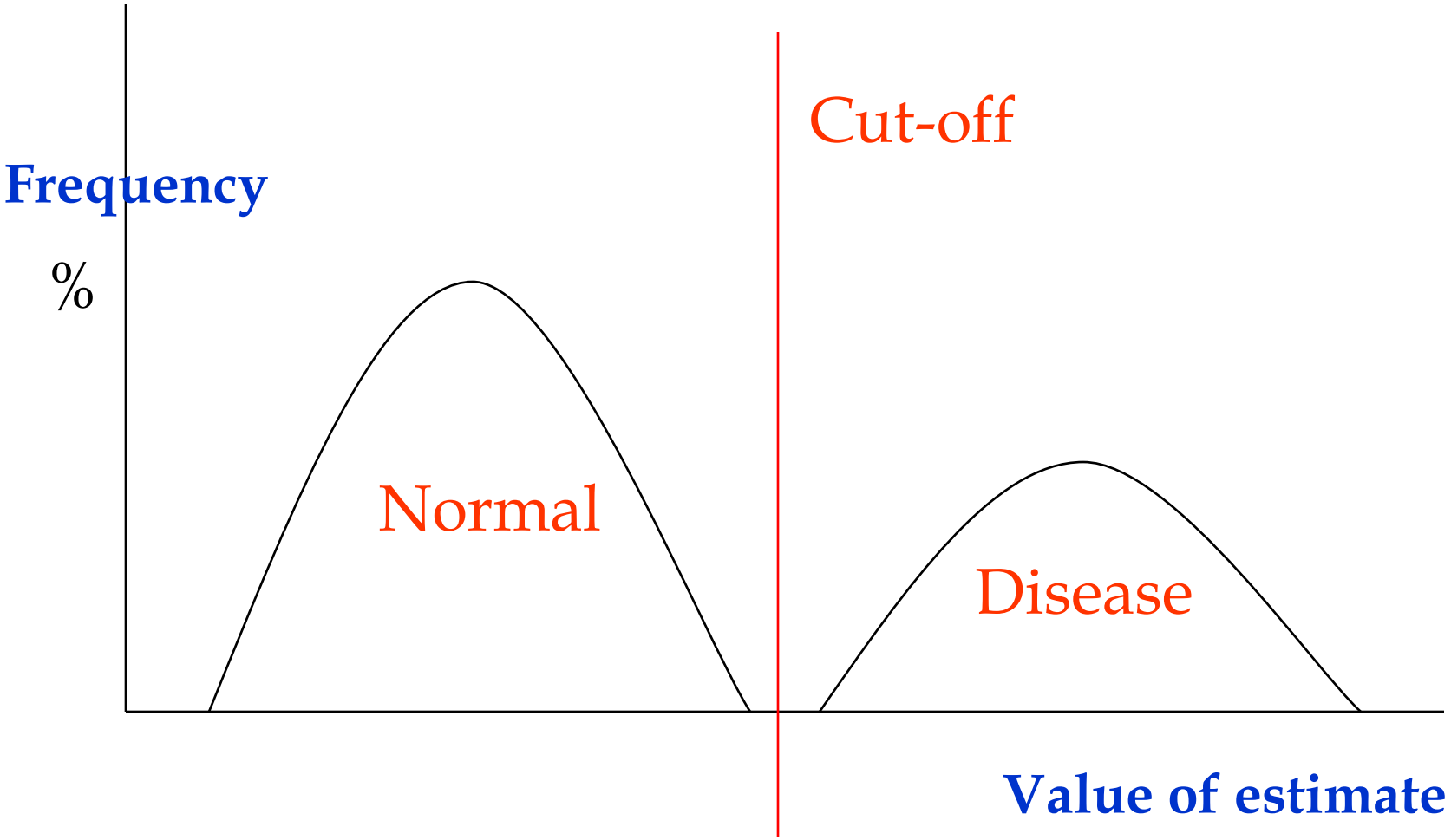

Critical Titer
Non diseased
Diseased
False Negatives
Relative Frequency
Titer
40
35
30
25
20
15
10
5
0


Critical Titer
Non diseased
Diseased
False Negatives
False Positives
Relative Frequency
Titer
40
35
30
25
20
15
10
5
0

Decreased Sensitivity – Increased Specificity

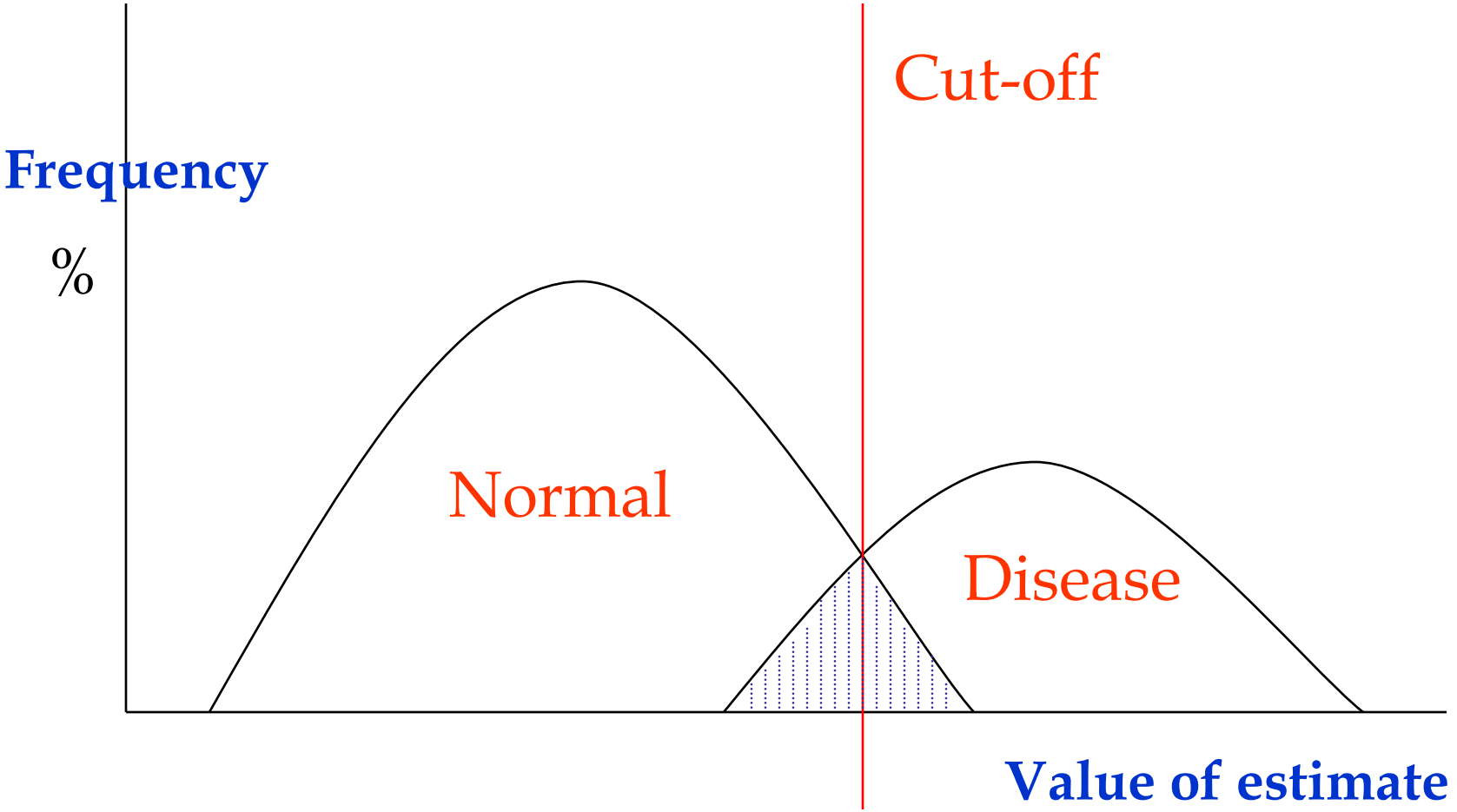


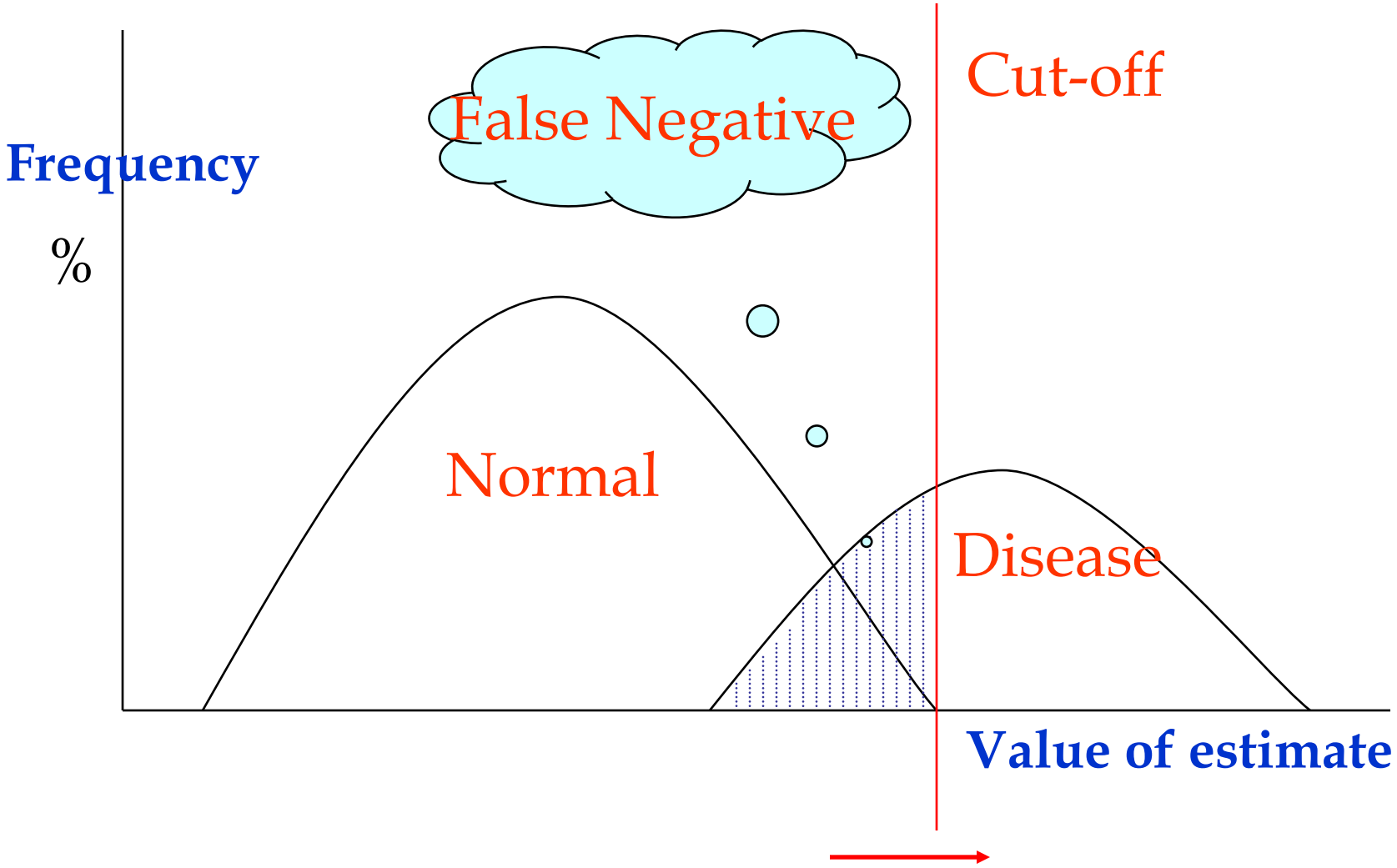
Ideal situation

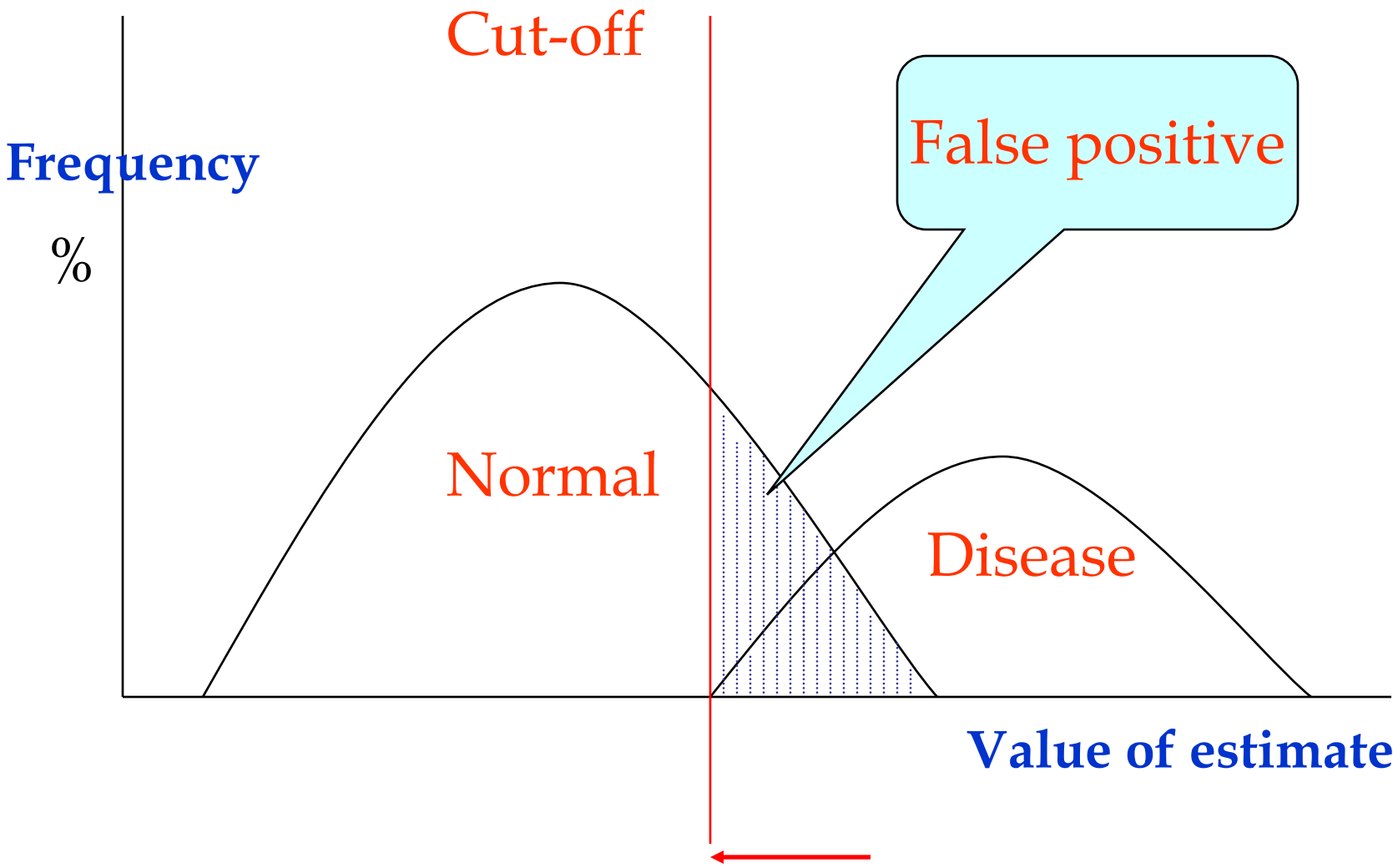


Cutoff point

Actual situation




Frequency
%
False Negative
Cut-off
Normal
Disease
Value of estimate


Cut-off
False positive
Frequency
%
Normal
Disease
Value of estimate

The results of the FNA test often are classified

(1) Insufficient material to adequately assess presence or absence of malignant cells

(2) Benign (no malignant cells present)

(3) Suspicious (atypical cells present but not definitely malignant)

(4) Malignant cells present.

- **The important decision concerning FNA is whether to classify women with suspicious or atypical results as positive or negative.**

- **Employs two different assumptions:**

  (1) suspicious FNA results were considered to be positive.

  (2) suspicious FNA results were considered to be negative.

Suspicious FNA results considered positive (Cutoff value ↓)

	Surgical biopsy		
FNA result	Cancer	No cancer	Total
Positive	113	15	128
Negative	8	181	189
Total	121	196	317

Prevalence =38%
Sensitivity =93%
Specificity =92%
PV^+ =88%
PV^- =96%

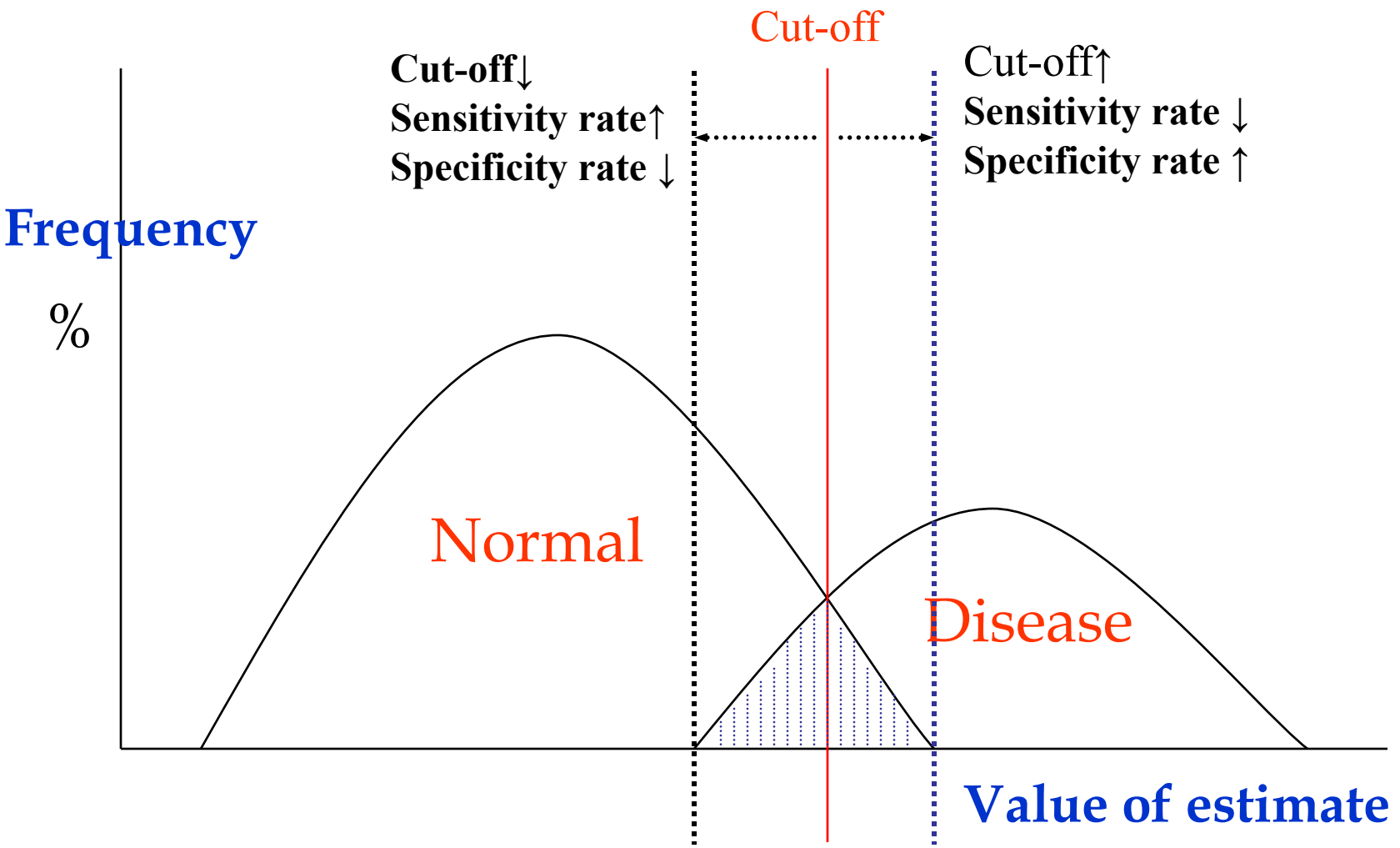
Suspicious FNA results considered negative (Cutoff value ↑)

	Surgical biopsy		
FNA result	Cancer	No cancer	Total
Positive	91	0	91
Negative	30	196	226
Total	121	196	317

Prevalence =38%
Sensitivity =75%
Specificity =100%
PV^+ =100%
PV^- =87%

- Note that the ***sensitivity*** of FNA ***decreased*** and the ***specificity increased*** when women with ***suspicious*** FNA findings were considered to have ***negative*** test results.

- Moving the **cutoff point** changes the test's sensitivity, specificity, positive and negative predictive values, and, hence, the way in which the test is used.


Cut-off
Cut-off↓
Sensitivity rate↑
Specificity rate ↓
Cut-off↑
Sensitivity rate ↓
Specificity rate ↑
Frequency
%
Normal
Disease
Value of estimate

Likelihood Ratios (LR)

- An LR is the probability of a particular test result for a person with the disease of interest divided by the probability of that test result for a person without the disease of interest.

- For the dichotomous results, LRs are defined for either positive or negative test results.

The likelihood ratio for a positive test (LR+)

- the probability of a positive test result for a person with the disease of interest divided by the probability of a positive test result for a person without the disease.

$$LR+ = \text{Sensitivity}/(1\text{-Specificity})$$

$$= \frac{\text{True positive rate}}{\text{False positive rate}}$$

The larger the value of the *LR*+, the stronger the association between having a positive test result and having the disease of interest.

The likelihood ratio for a negative test (LR－)

- the probability of a negative test result for a person with the disease of interest divided by the probability of a negative test result for a person without the disease.

$$LR- = (1 - \text{Sensitivity})/\text{Specificity}$$

$$= \frac{\text{False negative rate}}{\text{True negative rate}}$$

The smaller the value of the LR-, the stronger the association between having a negative test result and not having the disease of interest.

- For example:

$$LR+ = \text{Sensitivity}/(1-\text{Specificity})$$
$$= 0.93/(1-0.92) = 0.93/0.08 = 11.63$$

$$LR- = (1-\text{Sensitivity})/\text{Specificity}$$
$$= (1-0.93)/0.92 = 0.07/0.92 = 0.08$$

The likelihood ratios do not vary as a function of the prevalence of disease

For women without palpable breast masses, the prevalence of disease is 13%

$$LR+ = \text{Sensitivity}/(1-\text{Specificity})$$
$$= 0.93/(1-0.92) = 0.93/0.08 = 11.63$$

$$LR- = (1-\text{Sensitivity})/\text{Specificity}$$
$$= (1-0.93)/0.92 = 0.07/0.92 = 0.08$$

For women with palpable breast masses, the prevalence of disease is much greater--38%.

$$LR+ = \text{Sensitivity}/(1-\text{Specificity})$$
$$= 0.93/(1-0.92) = 0.93/0.08 = 11.63$$

$$LR- = (1-\text{Sensitivity})/\text{Specificity}$$
$$= (1-0.93)/0.92 = 0.07/0.92 = 0.08$$

Diagnostic value

- The larger the size of the *LR*+, the better the diagnostic value of the test.
- The smaller the size of the *LR*—, the better the diagnostic value of the test.

Receiver Operating Characteristic ROC curves

- Discrete categories of possible outcome

such as: evaluation of specimen (benign, suspicious, and malignant)

The cutoff point between positive and negative test results can be set to have either a restrictive definition of positive test results (malignant cells present) or a more inclusive definition of positive test results (either suspicious or malignant cells present).

- Continuous scale of measurement,

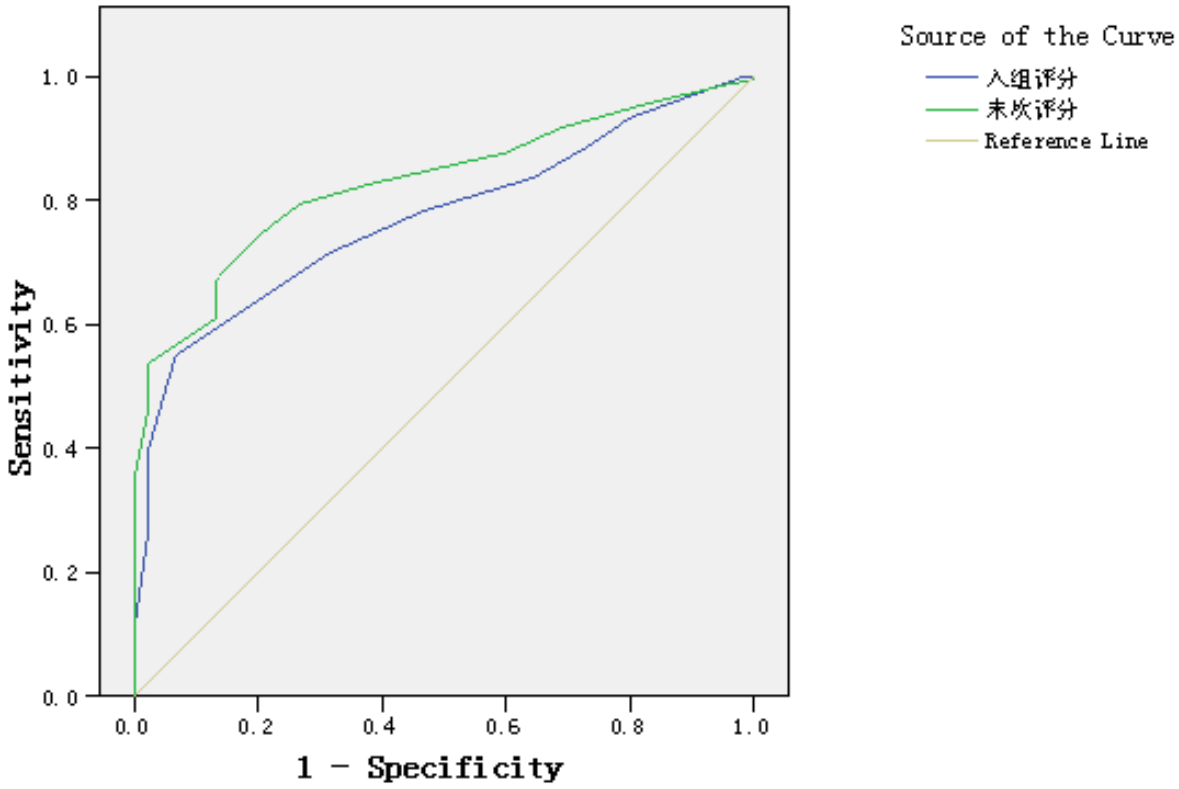
such as: serum levels of enzymes.

Different cutoff points will result in differing levels of sensitivity or specificity.

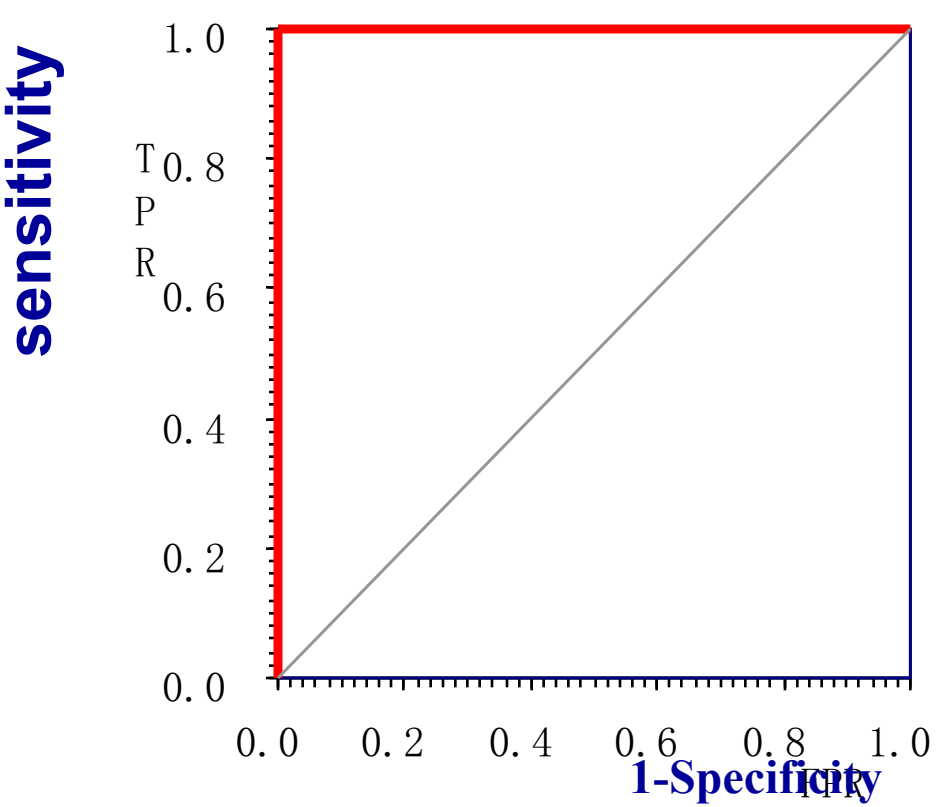
ROC curves

- For the purposes of diagnostic testing, a graph is constructed with sensitivity (sometimes labeled as the true-positive rate) on the vertical axis and 1 – specificity (sometimes labeled as the false-positive rate) on the horizontal axis.

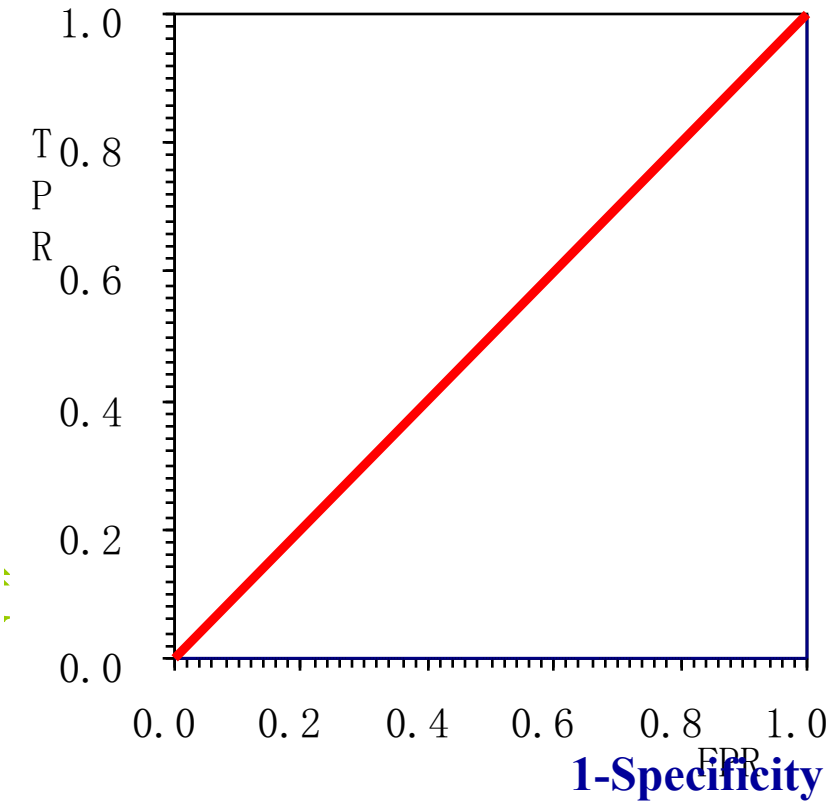
ROC curves



Cutoff	10	11	12	13	14	15	16	17	18	19	20
Se	26.7	35.6	53.3	68.9	80.0	93.3	95.6	97.8	97.8	97.8	100
Sp	88.7	83.6	78.3	71.3	63.8	55.0	47.5	39.7	33.5	26.5	11.0
1-Sp	11.3	16.4	21.7	28.7	36.2	45.0	52.5	60.3	66.5	73.5	89.0



Perfect ROC



Useless ROC

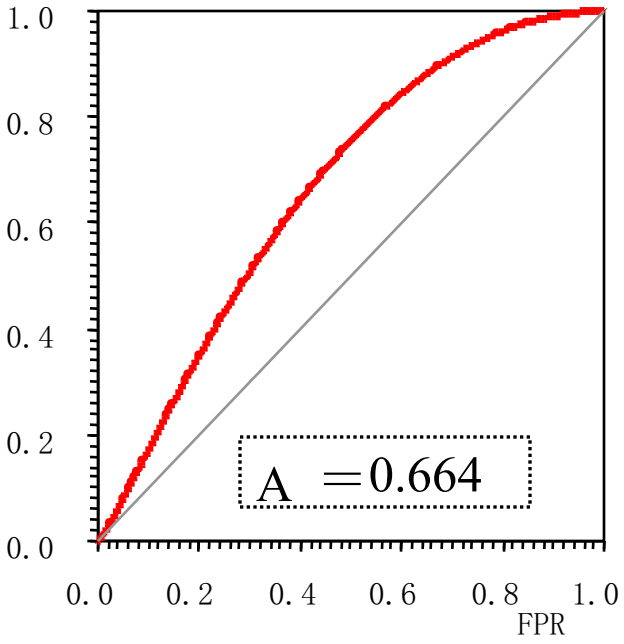
Perfect and useless ROC curve

Area of ROC curve and the accuracy of diagnosis

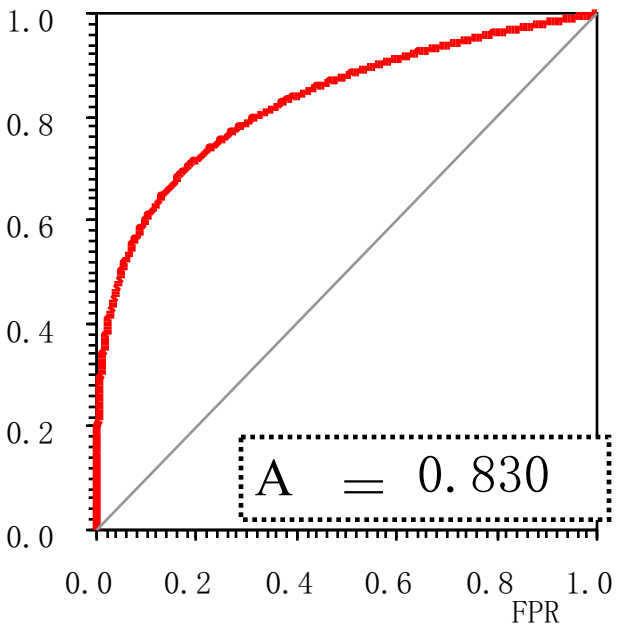
higher 0.90-1.00 = excellent (A)

moderate 0.80-0.90 = good (B) 0.70-0.80 = fair (C)

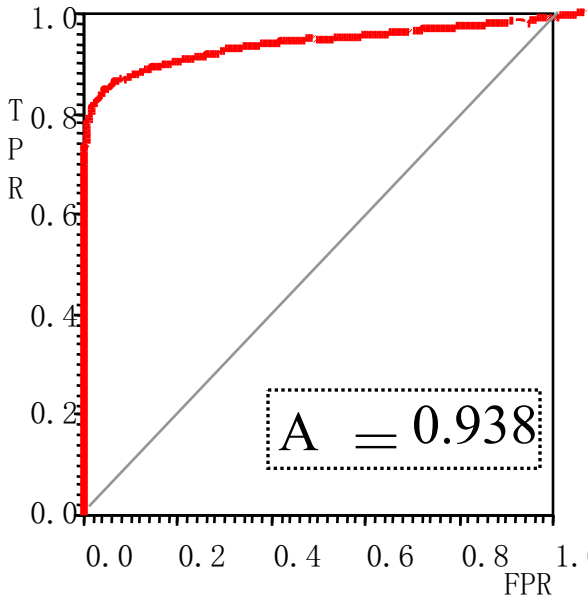
lower 0.60-0.70 = poor (D) 0.50-0.60 = fail (F)



Lower diagnosis accuracy（<0.7)



Moderate diagnosis accuracy (0.7~0.9)



Higher diagnosis accuracy (>0.9)

For example: New Diagnose Test for TB

- Sweating , appetite loss and chest pain are major symptoms. Tuberculosis can be easily transmitted via aerosol droplets.
- TB accounts for 2 million deaths a year, including the deaths of over 250,000 children.
- Tuberculosis is curable but access to those medications differ worldwide even though non-expensive treatments exist
- India is one of the most infected countries.

- It can now be diagnosed in just 30 minutes, using magnetic nanoparticles which identify *Mycobacterium tuberculosis* in sputum, even at very low concentrations.
- This involves growing larger colonies of the bacteria, which can take up to *two weeks,* delaying treatment and risking continued spread of the disease.

- **The new test, developed by Ralph Weissleder of Harvard Medical School, gives the answer in half an hour.**

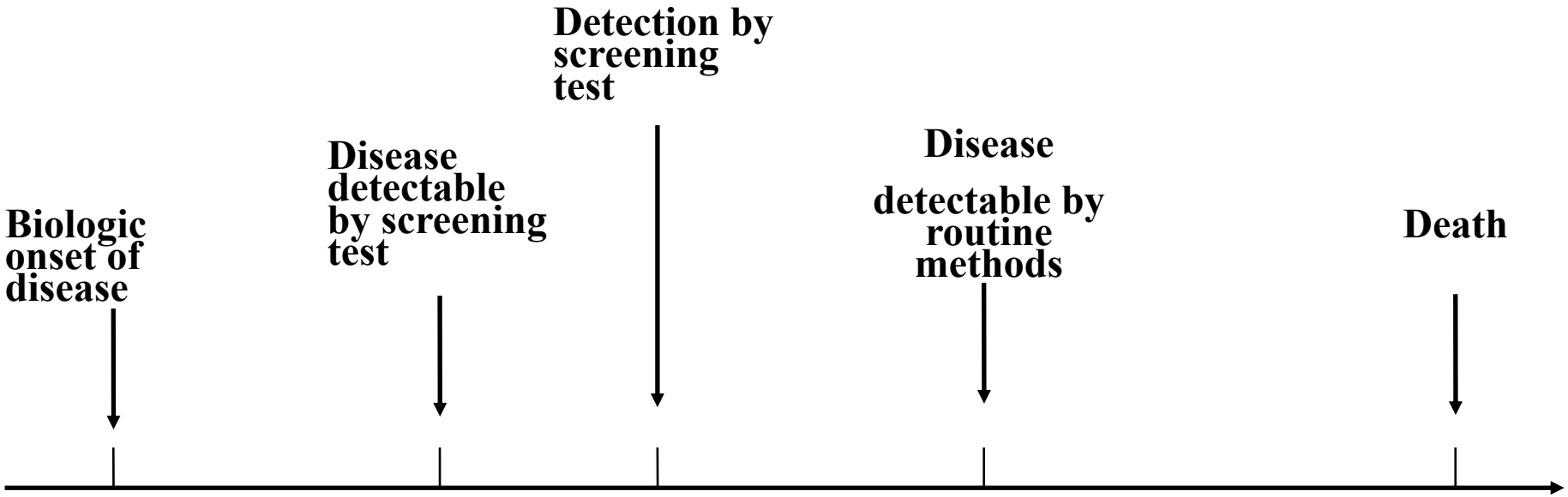
Screening Test

- **Screening** is defined as the use of **quick and simple testing procedures** to identify and separate persons who are **apparently well**, **but who may be at risk of a disease**, from those who probably do not have the disease.

- Screening is used to identify those persons suspected of having a disease so they may be sent for more definitive diagnostic studies and medical exams.

- In the Patient Profile, a **mammogram** was recommended to the patient as a **screening test** for breast cancer.

The natural history of a disease over time, including the preclinical stage in which a screening test can detect the presence of disease.

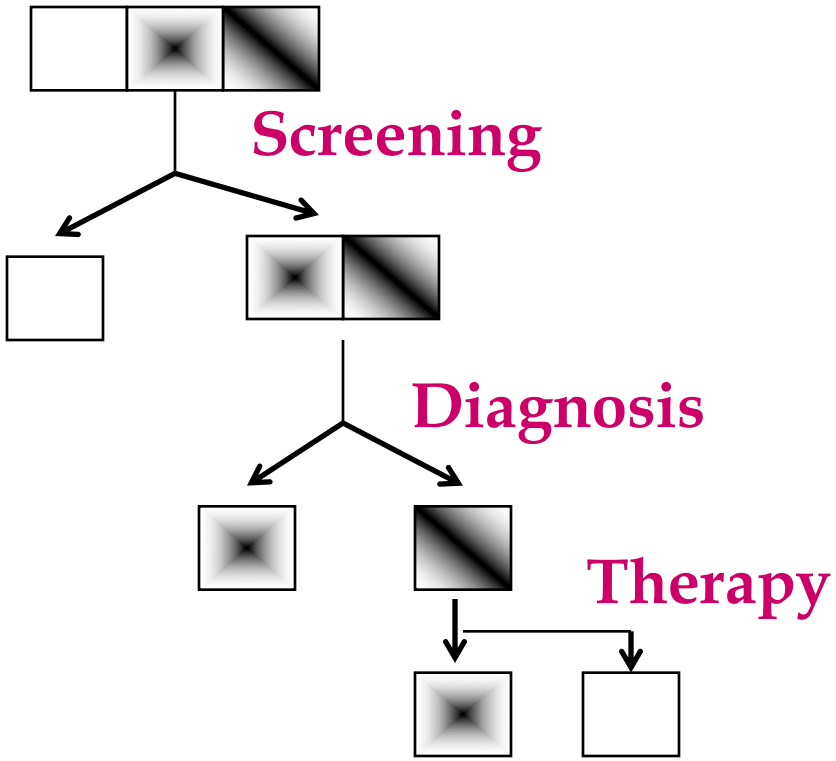


Two important concepts

- **(1) A screening test can identify individuals with a disease before the presence of disease is detected by routine diagnosis (eg, when symptoms occur).**

- **(2) Treatment at the time of detection by screening, as opposed to the time of routine diagnosis, results in an improved chance of survival.**

Screening and diagnosis

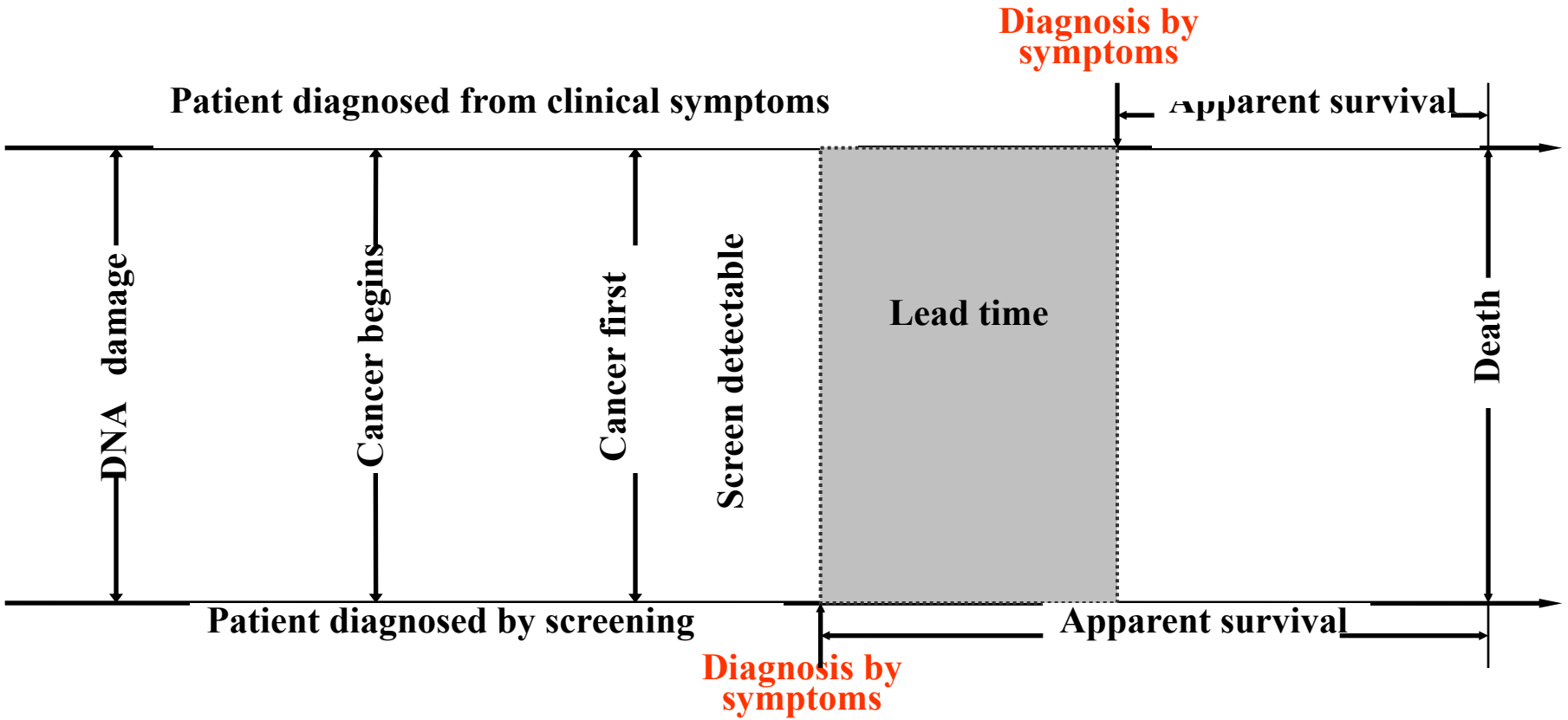


True free of disease

Probable disease

True disease

A comparison of a patient with a routine clinical diagnosis of disease and a patient with disease detected by screening.



Biases occur in the study

- **lead-time bias is an increase in survival as measured from detection of disease to death, without lengthening of life.**

- **Length-biased sampling occurs when disease detected by a screening program is less aggressive than disease detected without screening.**

Results of mammographic screening in the Health Insurance Plan (HIP) of New York study.

Mammography	Disease status		Total
	Cancer	No cancer	
Positive	132	985	1,117
Negative	47	62,295	62,342
Total	179	63,280	63,459

Prevalence =0.3%
Sensitivity =73.7%
Specificity =98.4%
PV^+ =11.8%
PV^- =99.9%

Criteria for a successful screening program.

Basis for Criteria	Criteria
Effect of morbidity and mortality on population	Morbidity or mortality of the disease must be a sufficient concern to public health.
	A high-risk population must exist.
	Effective early intervention must be known to reduce morbidity or mortality.
Screening test	The screening test should be sensitive and specific.
	The screening test must be acceptable to the target population.
	Minimal risk should be associated with the screening test.
	Diagnostic workup for a positive test result must have acceptable morbidity given the number of false-positive results.

Summary

- Diagnosis test
- Screening test
- Evaluation of the diagnosis and screening test


我图网 ooopic.com
NO.20090602221387
Thanks for your attention!