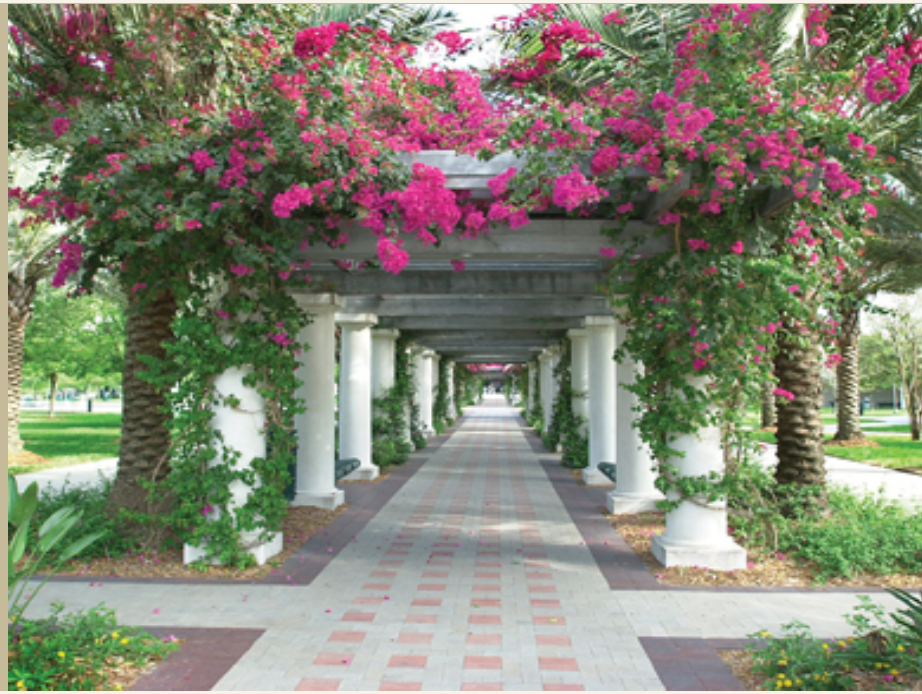
Fundamental Concepts of Study Design – Critical Appraisal: DIAGNOSIS



Ambuj Kumar, MD, MPH

Assistant Professor, Deputy Director

Division and Center for Evidence Based Medicine, Dept of Internal Medicine, College of Medicine, University of South Florida

Tampa, Florida USA

USF HEALTH CENTER FOR EVIDENCE BASED MEDICINE

Outline

- Introduction
  - Objectives
- Diagnostic Test Study Designs
  - Architecture of diagnostic research
- Summary

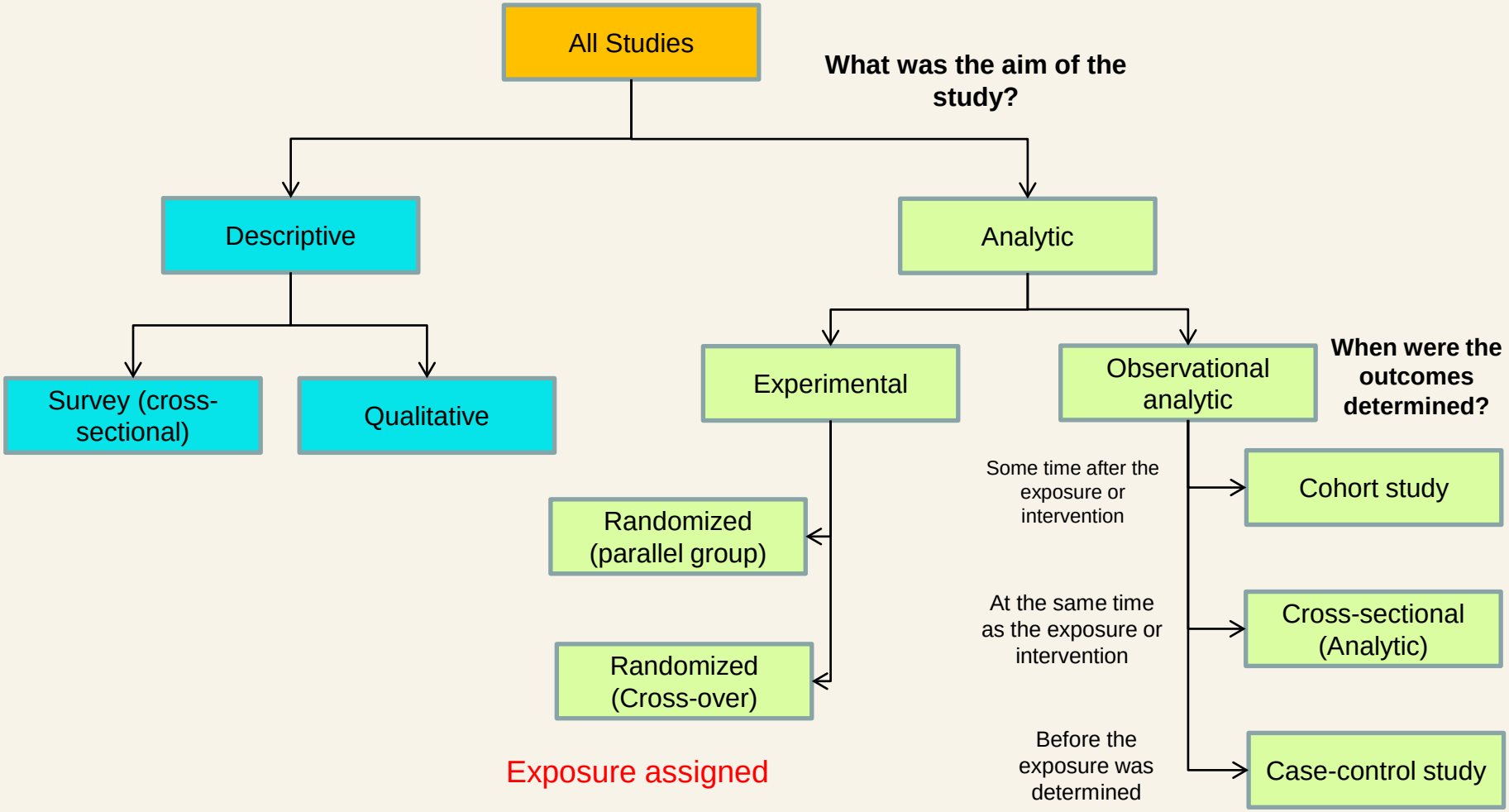
Objectives

- Be able to
  - Assess key study design features for evaluating diagnostic tests
    - Evidence Literacy
  - Define reference standard
    - Evidence Literacy
  - Recognize threats to claims that a test detects a health condition
    - Evidence Literacy
  - Understand measures of test performance: sensitivity, specificity, positive and negative predictive values, likelihood ratios
    - Evidence Numeracy

USF HEALTH CENTER FOR EVIDENCE BASED MEDICINE

What are the main study designs a clinician should be familiar with?

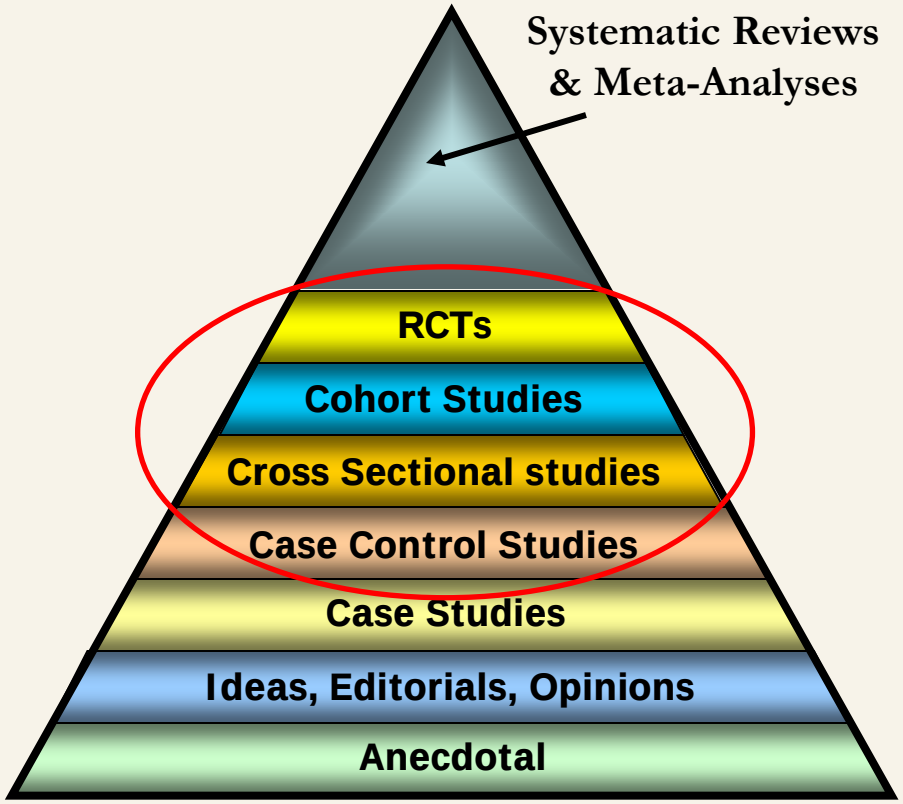
Spotting the study design



USF HEALTH CENTER FOR EVIDENCE BASED MEDICINE

What constitutes BEST Evidence?

- The Higher up a methodology is ranked, the more robust and closer to objective truth it is assumed to be.



USF HEALTH CENTER FOR EVIDENCE BASED MEDICINE

Architecture of diagnostic research

- Objectives of testing
  - *Increasing certainty*
    - *Presence*
    - *Absence*
  - *Supporting clinical management*
    - *E.g* surgery versus drugs
  - *Assessing prognosis*
    - As the starting point for clinical follow up and informing patients.
  - *Monitoring clinical course*
    - When a disease is untreated, or during or after treatment

Case

- You are in a busy primary care clinic seeing your last patient
- HISTORY
  - 70 yrs old male
  - Complains of cough and mild shortness of breath
  - Generally good health
  - Had cough for 2 weeks
  - Feels tired and slight shortness of breath when exercises
  - Lives alone and is a retired auto mechanic

Case

- Examination
  - Patient is comfortable
  - Temp of 38.0 degree Celsius
  - Respiratory rate of 15 and pulse of 98 (normal sinus rhythm)
  - Oxygen saturation in room air is 93%
  - Cardiac auscultation reveals a grade II systolic murmur
  - Lung exam mostly clear except for slight wheezing heard upon expiration
  - Rest of physical exam is completely normal

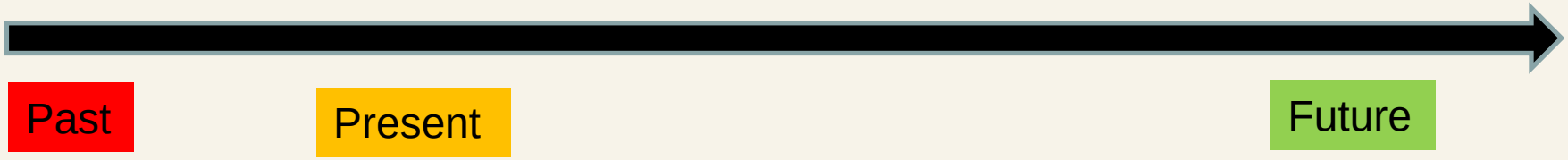
- Suspicion for pneumonia but you are concerned because patient lives alone
- You are unable to get chest x-rays at your clinic

Question/Wonder

- How accurate your assessment of this patient chances of having a pneumonia/How accurate history and physical exam is in detecting pneumonia?

Task

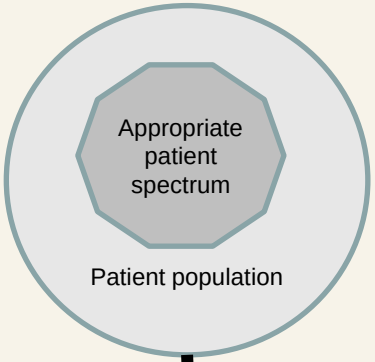
- Design a study on how to assess the diagnostic accuracy of X-ray compared with physical examination alone for the diagnosis of pneumonia
- Avoid jargon and explain in simple terms the process of study from beginning to end
- Time – 3 minutes



2 minutes

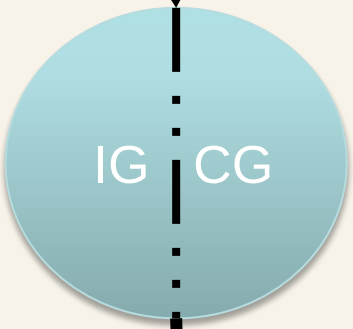
Validity

Participants



Representative

Index group (IG) &
Comparison Group (CG)
Gold/reference standard



Reproducible

	+	–
+	A	B
–	C	D

Outcome

Measurements
blind subjective? OR
objective?

USF HEALTH CENTER FOR EVIDENCE BASED MEDICINE

- Example:
  - assessment of the value of the plasma concentration of B-type natriuretic peptide (BNP) in the diagnosis of left ventricular dysfunction (LVD)

Do test results in patients with the target disorder differ from those in normal people?

- Investigators at a British university hospital measured concentrations of BNP precursor
  - in non-systematic ("convenience") samples
    - Normal controls
    - Patients (combinations of hypertension, ventricular hypertrophy, and left ventricular dysfunction)

Investigator's conclusion

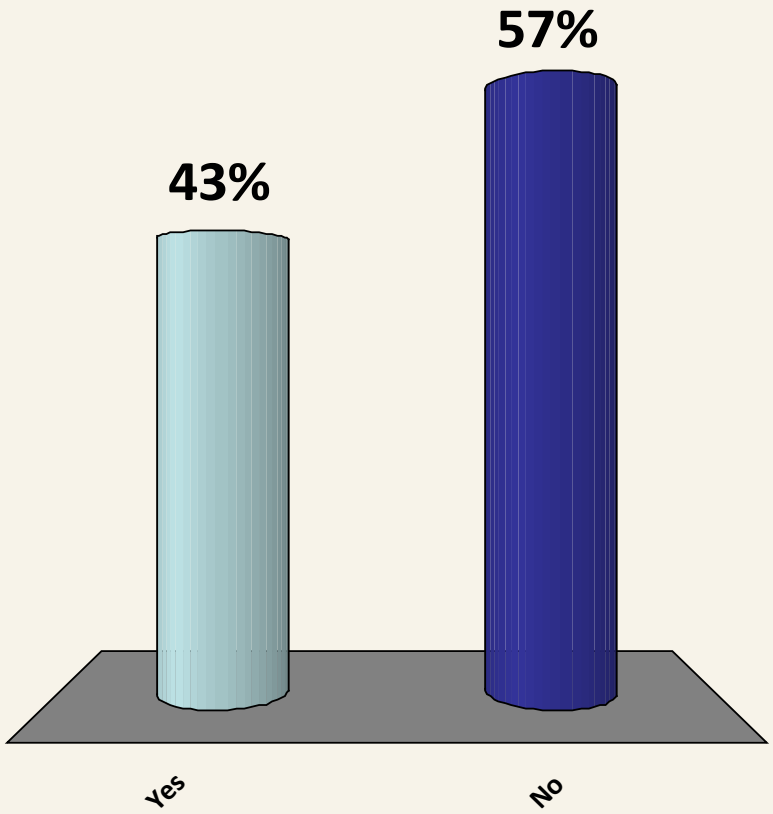
- Large differences in median concentrations of BNP precursors between the two groups (no overlap between the ranges)

	Patients known to have disorder	Normal controls
Median (range) concentration of BNP precursor (pg/ml)	493.5 (248.9-909.0)	129.4 (53.6-159.7)

- Testing for BNP concentration was "a useful diagnostic aid for left ventricular dysfunction."

Should the test be introduced in practice?

1. Yes
2. No



Threats to the validity of diagnostic studies

- Participants
- What is the right population?
  - One way of conceptualizing the right population is that it includes a broad spectrum of the diseased, from mild to severe.
- Spectrum bias (key concept)
  - Include subjects suspected of having the disease
    - Not known to be diseased or healthy

Should the test be introduced in practice?

- No!
  - compares test results of groups of patients who already have established diagnoses
    - rather than patients who are merely suspected of the target disorder
    - contrasts an extreme group of normal people with a group with severe disease
  - Tells us whether the test shows diagnostic promise under ideal conditions.

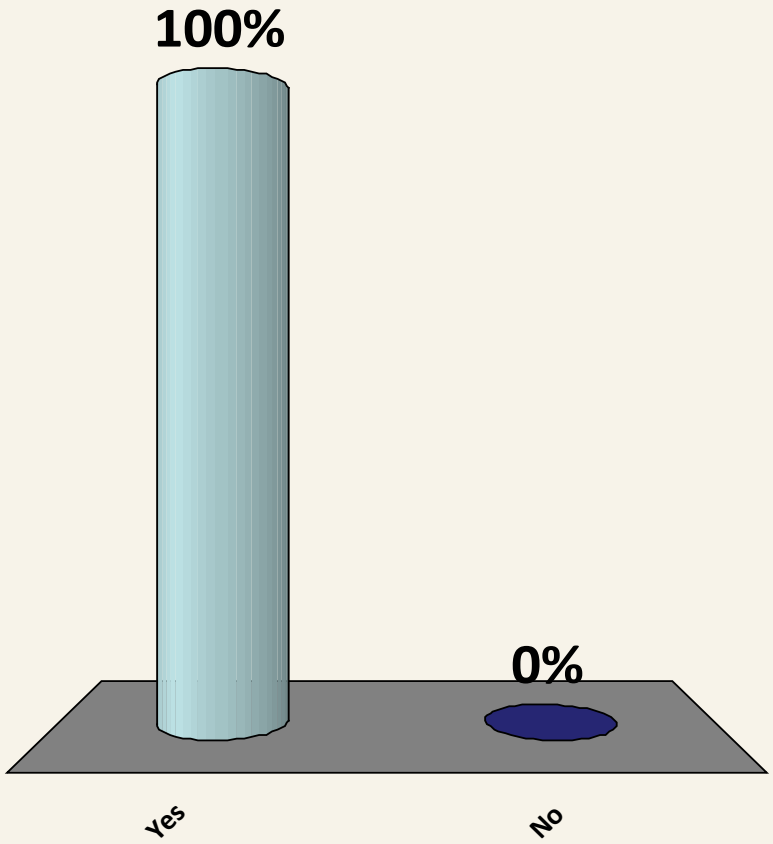
Example

- Among patients suspected of LVD, does the concentration of BNP distinguish patients with and without LVD?
- All subjects underwent confirmation of LVD using a reference standard of Echocardiography.
- Results

	Patients with LVD on echocardiography	Patients with normal results on echocardiography
Concentration of BNP		
High (>17.9 pg/ml)	35	5
Normal (<18 pg/ml)	5	29

Should the test be introduced in practice?

1. Yes
2. No



What changed?

- Referred patients (n=126) underwent independent, blind BNP measurements and echocardiography (gold standard).
- Conclusions
  - measurements of BNP concentration did not look nearly as promising when tested in the real world setting of routine clinical practice
  - introducing routine measurement [of BNP] would be unlikely to improve the diagnosis of symptomatic [left ventricular dysfunction] in the community

Threats to the validity of diagnostic studies

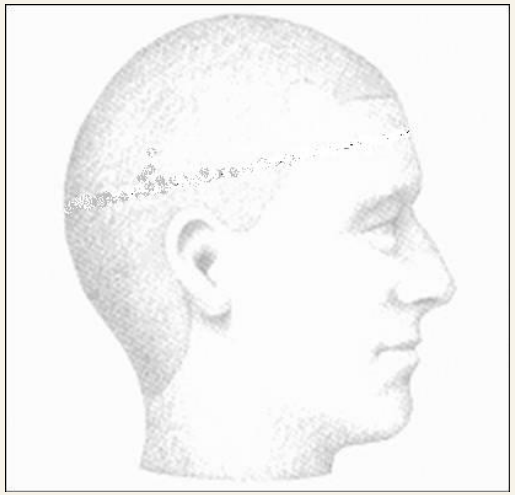
- Independent, blind comparison with a gold standard of diagnosis?
  - all study patients have undergone both the diagnostic test and the reference ("gold") standard evaluation
    - reference standard has been applied regardless of the result of the diagnostic test
    - When patients have a negative test result, Investigators may be tempted to forego applying the reference standard. Particularly when the standard is risky or invasive.
  - Blind
    - reference standard has been applied and interpreted in total ignorance of the diagnostic test result, and vice versa

Measurements blind subjective? OR objective?

- Investigators generally ensure that the individual interpreting the test does so without any information about the patient.
- To evaluate the reproducibility of a measurement technique the observers must be unaware of their previous measurement(s) on the same individual.
  - Example
    - When physicians are told of prior cardiac complications they begin to hear murmurs on auscultation

Figure 1: The authors: double blinded versus single blinded

Reproducible



Reproducibility (key concept)

- In pediatric practice following meningitis, a head circumference that increases by 7mm in a day will result in urgent head imaging
- In obstetrics measurements of the fundal height can vary by up to 5cm
  - the difference between having a baby delivered early due to IUGR or not when opposite occur
- The question is can you reproduce the test in your setting and will it perform as well in your setting

One measure of reproducibility is Kappa (key concept)

Kappa Value	Strength of agreement
<0.20	Poor
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Good
0.81 – 1.00	Very good

Interpreting Diagnostic Studies

USF HEALTH CENTER FOR EVIDENCE BASED MEDICINE

Key concepts Interpretation

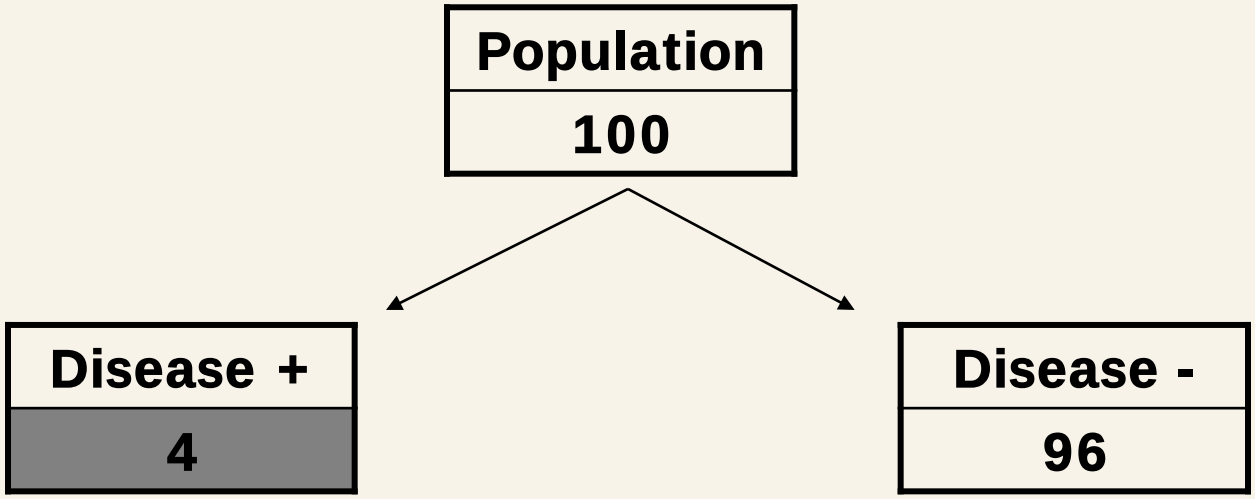
- Prevalence/Pre-test probability
- Sensitivity & Specificity
- Negative and positive predictive value
- Likelihood ratios
- Post-test probability

Natural frequencies provide a more graphic, easy to understand way to portray probabilities for both physicians and patients.

METHOD 1: NATURAL FREQUENCIES TREE

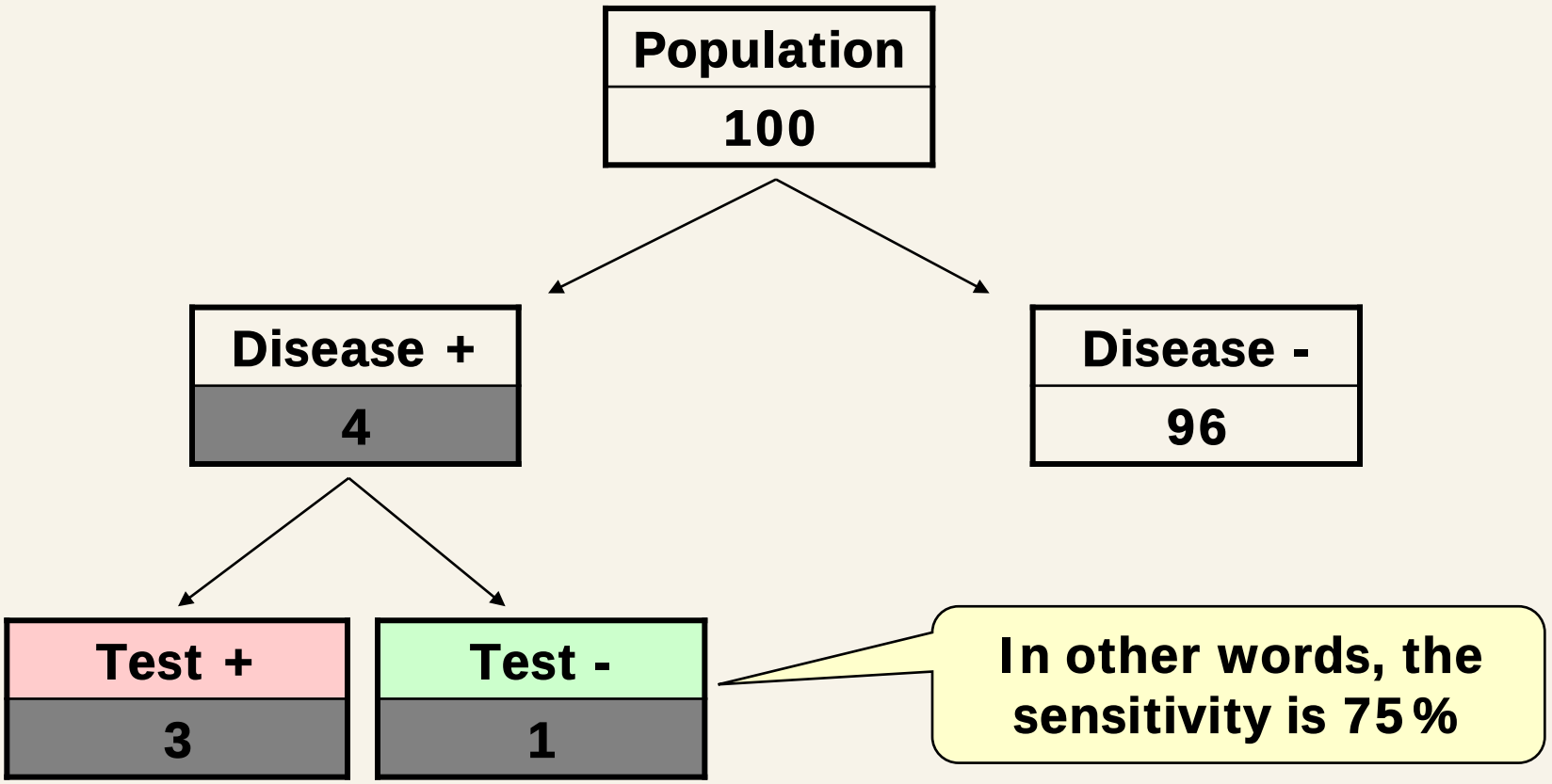
Population
100

IN EVERY 100 PEOPLE, 4 WILL HAVE THE DISEASE

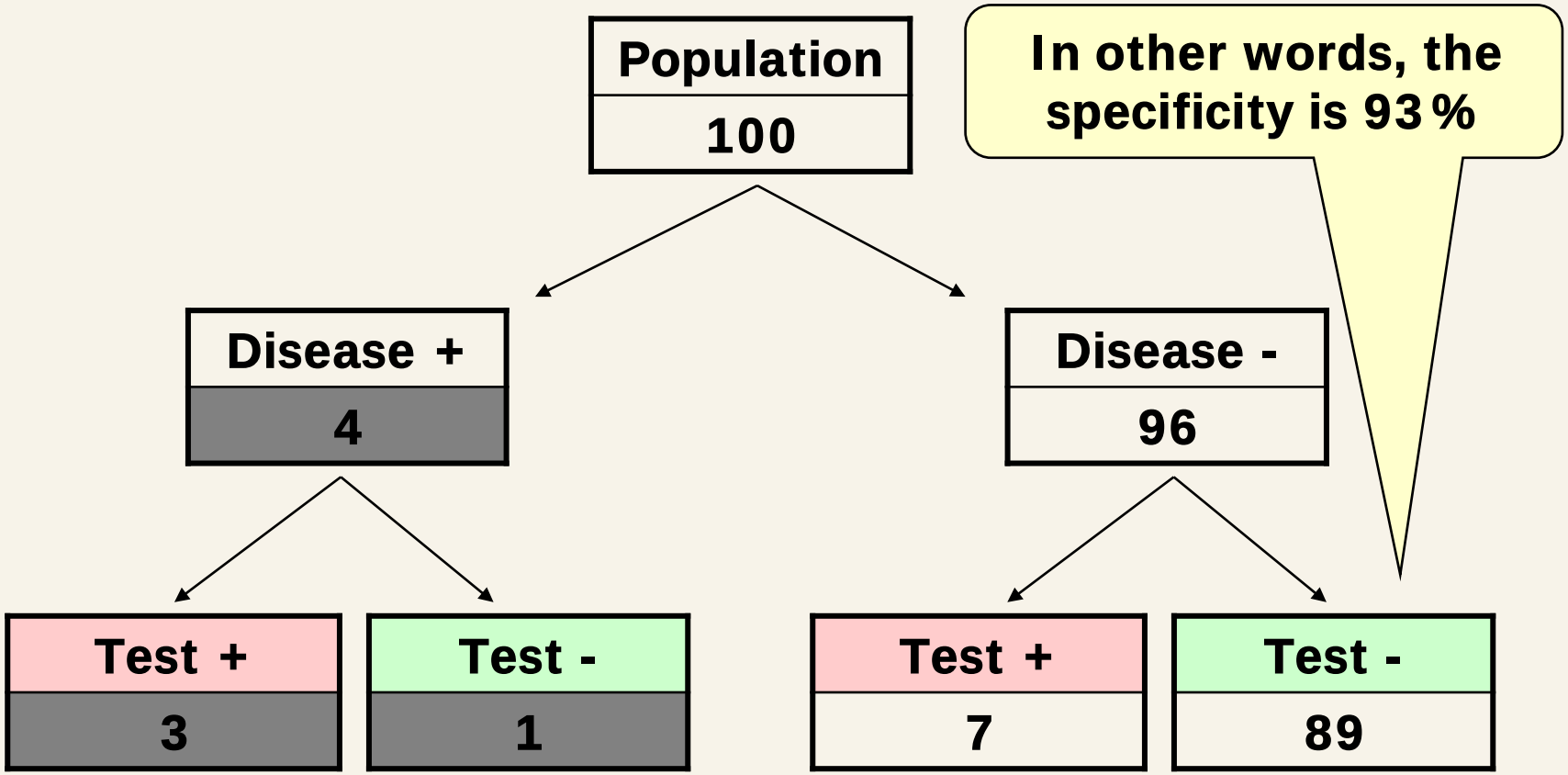


If these 100 people are representative of the population at risk, the assessed rate of those with the disease (4%) represents the PREVALENCE of the disease – it can also be considered the PRE-TEST PROBABILITY of having the disease

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

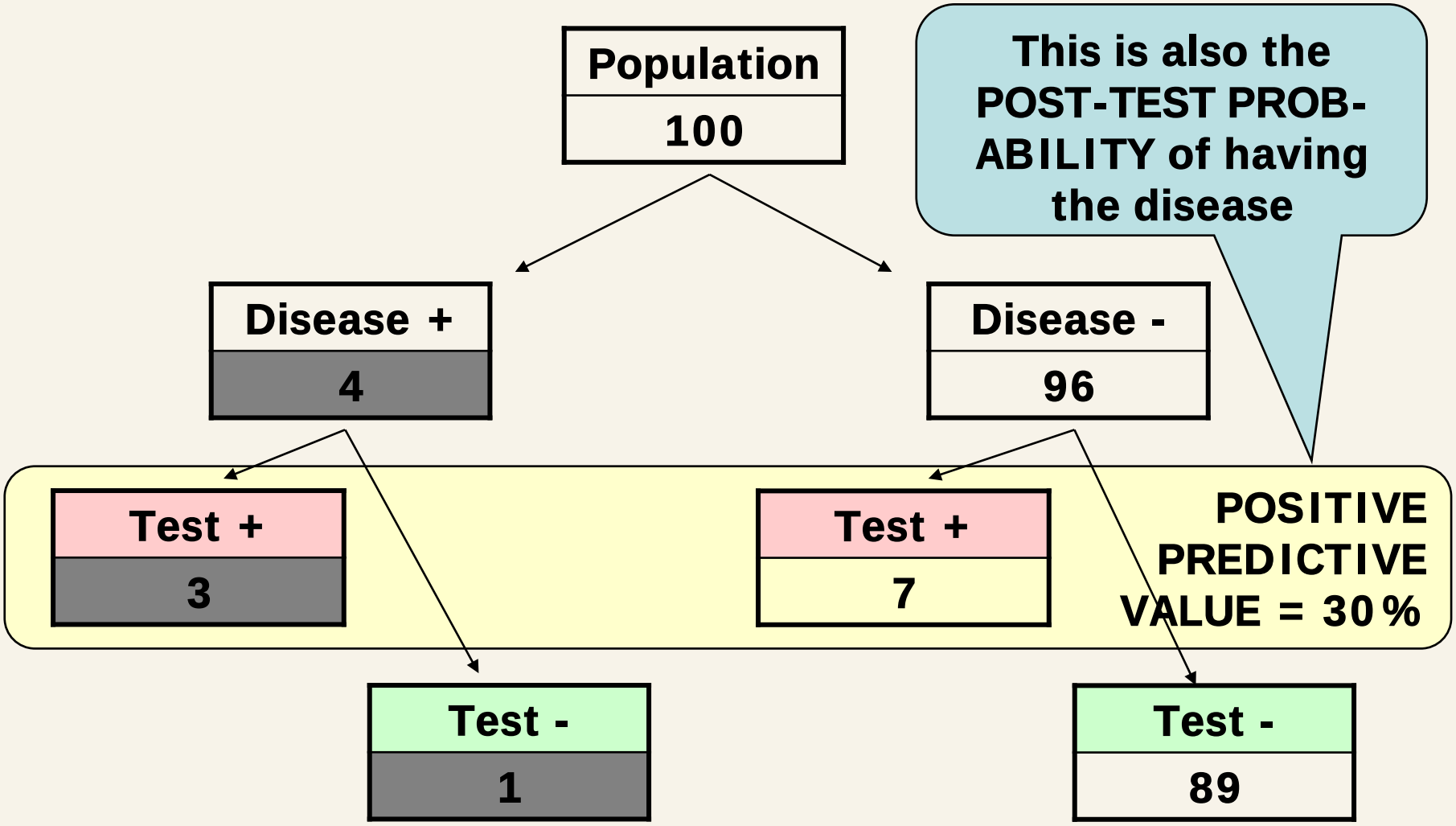


AMONG THE 96 PEOPLE WITHOUT THE DISEASE, 7 WILL TEST POSITIVE

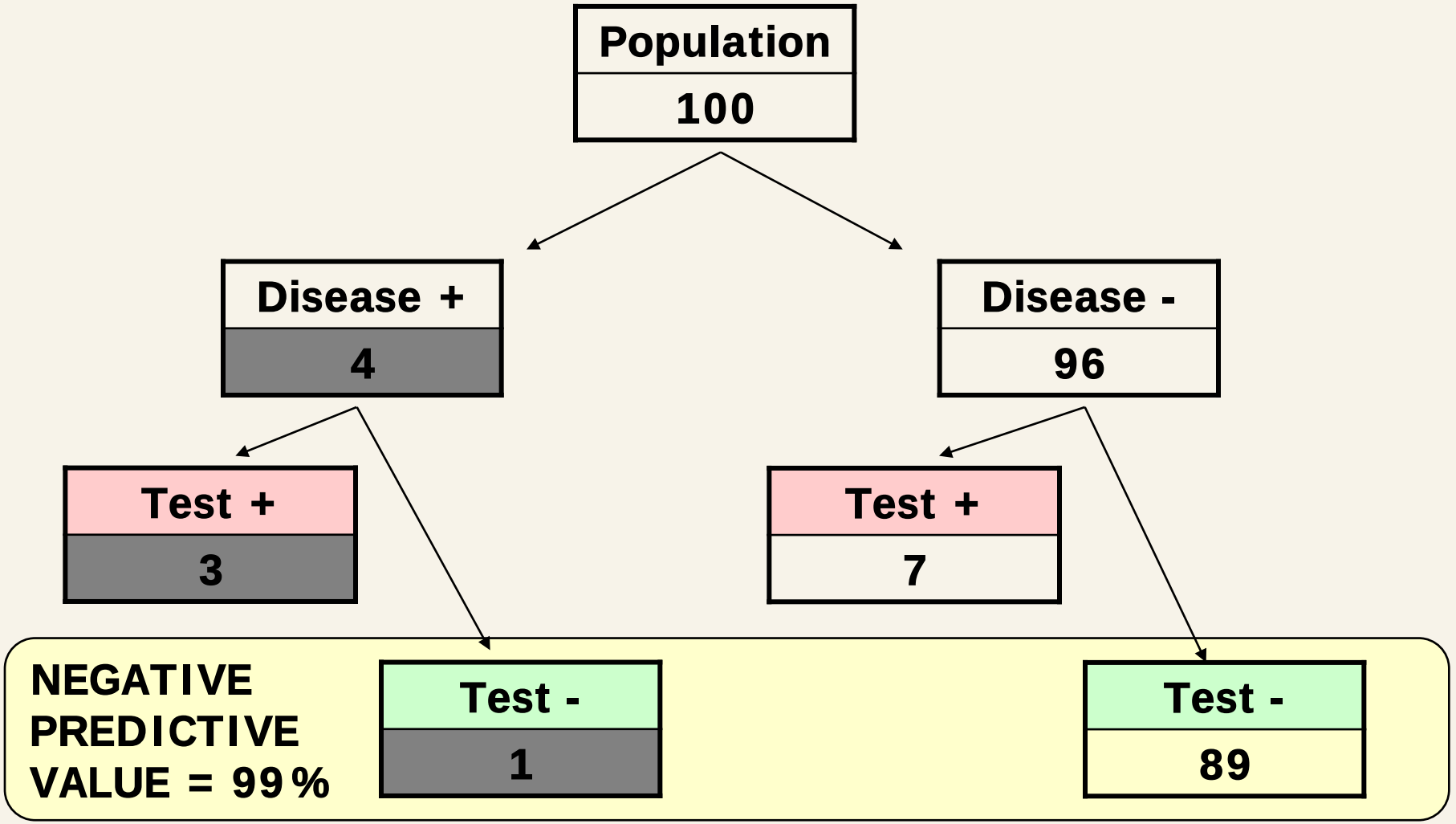


USF HEALTH CENTER FOR EVIDENCE BASED MEDICINE

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

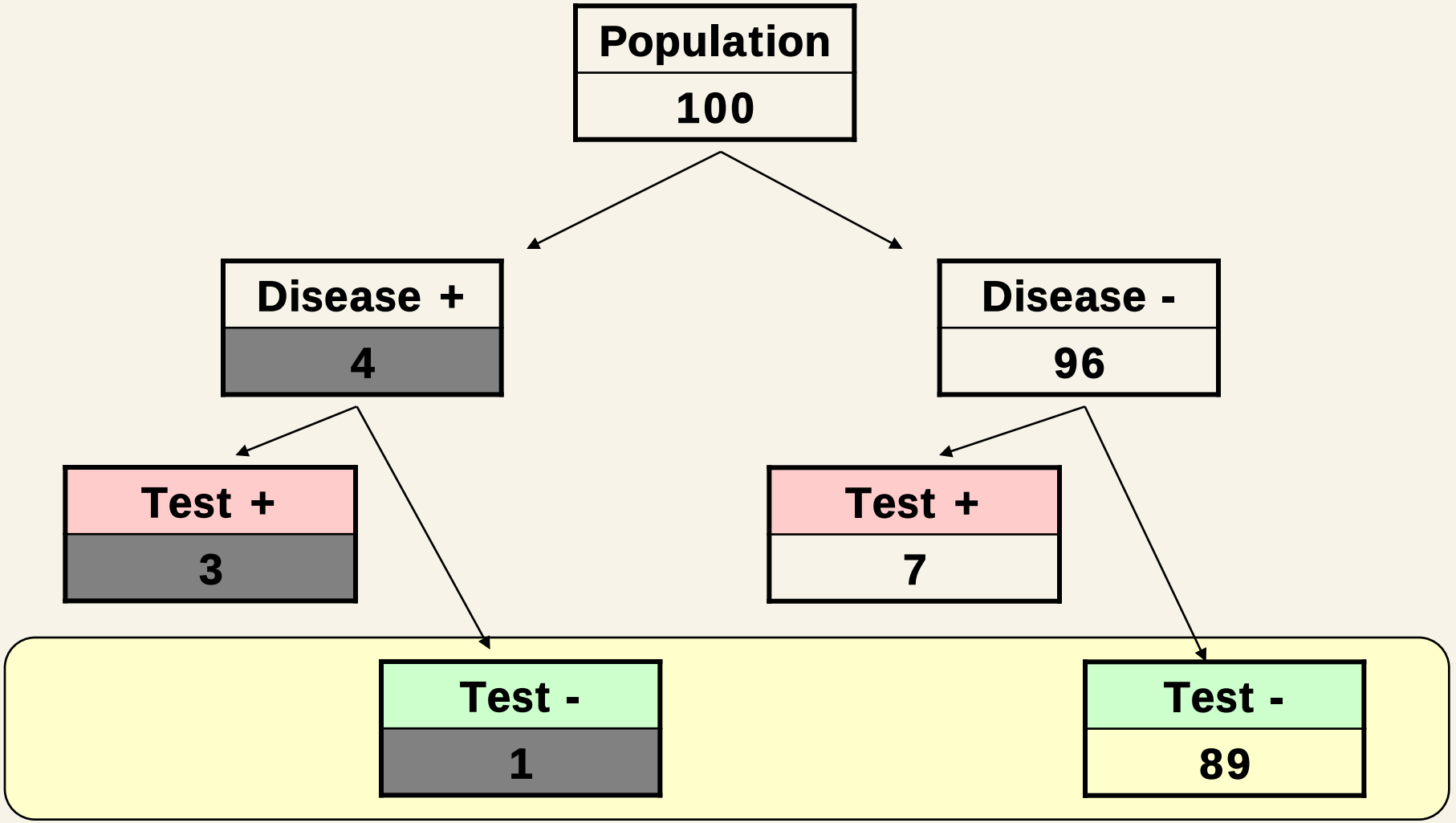


AMONG THOSE WHO TEST NEGATIVE, 89 OF 90 WILL NOT HAVE THE DISEASE



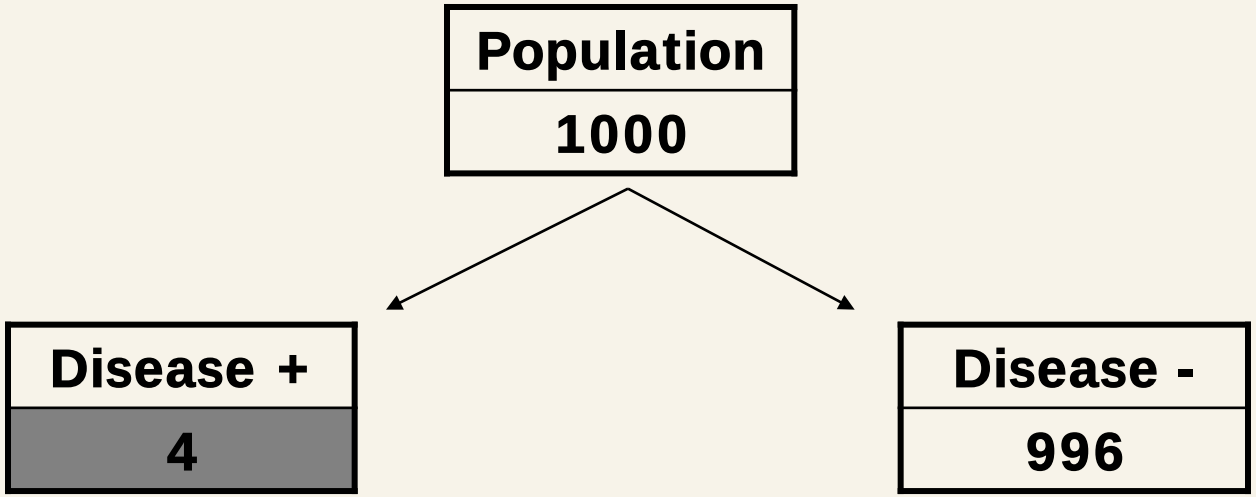
USF HEALTH CENTER FOR EVIDENCE BASED MEDICINE

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



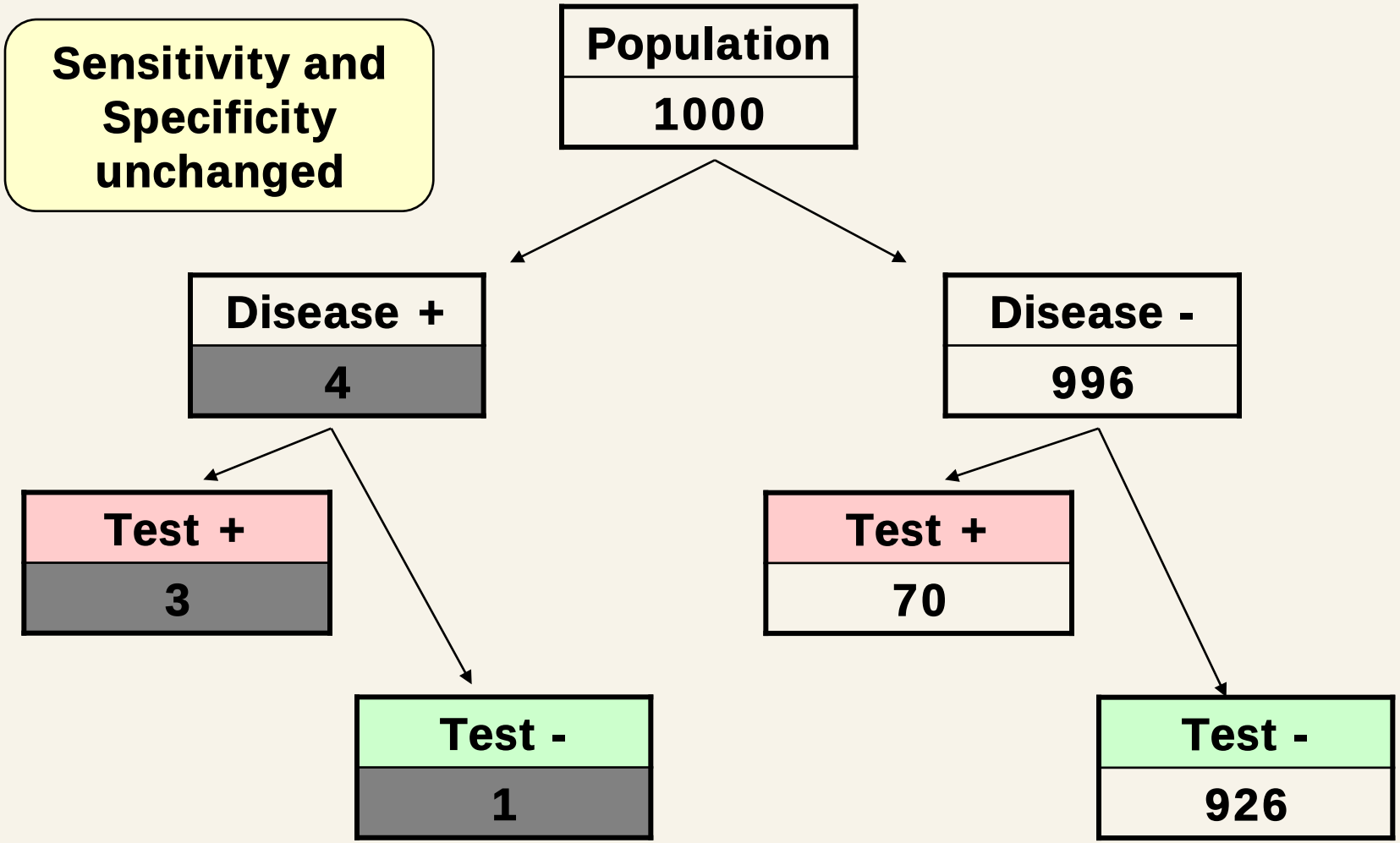
USF HEALTH CENTER FOR EVIDENCE BASED MEDICINE

PREDICTIVE VALUES AND CHANGING PREVALENCE

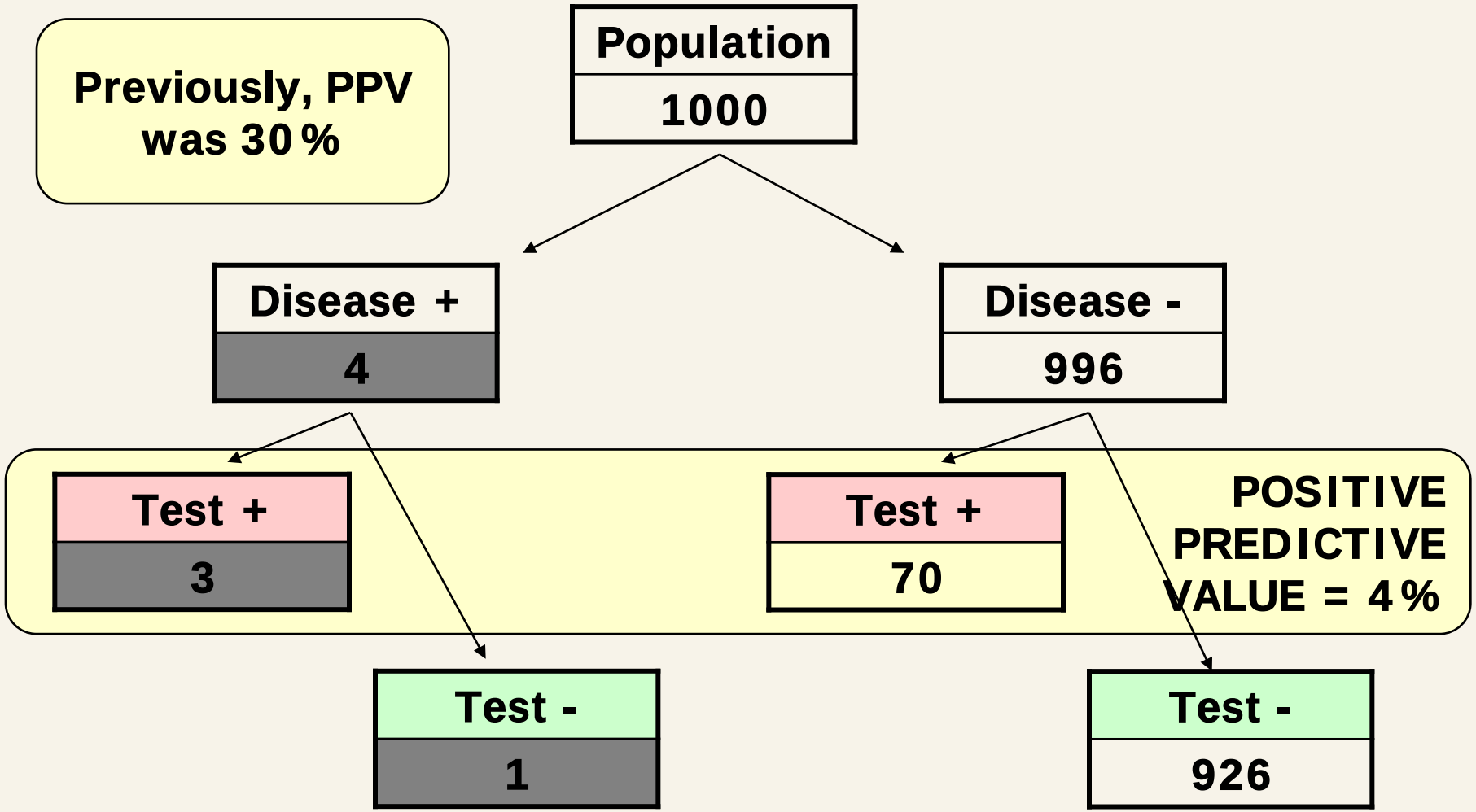


Prevalence reduced by an order of magnitude from 4% to 0.4%

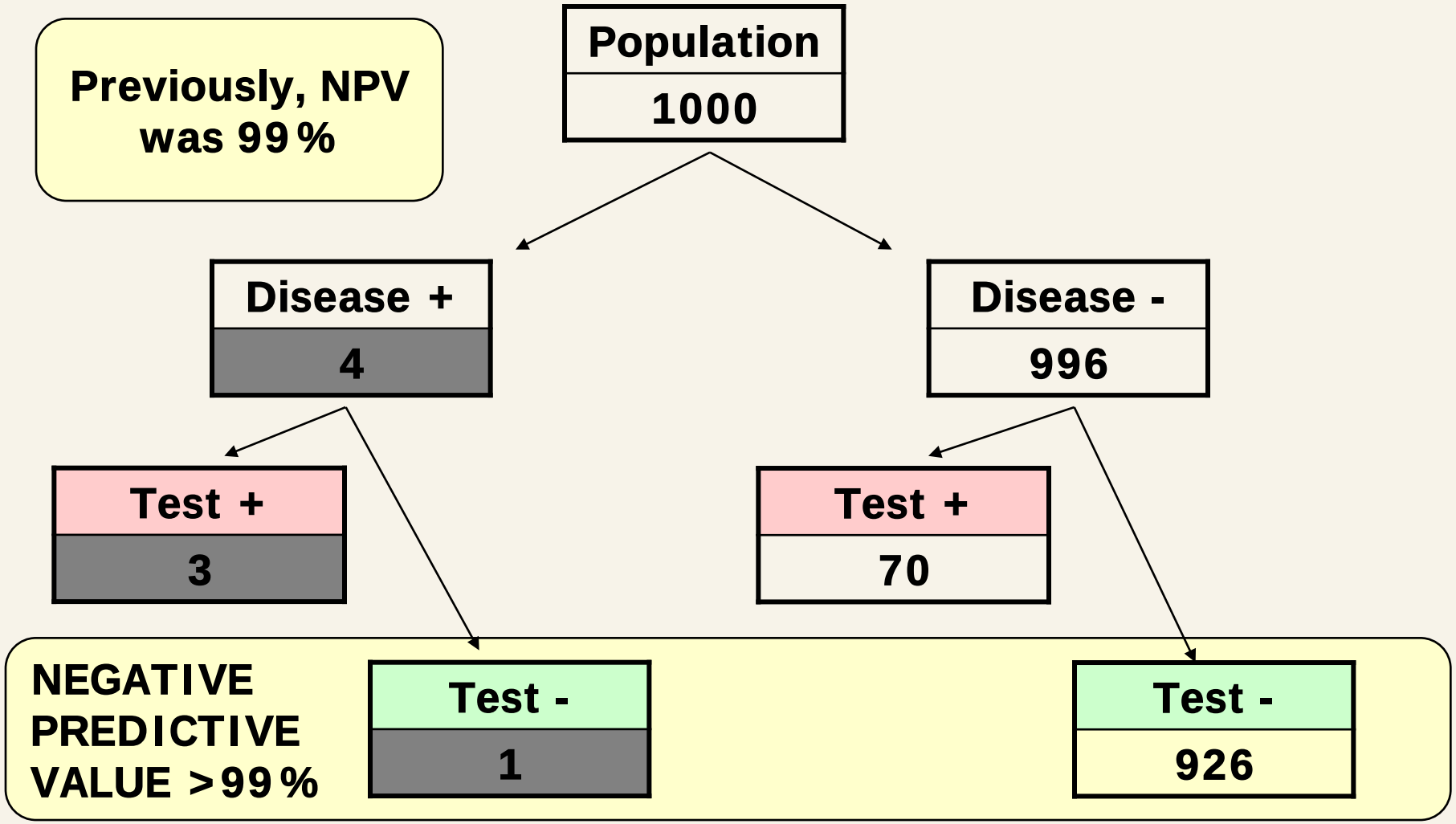
PREDICTIVE VALUE AND CHANGING PREVALENCE



POSITIVE PREDICTIVE VALUE AT LOW PREVALENCE



NEGATIVE PREDICTIVE VALUE AT LOW PREVALENCE



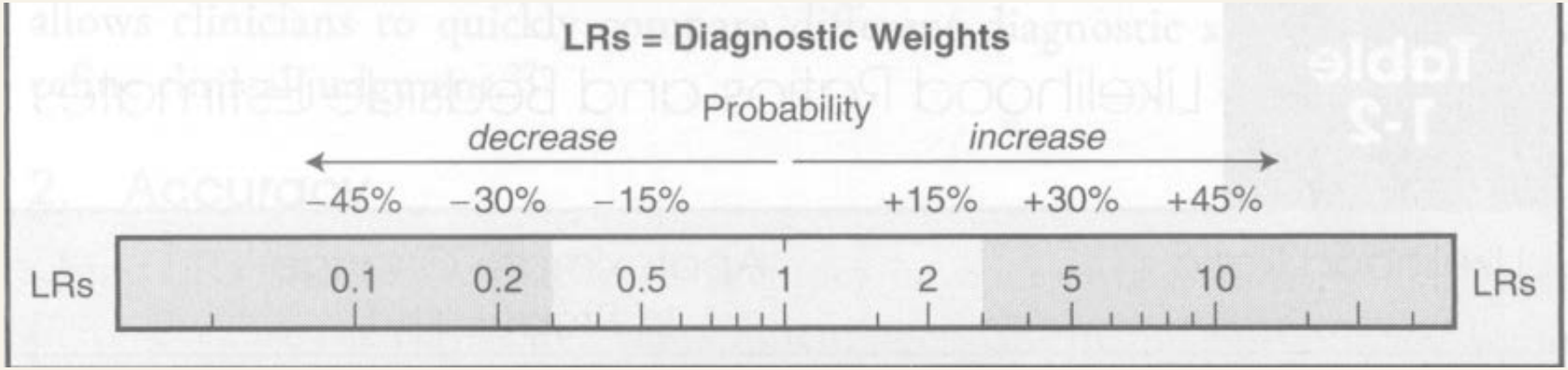
SENSITIVITY, SPECIFICITY AND PREDICTIVE VALUES (key concepts)

- For sensitivity and specificity, the reference variable (‘denominator) is the DISEASE
- For predictive value, the reference variable (‘denominator’) is the TEST

PREDICTION OF LOW PREVALENCE EVENTS (key concepts)

- Even highly specific tests, when applied to low prevalence events, yield a high number of false positive results
- Because of this, under such circumstances, the Positive Predictive Value of a test is low
- However, this has much less influence on the Negative Predictive Value

What do likelihood ratios mean?



LR<0.1 = strong negative test result

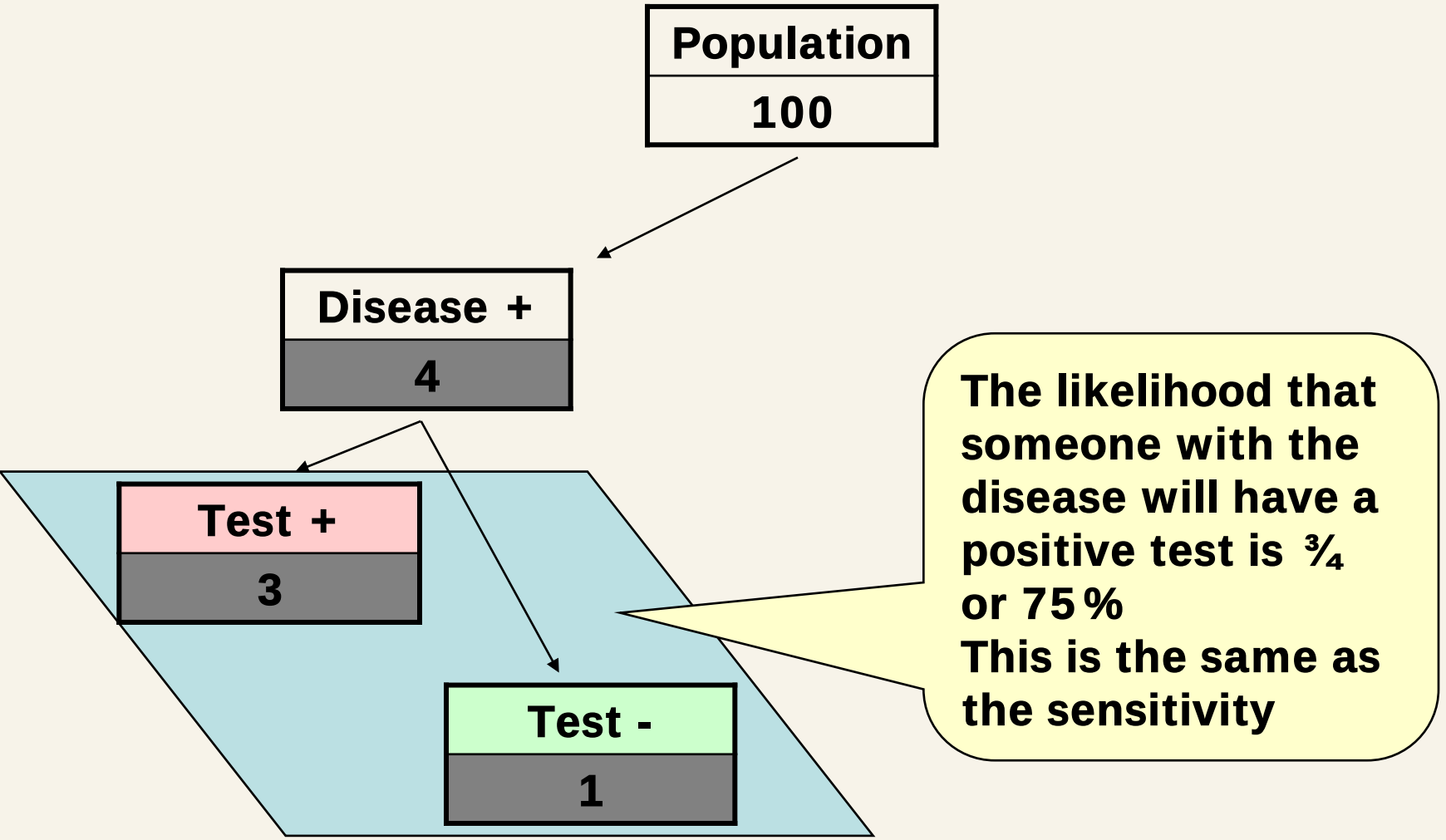
LR=1

No diagnostic value

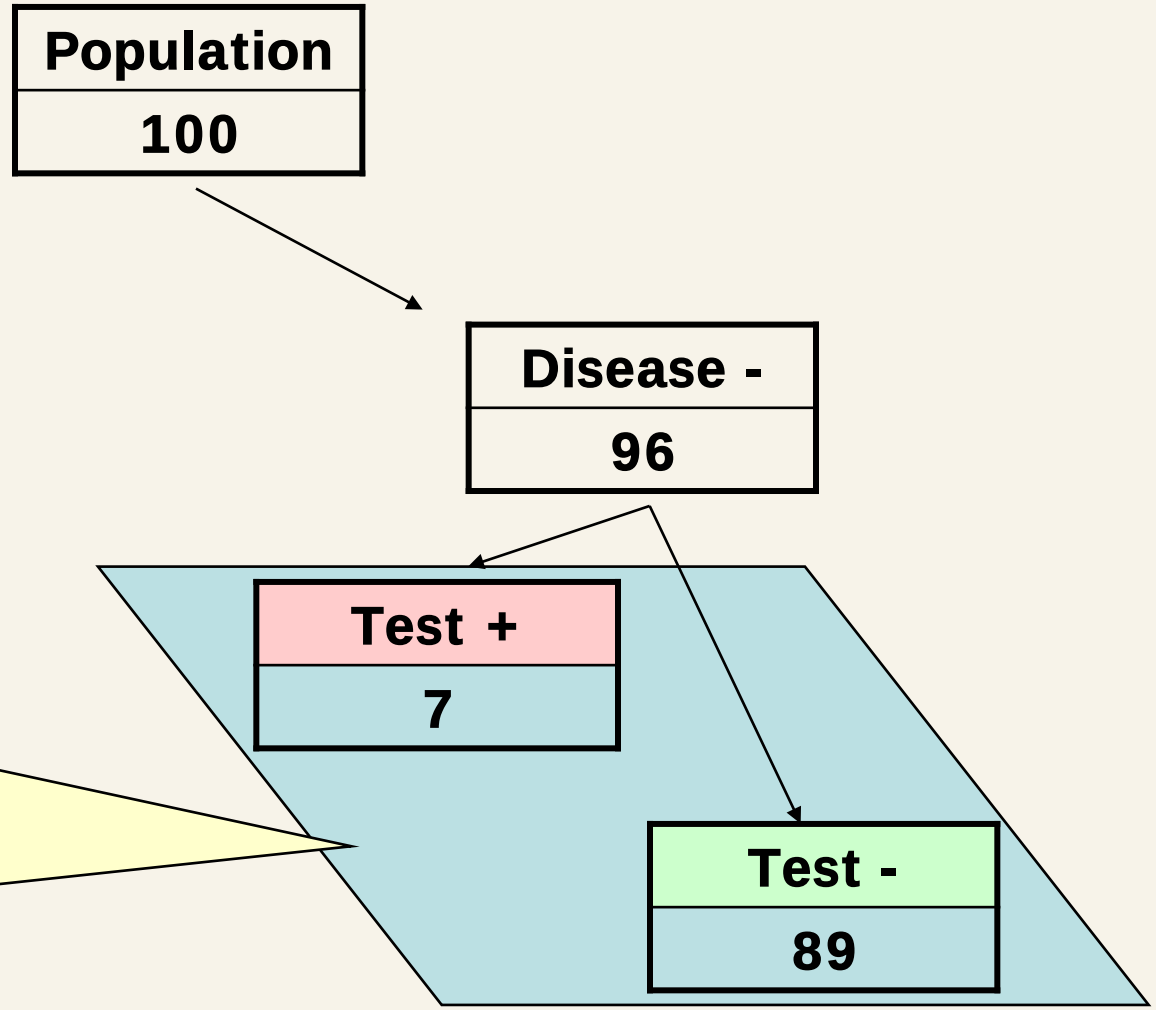
LR>10 = strong positive test result

USF HEALTH CENTER FOR EVIDENCE BASED MEDICINE

LIKELIHOOD


Population
100
Disease +
4
Test +
3
Test -
1
The likelihood that someone with the disease will have a positive test is ¾ or 75%
This is the same as the sensitivity

LIKELIHOOD II



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

$$\text{LIKELIHOOD RATIO} = \frac{\text{LIKELIHOOD OF POSITIVE TEST GIVEN THE DISEASE}}{\text{LIKELIHOOD OF POSITIVE TEST IN THE ABSENCE OF THE DISEASE}}$$

$$= \frac{\text{SENSITIVITY}}{\text{1- SPECIFICITY}} = \frac{0.75}{0.07} = 10.7$$

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

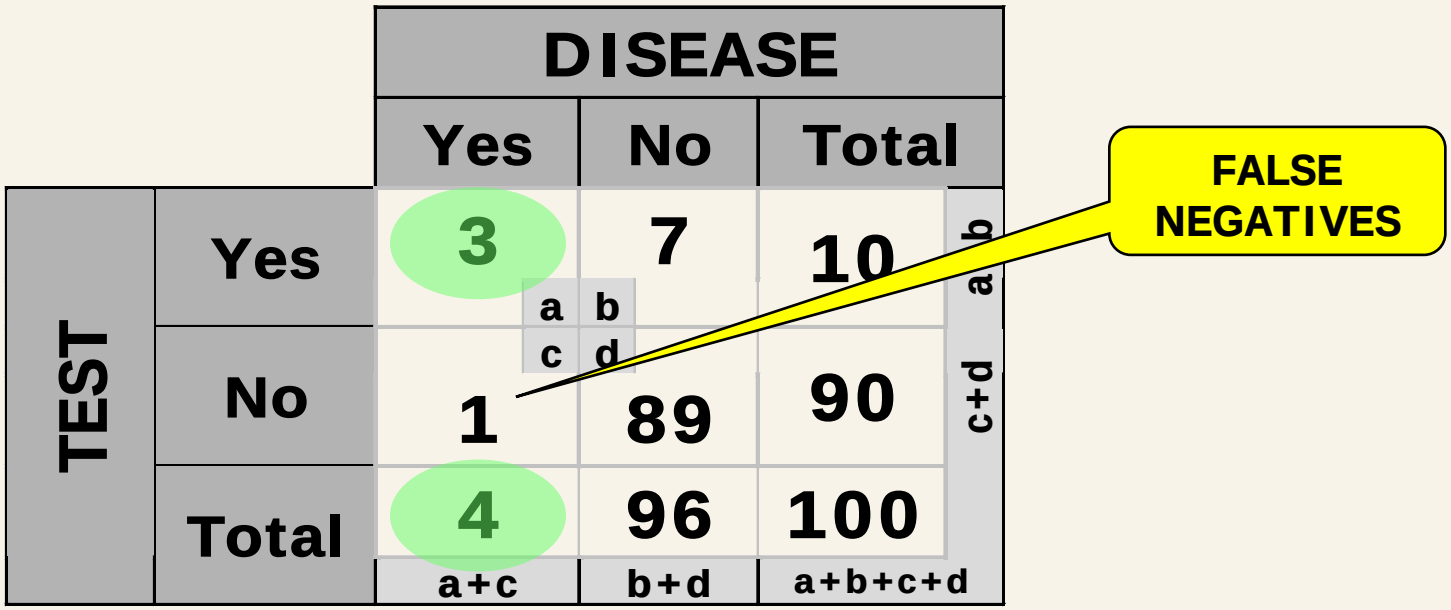
The higher the Likelihood Ratio, the better the test (other factors being equal)

USF HEALTH CENTER FOR EVIDENCE BASED MEDICINE

METHOD 3: 'TRADITIONAL' 2x2 TABLES

		DISEASE		
		Yes	No	Total
TEST	Yes	3 (a)	7 (b)	10 (a+b)
	No	1 (c)	89 (d)	90 (c+d)
	Total	4 (a+c)	96 (b+d)	100 (a+b+c+d)

SENSITIVITY

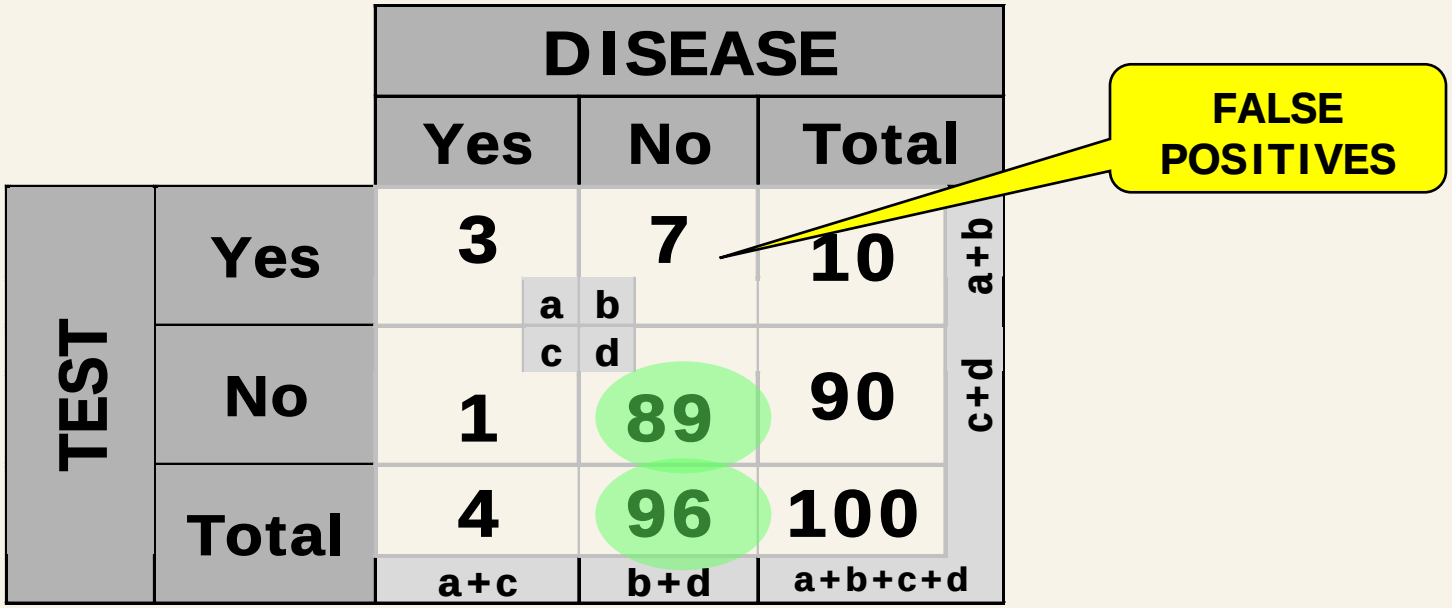


SENSITIVITY (SnOUT)

The proportion of people with the diagnosis (N=4) who are correctly identified (N=3)

Sensitivity = a/(a+c) = 3/4 = 75%

SPECIFICITY



SPECIFICITY (SpIN)

The proportion of people without the diagnosis (N=96) who are correctly identified (N=89)

Specificity = d/(b+d) = 89/96 = 93%

PRE-TEST ODDS

		DISEASE		
		Yes	No	Total
TEST	Yes	3 (a)	7 (b)	10 (a+b)
	No	1 (c)	89 (d)	90 (c+d)
	Total	4 (a+c)	96 (b+d)	100 (a+b+c+d)

In the sample as a whole, the odds of having the disease are 4 to 96 or 4% (the PRE-TEST ODDS)

POST-TEST ODDS

		DISEASE			
		Yes	No	Total	
TEST	Yes	3 (a)	7 (b)	10	a+b
	No	1 (c)	89 (d)	90	c+d
	Total	4	96	100	
		a+c	b+d	a+b+c+d	

In the sample as a whole, the odds of having the disease are 4 to 96 or 4% (the PRE-TEST ODDS)

In those who score positive on the test, the odds of having the disease are 3 to 7 or 43% (the POST-TEST ODDS)

POST-TEST ODDS

		DISEASE			
		Yes	No	Total	
TEST	Yes	3 (a)	7 (b)	10	a+b
	No	1 (c)	89 (d)	90	c+d
	Total	4	96	100	
		a+c	b+d	a+b+c+d	

In the sample as a whole, the odds of having the disease are 4 to 96 or 4% (the PRE-TEST ODDS)

In those who score positive on the test, the odds of having the disease are 3 to 7 or 43% (the POST-TEST ODDS)

In those who score negative on the test, the odds of having the disease are 1 to 89 or approximately 1%

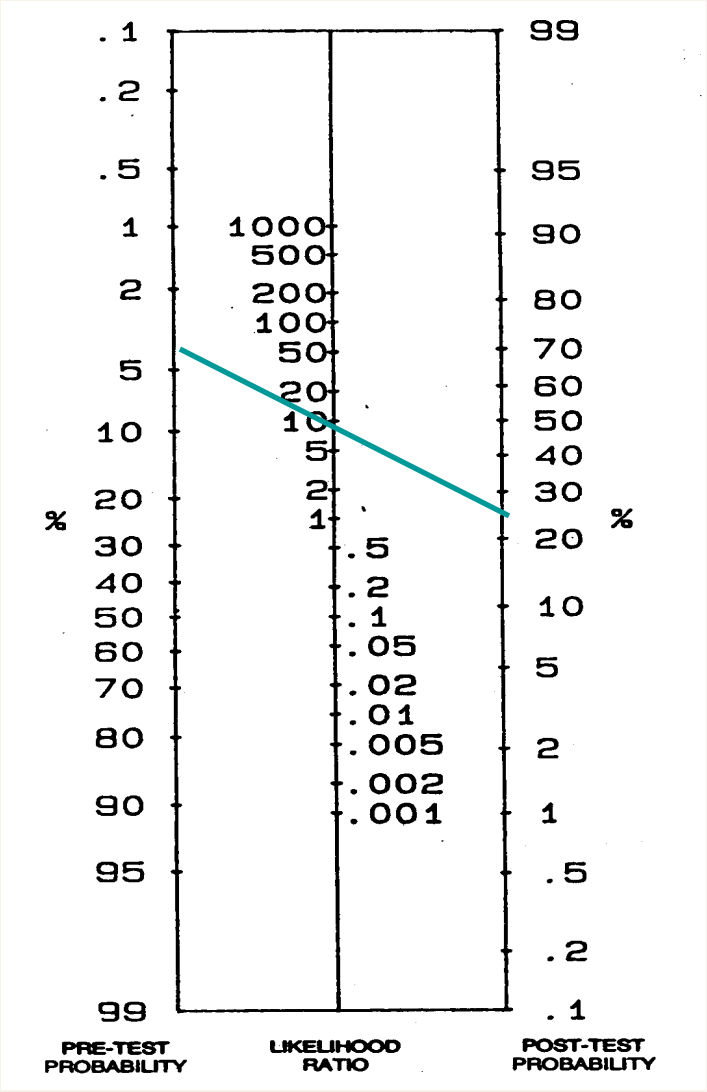
USF HEALTH CENTER FOR EVIDENCE BASED MEDICINE

BAYES THEOREM

POST-TEST ODDS =

LIKELIHOOD RATIO x PRE-TEST ODDS

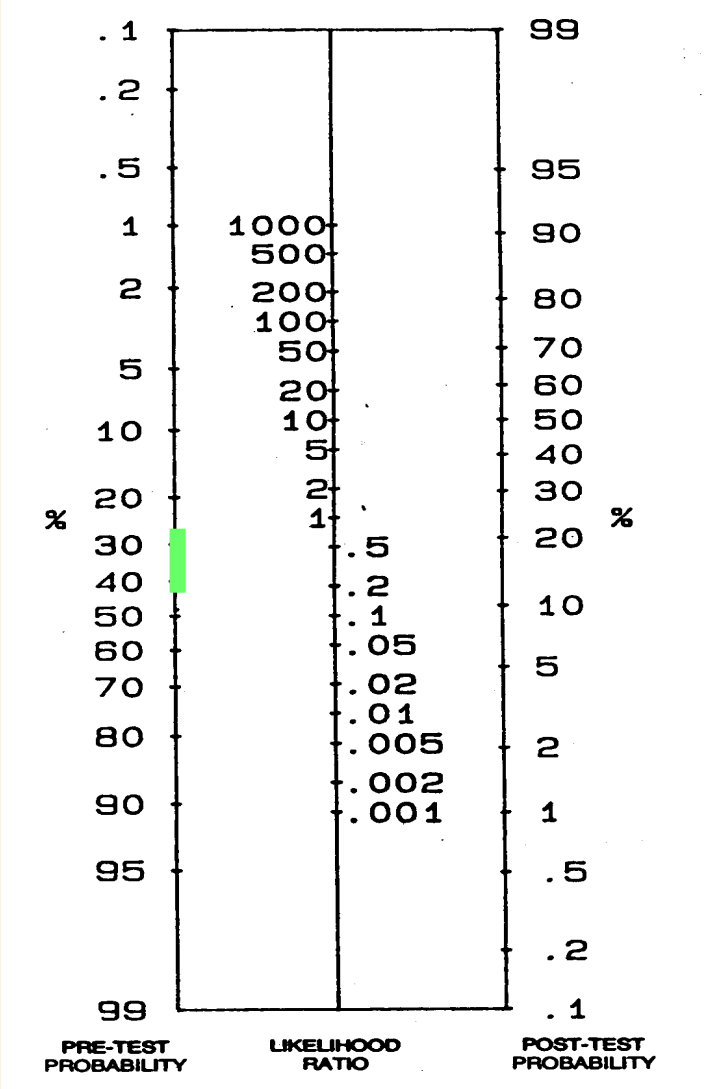
LIKELIHOOD RATIO AND PRE- AND POST-TEST PROBABILITIES



For a given test with a given likelihood ratio, the post-test probability will depend on the pre-test probability (that is, the prevalence of the condition in the sample being assessed)

USF HEALTH CENTER FOR EVIDENCE BASED MEDICINE

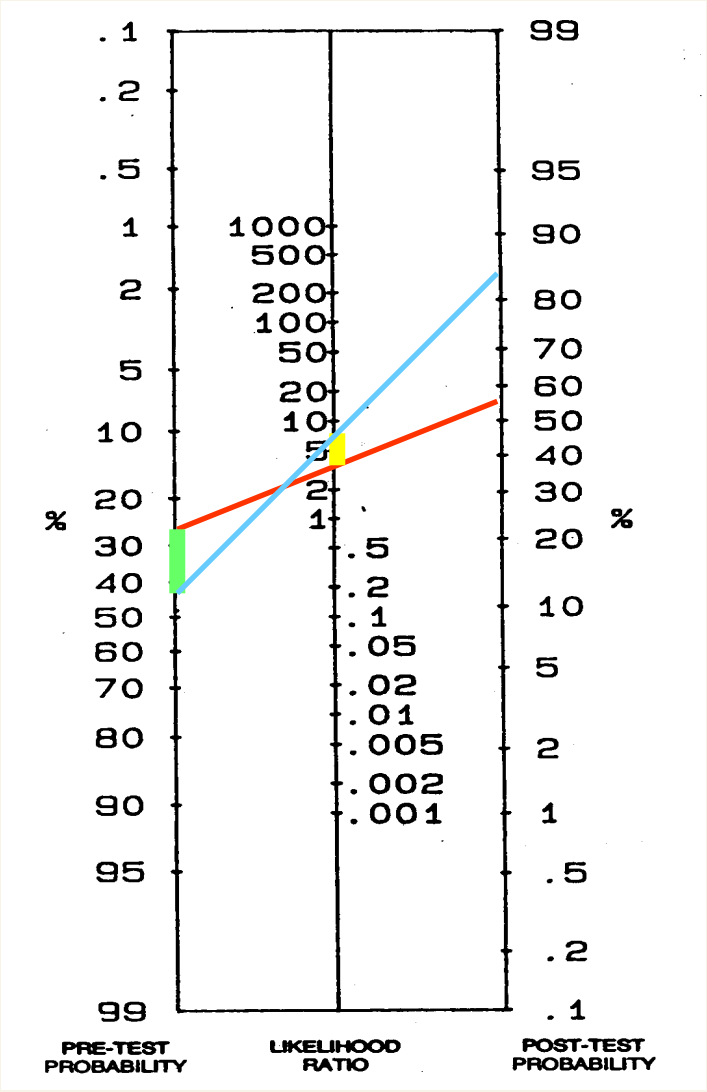
SENSITIVITY ANALYSIS OF A DIAGNOSTIC TEST



	Value	95% CI
Pre-test probability	**35%**	**26% to 44%**

USF HEALTH CENTER FOR EVIDENCE BASED MEDICINE

SENSITIVITY ANALYSIS OF A DIAGNOSTIC TEST



	Value	95% CI
Pre-test probability	35%	26% to 44%
Likelihood ratio	5.0	3.0 to 8.5

Applying the 95% confidence intervals above to the nomogram, the post-test probability is likely to lie in the range 55-85%

What was all this about?

FOB screening tests example

You find out that one of your patient undertook the FOB test and has a positive result. He/she asks…what are the chances of having cancer?

- Prevalence of disease is 10% must be diagnosed.
- Sensitivity of 50%
- False positive rate 3%.

62%

Doctors with an average of 14 yrs experience
Answers ranged from 1% to 99%
half of them estimating the probability as 50%
Gigerenzer G BMJ 2003;327:741-744

Explanation- 3 step approach

Information provided

Prevalence 10%

Sensitivity 50%

False positive rate 3%

Sample of patients (N=100)

Prevalence/pretest probability

Step 1

Probability of disease **10**

Probability of no disease **90**

Step 2

Sensitivity

50% of subjects who have the disease (prevalence) will test positive **5**

False positive rate

3% of the subjects who do not have the disease will incorrectly test positive **2.7**

Specificity 97%

No of positives all together (regardless of false or true) **8**

Step 3

Of the positives (all) how many will truly have the disease?

5÷8= **0.625 or 62%**

Summary

- Diagnostic studies should match methods to diagnostic questions
  - Do test results in affected patients differ from those in normal individuals?
  - Are patients with certain test results more likely to have the target disorder?
- Do test results distinguish patients with and without the target disorder among those in whom it is clinically sensible to suspect the disorder?
- Diagnostic test accuracy
  - Cross-sectional study
- Diagnostic test linked with treatment/outcome
  - RCTs

Summary

- Key concepts

  Pretest Probability (Prevalence)

  – The probability of the target condition being present before the results of a diagnostic test are known.

  Posttest Probability

  – The probability of the target condition being present after the results of a diagnostic test are known.

Summary

- Key concepts

  Reference Standard

  – You cannot decide if a test works unless you find or create a reference standard: the best available way of determining whether what you are looking for is present.

Summary

- Key concepts

  Likelihood Ratio

  – If you are trying to determine how well a test performs, find how much a given result will shift the belief (likelihood) that the problem exists.

Summary

- Key concepts

  Why use reference standards and LRs?

  – If you are trying to decide whether a test is worth the trouble, think about what your “ignore” and “act” thresholds are and if the test moves you from uncertainty into either zone.

Summary

- Key concepts

  Accurate ≠ Useful

  – Whether a test is useful depends on your perspective and values:

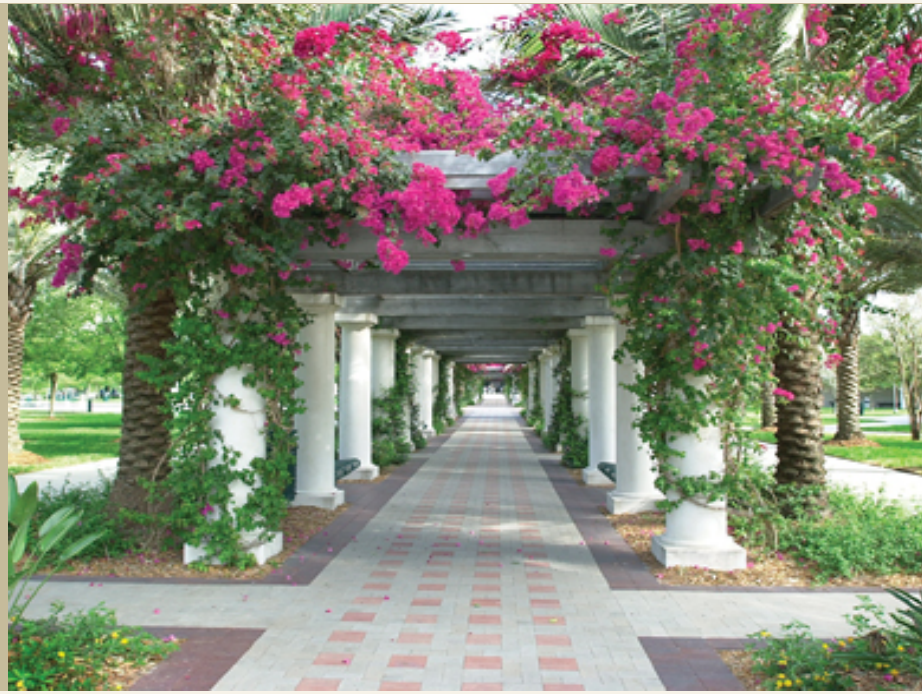
    - What would you do with an accurate negative result?
    - What would you do with an accurate positive result?

Thank you

Questions?

akumar1@health.usf.edu
813-396-9194
MDC 3122

Critical appraisal: Diagnostic study



Ambuj Kumar, MD, MPH

Associate Professor, Director

Division and Center for Evidence Based Medicine, Dept of Internal Medicine, College of Medicine, University of South Florida

Tampa, Florida USA

USF HEALTH CENTER FOR EVIDENCE BASED MEDICINE

Objective

- To learn how to apply the results of studies on diagnostic tests to clinical practice
  - Determine the validity of the study
  - Interpret the results
  - Apply the results to our patients

Assignment

- Read the clinical scenario, and then formulate a clinical question
- Read the article concerning a diagnostic issue
- Critically appraise the paper mentioned above using the critical appraisal questionnaire on diagnosis
- Decide how you would manage the patient described in the following scenario

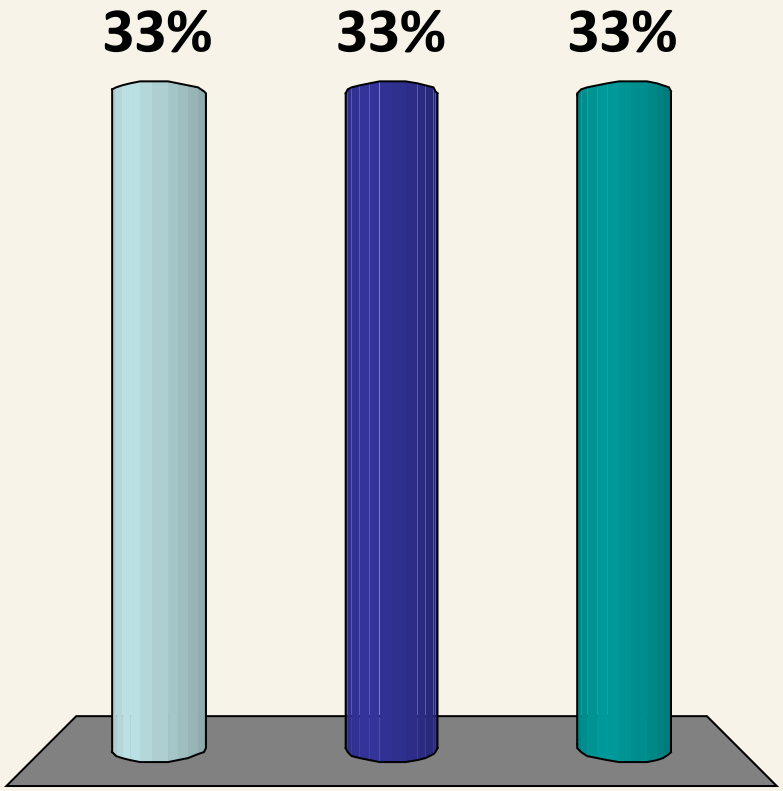
Clinical Scenario

- You are in a busy primary care clinic seeing your last patient of the day, a 70 year old male patient who is complaining of cough and mild shortness of breath, He has a history of hypertension and hypercholesterolemia, and is very active, walking two miles a day- i.e., he's generally in good health. He is currently taking Lipitor 20 mg daily and an ACE inhibitor.
- He has had a cough for two weeks and feels it is getting slightly better but he is frustrated with it not resolving quicker and feels he has Pneumonia. He has no fever but very productive cough, heavy green sputum. He feels tired and slightly short of breath when he exercises.
- He lives alone and is a retired auto mechanic.
- On examination you find the patient is comfortable and has a temperature of 38.0 ^{0} C., a respiratory rate of 15, pulse of 98 (normal sinus rhythm). Oxygen saturation on room air is 93%. Cardiac auscultation reveals a grade II systolic murmur. His lung exam is mostly clear except for slight wheezing heard upon expiration. The rest of his physical examination is completely normal.
- You are unable to get chest X-rays at your clinic and you have a low suspicion for pneumonia but you are concerned because the patient lives alone.

You wonder how accurate your assessment is of this patient's chances of having a pneumonia and in general how accurate history and physical exam is in detecting pneumonia?

Did participating patients present a diagnostic uncertainty?

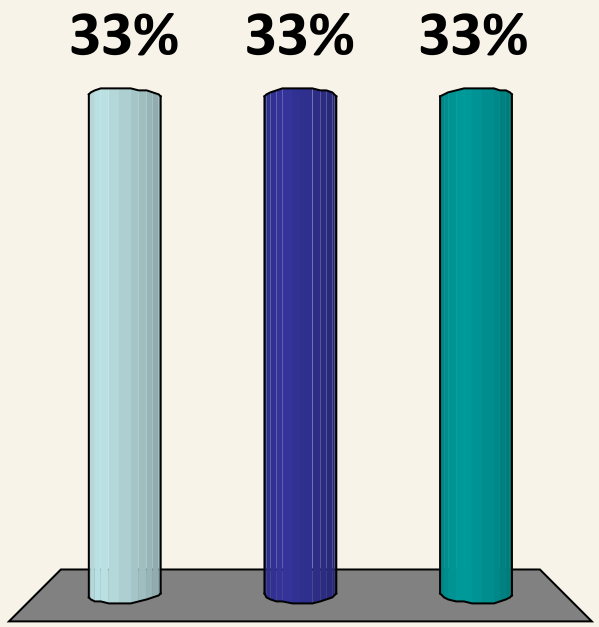
1. Yes
2. No
3. Not sure



USF HEALTH CENTER FOR EVIDENCE BASED MEDICINE

Diagnostic uncertainty is represented by which of the following in the given scenario?

1. Very sick pneumonia patients requiring hospitalization
2. Mildly sick patients not requiring hospitalization
3. Any patient suspected of pneumonia but not having a confirmed diagnosis



USF HEALTH
CENTER FOR EVIDENCE BASED MEDICINE

Are the Results Valid?

- Did participating patients present a diagnostic uncertainty?
  - Study patients presented from a variety of settings with cough and possible pneumonia, and were not sick enough to require hospital admission.

Are the Results Valid?

- Did investigators compare the test to an appropriate independent reference?
  - Investigators compared physicians' clinical diagnosis (prior to CXR) to a radiographic diagnosis. CXR is a reasonable reference standard for pneumonia
  - The final radiologic diagnosis of pneumonia required that both specialists independently determine that the acute CXR had an abnormality consistent with pneumonia and that the convalescent CXR showed some resolution.
  - Ideally if there was any disagreement the final determination would include a consensus of the 2 specialists (Investigators did not report the agreement between these specialists).

Are the Results Valid?

- Were those interpreting the test and reference standard blind to the other results?
  - Yes
- Did investigators perform the same reference standard to all patients regardless of the results of the test under investigation?
  - Yes

What are the Results?

- What likelihood ratios are associated with the range of possible test results?
  - The investigators did not calculate likelihood ratios
  - They did report sensitivity and specificity along with 95% confidence intervals.
  - The likelihood ratio associated with a positive test was 4.6 and the likelihood ratio associated with negative test was 0.3.

What are the results?

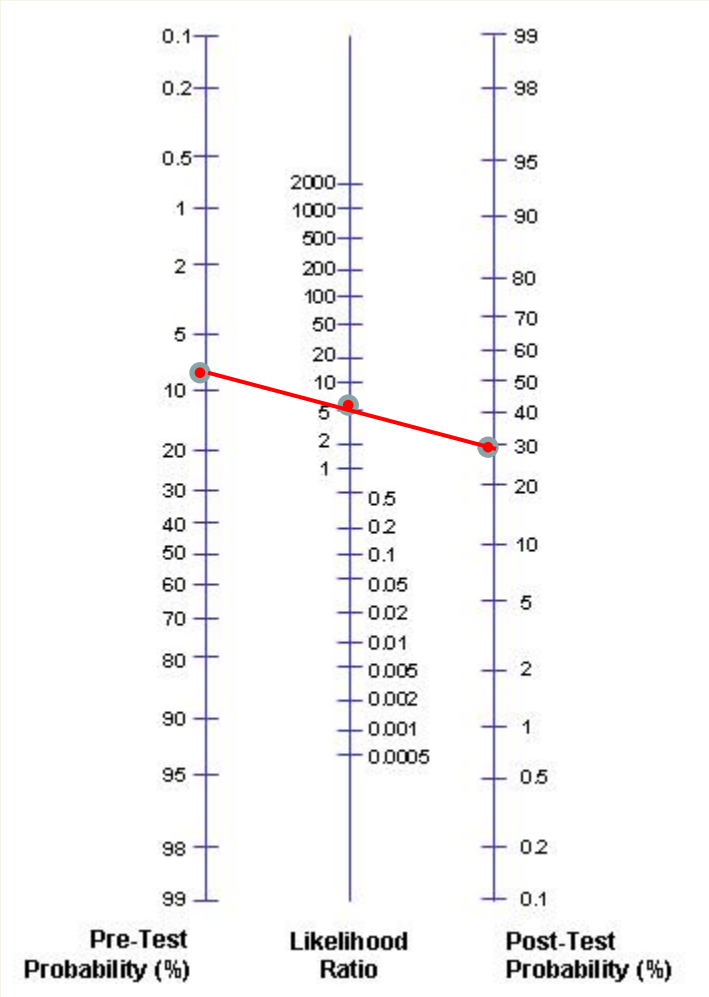
		Chest X-ray		Totals
		Positive	Negative	
Physician Judgment	Positive	14	37	51
		a (TP)	b (FP)	a+b
	Negative	5	194	199
		c (FN)	d (TN)	c+d
Totals		19	231	250
		a+c	b+d	a+b+c+d

What are the results?

Terminology	Formula	Results	
Sensitivity	= a/(a+c)	0.74	74%
Specificity	=d/(b+d)	0.84	84%
Likelihood ratio for a positive test result	=sensitivity/1-specificty	4.60	4.60%
Likelihood ratio for a negative test result	=(1-sensitivity)/specificty	0.31	31%
Positive predictive value	=a/(a+b)	0.27	27%
Negative predictive value	=d/(c+d)	0.97	97%
Pre-test probability (prevalence)	(a+c)/(a+b+c+d)	0.08	8%
Pre-test odds	= prevalence/(1-prevalence)	0.08	8%
Post-test odd	=pre-test odds x likelihood ratio	0.38	38%
Post-test probability	=post-test odds/(post-test odds+1)	0.27	27%

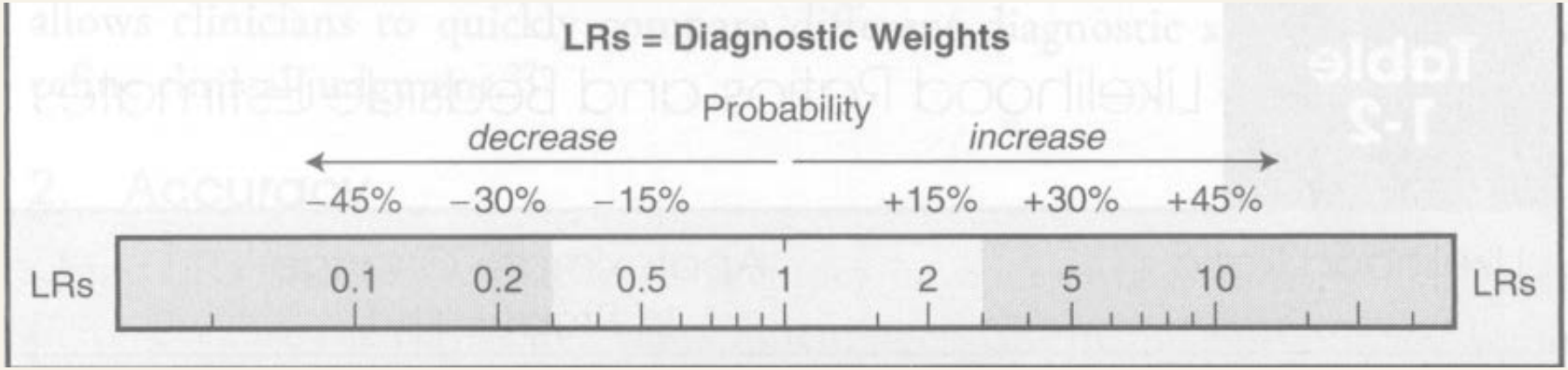
Nomogram

Pre-test probability – 8%
Likelihood ratio – 4.6



Post-test probability – 27%

What do likelihood ratios mean?



LR<0.1 = strong negative test result

LR=1

No diagnostic value

LR>10 = strong positive test result

USF HEALTH CENTER FOR EVIDENCE BASED MEDICINE

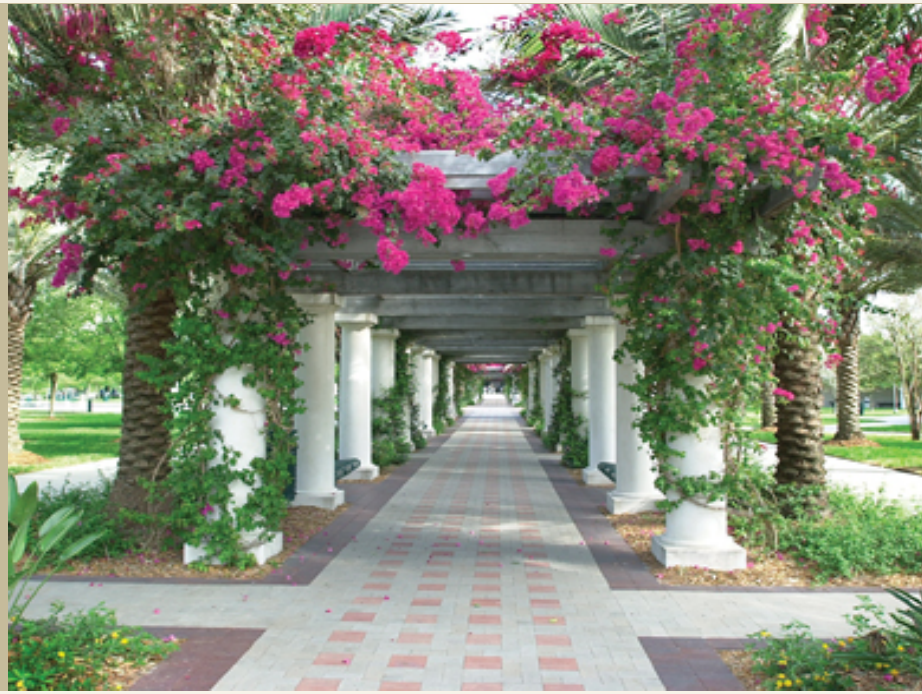
How can I apply the results to patient care?

- Will the reproducibility of the test result and its interpretation be satisfactory in my setting?
  - Unknown. Investigators did not evaluate agreement between independent physicians on the clinical diagnosis of pneumonia.
  - They reported (but did not provide data) that the performance of clinical examination (the 'test') did not vary among patients taking themselves to the emergency room versus those referred to the emergency room or presenting at one of 3 clinics. This is a significant flaw of the study.

How can I apply the results to patient care?

- Are the results applicable to patients in my practice?
  - Yes
  - Patients similar to mine are enrolled in the study as the study enrolled all consecutive patients presenting to the ER

Critical appraisal: PROGNOSTIC studies



Ambuj Kumar, MD, MPH

Associate Professor, Director

Division and Center for Evidence Based Medicine, Dept of Internal Medicine, College of Medicine, University of South Florida

Tampa, Florida USA

USF HEALTH CENTER FOR EVIDENCE BASED MEDICINE

Process of EBP

Patient dilemma

Ask

Acquire

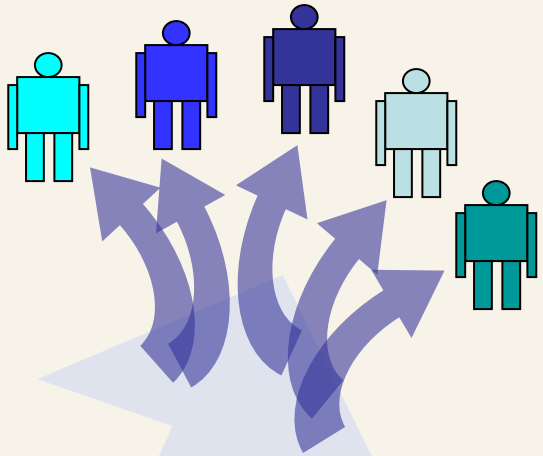
Appraise

Hierarchy of evidence

Apply

Principles of evidence-based practice

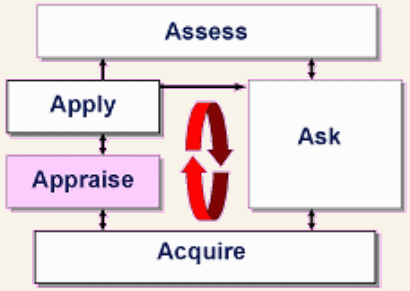
Act & Assess



Evidence alone does not decide – combine with other knowledge and values

Courtesy: Dr. Richardson

Outline



- Introduction
  - Objectives
  - Design considerations
- Prognosis Guides
  - Validity
  - Importance
  - Applicability
- Summary

Objectives

- Be able to
  - Define prognostic factors
    - Evidence Literacy
  - Recognize threats to claims that an attribute predicts future outcomes
    - Evidence Literacy
  - Understand measures of outcome over time: survival curves
    - Evidence Numeracy

Key Term

- Prognosis

  - Possible outcomes of a disease and the probability with which they can be expected to occur.

Key Concepts

- Prognosis Issues
  - Possibilities (Qualitative)
    - Which outcomes could happen?
  - Probabilities (Quantitative)
    - How likely are they to happen?
  - Periods (Temporal)
    - Over what time period?

Key Concepts

- Prognosis Uses
  - Predictive
    - What the future is likely to hold
  - Prescriptive
    - To select treatment that does more good than harm, anticipate future state without treatment
  - Comparative
    - To compare 2 populations for a given outcome, need to be able to adjust for prognosis (baseline likelihood of the outcome)

Key Terms

- Prognostic Factor
  - Patient characteristic(s) that confer increased or decreased risk of an outcome from a disease

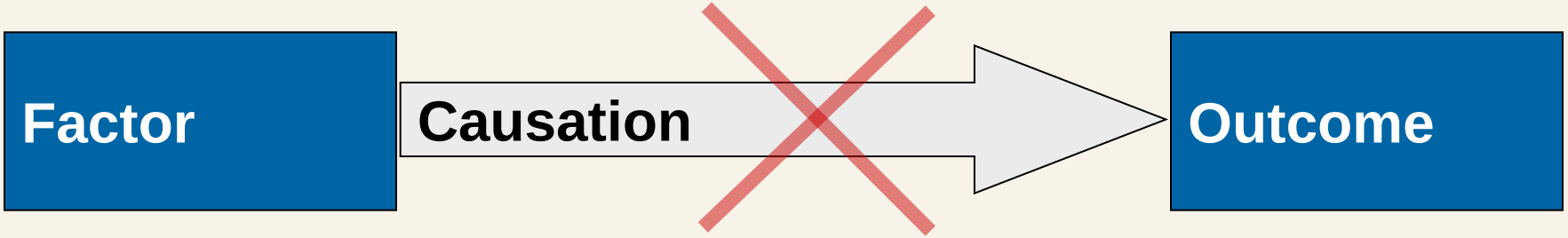
- Risk Factor
  - Patient characteristic(s) associated with development of the disease in the first place

Design Considerations

- Strength of Association

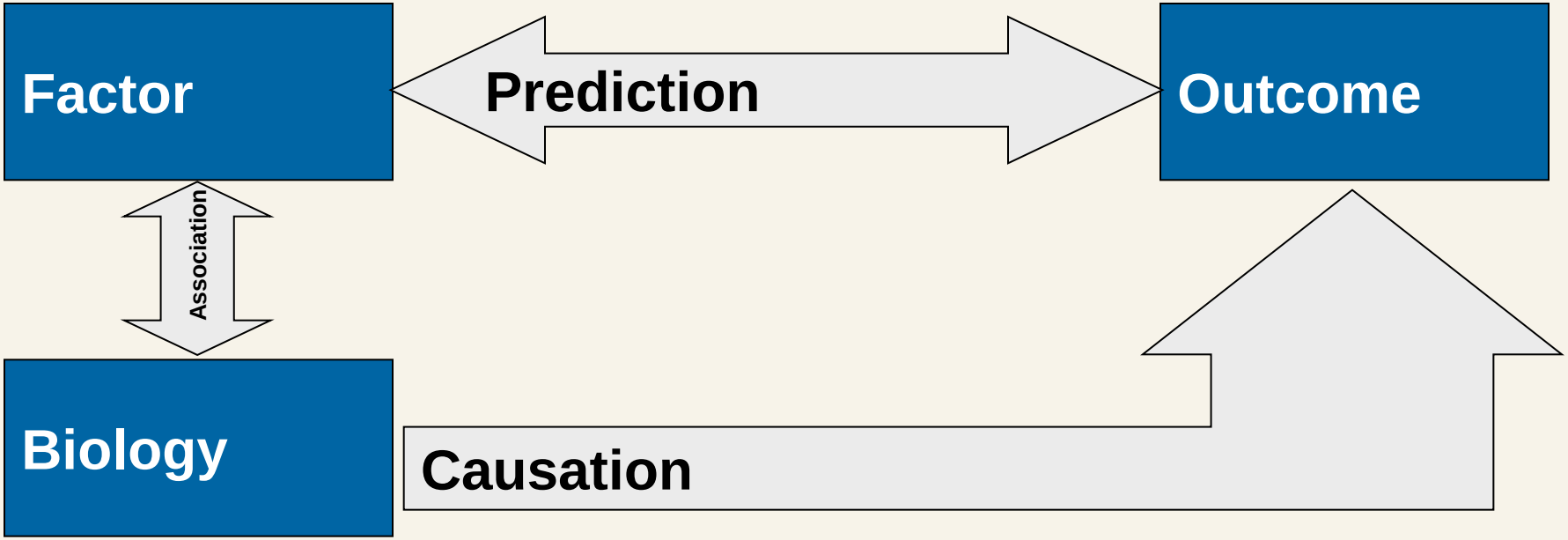
Design Considerations

- Strength of Association

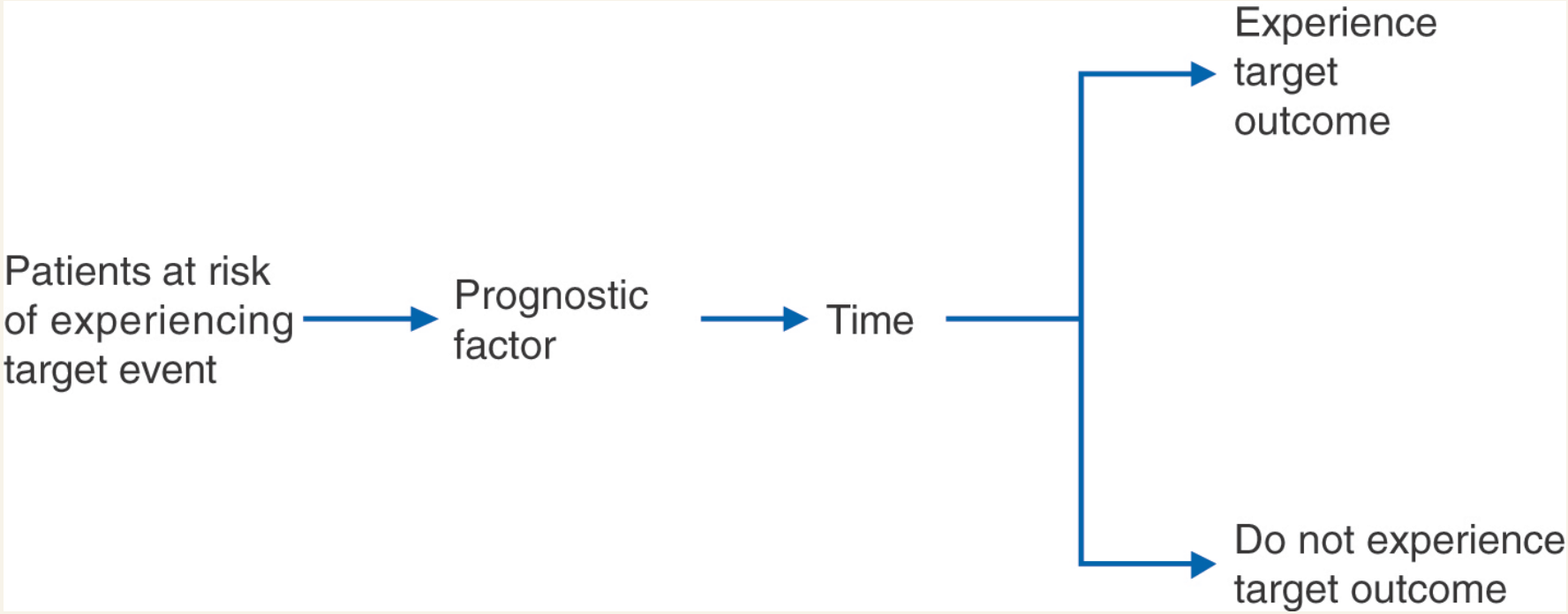


Design Considerations

- Strength of Association

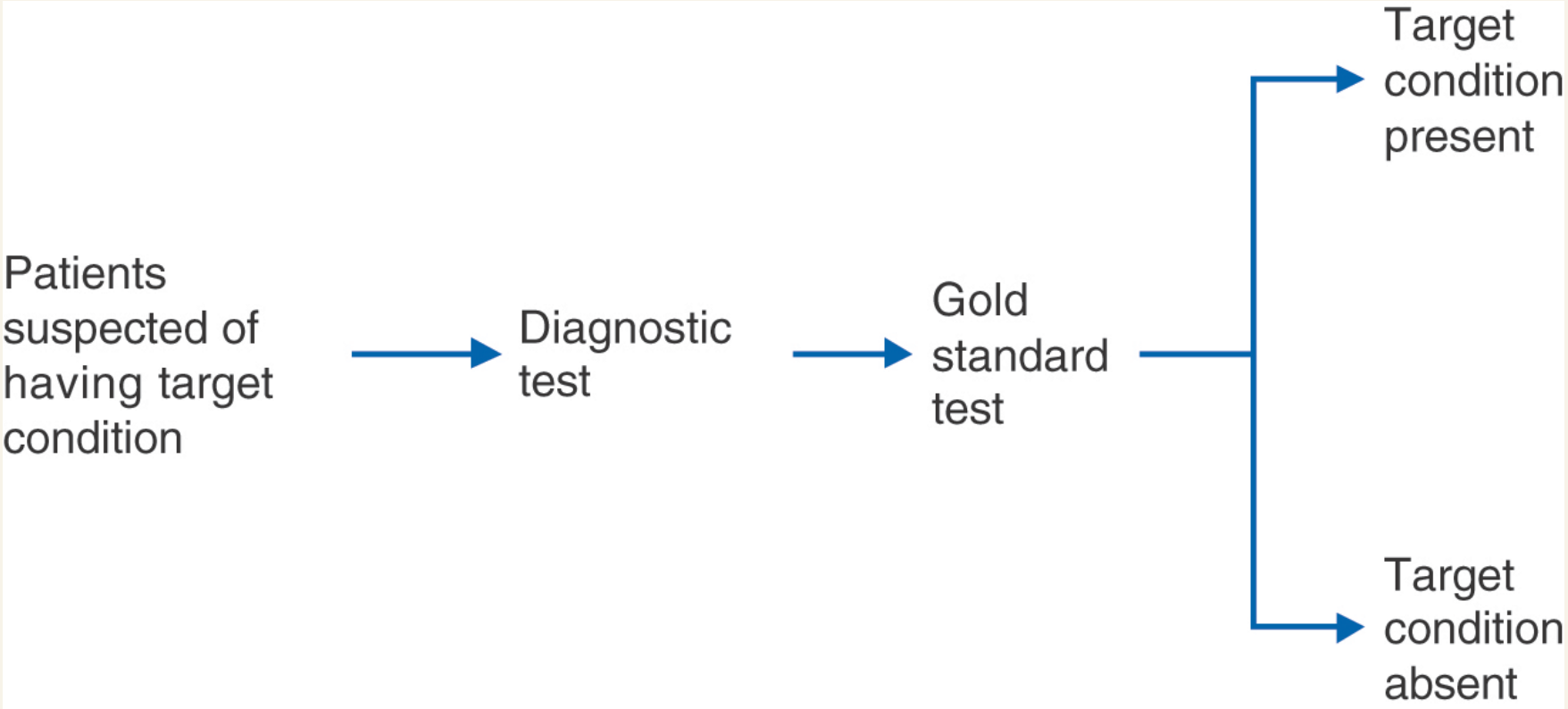


Prognosis Design



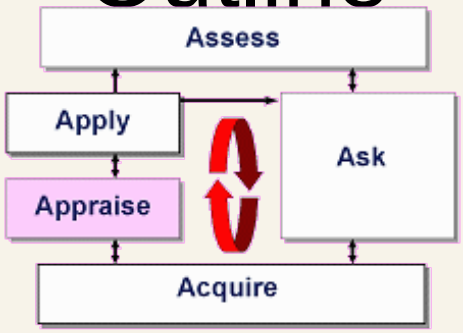
- **Cohort and case-control designs are used to explore determinants of outcomes.**
- **If controls are used, they are patients with different prognostic factors.**

vs. Diagnosis Design



- **The rules of evidence for judging prognosis studies are similar to those for studies of diagnostic and screening tests.**
- **In a prognostic study, time is the gold standard test.**

Outline



- Introduction
  - Objectives
  - Design considerations
- Prognosis Guides
  - Validity
  - Importance
  - Applicability
- Summary

Assess

Appraise: Interpret

- Are the results valid?
- What are the results?
- How can the results be applied?

Appraise

Acquire

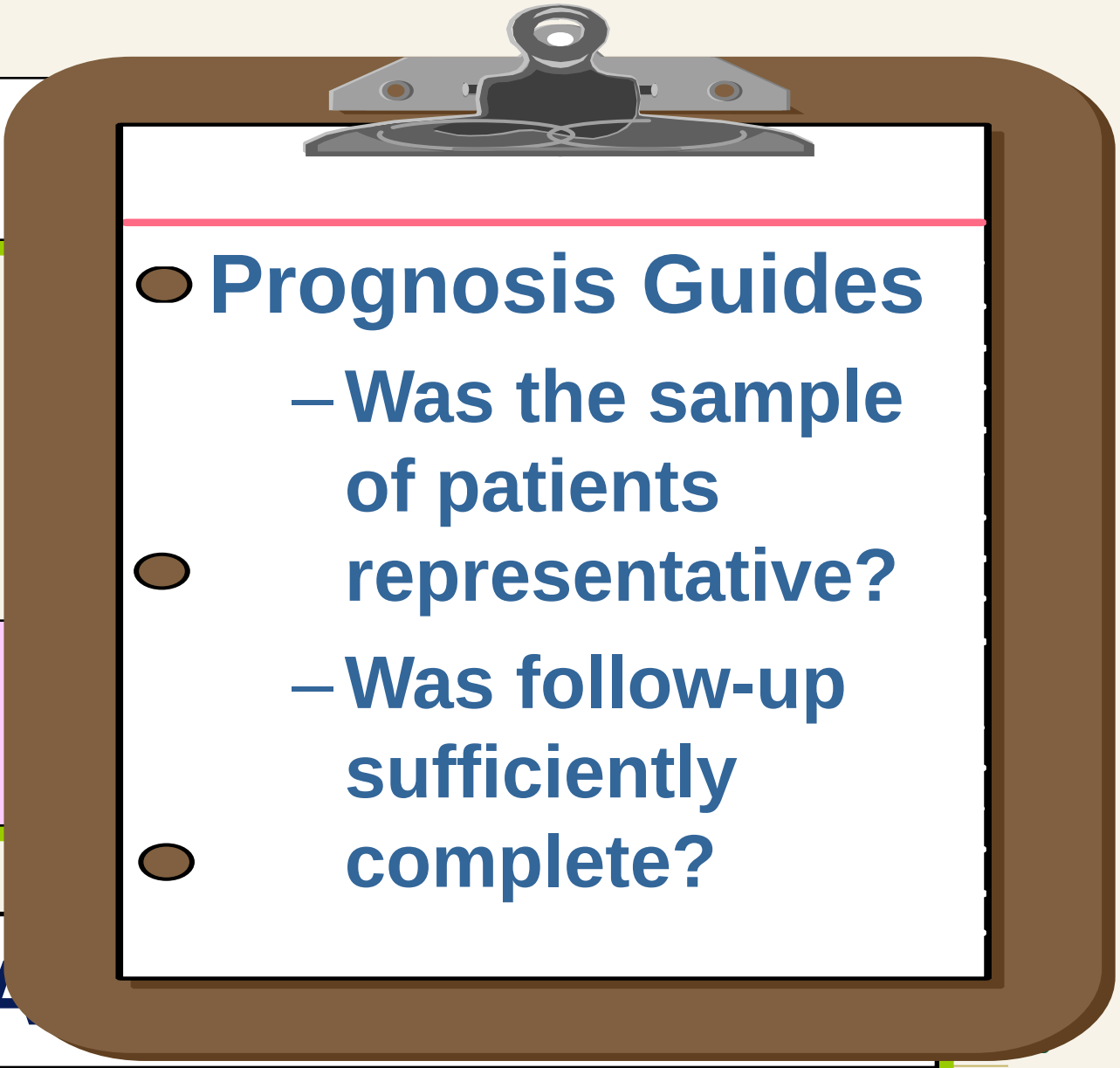
Assess

Appraise

- **Validity**
  Can I trust the information?
- **Importance**
  Will the information, if true, make an important difference?
- **Applicability**
  Can I use this information?

Appraise

A


Prognosis Guides
– Was the sample of patients representative?
– Was follow-up sufficiently complete?

Appraise

Guides

Users' Guides to an Article About Prognosis

Are the results valid?

- Was the sample of patients representative?
- Were the patients sufficiently homogeneous with respect to prognostic risk?
- Was follow-up sufficiently complete?
- Were outcome criteria objective and unbiased?

What are the results?

- How likely are the outcomes over time?
- How precise are the estimates of likelihood?

How can I apply the results to patient care?

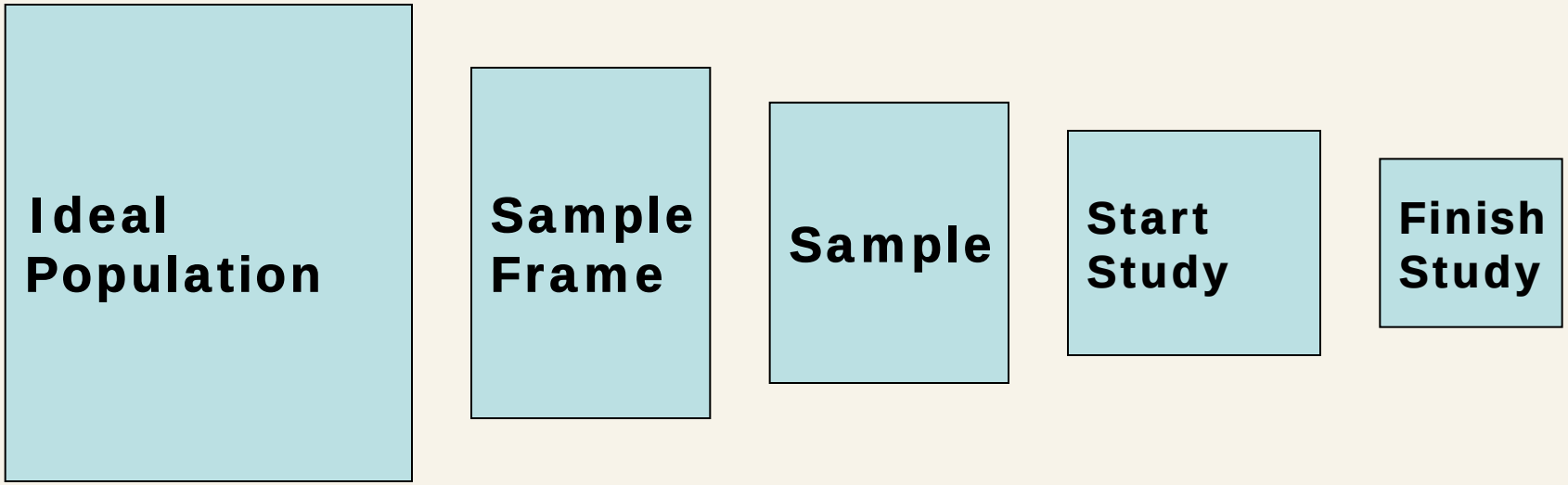
- Were the study patients and their management similar to those in my practice?
- Was the follow-up sufficiently long?
- Can I use the results in the management of patients in my practice?

Validity

- Are the results valid?
  1. Was the sample of patients representative?
  2. Were the patients sufficiently homogeneous with respect to prognostic risk?
  3. Was follow-up sufficiently complete?
  4. Were outcome criteria objective and unbiased?

Validity

- Was the sample of patients representative?



Validity

- Was the sample representative?
  - Expectation
    - People with a particular condition will, on average, experience similar outcomes as a population of similar people
  - Requirements
    - Ability to define and observe experiences in populations over time
    - Enough information to match people to populations

Validity

- Was the sample representative?
  - Matching people to populations
    - Demographics
      - Age, gender, socioeconomic status, etc.
    - Diseases
      - Severity, subtype
    - Disorders
      - Other illnesses or relevant conditions

Validity

- Was the sample representative?
  - Threats
    - Referral bias
    - Failure to clearly define study patients
    - Lack of objective criteria for defining demographics, diseases, or disorders

Example: risk of recurrent childhood seizures

- **1%-5% in family practice**
- **3%-75% in neurology clinics**

Validity

- Was the sample representative?
  - Remedies
    - Report of explicit or implicit filters passed before entering the study
      - Primary → Secondary → Tertiary Care
    - Clear description of which patients were included and which were excluded from study

Validity

- Were the patients sufficiently homogeneous with respect to prognostic risk?
  - Expectation
    - The outcome for the group should be applicable to each member of group.
  - Requirement
    - Subjects should be at a similar point in the disease process.

Validity

- Was the sample homogeneous?
  - Threats
    - Different demographics
    - Different disease(s)
      - Stage
      - Severity
    - Different disorders
      - Comorbidities that may define subgroups with different prognoses

Validity

- Was the sample homogeneous?
  - Protections
    - Define and track any subgroups
    - Adjust for demographics, disease (stage, severity), and disorders in analysis
    - Use clinical common sense
      - Have investigators missed important subgroups with different prognoses?

Validity

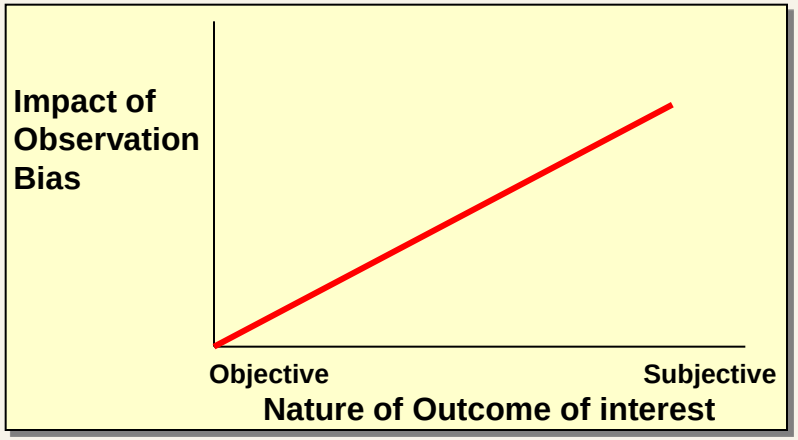
- Was follow-up sufficiently complete?
  - Threats
    - Important outcomes occurring beyond the end of the study
    - Significant losses to follow-up
      - Increased likelihood that those lost have significantly different outcomes

Validity

- Was follow-up sufficiently complete?
  - Protections
    - Simple sensitivity analysis
      - Recalculate risk based on “best case” scenario where all losses were free of adverse outcome
      - Recalculate risk based on “worst case” scenario where all losses suffered the adverse outcome
      - Compare these recalculations to gauge the potential impact of losses
    - “5%” and “20%” rule

Validity

- Were outcome criteria objective and unbiased?
  - As the subjectivity of outcome determination increases, the importance of blinding to prognostic factors increases



Guides

Users' Guides to an Article About Prognosis

Are the results valid?

- Was the sample of patients representative?
- Were the patients sufficiently homogeneous with respect to prognostic risk?
- Was follow-up sufficiently complete?
- Were outcome criteria objective and unbiased?

What are the results?

- How likely are the outcomes over time?
- How precise are the estimates of likelihood?

How can I apply the results to patient care?

- Were the study patients and their management similar to those in my practice?
- Was the follow-up sufficiently long?
- Can I use the results in the management of patients in my practice?

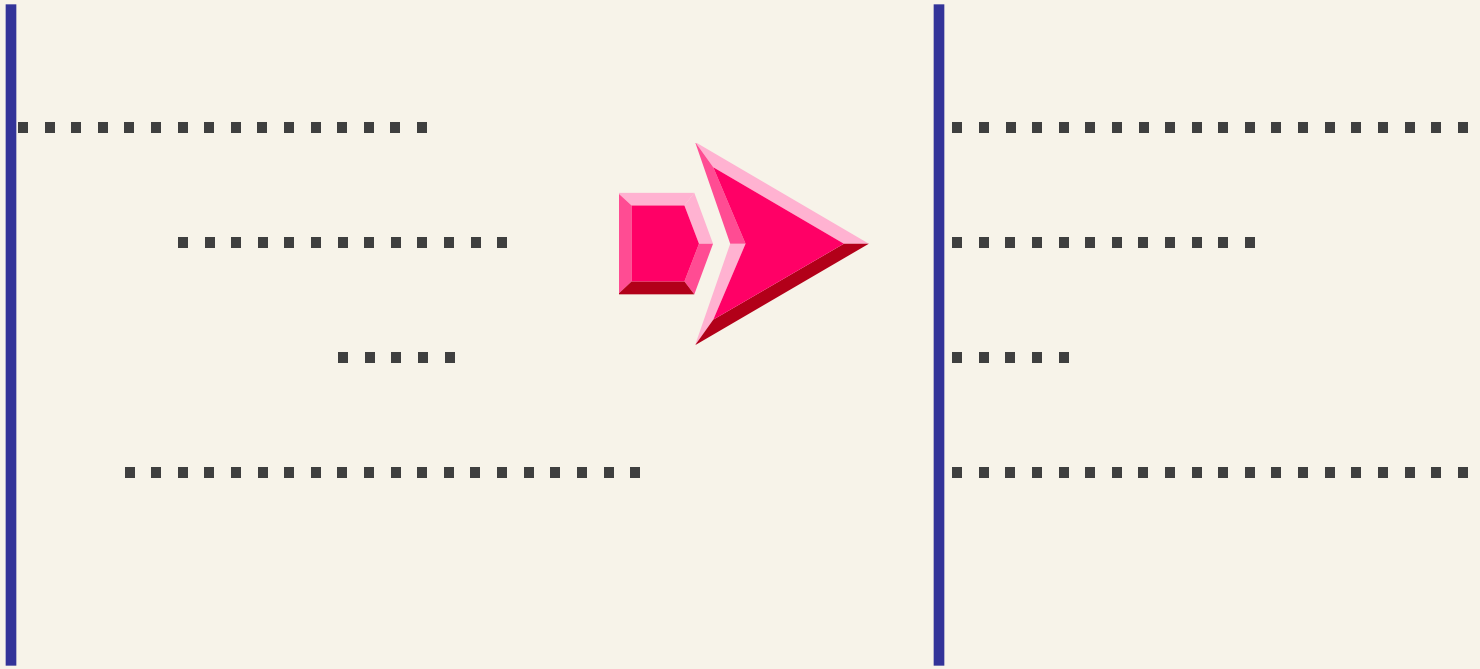
Importance

- What are the results?
  1. How likely are the outcomes over time?
  2. How precise are the estimates of likelihood?

Importance

- How likely are the outcomes over time?
  - Measures that relate events to time
    - **Survival rate**
      - Percent surviving at a given time
    - **Median survival**
      - Time at which 50% still surviving
    - **Survival curve**
      - Percent of original sample who have not yet had outcome of interest
      - Where events are discrete and time of event is precisely known

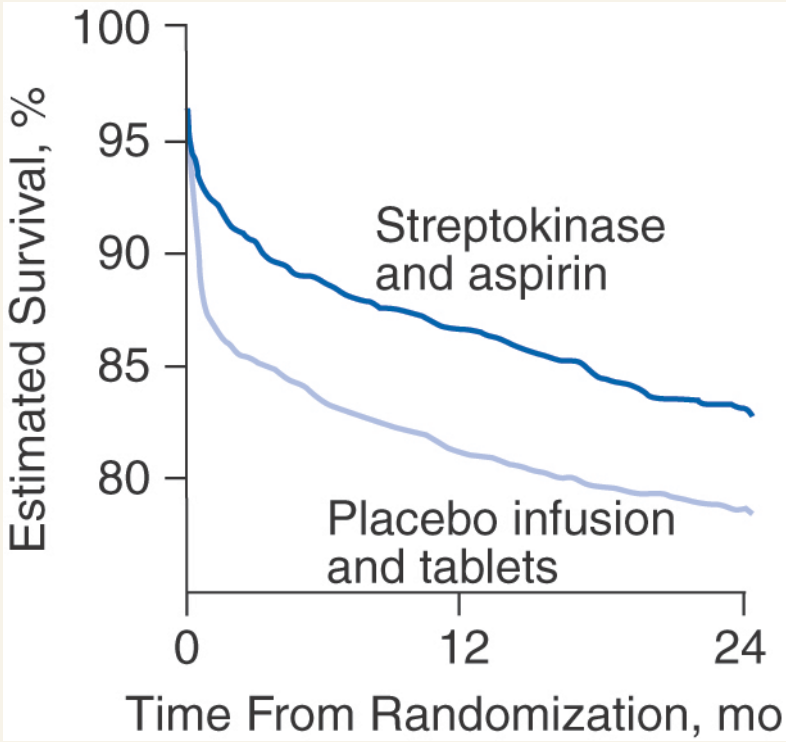
Common Starting Point Life Table



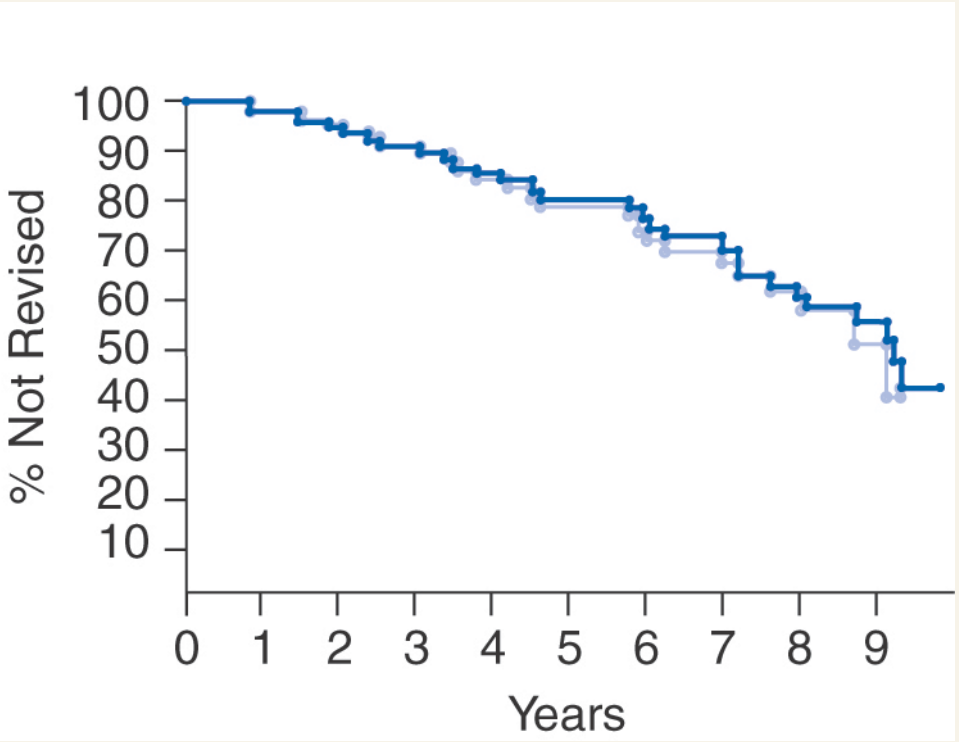
Start of Study

Patient Enters Study

Survival Curves



Survival after MI



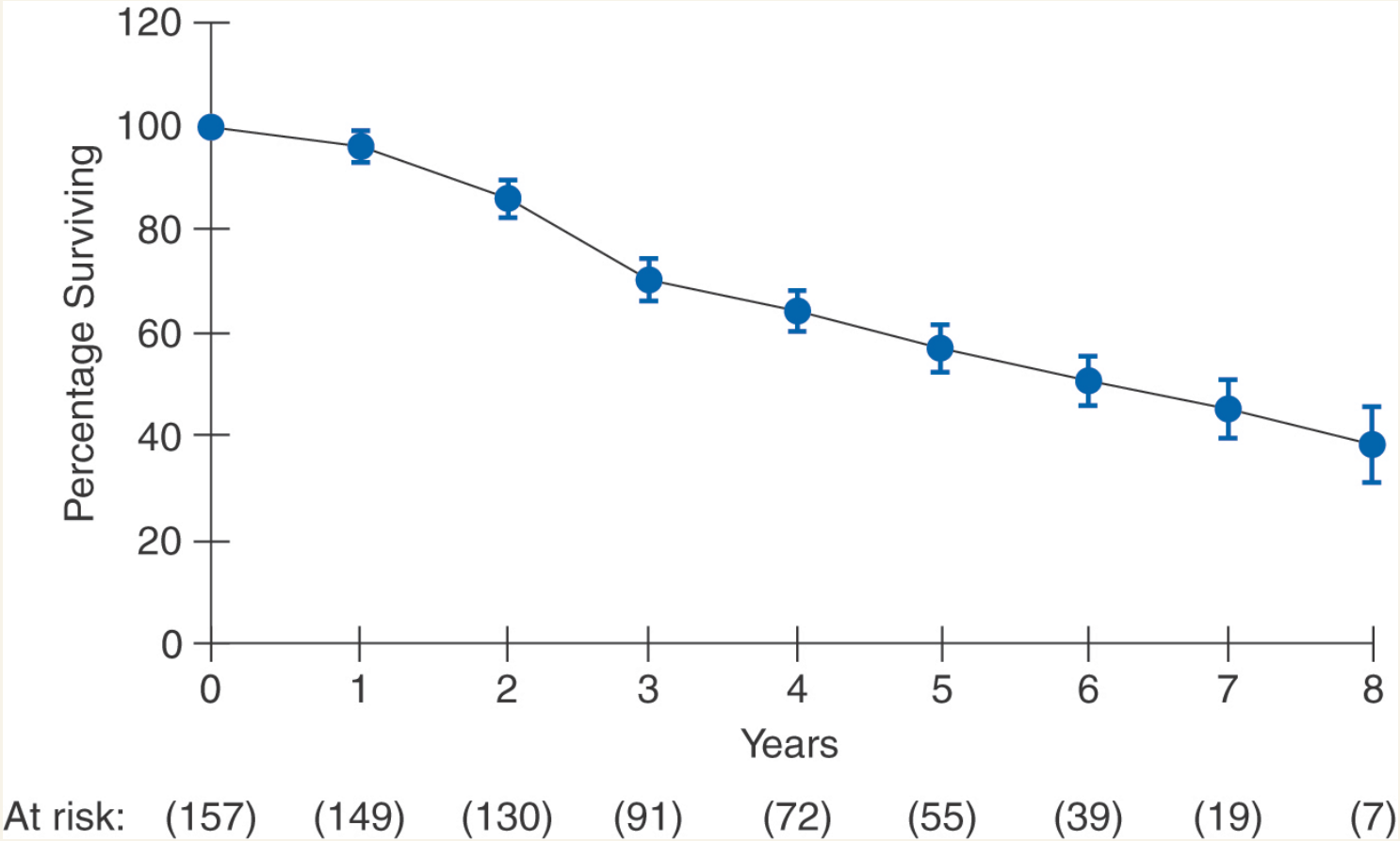
Complications after total hip arthroplasty

USF HEALTH CENTER FOR EVIDENCE BASED MEDICINE

Importance

- How precise are the estimates of likelihood?
  - Confidence intervals
    - Range within which it is likely that the true point estimate lies
    - Precision drops as time from exposure increases
      - Losses to follow-up and outcome assessment

Survival Curve Precision



Guides

Users' Guides to an Article About Prognosis

Are the results valid?

- Was the sample of patients representative?
- Were the patients sufficiently homogeneous with respect to prognostic risk?
- Was follow-up sufficiently complete?
- Were outcome criteria objective and unbiased?

What are the results?

- How likely are the outcomes over time?
- How precise are the estimates of likelihood?

How can I apply the results to patient care?

- Were the study patients and their management similar to those in my practice?
- Was the follow-up sufficiently long?
- Can I use the results in the management of patients in my practice?

Applicability

- How can I apply the results to patient care?
  1. Were the study patients and their management similar to those in my practice?
  2. Was follow-up long enough?
  3. Will using this evidence make a clinically important impact in your context?

Applicability

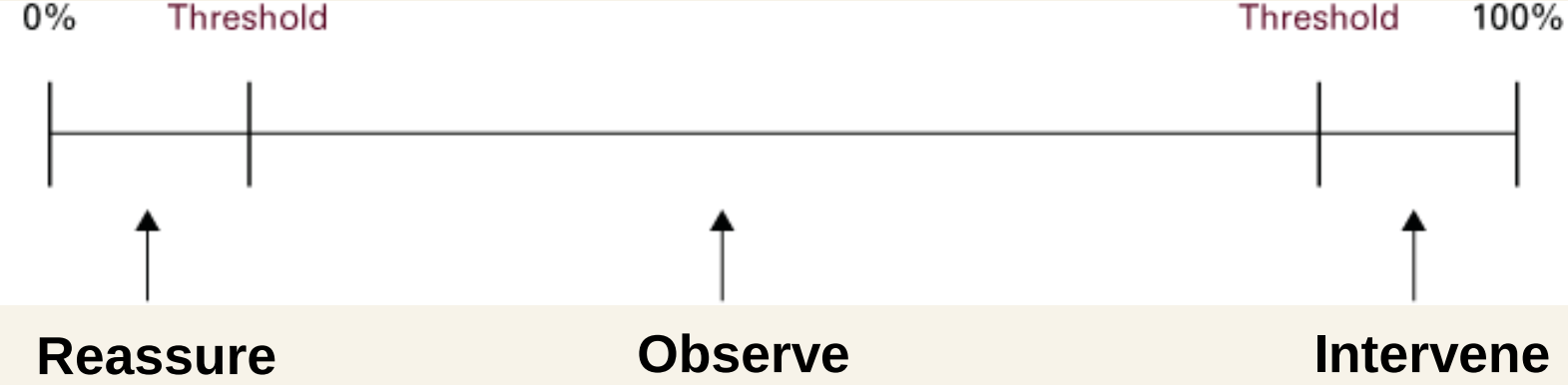
- Were the study patients and their management similar to mine?
  - Threats
    - Uneven application of therapies to different subgroups
    - Uneven application of therapies over time

Applicability

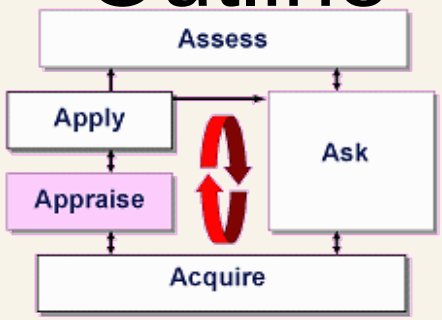
- Was the follow-up sufficiently long?
  - Threats
    - Important outcomes outlast the study duration

Applicability

- Can I use the results in the management of my practice?
  - Does the effect of the prognostic factor cross a decision threshold?



Outline

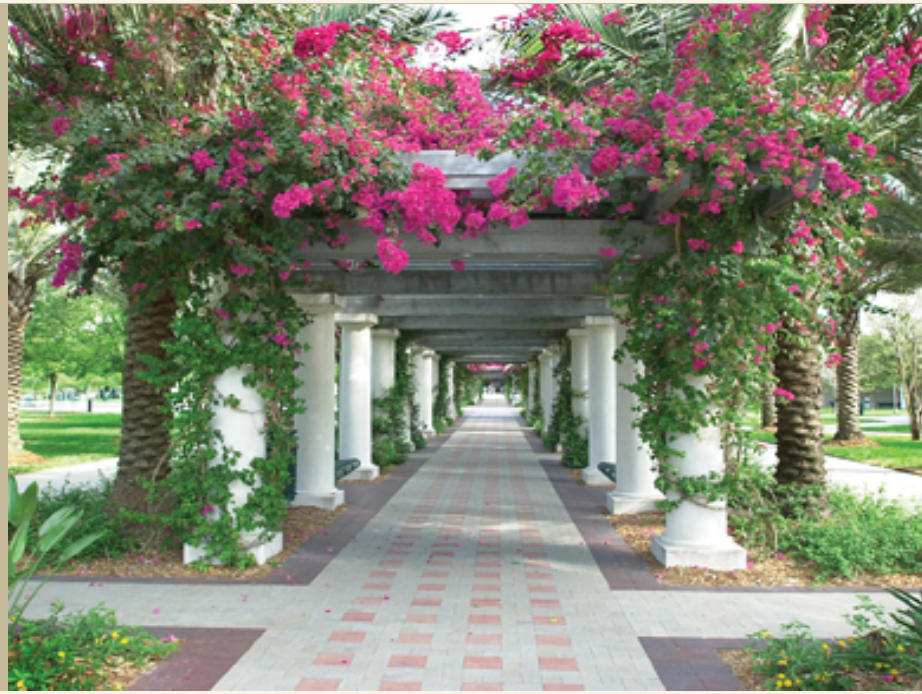


- Introduction
  - Objectives
  - Design considerations
- Prognosis Guides
  - Validity
  - Importance
  - Applicability
- Summary

Summary

- Key concept
  - The design goal of a study of prognosis is to avoid systematic overestimation or underestimation of the likelihood of outcome events in the patients under study—to make the population representative.

Critical appraisal: Prognostic study



Ambuj Kumar, MD, MPH

Associate Professor, Director

Division and Center for Evidence Based Medicine, Dept of Internal Medicine, College of Medicine, University of South Florida

Tampa, Florida USA

USF HEALTH CENTER FOR EVIDENCE BASED MEDICINE

Objective

- To learn how to apply the results of studies on diagnostic tests to clinical practice
  - Determine the validity of the study
  - Interpret the results
  - Apply the results to our patients

Assignment

- Read the clinical scenario, and then formulate a clinical question
- Read the article concerning a prognosis issue
- Critically appraise the paper mentioned above using the critical appraisal questionnaire on prognosis
- Decide how you would manage the patient described in the following scenario

Clinical Scenario

- You are about to see a 76 year old retired school teacher for the second time. You first saw her in the clinic a month ago because of cognitive problems. Your evaluation at that time included a Standardized Mini Mental State Examination, on which she scored 18 out of a possible 30, and a physical examination which was normal including no focal neurological signs. You arranged investigations for the treatable causes of dementia which were negative, and you thus feel she has probable Alzheimer's disease.
- She has lived with her son since her husband died six years ago. Her son thinks that she first developed significant problems with her memory about three years ago. However, she has become increasingly agitated and paranoid during the last year. She has refused to allow him to look after her financial affairs, despite the fact that she owns three pieces of property and isn't able to manage them herself. Her son asked you about her prognosis, and whether she is likely to die soon from the dementia.

You indicated that you would discuss this with him at the second visit once the results of all the tests are available.

Search

- Hoping to provide the son with the most specific information possible about his mother's prognosis, after the initial visit you searched the medical library's MedLine system via the INTERNET.
- "*Alzheimer's Disease", which yielded 3687 articles from 1990 onward.
- “prognosis", which yielded 23,004 articles;
- crossing the two sets yielded 27 articles.
- On target: "Survival of outpatients with Alzheimer-type dementia" by Walsh and colleagues (Walsh JS, Welch G, Larson EB, Welch HG. Survival of outpatients with Alzheimer-type dementia. Ann Intern Med 113. 429-34 1990).

Are the Results Valid?

- What was the sample patients' representative?
  - Walsh and colleagues studied 126 outpatients with Alzheimer's disease who were **consecutively** referred to a multidisciplinary clinic for evaluation between 1980 and 1982.

Are the Results Valid?

- Were patients sufficiently homogeneous with respect to prognostic risk?
  - The diagnosis was made by consensus by a group consisting of an internist, psychiatrist, psychologist, neurologist or neuropathologist, and research nurse using the conventional DSM-III criteria for dementia.
  - The tests used to exclude other causes of dementia were not described. However, given the multidisciplinary nature and expertise of the group, it seems reasonable to assume that the appropriate tests were done to exclude disorders such as hypothyroidism, depression and space occupying lesions of the brain that can be confused with Alzheimer's disease.
  - reported survival from two different points in time:
    - referral to the clinic and
    - the point at which symptoms of memory loss were first noticed.
- Former is a more certain point in time, but suffers from the disadvantage that patients come to medical attention at different stages in the progression of their disease.
- The latter provides a more uniform starting point, but is potentially imprecise because dementia develops insidiously and the time of onset is identified retrospectively. Survival after presentation to clinic is probably more relevant for your patient's son.

Are the Results Valid?

- Was follow up complete?
  - Patients in the dementia study were enrolled between 1980 and 1982 and followed until 1988 or their death. Thus the follow-up was quite long, and 61 percent of the cohort died during this time.
- Were outcome criteria objective and unbiased?
  - In Walsh's study, the method and intensity of follow-up was not described. However, all patients were accounted for at the end of the study, and the date of death was known in those who died.

What are the Results?

- **How Likely are the outcomes over time?**
- The patient's son asked "What are the chances that my mother will still be alive in five years?"
  - A simple and direct answer in absolute terms.
    - Five years after presentation to the clinic about one half the patients (50%) had died. Thus there is about a 50:50 chance that his mother will be alive in five years.

What are the Results?

- **How Likely are the outcomes over time?**
- On the basis of family history “he knows with Alzheimer's disease is a 65 year old uncle who was diagnosed 10 years ago and is still living”. He is surprised that his mother's chance of dying in the next five years is so high.
  - Prognostic factors for death in patients with Alzheimer's disease.
  - The statistically significant prognostic factors for death were increasing
    - Age
    - Dementia severity
    - Behavioral problems, and hearing loss.

What are the Results?

- **How Likely are the outcomes over time?**
- You explain that his mother is considerably older than his uncle was at the time of diagnosis
- Her age is almost identical to the mean age of the cohort studied by Walsh and colleagues.
- However, her Mini Mental State Examination score is quite low (indicating more severe dementia) and her behavioral problems also suggest that she is at higher risk than the average patient in Walsh's study.
- No table or formula was presented which allows to combine all of these factors and estimate a risk of mortality that is specific for your patient.
- However, you can feel confident in telling her son that his mother's chances of dying are at least 50% during the next five years, and probably greater.

What are the Results?

- **How precise are the estimates of likelihood?**
- 95% confidence intervals for the relative risk associated with each prognostic factor.
  - The relative risk associated with a behavioral problem was 1.5 with a 95% confidence interval of 1.0 to 2.5.
  - This means that the best estimate is that a patient with a behavioral problem is 1.5 times more likely to die than an individual without a behavioral problem.
  - The probability that the true relative risk is between 1.0 (ie no effect) and 2.5 is 95%.

How can I apply the results to patient care?

- **Were the study patients and their management similar to my own?**
  - The characteristics of the study patients were quite similar to your patient.
- **Was the follow-up sufficiently long?**
  - Patients in the dementia study were enrolled between 1980 and 1982 and followed until 1988 or their death. Thus the follow-up was quite long, and 61 percent of the cohort died during this time.

How can I apply the results to patient care?

- **Is the prognostic information helpful in choosing therapy, or counseling my patients?**
  - Information on the likelihood of death will be useful to the son and his family as they plan the future care of his mother. Of course other prognostic information about the rate of progression of the dementing process and the need for intensive nursing care would be also be useful