غربالگری Screening

دکترمسعود امیری

m.amiri.skums@gmail.com

گروه اپیدمیولوژی و آمارزیستی ، دانشکده بهداشت شهرکرد

موضوعات مورد بحث

- محدودیت های تشخیصی در پزشکی سنتی
- تنوع تست های تشخیصی در پزشکی نوین
- قابلیت تکرار و اعتبار
- درصد توافق – ضریب کاپا
- غربالگری Screening
- قابلیت اعتماد و صحت
- ارزیابی تست های تشخیصی
- حساسیت و ویژگی، مثبت و منفی کاذب

محدوديت‌هاي تشخيصي در پزشکی سنتی

- تشخيص‌هاي باليني و اپيدميولوژيك
- به عنوان دقيق‌ترين حربه‌هاي تشخيصي
- توجّه به رنگ، بو، طعم ادرار
- به منظور تاييد تشخيص‌هاي باليني

محدوديت‌هاي تشخيصي در پزشکی سنتی

- محمّد زكرياي رازي

موفق به افتراق باليني آبله از سرخك مي‌شود

- ابن سينا

اين شبهه را مطرح مي‌كند كه ممكن است دراكونكوليازيس، نوعي بيماري عروقي به اسم "رَگِ مدينه‌اي" = دراكونكولوس مدينسيس) (عِرقُ‌المَدينَه) نباشد! بلكه موجود زنده‌اي باشد كه در زير پوست انسان حركت مي‌كند

محدوديت‌هاي تشخيصي در پزشکی سنتی

درمان‌پذيري هاري تا قبل از ظهور Hydrophobia ؟؟

محدوديت دامنة تعداد و كارايي
ابزارهاي تشخيصي در آن زمان

تنوع آزمون‌هاي تشخيصي در پزشکی نوين

نياز به ارزيابي قابليت اعتماد يا اعتبار (Validity) و صحت يا درستي (Accuracy) آزمون‌ها

قابليت اعتماد يا اعتبار يك آزمون، بيان كننده توانايي آن در تشخيص و جدا كردن موارد بيمار از غيربيمار است

قابليت اعتماد

قابليت اعتماد داراي دو جزء حساسيت (Sensitivity) و ويژگي (Specificity) است

صحت آزمون

صحت آزمون به درجه نزديك بودن مقدار اندازه‌گيري شده به مقدار واقعي و يا *درجه انطباق نتايج با واقعيت*، اطلاق مي‌شود

Reproducibility and Validity

- Reproducibility = agreement upon repetition
- Validity = ability to discriminate those with and without disease
- A test can be reproducible but not valid
- A test can be valid but not reproducible

Reproducibility Statistics

- Overall agreement
  - Percentage of diagnoses that agree
  - Some agreement due to chance (e.g., two coins flipped simultaneously will agree 50% of time)

- Kappa statistic - quantifies percent agreement above chance

Kappa Statistic (Formula)

	Rater 2		
Rater 1	+	-	Total
+	a	b	p_1
-	c	d	q_1
Total	p_2	q_2	N

Text says to convert counts to proportions, but this is not necessary

$$\kappa = \frac{2(ad - bc)}{p_1 q_2 + p_2 q_1}$$

Kappa Statistic (Illustration)

	Rater 2		
Rater 1	+	-	Total
+	20	4	24
-	5	71	76
Total	25	75	100

$$\kappa = \frac{2\left[(20)(71)-(4)(5)\right]}{\left[(24)(75)+(25)(76)\right]} = 0.76$$

Interpretation of Kappa

- Percent agreement above chance
- The closer to 1, better agreement

Range of Kappa	Interpretation
> .75	Excellent agreement
.40 to .75	Good agreement
< .40	Poor agreement

Evaluation of Diagnostic Tests

- True Positives (= A) ,
- False Negatives (= B)
- False Positives (= C)
- True Negatives, (= D)
  - Sensitivity
  - Specificity
  - Positive Predictive Value (PPV) and
  - Negative Predictive Value (NPV)

بیماران

افراد سالم

Evaluation of Diagnostic Tests

- True Positives (= A) , بیماران تست مثبت
- False Negatives (= B) بیماران تست منفي
- False Positives (= C) افراد سالم تست مثبت
- True Negatives, (= D) افراد سالم تست منفي

جامعه تحت بررسی	آزمون‌های مثبت	آزمون‌های منفی
افراد بیمار	A = مثبت‌های واقعی	B = منفـــــی کاذب
افراد سالم	C = مثبت‌های کاذب	D = منفی‌های واقعی
	کل آزمون‌های مثبت	کل آزمون‌های منفی

Evaluation of Diagnostic Tools

- Case-Control
- Golden standard

کشت مثبت

اسمیر مثبت

آزمون مولکولي مثبت

تست اختصاصي مثبت

...

ShahidBeheshti University of Medical Sciences
School of Health / MPH Department

EVALUATION OF DIAGNOSTIC /TOOLS OF M. MENINGITIS

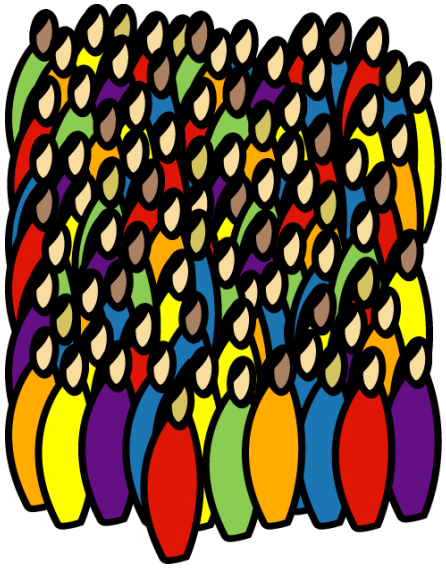
N= 100	Cases=50		Cont.= 50		Diagnostic Tests Charactristics					
WIDAL	POS. A	NEG. B	POS. C	NEG. D	SENSITIVITY	SPECIFICITY	PPV	NPV	LR+	LR-
Leukocyt	44	6	6	44	88.0%	88.0%	88.0%	88.0%	7.3%	0.14%
Glucose	44	6	5	45	88.0%	90.0%	89.8%	88.2%	8.8%	0.13%
Gram st.	42	8	0	50	84.0%	100.0%	100.0%	86.2%	0.0%	0.16%
LatexAg.	30	20	0	50	60.0%	100.0%	100.0%	71.4%	0.0%	0.40%
PCR	45	5	0	50	90.0%	100.0%	100.0%	90.9%	0.0%	0.10%
tool6	44	6	38	12	88.0%	24.0%	53.7%	66.7%	1.2%	0.50%
Culture	45	5	0	50	90.0%	100.0%	100.0%	90.9%	0.0%	0.10%
fever	50	0	50	0	100.0%	0.0%	50.0%	0.0%	1.0%	0.00%
Headache	50	0	50	0	100.0%	0.0%	50.0%	0.0%	1.0%	0.00%
Redor	50	0	30	20	100.0%	40.0%	62.5%	100.0%	1.7%	0.00%
Kernig	45	5	25	25	90.0%	50.0%	64.3%	83.3%	1.8%	0.20%
Brodzins	35	15	20	30	70.0%	60.0%	63.6%	66.7%	1.8%	0.50%

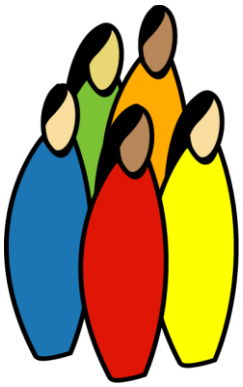
Cont. = control, Pos. = positive, Neg. = Negative, Sensit. = sensitivity, Specif. = specificity, PPV & NPV = positive & negative predictive value, LR+ & LR- = positive & negative likelihood ratio

Editd by: H. Hatami M.D.

SENSITIVITY:

SENSITIVITY =*A / (A+B)*100=* 95%





بيماران تست منفي = ۵

(منفي‌هاي كاذب)

B = 5

افراد بيمار = ۱۰۰ نفر

A=بيماران تست مثبت = ۹۵

Cases (A+B)=100

حساسيت، عبارت است از: ميزان مثبت بودن آزمون در افراد واقعا بيمار

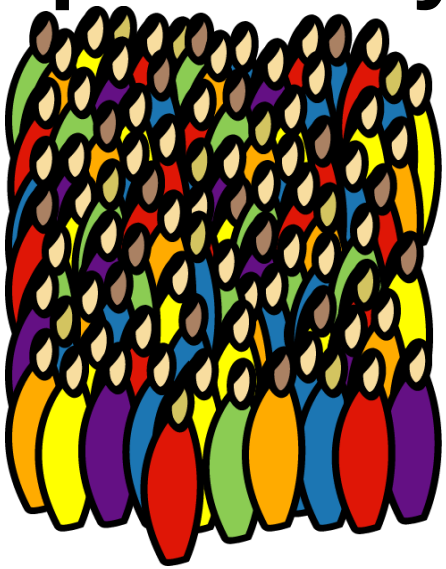
حساسیت، عبارت است از: میزان مثبت بودن آزمون در افراد واقعا بیمار

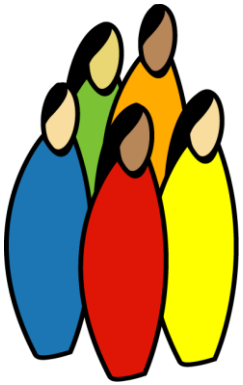
آزمون‌های منفی	آزمون‌های مثبت	جامعه تحت بررسی
B = منفـــــی کاذب	A = مثبت‌های واقعی	**افراد بیمار**
D = منفی‌های واقعی	C = مثبت‌های کاذب	**افراد سالم**
کل آزمون‌های منفی	کل آزمون‌های مثبت	

Specificity = D / (D + C) = ویـــــژگی	Sensitivity = A / (A + B) = حساسیـــت
NPV = D / (B + D) = ارزش اخباری منفی	PPV = A / (A + C) = ارزش اخباری مثبت
LR- درستنمایی منفی = B / (A + B) / D / (C + D)	LR+ درستنمایی مثبت = A / (A + B) / C / (C + D)

SPECIFICITY:

Specificity = D/(D+C)*100= 95%





افراد سالم تست مثبت = ۵

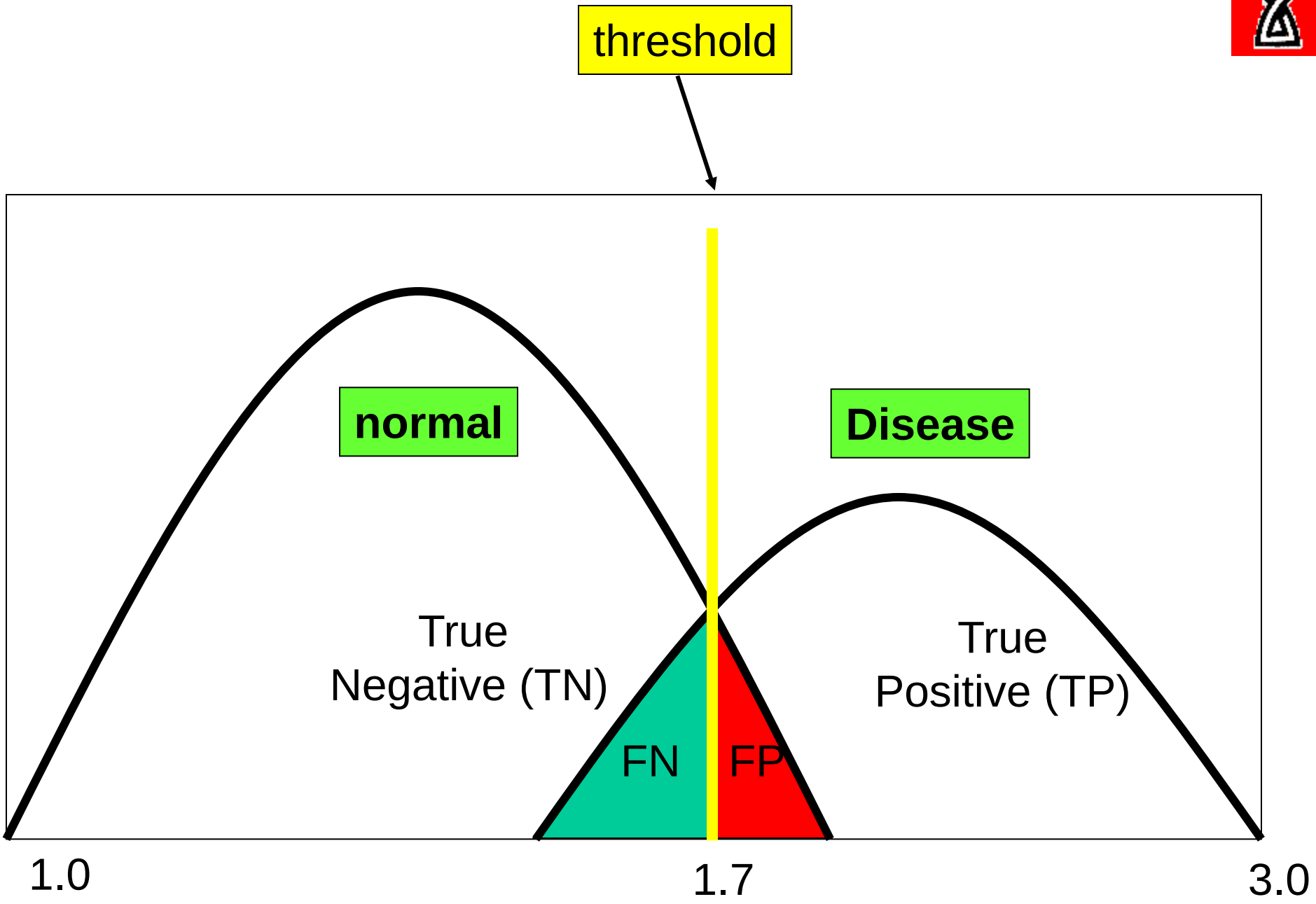
(False positive)

افراد سالم = ۱۰۰ نفر

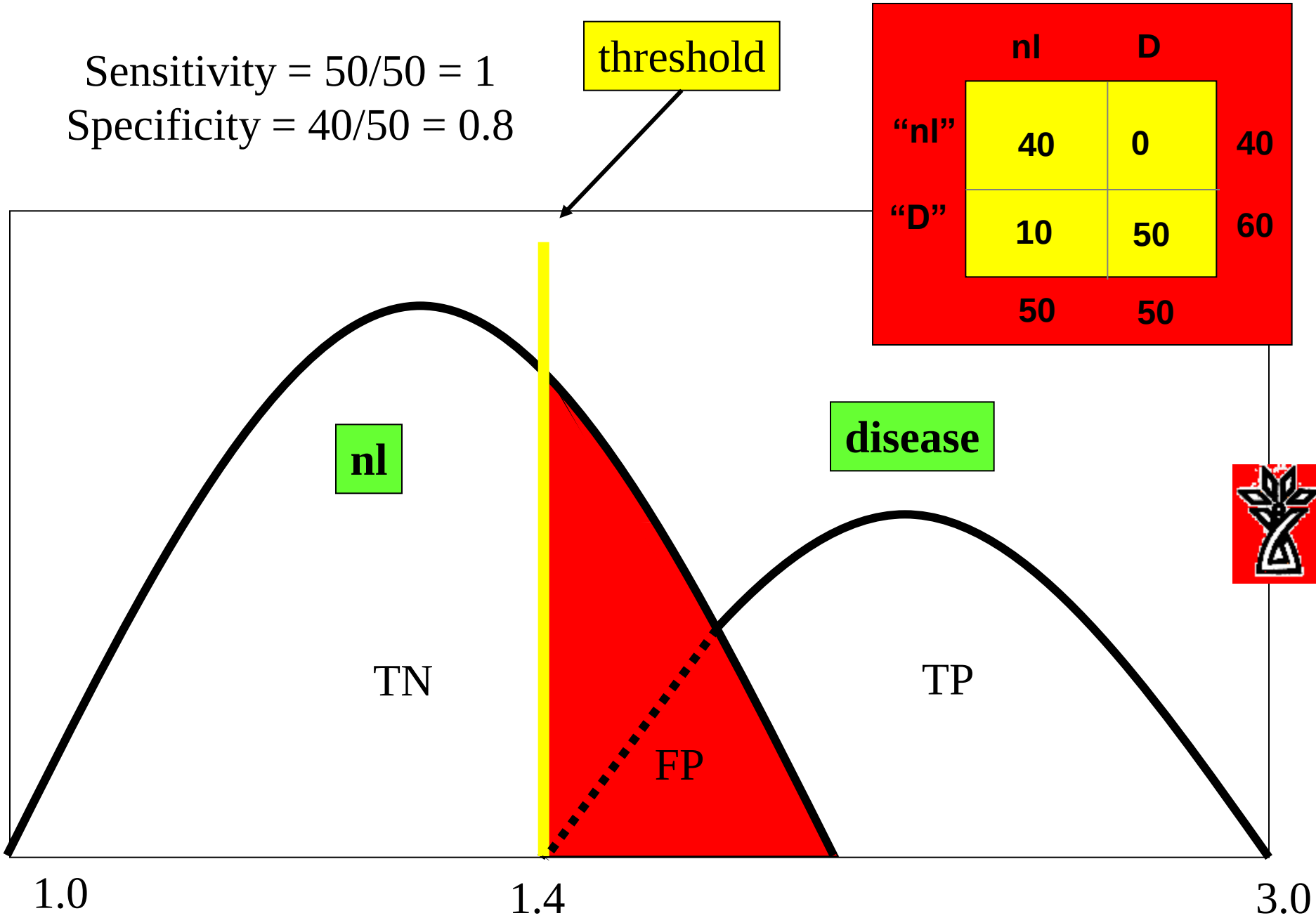
Controls (C+D)=100

C = 5

ويژگي، عبارت است از: ميزان منفي بودن آزمون در افراد غير بيمار

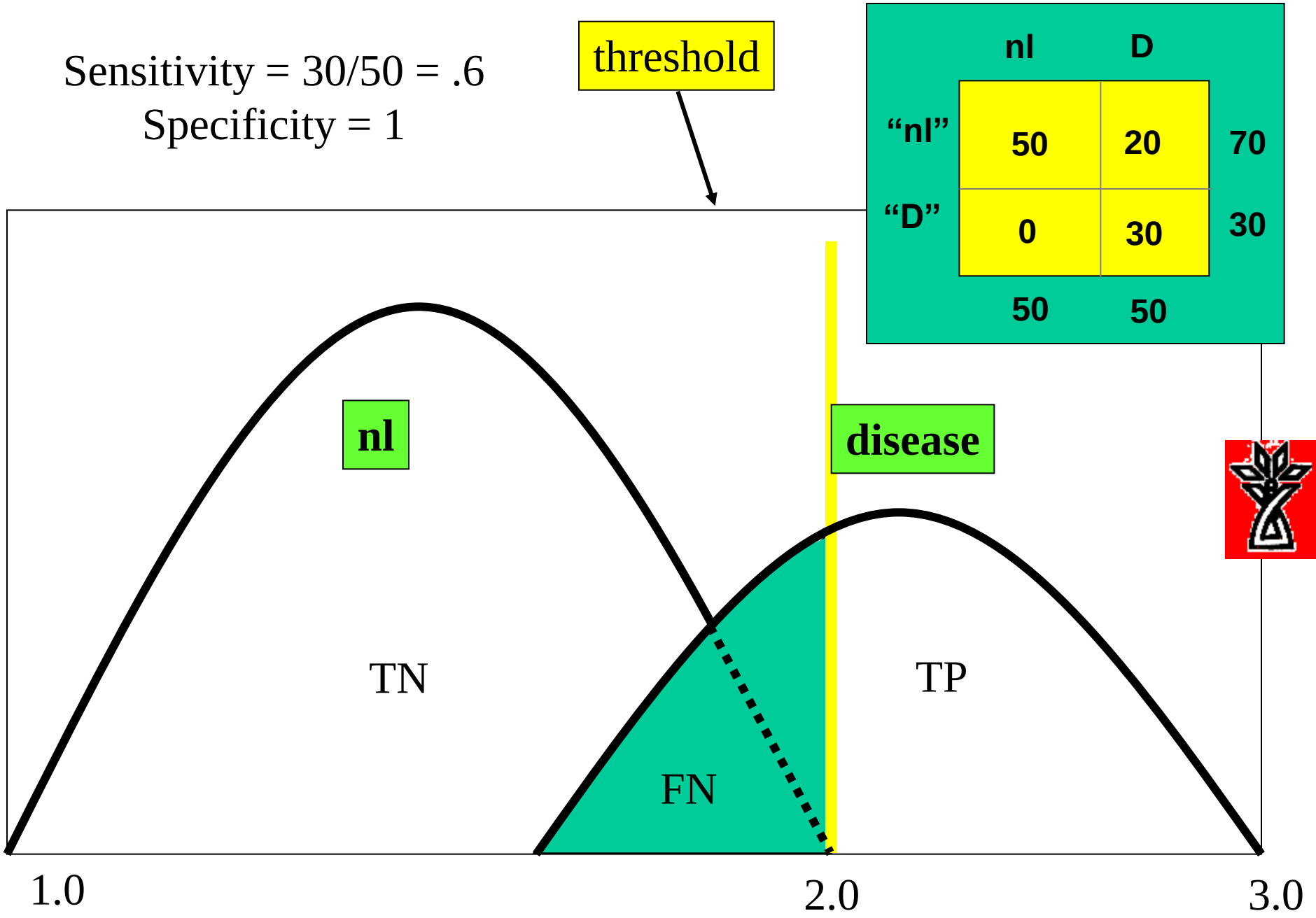

threshold
normal
Disease
True
Negative (TN)
True
Positive (TP)
FN
FP
1.0
1.7
3.0

Sensitivity = 50/50 = 1
Specificity = 40/50 = 0.8



حساسیت، عبارت است از: میزان مثبت بودن آزمون در افراد واقعا بیمار

Sensitivity = 30/50 = .6
Specificity = 1



ويژگي، عبارت است از: ميزان منفي بودن آزمون در افراد غير بيمار

ShahidBeheshti University of Medical Sciences

School of Health / MPH Department

EVALUATION OF DIAGNOSTIC TESTS/TOOLS (Widal Test)

N= 288	Cases=144		Cont.= 144		Diagnostic Tests Charactristics					
WIDAL	POS. A	NEG. B	POS. C	NEG. D	SENSITIVITY	SPECIFICITY	PPV	NPV	LR+	LR-
1:40	3	141	3	141	2.1%	97.9%	50.0%	50.0%	1.0%	1.00%
1:80	16	128	1	143	11.1%	99.3%	94.1%	52.8%	16.0%	0.90%
1:160	21	123	4	140	14.6%	97.2%	84.0%	53.2%	5.3%	0.88%
1:320	15	129	1	143	10.4%	99.3%	93.8%	52.6%	15.0%	0.90%
1:640	3	141	0	144	2.1%	100.0%	100.0%	50.5%	0.0%	0.98%

Cont. = control, Pos. = positive, Neg. = Negative, Sensit. = sensitivity, Specif. = specificity, PPV & NPV = positive & negative predictive value, LR+ & LR- = positive & negative likelihood ratio

Editd by: H. Hatami M.D

مفاهيم عملي حساسيت و ويژگي

حساسيت پايين آزمايش يعني تعدادي از افراد بيمار ناديده گرفته شده و سالم در نظر گرفته مي‌شوند

ويژگي پايين يعني تعداد زيادي از افراد سالم جزو بيماران طبقه بندي مي‌شوند

TABLE 3.1 Trade-Off Between Sensitivity and Specificity When Using PSA Levels to Detect Prostate Cancer in Black Men Ages 70–79

PSA Level (ng/ml)	Sensitivity	Specificity
1.0	100	21
2.0	100	48
3.0	100	60
4.0	99	73
5.0	96	76
6.0	94	79
7.0	90	83
8.0	90	88
9.0	68	90
10.0	54	93
11.0	47	94
12.0	30	95
13.0	23	96
14.0	17	97
15.0	11	97

ارزش اخباري مثبت = PPV

$$\frac{\text{مثبت‌هاي واقعي}}{\text{كل آزمون‌هاي مثبت}}$$

جامعه تحت بررسی	آزمون‌های مثبت	آزمون‌های منفی	جمع
افراد بیمار	A = مثبت‌های واقعی	B = منفـــــی کاذب	کل بیماران
افراد سالم	C = مثبت‌های کاذب	D = منفی‌های واقعی	کل افراد سالم
	کل آزمون‌های مثبت	کل آزمون‌های منفی	

Sensitivity = A / (A + B) = حساسیـــــت	**Specificity = D / (D + C) = ویـــــــژگی**
PPV = A / (A + C) = ارزش اخباری مثبت	**NPV = D / (B + D) = ارزش اخباری منفی**
LR+ درستنمایی مثبت $= \frac{A / (A + B)}{C / (C + D)}$	**LR- درستنمایی منفی** $= \frac{B / (A + B)}{D / (C + D)}$

ارزش اخباري منفي = NPV

$$\frac{\text{منفي‌هاي واقعي}}{\text{كل آزمون‌هاي منفي}}$$

جامعه تحت بررسی	آزمون‌های مثبت	آزمون‌های منفی	جمع
افراد بیمار	A = مثبت‌های واقعی	B = منفـــــی کاذب	کل بیماران
افراد سالم	C = مثبت‌های کاذب	D = منفی‌های واقعی	کل افراد سالم
	کل آزمون‌های مثبت	کل آزمون‌های منفی	

Sensitivity = A / (A + B) = حساسیـــــت	**Specificity = D / (D + C) = ویـــــــژگی**
PPV = A / (A + C) = ارزش اخباری مثبت	**NPV = D / (B + D) = ارزش اخباری منفی**
LR+ درستنمایی مثبت $= \dfrac{A / (A + B)}{C / (C + D)}$	**LR- درستنمایی منفی** $= \dfrac{B / (A + B)}{D / (C + D)}$

ارزش اخباري
Predictive Value

- *ارزش اخباري، بازتاب توان تشخيصي آزمون است*
- دقت ارزش اخباري به عواملي نظير حساسيت، ويژگي و شيوع بيماري بستگي دارد
- شامل ارزش اخباري مثبت و منفي مي‌باشد.

Predictive Value

- **The probability of the presence or absence of disease given the test result**
  - **PPV is the probability of disease in a patient with a positive test result**
  - **NPV is the probability of not having disease when the test result is negative.**
- **Determined by**
  - **sensitivity and specificity of the test**
  - **prevalence of disease in the population being tested**

Predictive Value of a test

PPV = (True positive) / (True + False positives)

Example :

Low Prevalence Disease of 1 per 100

Test efficacy: Sensitivity: 100%, Specificity: 99% (1 false positive in 100)

Test 1000 patients

- **True positives: 1 per 100 (1% Prevalence)**
- **False positives: 1 per 100 (99% Test Specificity)**
- **PPV: 1 true positive / 2 total positives = 50%**

Test 1000 People for Influenza

Test Result	True Influenza	No Influenza	
P	A=80	C=100	180 کل مثبت‌ها
N	B=20	D=800	820 کل منفي‌ها
Total	100	900	

PPV=80/180= 44% NPV=800/820= 98%

مثبت‌هاي واقعي
کل آزمون‌هاي مثبت

منفي‌هاي واقعي
کل آزمون‌هاي منفي

Hypothetical Influenza Test Performance

- **Prevalence = 1 in 5**
  - **Sensitivity = 95%**
  - **Specificity = 96%**
  - **PPV = 85.6%**
  - **NPV = 96%**
- **Prevalence = 1 in 100**
  - **Sensitivity = 95%**
  - **Specificity = 96%**
  - **PPV = 19.2%**
  - **NPV = 99.9%**

Which test is best? Know the test performance characteristics and the epidemiology of the disease

$$\text{Likelihood Ratio (positive)} = \frac{A/(A+B)}{C/(C+D)} = \frac{A*(C+D)}{C*(A+B)} \times 100$$

(Dorost namaeie mosbat)

$$\text{Likelihood Ratio (negative)} = \frac{B/(B+A)}{D/(D+C)} = \frac{B*(C+D)}{D*(A+B)} \times 100$$

(Dorost namaeie manfi)

Performance Characteristics

Test	Sensitivity	Specificity
Directigen Flu A	91% (A)	95%(A)
Directigen Flu A & B	86%(A) 81%(B)	91%(A) 99.5%(B)
Flu OIA	62-88% (A/B)	52%-80%(A/B)
Flu OIA A/B	62%-88%(A/B)	52%-80%(A/B)
NOW Flu A & B	78%-82%(A) 58-71%(B)	92-94%(A) 97%(B)

Performance Characteristics

Test	Sensitivity	Specificity
QuickVue Influenza	73-81% (A/B)	95%(A)
QuickVue Influenza A&B	72-77%(A) 73-82% (B)	91%(A) 99.5%(B)
SAS Influenza A & Influenza B	76%(A) 91%(B)	98%(A) 100%(B)
Xpect Flu A & B	92%(A) 98%(B)	100%(A) 100%(B)
ZstatFlu	58-65% (A/B)	98-100%(A/B)

نسبت احتمال یا درستنمایي مثبت

Likelihood Ratio+

$$LR+ = \frac{\text{نسبت پاسخ‌هاي مثبت واقعي}}{\text{نسبت مثبت‌هاي كاذب}}$$

جامعه تحت بررسی	آزمون‌های مثبت	آزمون‌های منفی	جمع
افراد بیمار	A = مثبت‌های واقعی	B = منفـــــی کاذب	کل بیماران
افراد سالم	C = مثبت‌های کاذب	D = منفی‌های واقعی	کل افراد سالم
	کل آزمون‌های مثبت	کل آزمون‌های منفی	

Sensitivity = A / (A + B) = حساسیـــــت	Specificity = D / (D + C) = ویـــــــژگی
PPV = A / (A + C) = ارزش اخباری مثبت	NPV = D / (B + D) = ارزش اخباری منفی
LR+ درستنمایی مثبت = A / (A + B) ---------- / C / (C + D)	LR- درستنمایی منفی = B / (A + B) ---------- / D / (C + D)

نسبت احتمال يا درستنمايي منفي

Likelihood Ratio-

$$LR- = \frac{\text{نسبت پاسخ‌هاي منفي كاذب}}{\text{نسبت منفي‌هاي واقعي}}$$

جامعه تحت بررسی	آزمون‌های مثبت	آزمون‌های منفی	جمع
افراد بیمار	A = مثبت‌های واقعی	B = منفـــــی کاذب	کل بیماران
افراد سالم	C = مثبت‌های کاذب	D = منفی‌های واقعی	کل افراد سالم
	کل آزمون‌های مثبت	کل آزمون‌های منفی	

Sensitivity = A / (A + B) = حساسیـــــت	Specificity = D / (D + C) = ویـــــــژگی
PPV = A / (A + C) = ارزش اخباری مثبت	NPV = D / (B + D) = ارزش اخباری منفی
LR+ درستنمایی مثبت = A / (A + B) ---------- / C / (C + D)	LR- درستنمایی منفی = B / (A + B) ---------- / D / (C + D)

نسبت احتمال يا درستنمايي مثبت Likelihood Ratio+

- نسبت مثبت‌هاي واقعي در بيماران، تقسيم بر نسبت مثبت‌هاي كاذب در افراد غيربيمار
- برخلاف ارزش اخباري، تحت تاثير شيوع بيماري مربوطه قرار نمي‌گيرد
- هرچه از عدد يك بيشتر باشد دليل بر مناسب بودن آن آزمايش است
- صورت اين كسر را حساسيت آزمون، تشكيل مي‌هد.

نسبت احتمال يا درستنمايي منفي Likelihood Ratio-

- نسبت منفي‌هاي كاذب در بيماران تقسيم بر نسبت منفي‌هاي واقعي در افراد غيربيمار
- هرچه به صفر، نزديك‌تر باشد دليل بر مناسب بودن آزمايش است
- مخرج اين كسر را ويژگي آزمون، تشكيل مي‌دهد. در مجموع، هرچه درستنمايي مثبت يك آزمايش، بيشتر و درستنمايي منفي آن كمتر باشد دليل بر مناسب بودن آن آزمايش است.

نسبت احتمال يا درستنمايي
Likelihood Ratio

در مجموع، هرچه درستنمايي مثبت يك آزمايش، بيشتر و درستنمايي منفي آن كمتر باشد دليل بر مناسب بودن آن آزمايش است

Screening غربالگری یا

Definition

- Examination of asymptomatic people in order to classify them as likely or unlikely to have the disease
- Organized application of early diagnosis and treatment activities in large groups

Goal

Reduce morbidity or mortality among the people screened

Screening غربالگری یا

Definition

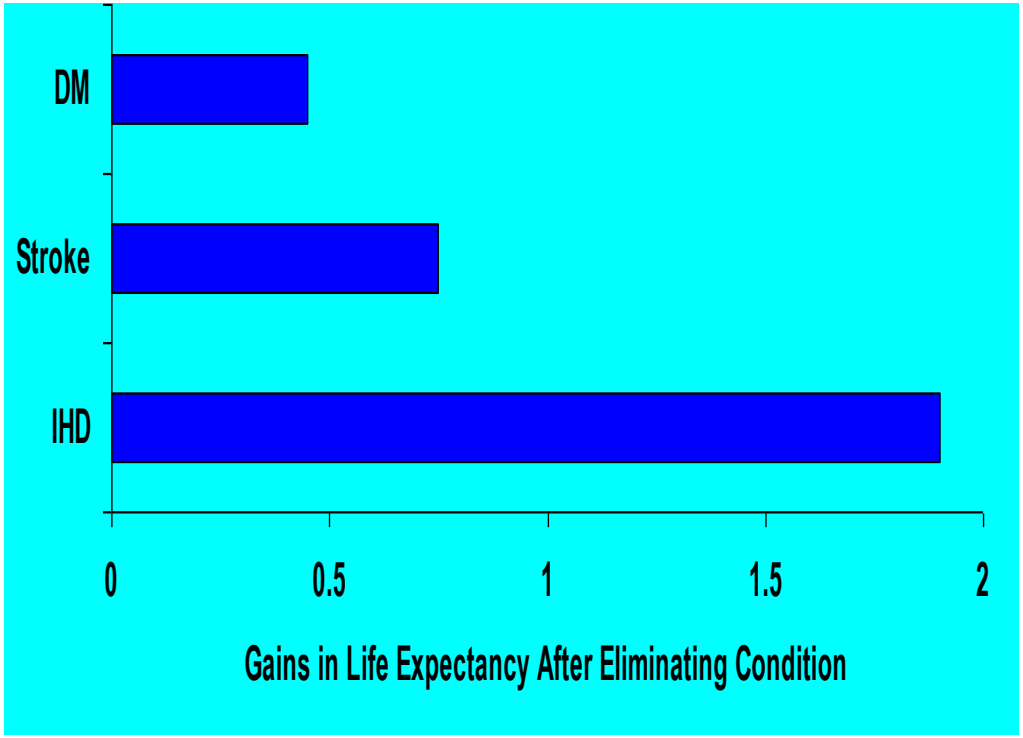
- Examination of asymptomatic people in order to classify them as likely or unlikely to have the disease
- Organized application of early diagnosis and treatment activities in large groups

Goal

Reduce morbidity or mortality among the people screened

Shifting the Paradigm to Screening

- **Early detection leads to:**
  - improved life expectancy.
  - less costly
  - less invasive care
  - less hospitalizations
  - improved social productivity



Pasternak RC, Abrams J, Greenland P, Smaha LA, Wilson PW, Houston-Miller N. 34th Bethesda Conference: Identification of coronary heart disease risk: is there a detection gap? J Am Coll Cardiol 2003 Jun 4; 41(11): 1863-74.

Screening غربالگری یا (ادامه)

Success of a screening program depends on

- **Disease**
  Must be detectable before symptoms.
  Must have diagnostic test and be treatable.

- **Screening**
  Must be convenient and free of discomfort or risk, attractive, efficient, cost-effective.
  High level of case detection and low level of false-positive

- **Treatment**
  Must be available and effective.

Disease must pass through a preclinical phase (undiagnosed but detectable) and
early treatment must offer some advantage over later treatment.

Cole and Morison, 1980

Screening غربالگری یا (ادامه)

Utility of screening depends on:

1. Prevalence of detectable preclinical disease
2. Average length of preclinical phase *(chronic disease)*
3. History of screening program *(Rescreening, Acute disease)*

Screening غربالگری یا (ادامه)

- Since known cases are not eligible to be screened, death in screened population is lower than general population.

- This death rate difference is especially large during the early years of a screening program.

- This difference should be take into account deriving the expected number of death to be compared with that observed in a screened group.

Expected number of death in the absence of screening depends on:

1. Incidence of the disease
2. Distribution of duration
3. Case fatality

Morrison et al, 1979,1988, 1991

Screening غربالگری یا (ادامه)

The results of screening reflects :

(X-ray, Mammography, Cytology)

1. Test (False +/-)
2. Procedure for Interpreting the findings

Measures of test performance

- **Sensitivity** (Percent with the disease who test positive)
- **Specificity** (Percent healthy who test negative)

Problems

- Impossibility of applying a definitive diagnostic test
- Variability with the chosen case series

Sensitivity and specificity

Ability of a test to identify correctly diseased and nondiseased

Reliability

Capacity to give the same results - positive or negative, whether correct or incorrect – on repeated application

- Reliability dose not guarantee high sensitivity and specificity
- Preliminary assessment of a screening teat may include studies of its reliability
- Highly sensitive test must be highly reliable to a diseased person
- Highly specific test must be highly reliable to a non-diseased person

Common Screening Tests

Table 1. Effectiveness of the ABI vs Other Common Screening Tests

Diagnostic Test	Sensitivity, %	Specificity, %
Pap smear[37]	30-87	86-100
Fecal occult blood test[38]	37-78	87-98
Mammography[39]	75-90	90-95
ABI[5,35,36]	95	100

Abbreviation: ABI, ankle-brachial index.

Belch JJF, Topol EJ, Agnelli G, et al. Critical issues in peripheral arterial disease detection and management. Arch Int Med 2003;163:884-892.

Validity

- Compare test results to definitive diagnostic procedure (“gold standard”)

- Each case classified as
  - TP = true positive
  - TN = true negative
  - FP = false positive
  - FN = false negative

Organize Data in Table

	Gold Standard		
	+	-	Total
Test Result			
+	TP	FP	TP+FP
-	FN	TN	FN+TN
Total	TP+FN	FP+TN	*n*

Definitions

- Notation: Pr (T|D) = probability test result *given* disease status
- Sensitivity = Pr (T+|D+)
- Specificity = Pr (T-|D-)
- Positive Predictive Value= Pr (D+|T+)
- Negative Predictive Value= Pr (D-|T-)
- Prevalence = Pr (D+)

Formulas

- SENSITIVITY = TP / (TP + FN)
- SPECIFICITY = TN / (TN + FP)
- PPV = TP / (TP + FP)
- NPV = TN / (TN + FN)

Example (HIV Screening Test)
Prevalence = 1000 / 1,000,000 = .001

Screening Test	HIV +	HIV -	Total
+	990	9,990	10,980
-	10	989,010	989,020
Total	1000	999,000	1,000,000

- SEN = 990 / 1000 = .99
- SPEC = 989,010 / 999,000 = .99
- PPV = 990 / 10,980 = .09
- NPV = 989,010 / 989,020 ≅ 1.000

Example (HIV Screening Test)
Prevalence = 10,000 / 1,000,000 = .01

Screening Test	HIV +	HIV -	Total
+	99,000	9,000	108,000
-	1,000	891,000	892,000
Total	100,000	900,000	1,000,000

- SEN = 99,000 / 100,000 = .99
- SPEC = 891,000 / 900,000 = .99
- PPV = 99,000 / 108,000 = .92
- NPV = 891,000 / 900,000 = .99

Conclusion

- A test that is 99% sensitive and 99% specific is used in two populations
- In low prevalence population PPV = .09
- In high prevalence population, PPV = .92
- Therefore, PPV is a function of prevalence

Relation Between Predictive Value and Prevalence

- PPV is a function of
  - Test SEN
  - Test SPEC
  - prior probability (prevalence) of disease

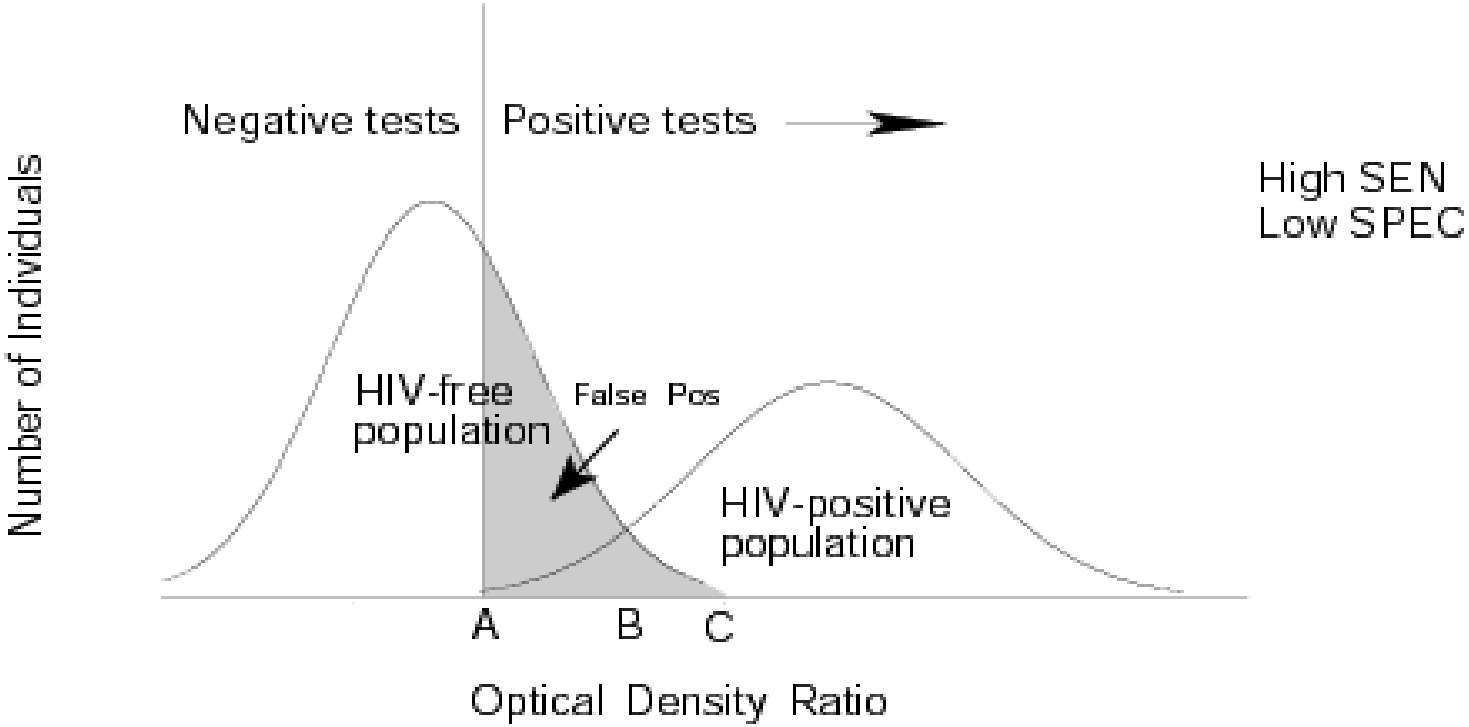
Selecting a Cutoff for a Test

- HIV screening test detects color change (Optical Density Ratio)
- Can change cutoff for amount of color change to classify as “positive”

Low Cutoff

- no false negatives (high SEN)
- many false positive (low SPEC)

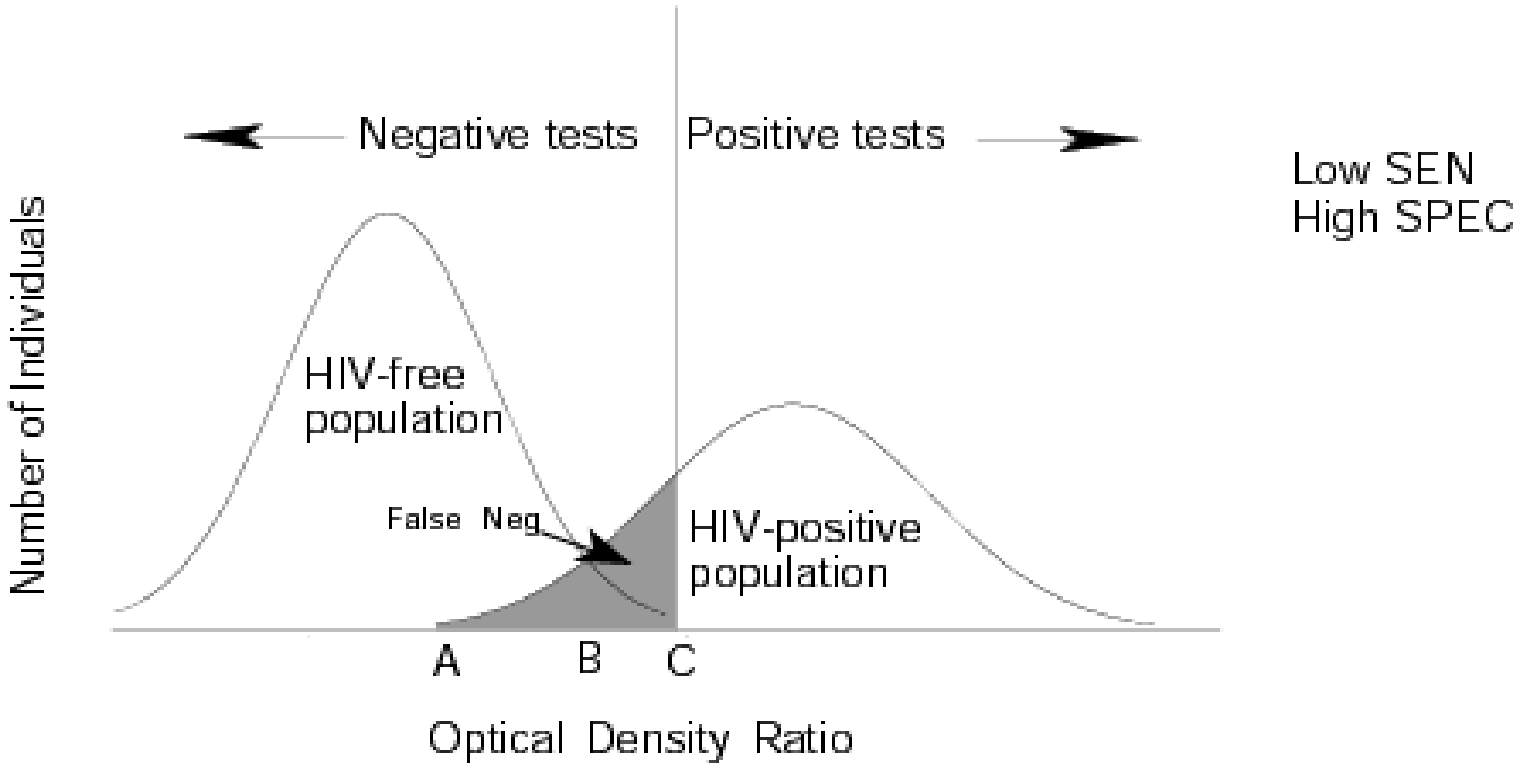
A.



High Cutoff

- no false positive (high SPEC)
- many false negative (low SEN)

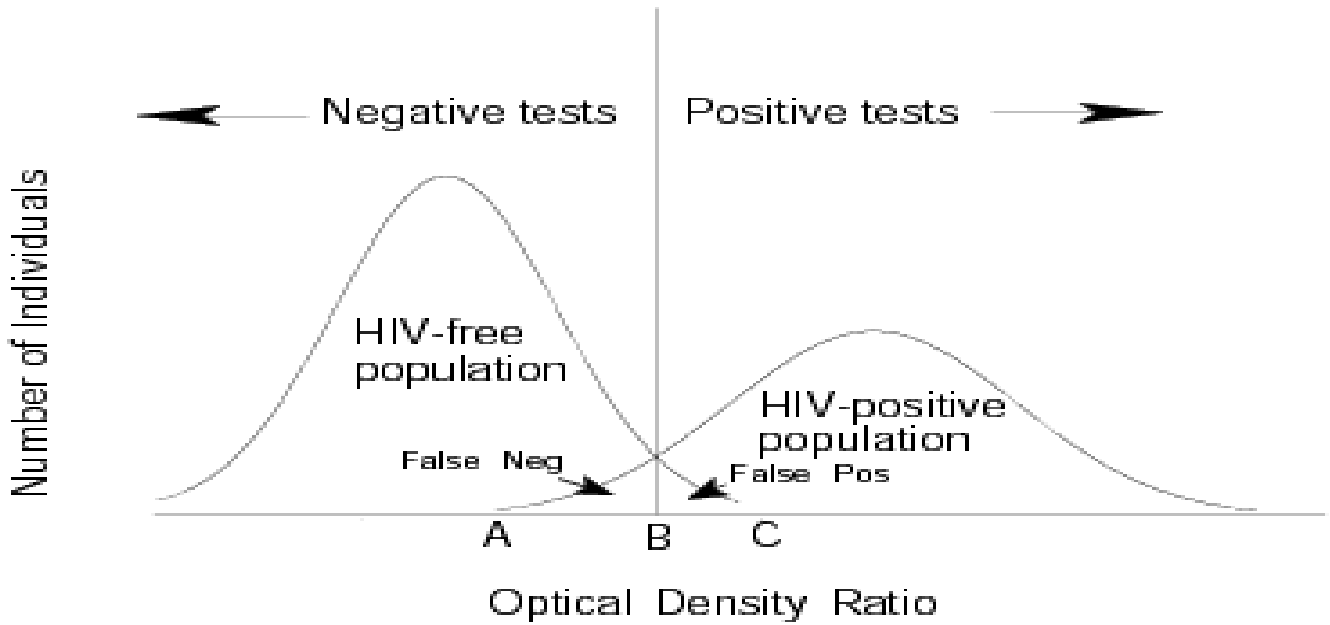
C.



In-Between Cutoff

- medium no. of false positive (intermediate SPEC)
- medium no. of false negative (intermediate SEN)

B.



Int SEN
Int SPEC