

Critical Appraisal

Dr. M. Ebrahimi-Mameghani
Dep. Community Nutrition

Email:

ebrahimimamagani@tbzmed.ac.ir

Critical appraisal

- This process provides clinicians with the means to interpret and determine the applicability of results to their particular patients
- Clinicians should consider multiple explanations for any effect reported in a study which is obtained from:
 - *Quantitative studies*
 - *Qualitative studies*

Critical Appraisal

- The assessment of evidence by systematically reviewing its relevance, validity and results to specific situations.

What is 'best' evidence?

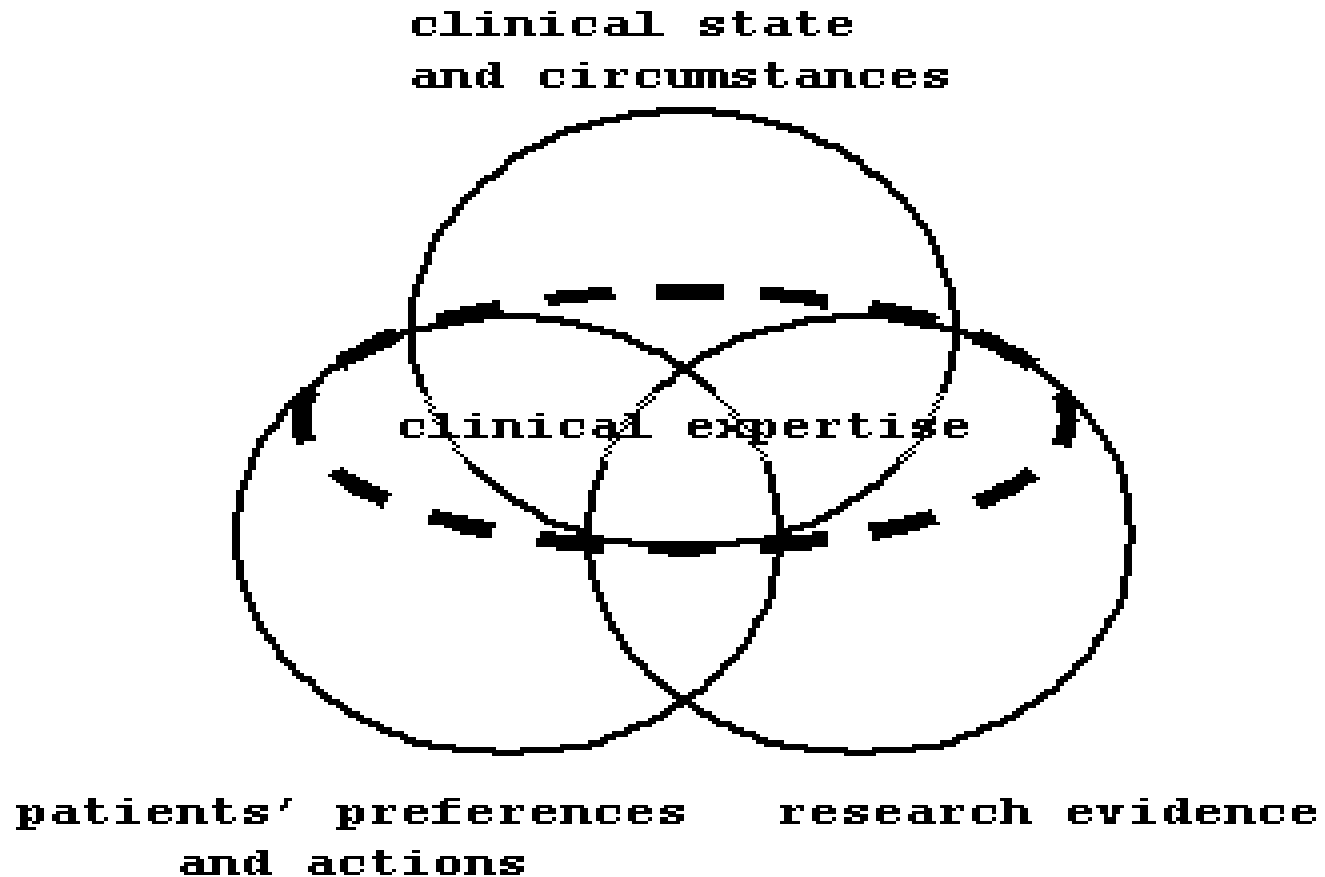
- Using critical appraisal skills you can understand the methods and results of the research and then be able to assess the quality of the research.

Information Mastery



Slawson and Shaughnessy Formula:

$$\frac{\text{Usefulness of Medical Information} = \text{Relevance} \times \text{Validity}}{\text{Work to Access}}$$



Basic elements of clinical decision making

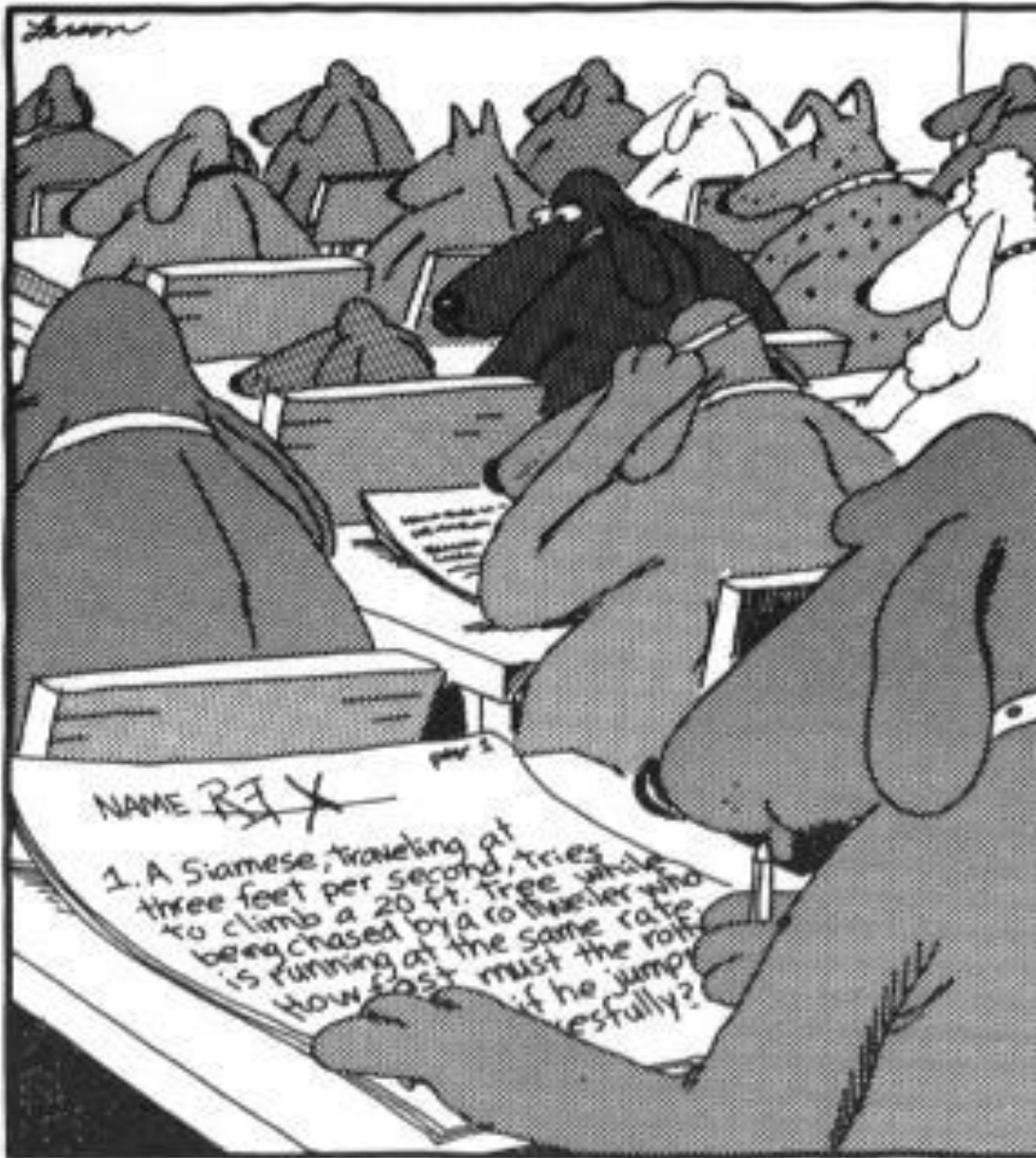
Haynes RB. Loose connections between peer-reviewed clinical journals and clinical practice. Ann Intern Med 1990;113:724-8.

Why Critical Appraise?

Published research is not always
reliable

Published research is not always
relevant

To improve clinical effectiveness, we
need a systematic framework to
interpret research



Before their admission to any canine university, dogs must first do well on the CATs.

CATs:

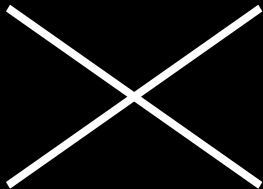
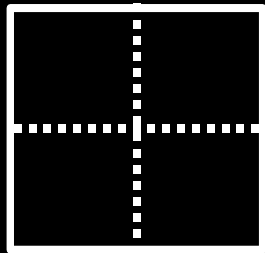
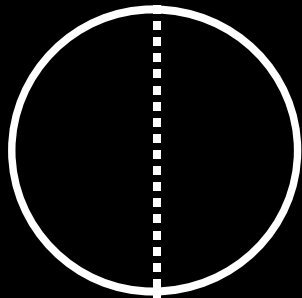
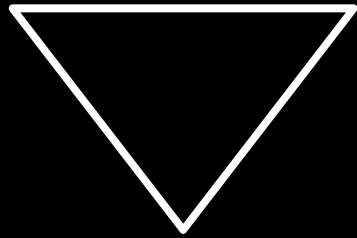


Critically Appraised Topics:

'a tool for modeling the steps of EBP'

www.epiq.co.nz

The GATE Approach: every epidemiological study hangs on the GATE frame

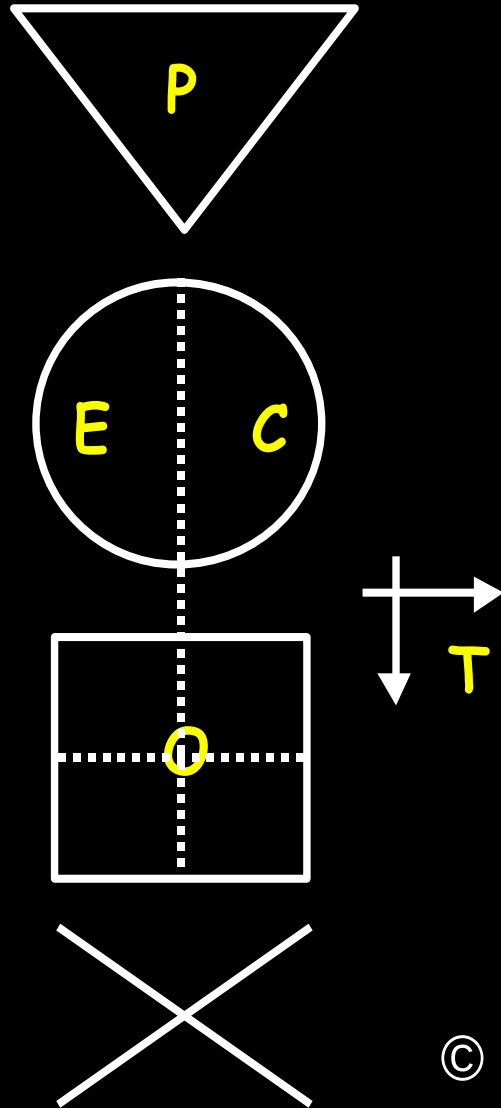


there is only one study design:

- RCT - interventions
- Cohort studies - prognosis / interv./ aetiology
- Cross-sectional studies - diagnosis
- Case-control studies - interv./aetiol.

GATE: Graphic Appraisal Tool for Epidemiology

GATE Frame: PECOT



- **P**opulation
- **E**xposure
- **C**omparison
- **O**utcome
- **T**ime

©

GATE: Graphic Appraisal Tool for Epidemiology

Difficulties with Critical Appraisal

- Can be time consuming initially
- Doesn't provide an “easy” answer
- It could show a lack of good evidence in a particular topic

Key Steps to Effective Critical Appraisal



1. Are the results valid?
2. What are the results?
3. How will these results be relevant to the patient?

Validity

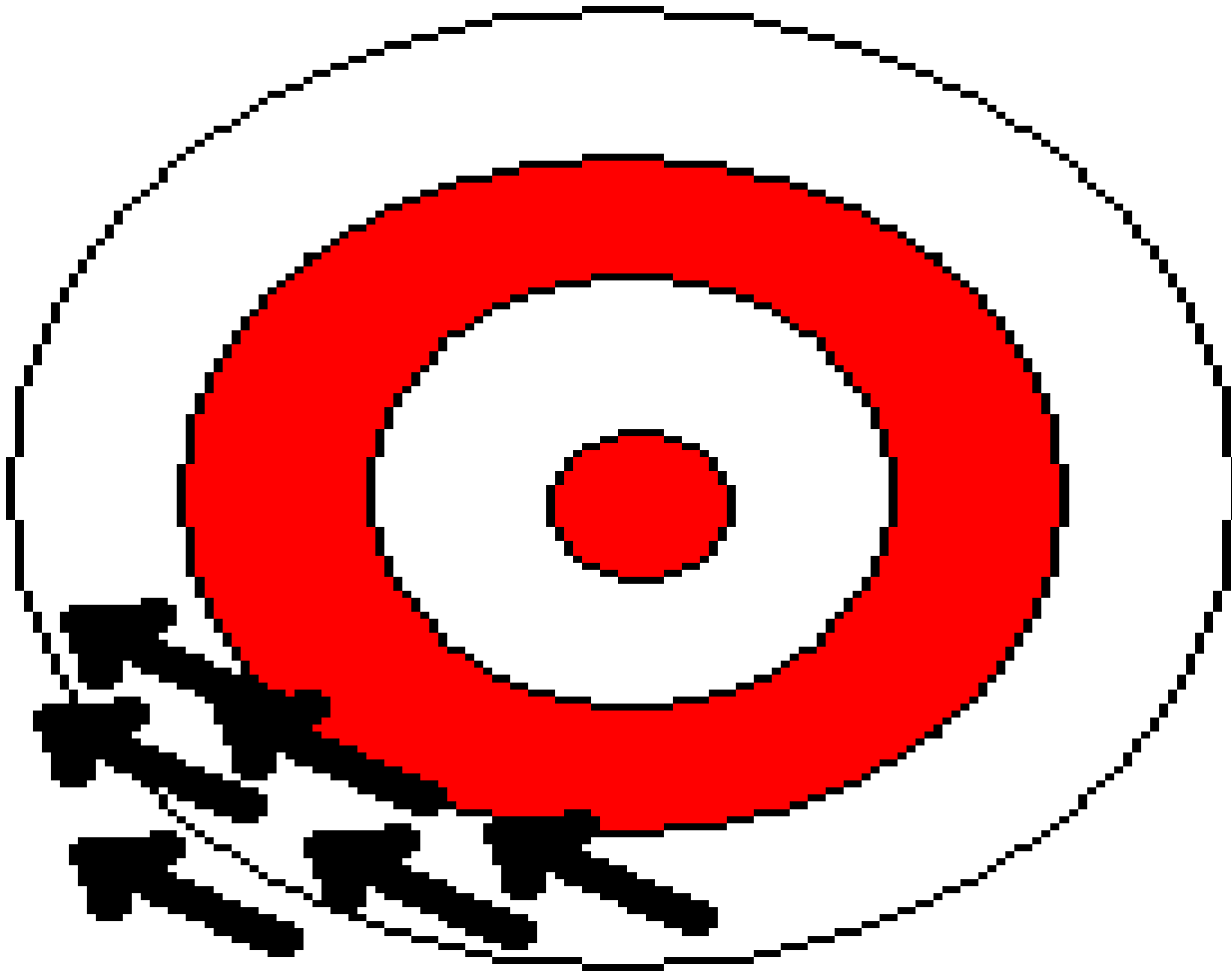
- A test is valid when it measures what it's supposed to.
- If a test is reliable, it yields consistent results.
- A test can be both reliable and valid, one or the other, or neither.

Reliability

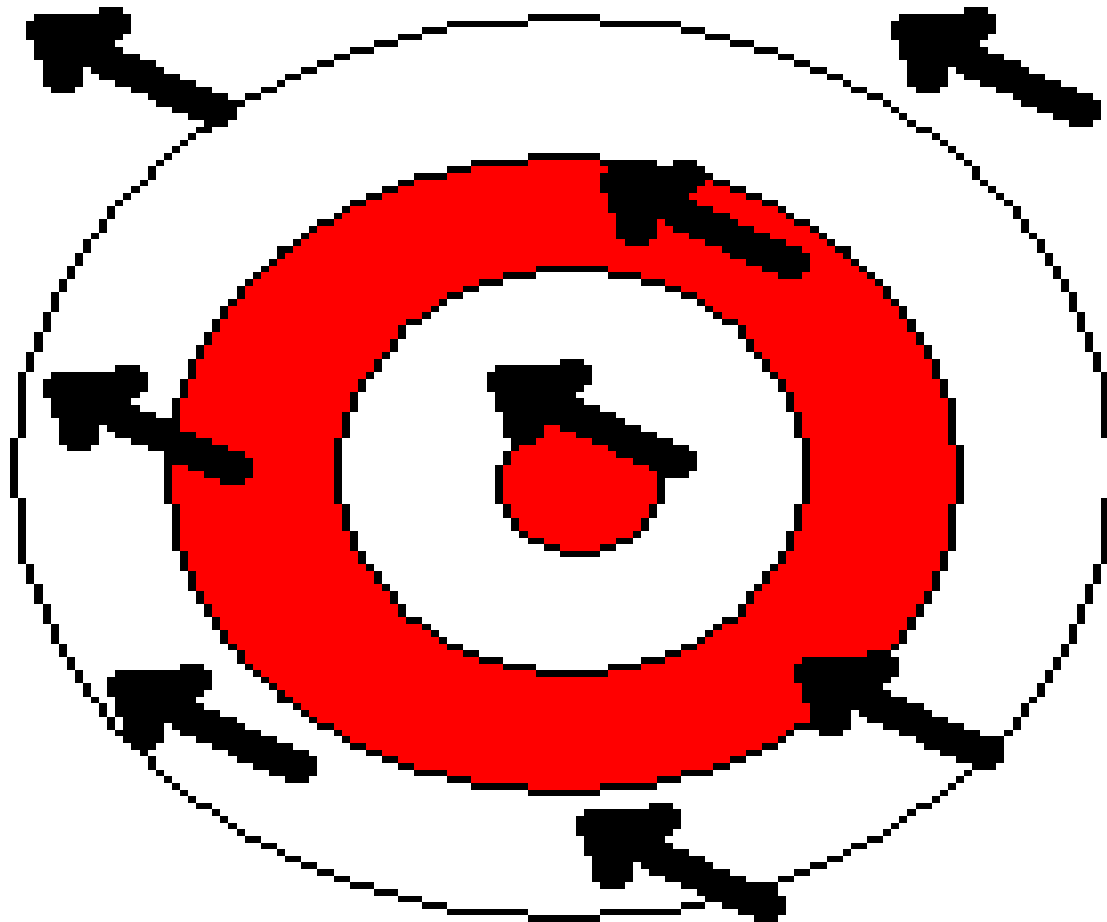


a prerequisite for measurement
validity

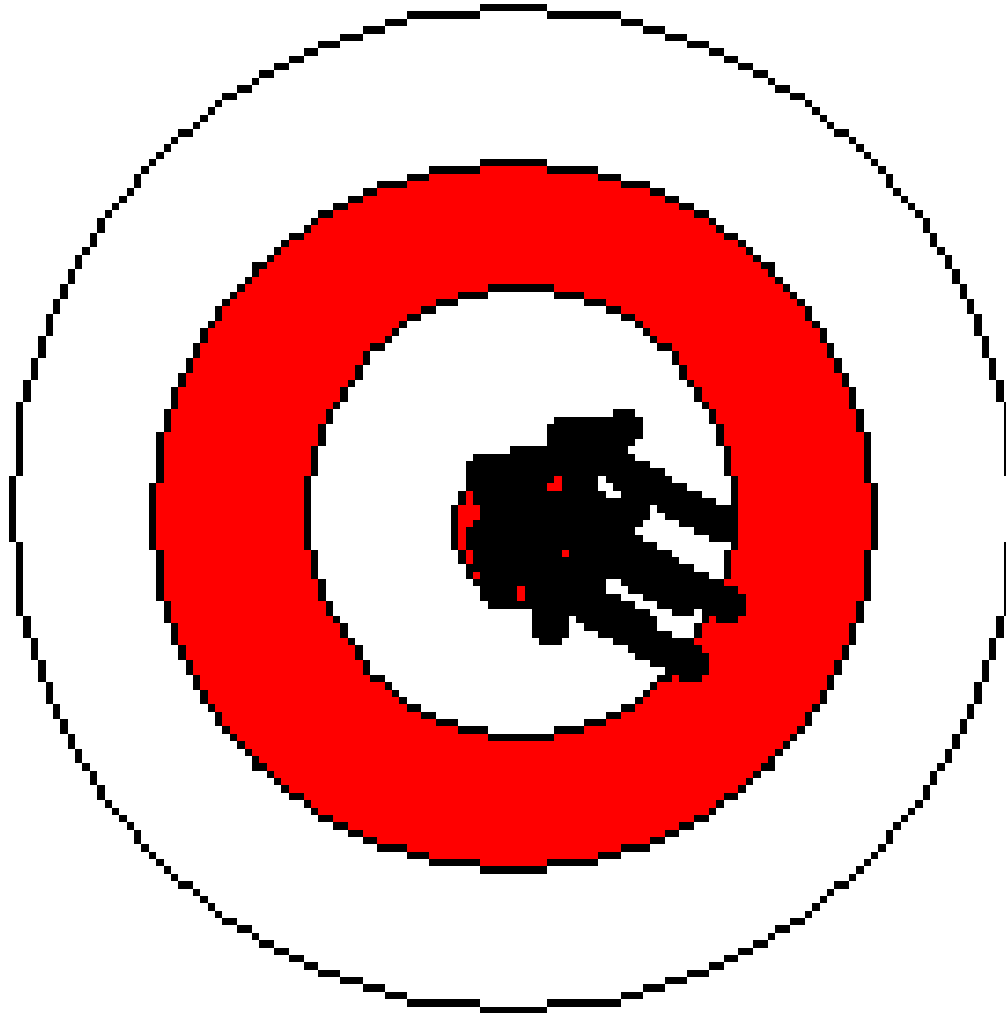
Reliable, but Not Valid!



Not Reliable, Not Valid!



Reliable and Valid



Questions in critical appraisal

□ **Validity:**

- ▣ Bias (reference pop., non-homogenous sample, sample selection, measurements, contamination, study attrition or drop-out rate)
- ▣ Confounding variables

Reliability

- ▣ Size of the effect
- ▣ Precision of the effect
- ▣ Magnitude of the effect (i.e. prevalence, incidence)
- ▣ Effect measures (e.g. risk, odds) and its precision
- ▣ Strength of the association (e.g. RRR, NNT)
- ▣ Measures of clinical significance
- ▣ Power (accuracy and representativeness of the results, i.e. type I and II errors or α and β error)
- ▣ Sensitivity and specificity

Applicability

Will the results help locally? 

Type of articles

- Diagnosis
- Etiology/harm
- Prognosis
- Therapy

مطالعات تشخیصی

- پزشکان به نحوی فرآیند دستور انجام آزمایشهای تشخیصی و غربالگری را صادر می کنند و استفاده کنندگان از خدمات بهداشتی هم همواره انتظار انجام این آزمایشها را دارند. اما آزمایش ها هرگز ۱۰۰٪ دقیق نیستند و نتیجه " مثبت کاذب" و " منفی کاذب" متضمن حالت ایتلاي خود هستند.
- هدف این بخش ایجاد توانایی در ارزیابی يك مطالعه مربوط به تعیین عملکرد آزمایش تشخیصی یا غربالگری نسبت به يك معیار اساسی (استاندارد طلایی Gold Standard)
- مشخص نشان بدهند و در مفید بودن علمی آن تصمیم بگیرند شرکت کنندگان باید احساس اطمینان کنند که می توانند به دیگران هم کمک کنند تا این توانایی را نشان بدهند.
- مطالعاتی که صحت آزمونهایی تشخیصی را بررسی می کنند ممکن است از نوع مقطعی یا مورد شاهدهی باشند.

□ در حالی که مطالعاتی که صحت آزمونهایی پیش بینی کننده وضعیت بیماری را ارزیابی می کنند، معمولاً از نوع هم گروهی هستند. البته مطالعات مورد - شاهدهی در ارزیابی صحت آزمون های تشخیصی محدود به بیماریهای نادر می شود که هیچ طرح مطالعه دیگری برای آنها نمی توان انجام داد یا انجام آن بسیار دشوار است.

در تعیین صحت آزمایش یا روش تشخیصی دو نکته را باید در نظر گرفت

□ میزان کارایی روش تشخیصی در وضعیت مورد نظر چقدر است؟ (چند درصد از بیمارانی که دارای نتیجه آزمایش مثبت هستند، واقعاً مبتلا به بیماری هستند؟) (حساسیت = میزان مثبت واقعی)

□ میزان شناسایی صحیح آن دسته از افرادی است که به بیماری مورد نظر مبتلا نیستند چقدر است؟ (ویژگی = میزان منفی واقعی)

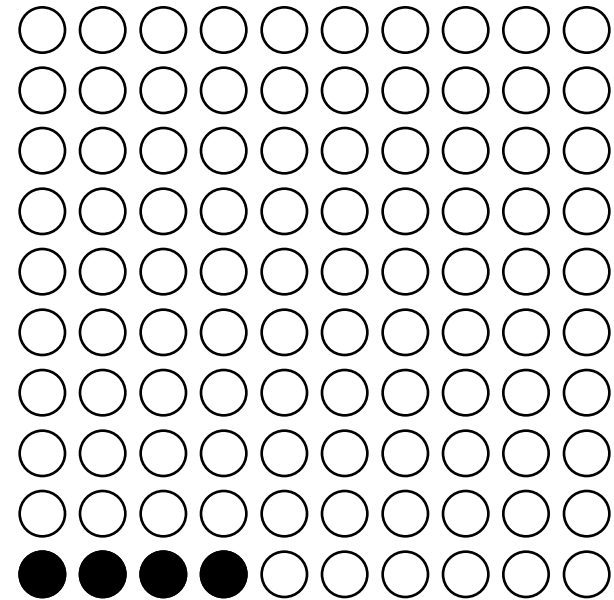
- علاوه بر **حساسیت و ویژگی**، دو میزان دیگر در تعیین اعتبار يك آزمون تشخیصی اهمیت دارند حساسیت و ویژگی را می توان بطور مستقل با مطالعه جداگانه گروههای افراد بیمار و سالم تعیین نمود. با وجود این مفید بودن يك آزمون تشخیصی به میزان **شیوع** حقیقی بیماری مورد نظر در جمعیت مورد مطالعه نیز بستگی دارد.
- **ارزش (مقدار) پیش بینی** يك آزمون پارامتر مهمتری است. مقدار پیش بینی يك آزمون نسبت موارد واقعی در میان کل افرادی است که نتیجه آزمون در آنها مثبت است .
- میزان شیوع حقیقی بیماری با مقدار پیش بینی يك آزمون به ازای هر مقدار حساسیت و ویژگی ارتباط مستقیم دارد ،

- با افزایش شیوع بیماری "مقدار پیش بینی" افزایش می یابد.
- برای مثال، یک آزمون در یک محیط بیمارستانی که انتظار می رود شیوع هر وضعیت زیاد باشد، مقدار پیش بینی زیادی خواهد داشت، در حالیکه وقتی همین آزمون برای جمعیت عمومی بکار می رود مقدار پیش بینی (Predictive value) کمی داشته باشد که عملاً کارایی نخواهد داشت.

□ "مقدار پیش بینی" = "ارزش اخباری"

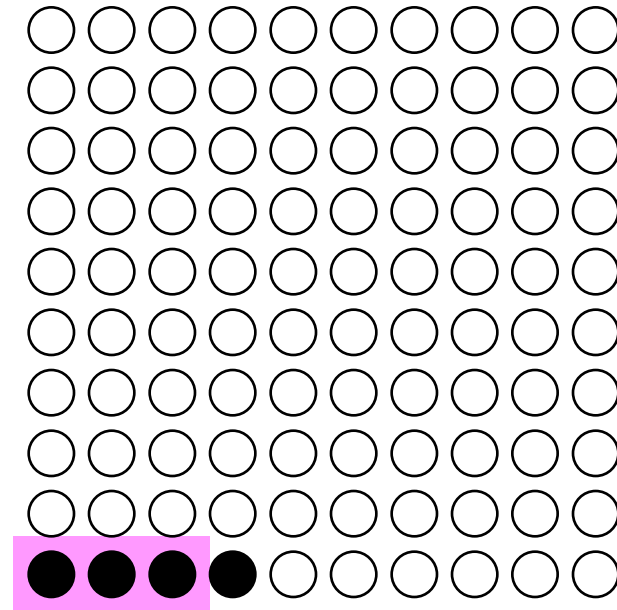
Person without the disease	○
Person with the disease	●

**Assume that the
prevalence of the disease
is 4%**



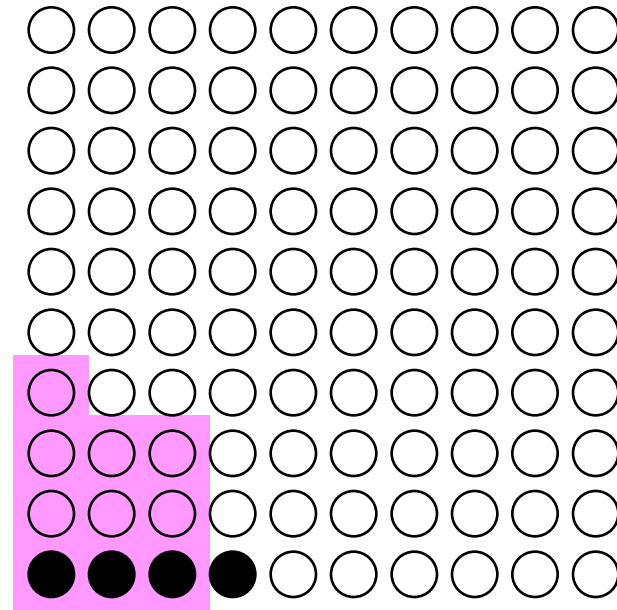
Assume that of the 4 people with the disease, 3 are picked up by the test

Person without the disease	○
Person with the disease	●
Person who tests positive	■
Person who tests negative	■



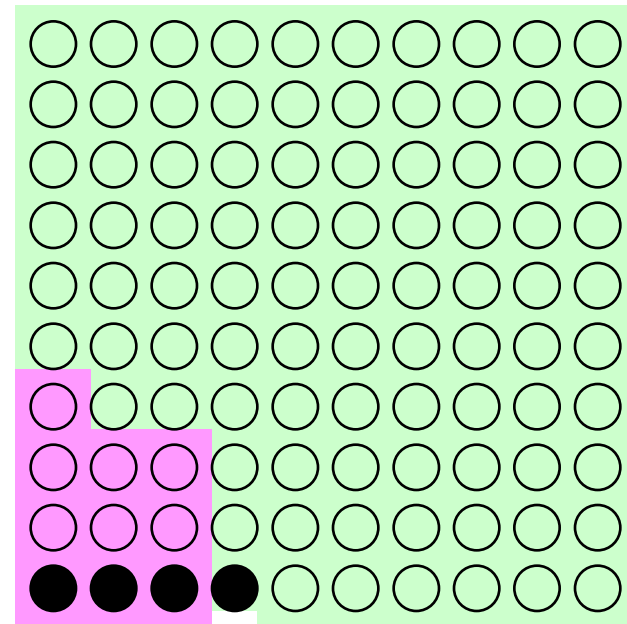
Assume that of the test is positive for a further 7 people who don't have the disease

Person without the disease	○
Person with the disease	●
Person who tests positive	■
Person who tests negative	■
True positive on the test	●
False positive on the test	○



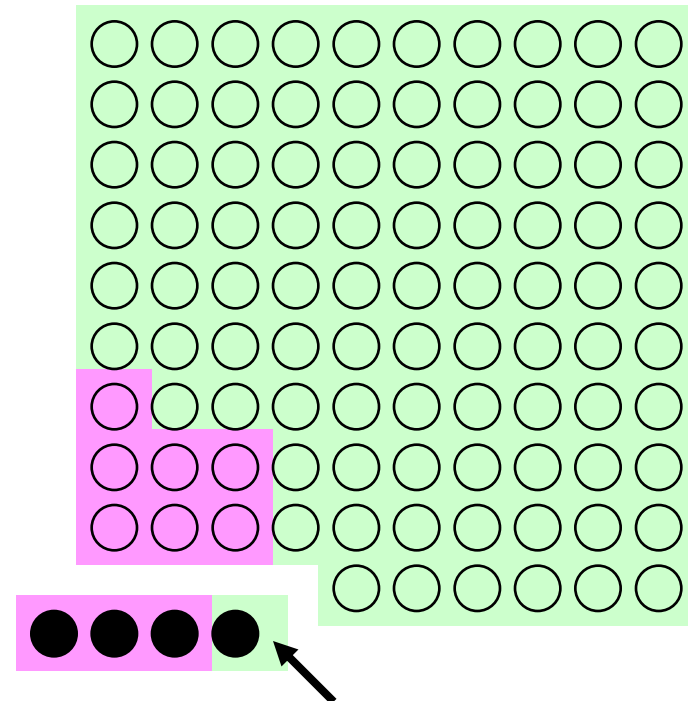
The remainder of the sample are negative on the test

Person without the disease	
Person with the disease	
Person who tests positive	
Person who tests negative	
True positive on the test	
False positive on the test	
True negative on the test	
False negative on the test	



SENSITIVITY

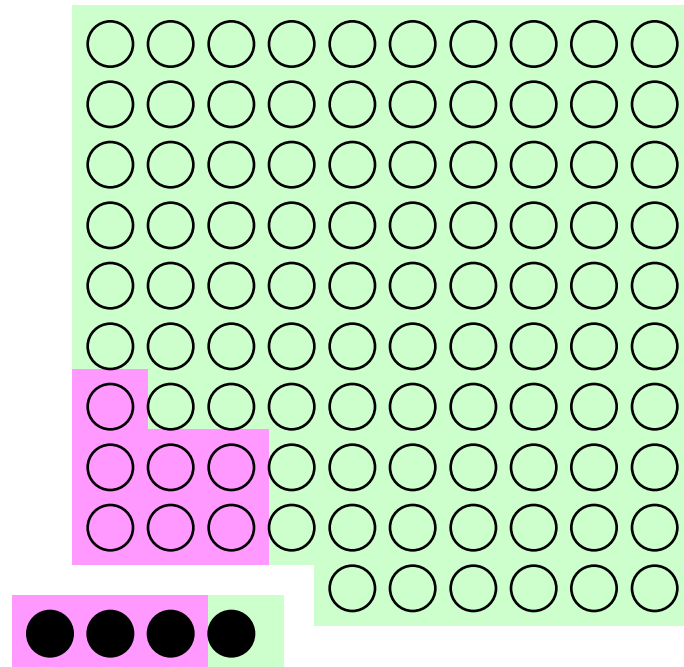
Person without the disease	○
Person with the disease	●
Person who tests positive	■
Person who tests negative	□
True positive on the test	●■
False positive on the test	○■
True negative on the test	○□
False negative on the test	●□



- **SENSITIVITY** is the proportion of people with the disease correctly identified by the test
- It measures the proportion of false **NEGATIVES**

SENSITIVITY

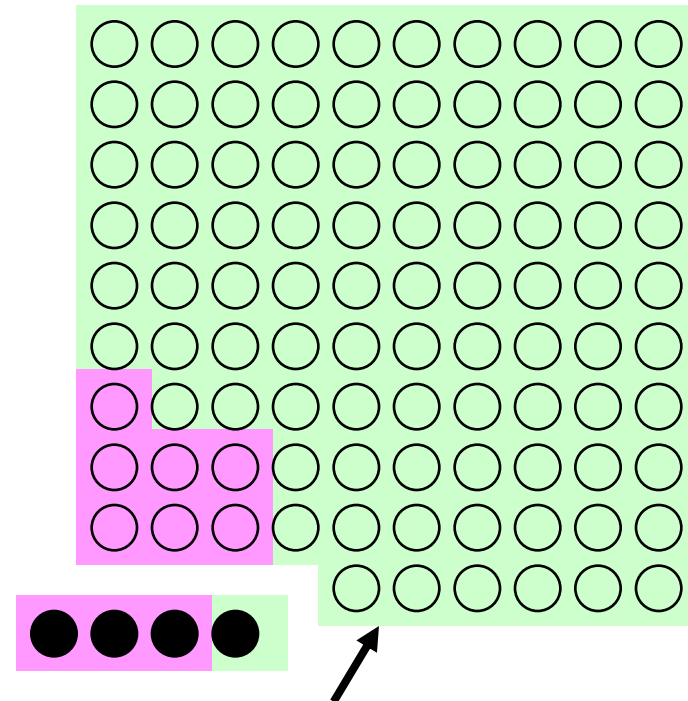
Person without the disease	○
Person with the disease	●
Person who tests positive	■
Person who tests negative	□
True positive on the test	●■
False positive on the test	○■
True negative on the test	○□
False negative on the test	●□



In this case, sensitivity is $\frac{3}{4}$ or 75%

SPECIFICITY

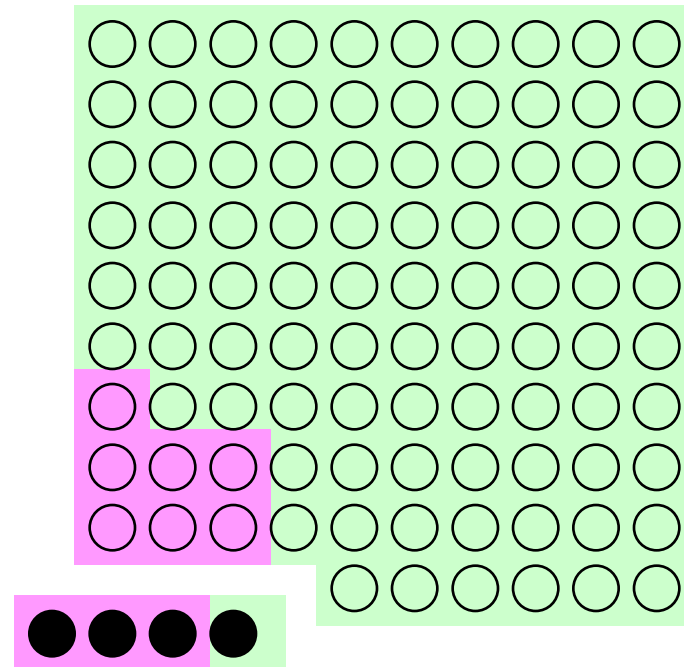
Person without the disease	○
Person with the disease	●
Person who tests positive	■
Person who tests negative	□
True positive on the test	●■
False positive on the test	○■
True negative on the test	○□
False negative on the test	●□



- **SPECIFICITY** is the proportion of people without the disease correctly identified by the test
- It measures the proportion of false **POSITIVES**

SPECIFICITY

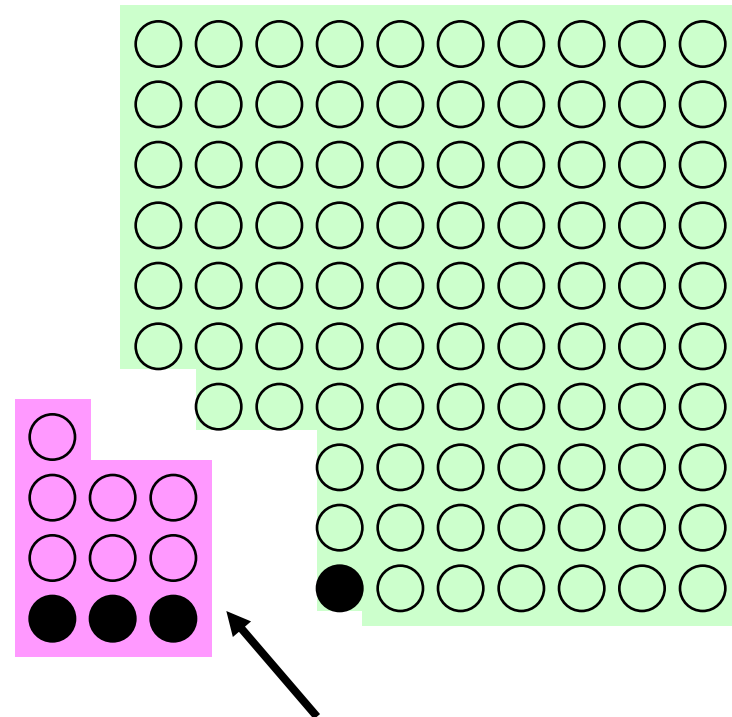
Person without the disease	○
Person with the disease	●
Person who tests positive	■
Person who tests negative	□
True positive on the test	●■
False positive on the test	○■
True negative on the test	○□
False negative on the test	●□



In this case, specificity is $(96-7)/96$ or 93%

If someone is positive on the test, what are the chances that he/she has the disease?

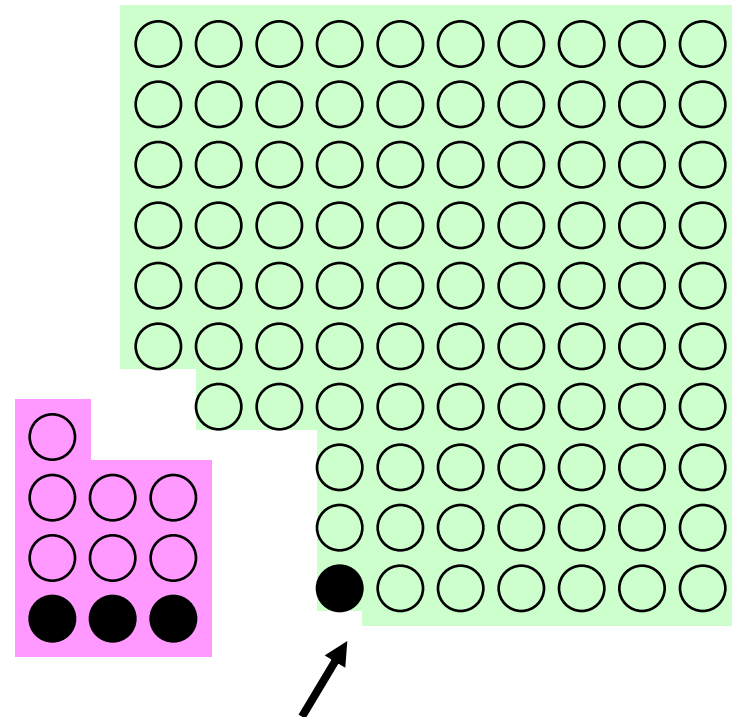
Person without the disease	○
Person with the disease	●
Person who tests positive	■
Person who tests negative	□
True positive on the test	●
False positive on the test	○
True negative on the test	○
False negative on the test	●



- **Probability = $3/10 = 30\%$**
- **This is the POSITIVE PREDICTIVE VALUE (the value of the test in predicting a positive result)**

If someone is negative on the test, what are the chances that he/she does not have the disease?

Person without the disease	○
Person with the disease	●
Person who tests positive	■
Person who tests negative	□
True positive on the test	●■
False positive on the test	○■
True negative on the test	○□
False negative on the test	●□



- **Probability = $89/90 = 99\%$**
- **This is the NEGATIVE PREDICTIVE VALUE (the value of the test in predicting a negative result)**

Diagnosis

Validity:

- What is the **defined** question?
- How could be compared with **gold standard**?
- How about **Spectrum**?
- Application for **all patients**?

Diagnosis

Reliability:

- ▣ Sensitivity
- ▣ Specificity
- ▣ Positive/Negative predictive value (PPV/NPV)
- ▣ Positive/Negative likelihood ratio
- ▣ Prevalence

Definition

- **Sensitivity** is the probability of a positive test in a diseased person
- **Specificity** is the probability of a negative test in a non-diseased person
- **POSITIVE PREDICTIVE VALUE (PPV)** is the value of the test in predicting a positive result
- **NEGATIVE PREDICTIVE VALUE (NPV)** is the value of the test in predicting a negative result



Sensitivity: % patients with disease and positive test

Specificity: % people without disease and with negative test

PPV: % people with positive test who are really diseased

NPV: % people with negative test who are really healthy

		Target	disorder	total
		+	-	
Diagnostic	+	A	B	A+B
test result	-	C	D	C+D
total		A+C	B+D	A+B+C+D

Sensitivity = $a/(a+c)$

Specificity = $d/(b+d)$

Positive Predictive Value = $a/(a+b)$

Negative Predictive Value = $d/(c+d)$

- 
- Pre-test probability (prevalence) = $(a+c)/(a+b+c+d)$
 - Pre-test odds = $\text{prevalence}/(1-\text{prevalence})$
 - Post-test odds = pre-test odds $\times$ LR
 - Post-test probability = $\text{post-test odds}/(\text{post-test odds} + 1)$

- Likelihood ratio for a positive test result :

$$\text{LR}(+) = \text{sensitivity} / (1 - \text{specificity})$$

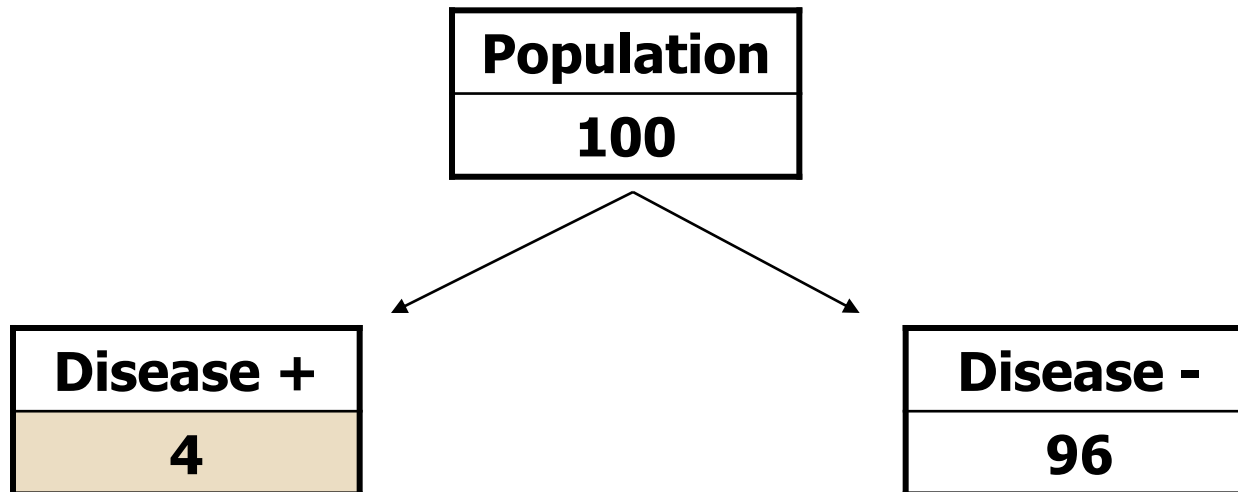
- Likelihood ratio for a negative test result :

$$\text{LR}(-) = (1 - \text{sensitivity}) / \text{specificity}$$

NATURAL FREQUENCIES TREE

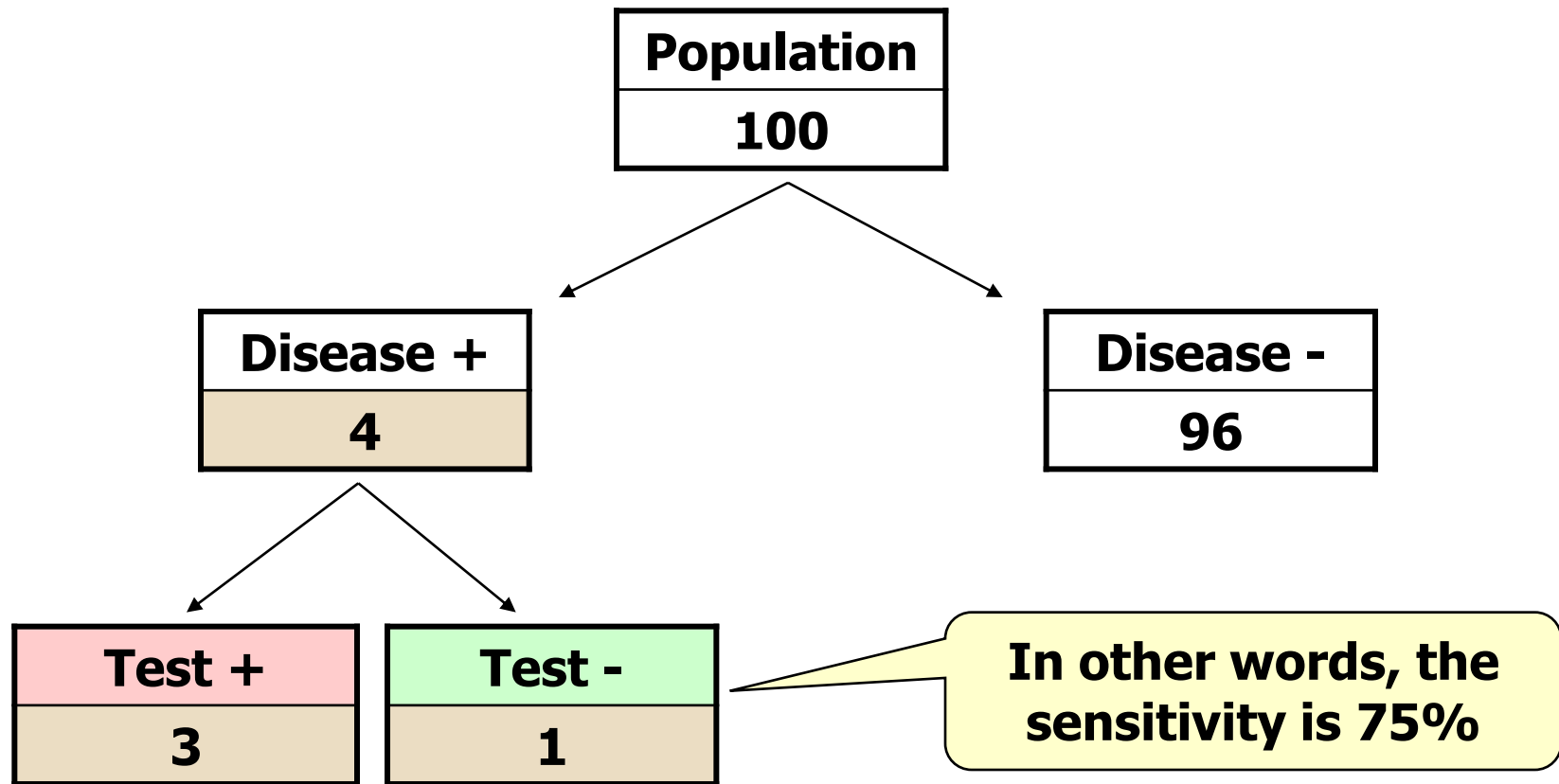
Population
100

IN EVERY 100 PEOPLE, 4 WILL HAVE THE DISEASE

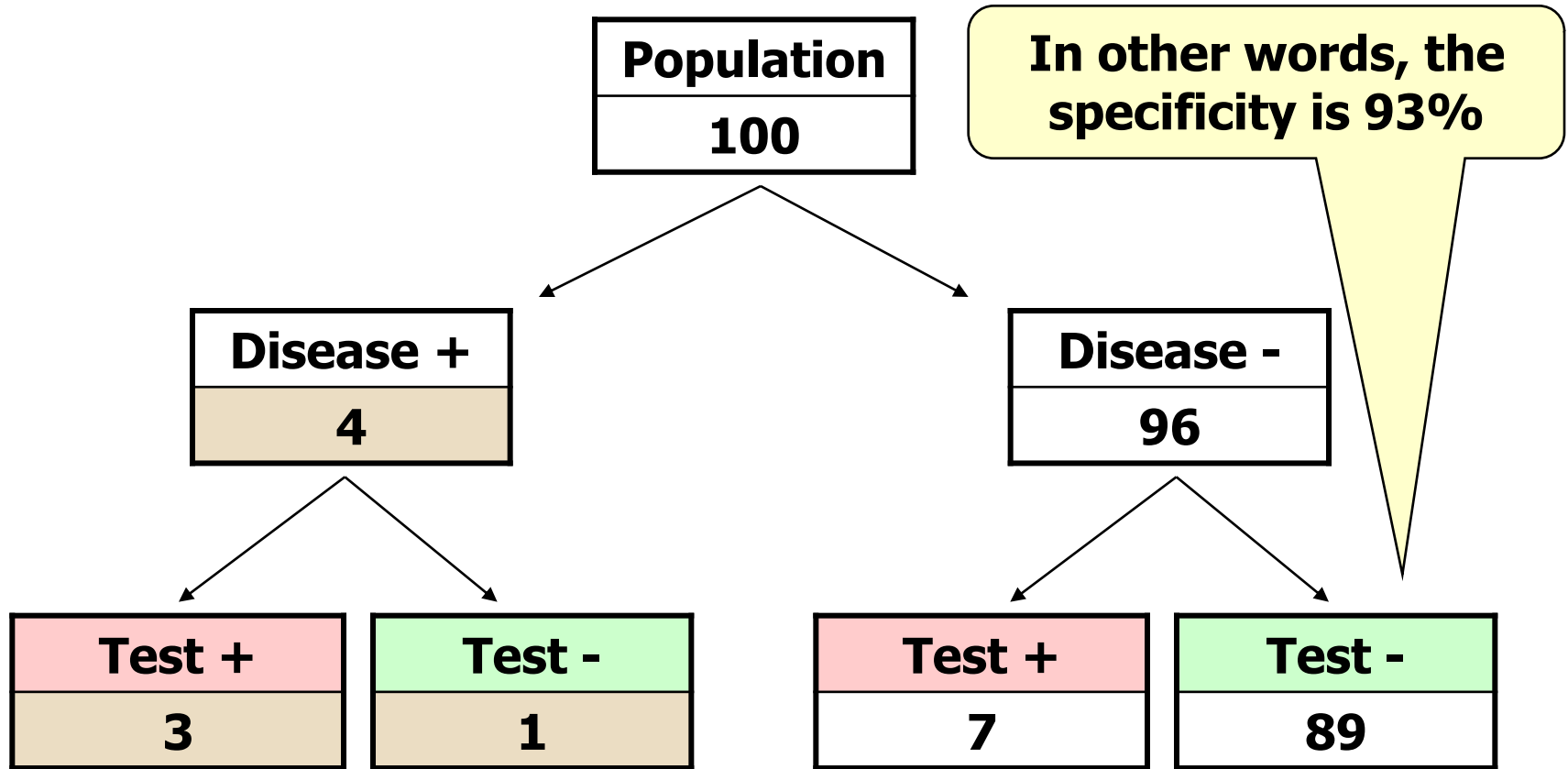


If these 100 people are representative of the population at risk, the assessed rate of those with the disease (4%) represents the PREVALENCE of the disease – it can also be considered the PRE-TEST PROBABILITY of having the disease

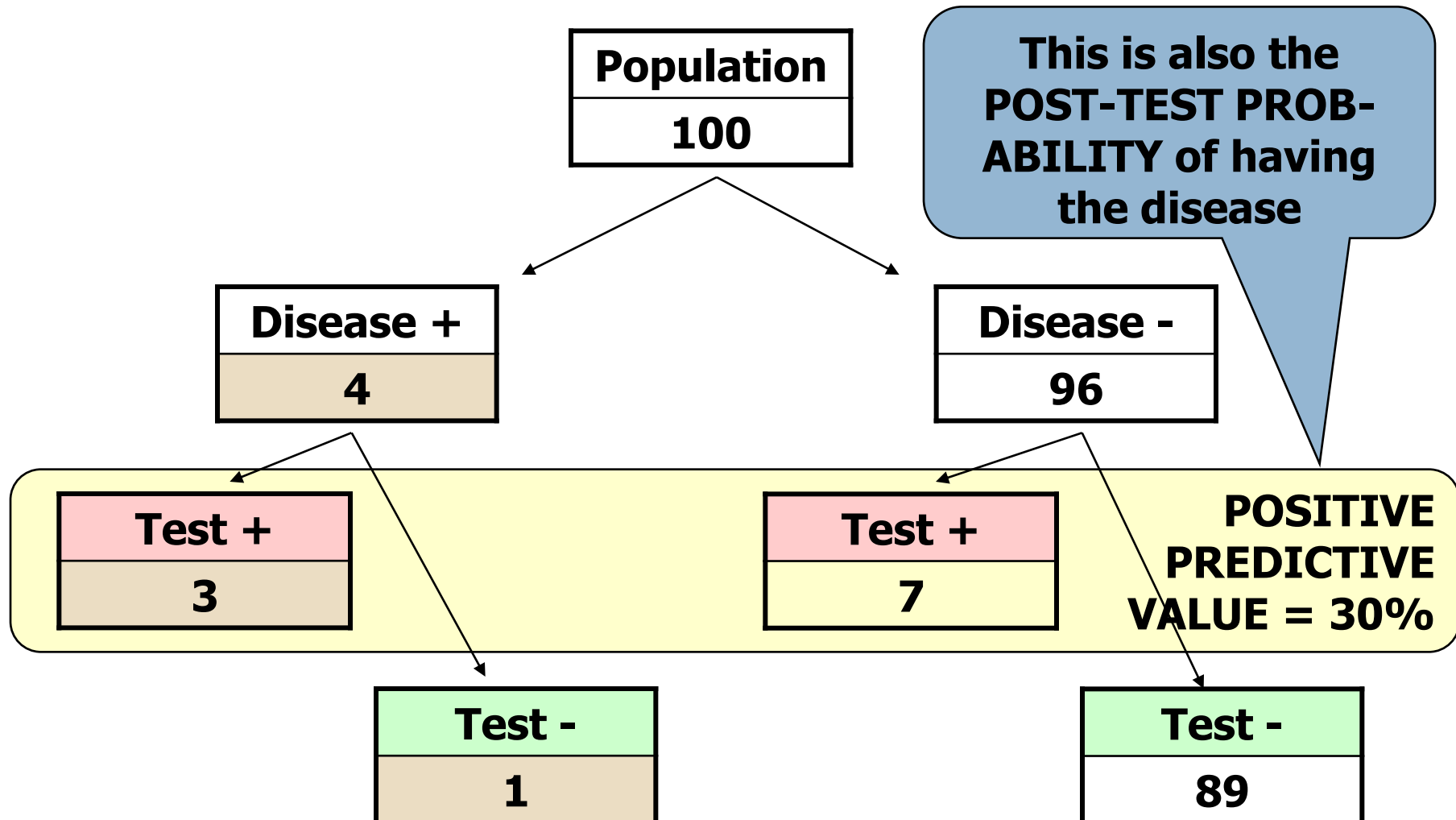
OF THE 4 PEOPLE WITH THE DISEASE, THE TEST
WILL DETECT 3



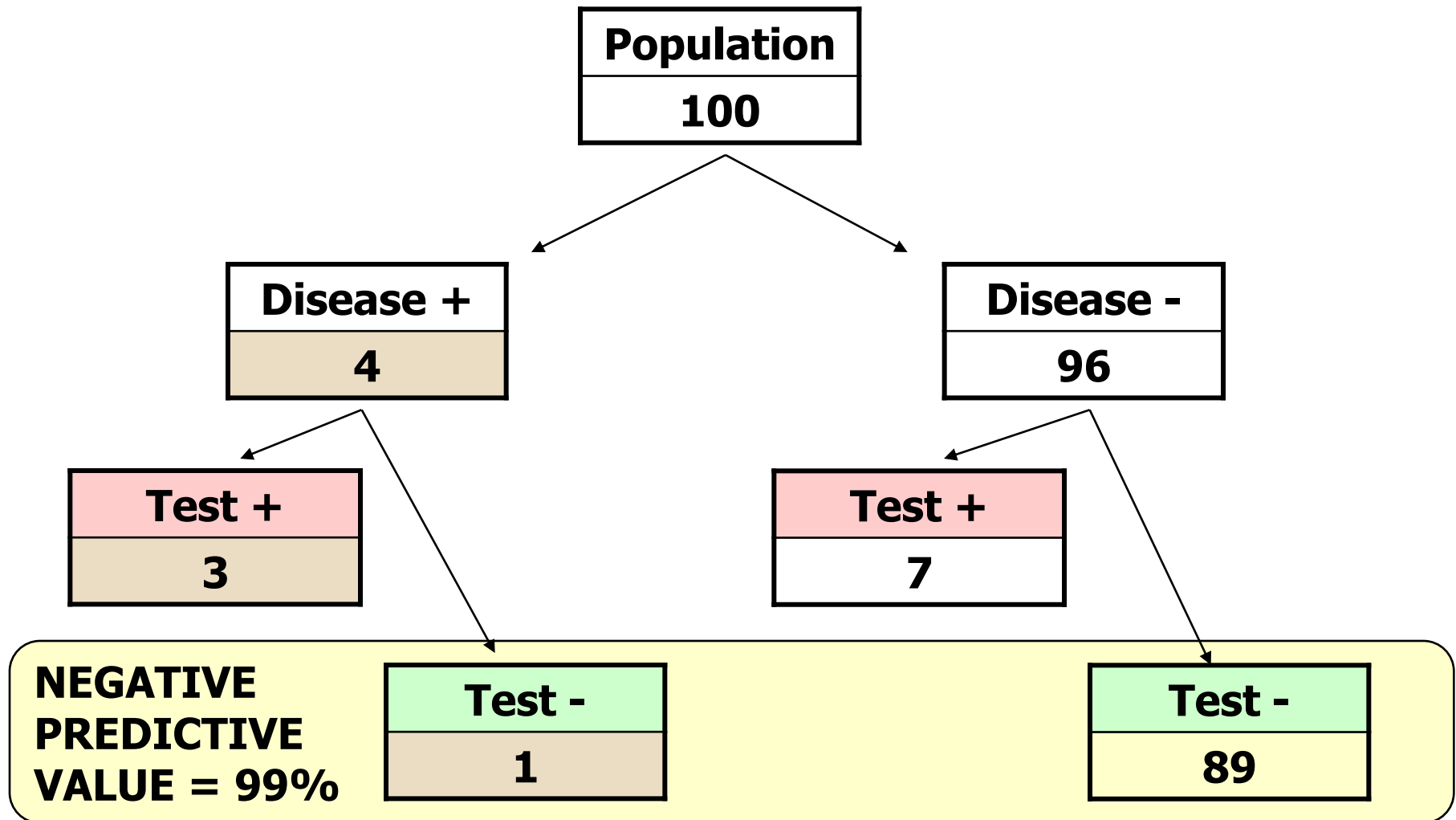
AMONG THE 96 PEOPLE WITHOUT THE DISEASE, 7
WILL TEST POSITIVE



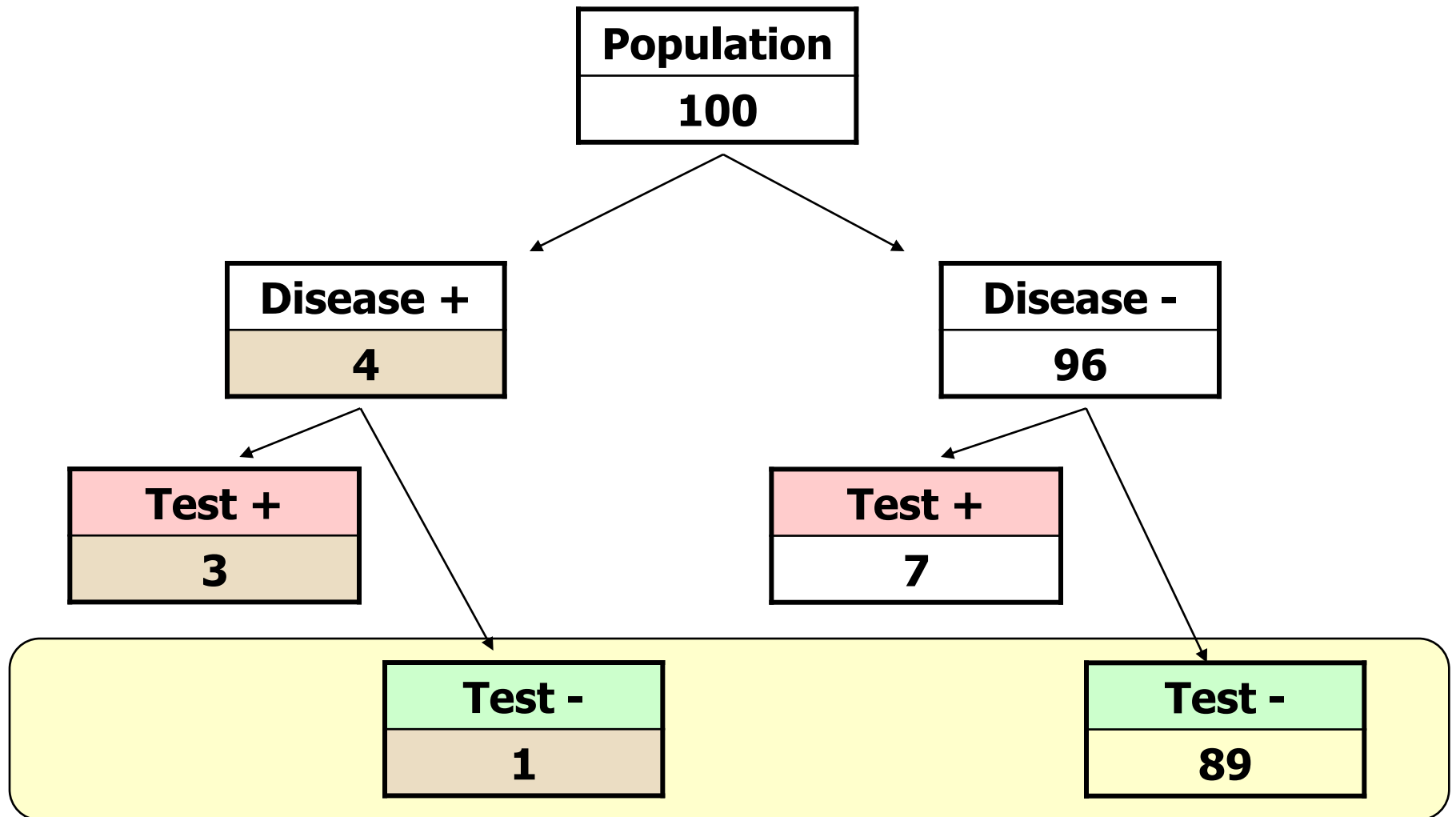
AMONG THOSE WHO TEST POSITIVE, 3 IN 10 WILL ACTUALLY HAVE THE DISEASE



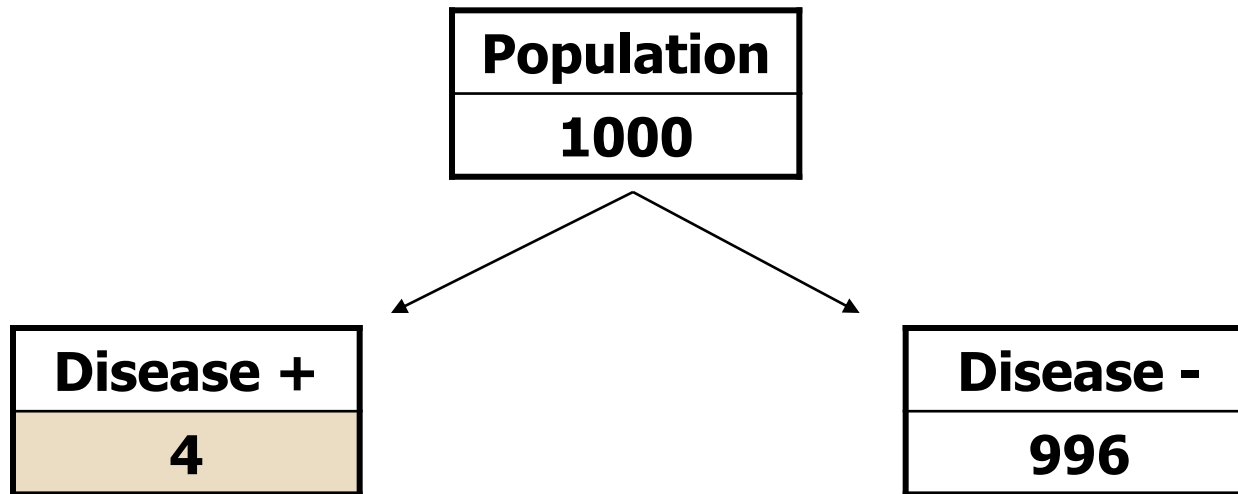
AMONG THOSE WHO TEST NEGATIVE, 89 OF 90 WILL NOT HAVE THE DISEASE



CONVERSELY, IF SOMEONE TESTS NEGATIVE, THE CHANCE OF HAVING THE DISEASE IS ONLY 1 IN 90

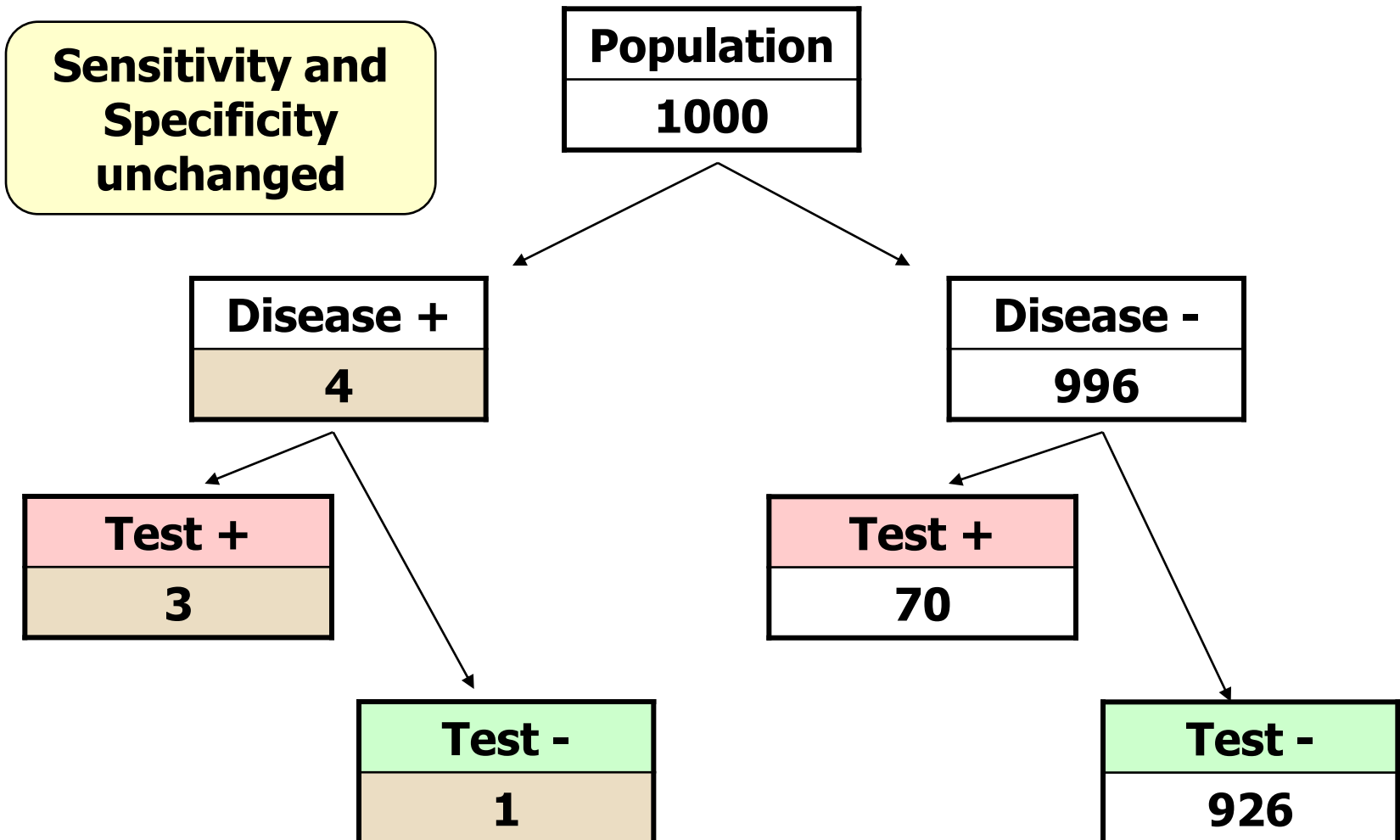


PREDICTIVE VALUES AND CHANGING PREVALENCE

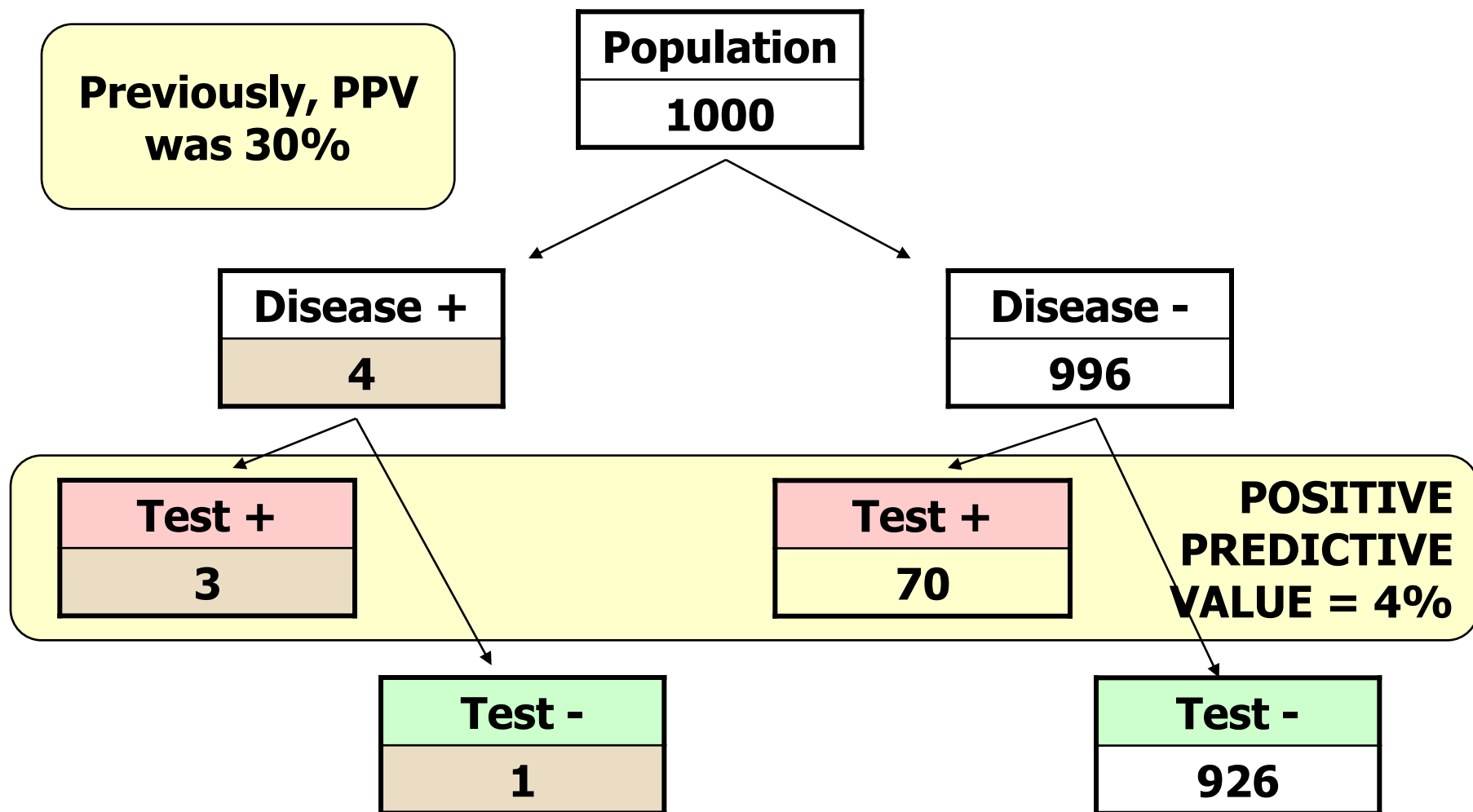


Prevalence reduced by an order of magnitude from 4% to 0.4%

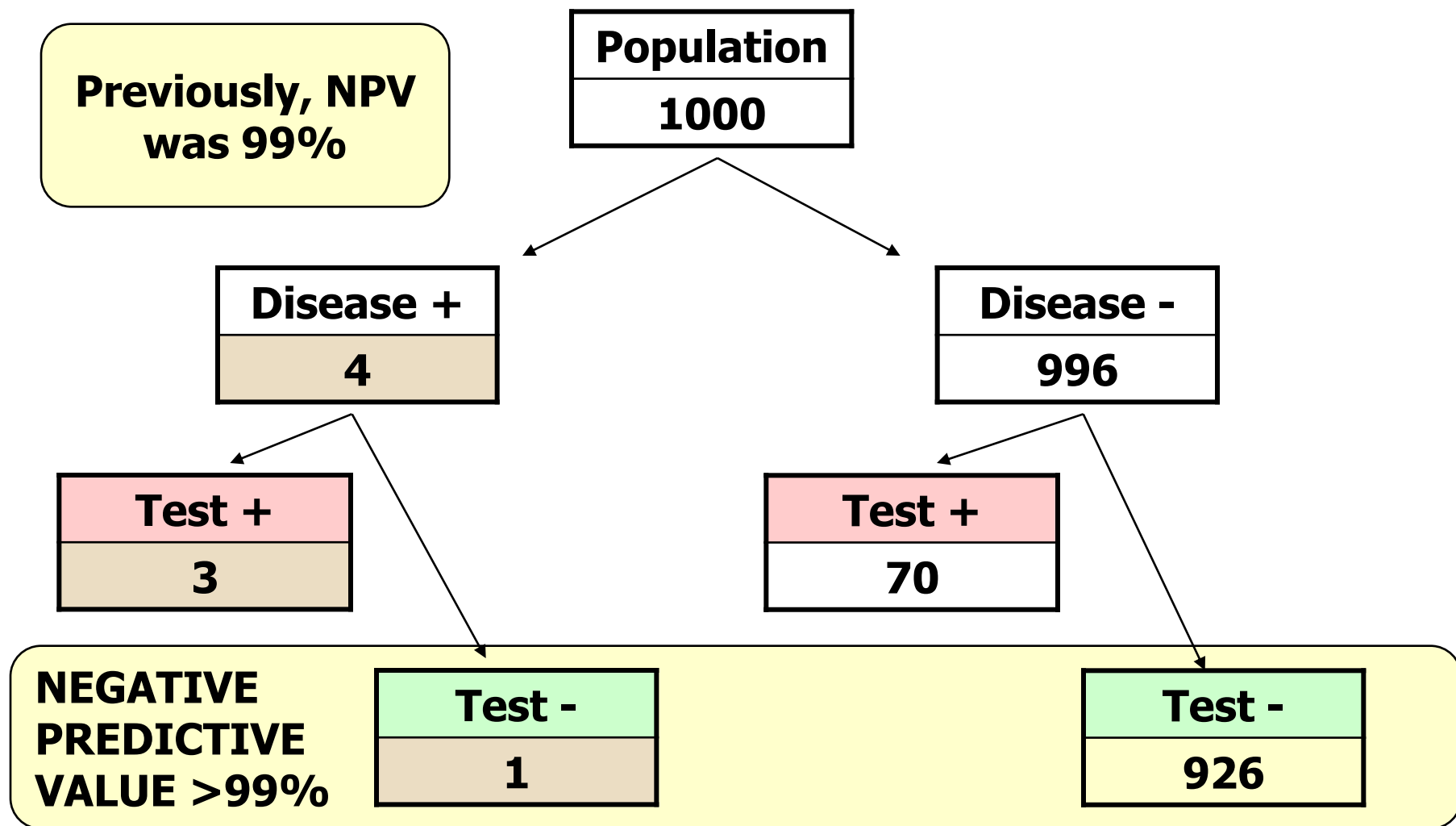
PREDICTIVE VALUE AND CHANGING PREVALENCE



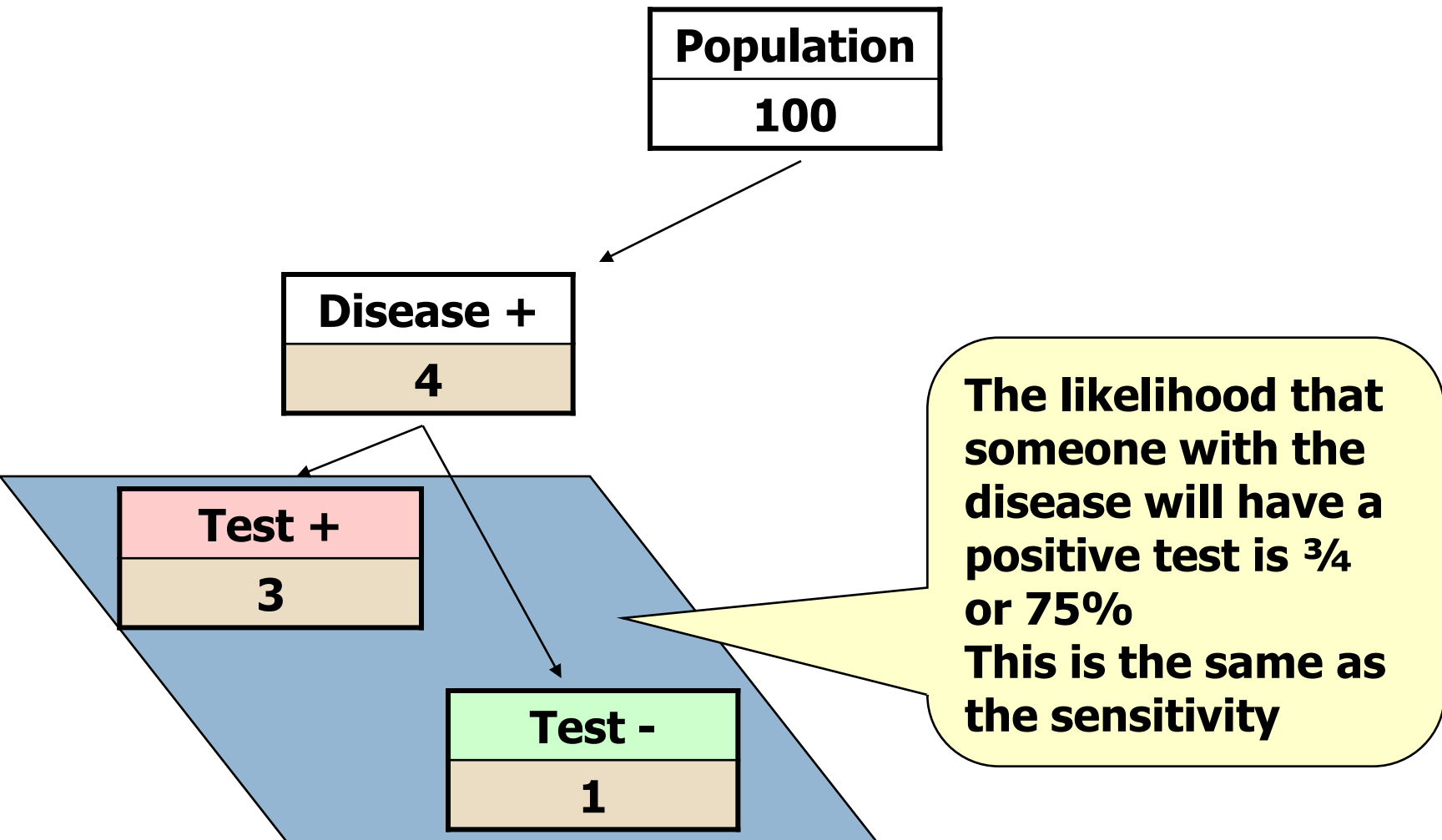
POSITIVE PREDICTIVE VALUE AT LOW PREVALENCE



NEGATIVE PREDICTIVE VALUE AT LOW PREVALENCE



LIKELIHOOD



LIKELIHOOD II

Population
100

Disease -
96

Test +
7

Test -
89

The likelihood that someone without the disease will have a positive test is $7/96$ or 7%. This is the same as the (1-specificity)

LIKELIHOOD RATIO

$$\begin{aligned}\text{LIKELIHOOD RATIO} &= \frac{\text{LIKELIHOOD OF POSITIVE TEST GIVEN THE DISEASE}}{\text{LIKELIHOOD OF POSITIVE TEST IN THE ABSENCE OF THE DISEASE}} \\ &= \frac{\text{SENSITIVITY}}{1 - \text{SPECIFICITY}} = \frac{0.75}{0.07} = 10.7\end{aligned}$$

A Likelihood Ratio of 1.0 indicates an uninformative test (occurs when sensitivity and specificity are both 50%)

The higher the Likelihood Ratio, the better the test (other factors being equal)

odds

- Pre-test probability (prevalence) $= (a+c)/(a+b+c+d)$

$$\text{Prevalence} = 4/100 = 4\%$$

- Pre-test odds $= \text{prevalence}/(1-\text{prevalence})$

$$\text{Pre-test odds} = 4/(1-4) = 0.042$$

- 
- Post-test odds = pre-test odds * LR

$$\text{Post-test odds} = 0.042 * 10.7 = 0.446$$

- Post-test probability = post-test odds / (post-test odds + 1)

$$\text{Post-test probability} = 0.446 / 1.446 = 0.309$$

Is the test helpful (valid)?

The Youden Index

$$\text{Youden Index} = (\text{Sensitivity} + \text{Specificity}) - 1$$

For a test to be useful, then

$$\text{sensitivity} + \text{specificity} > 1$$

$$(\text{Youden Index} > 0)$$

Examples:

$$\text{sensitivity} = 0.5$$

$$\text{specificity} = 0.5$$

$$\text{Youden index} = 0$$

Example

- Suppose that we are working up a patient with anemia. The probability that she has IDA is 50% (i.e. odds=50:50).
- We are looking for the usefulness of performing a serum **Ferritin** on our patient as a mean for detecting IDA.
- The reference standard of a **bone marrow stain** for Fe is our GOLD STANDARD.

		<u>Bone</u>	<u>Marrow</u>	
		Positive	Negative	
Test result	Positive	a=731	b=270	a+b=1001
(Ferritin)	Negative	c=78	d=1500	c+d=1578
		a+c=809	b+d=1770	A+b+c+d = 2579

Results

- Sensitivity: $a/(a+c)=731/809=90\%$
- Specificity: $d/(b+d) = 1500/1770 = 85\%$
- $LR(+) = 0.90/(1-0.85)= 6:1$



Thank you