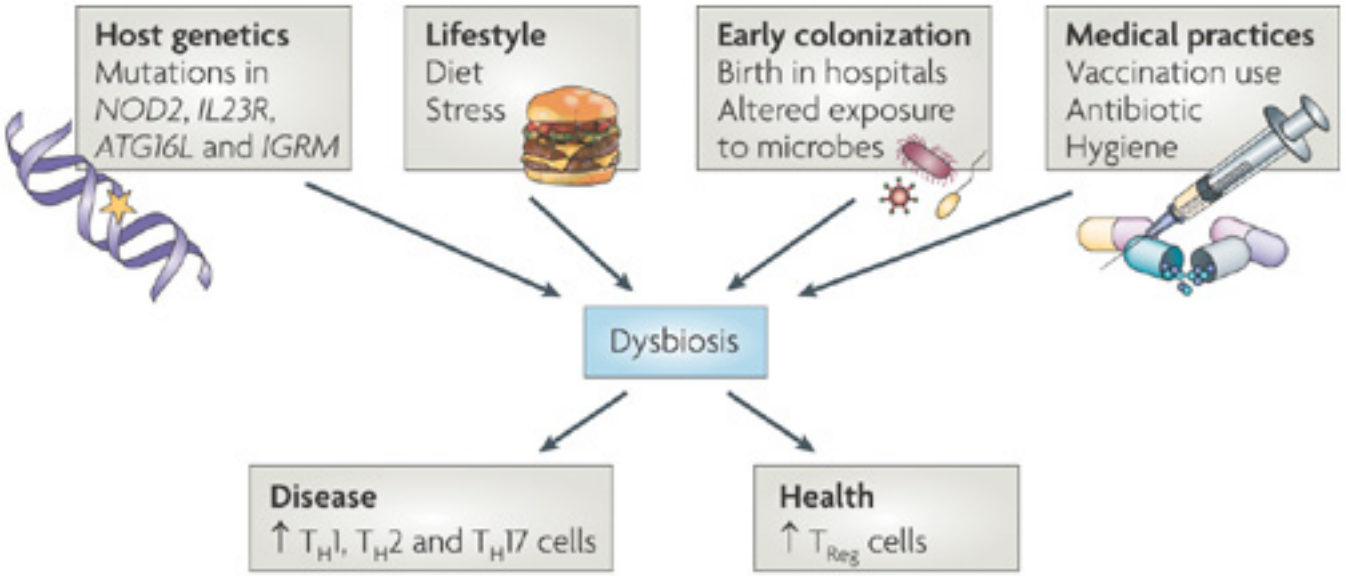
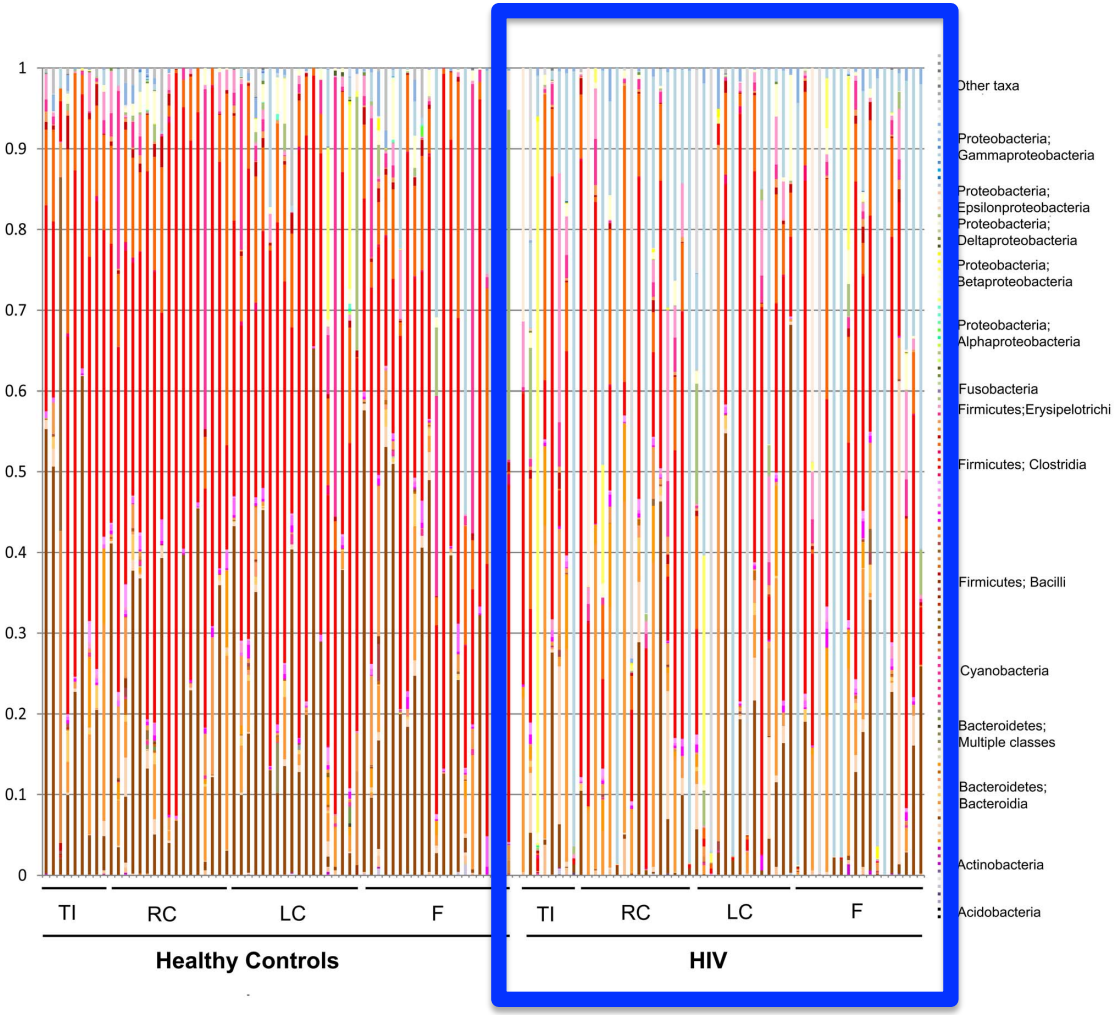
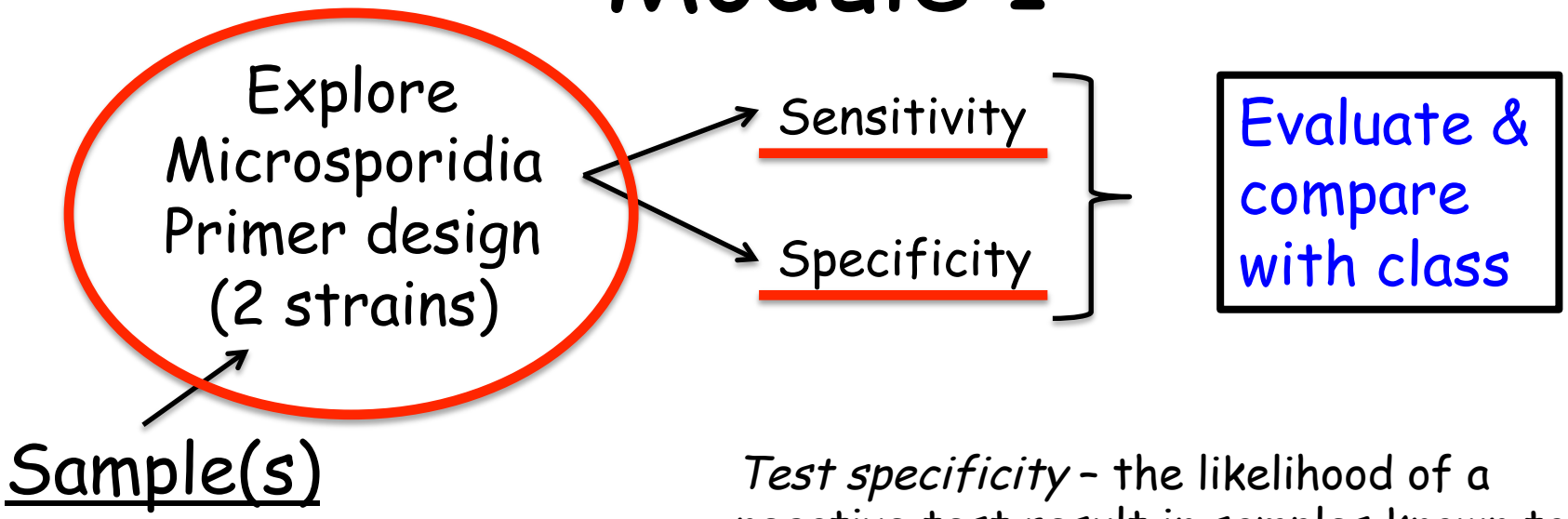


Nature Reviews | Immunology

HIV vs Control differences in the microbiome



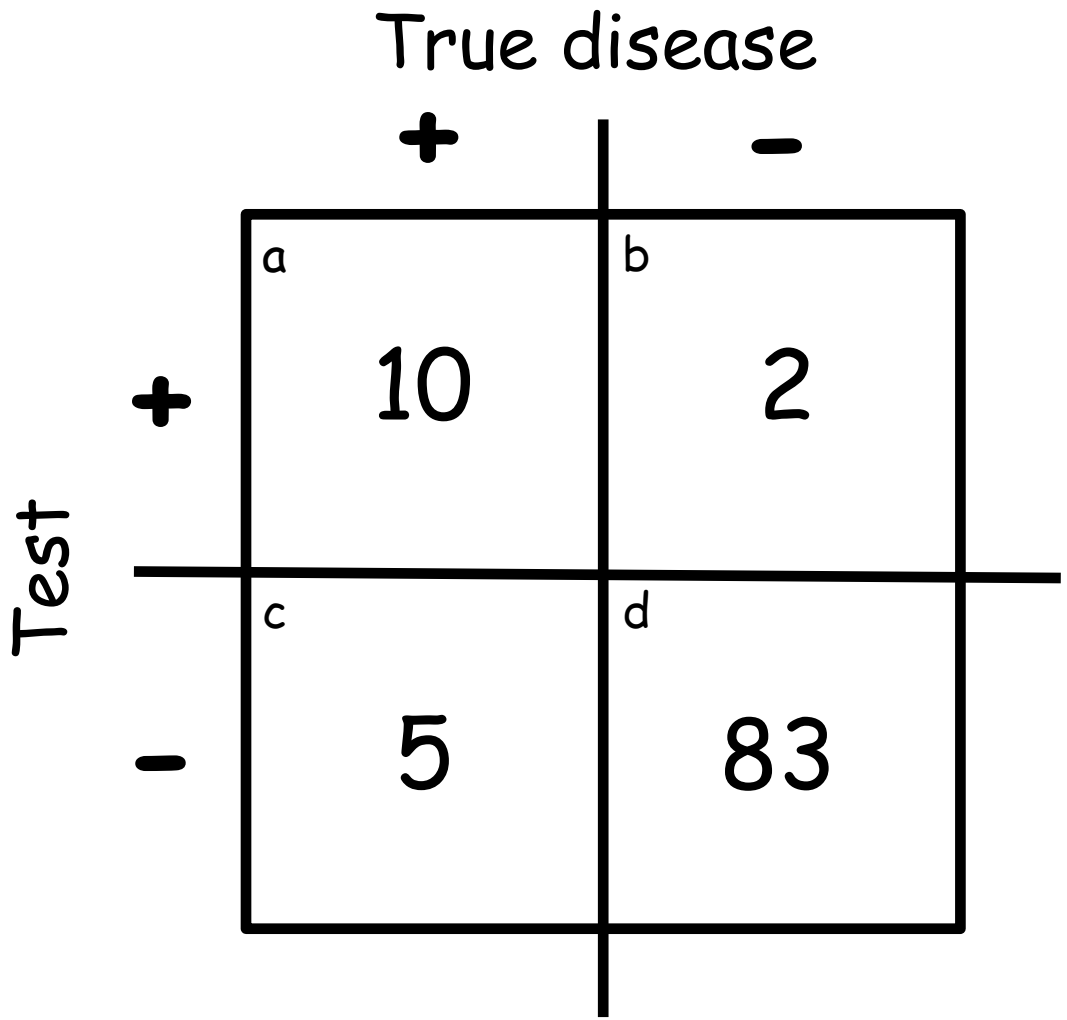
Module 1



Test specificity – the likelihood of a negative test result in samples known to be free of the microbe (pT-/D-).
aka - "true negative rate"

Test sensitivity – the likelihood of a positive test result in patients known to have the disease (pT+/D+).
aka – "true-positive rate" or "operational sensitivity"

Calculating sensitivity and specificity



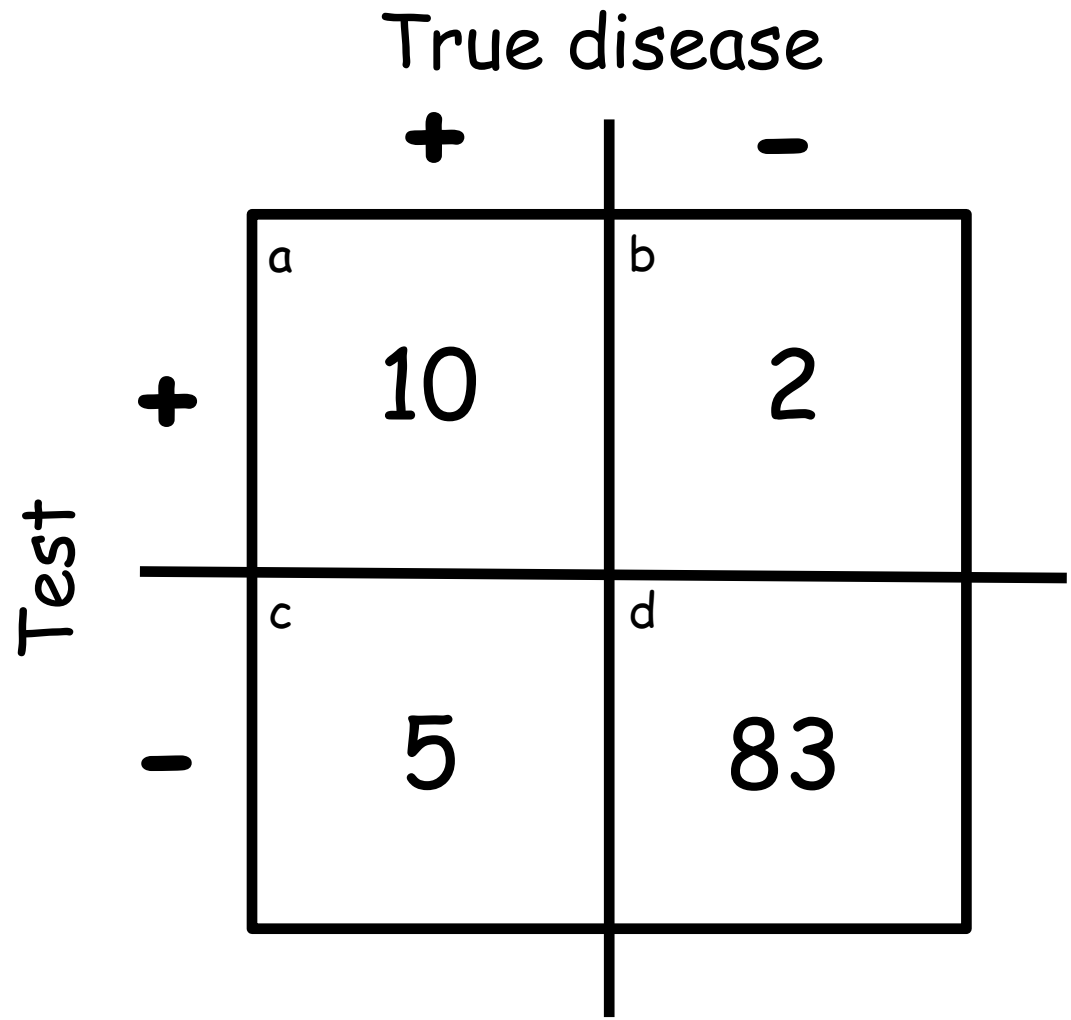
Sensitivity = 10/(10+5) = 0.67

Specificity = 83/(2+83) = 0.98

Savage Chickens

by Doug Savage

www.savagechickens.com

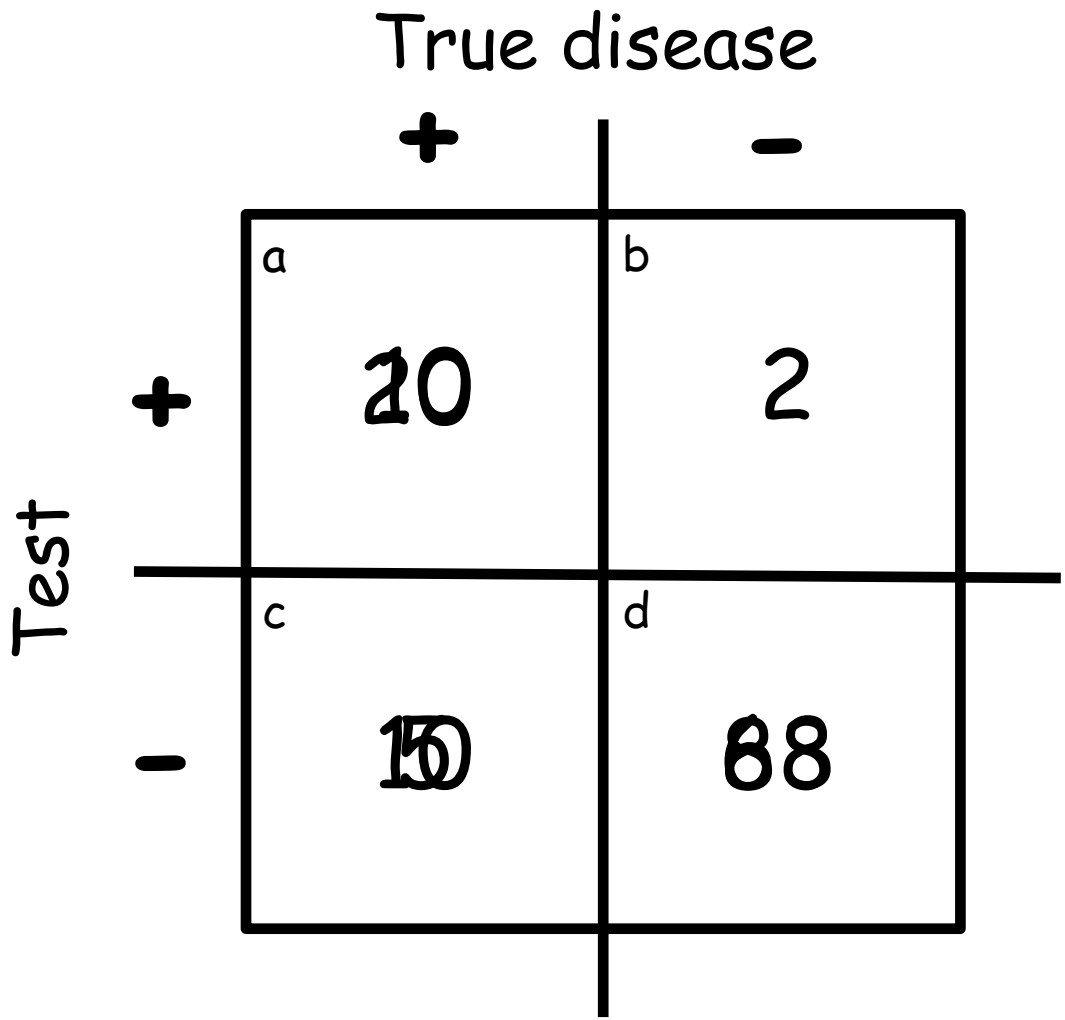


Positive Predictive Value = 10/(10+2)

Negative Predictive Value = 83/(83+5)

Sensitivity = 10/15 83/85 = Specificity

Prevalence and predictive values

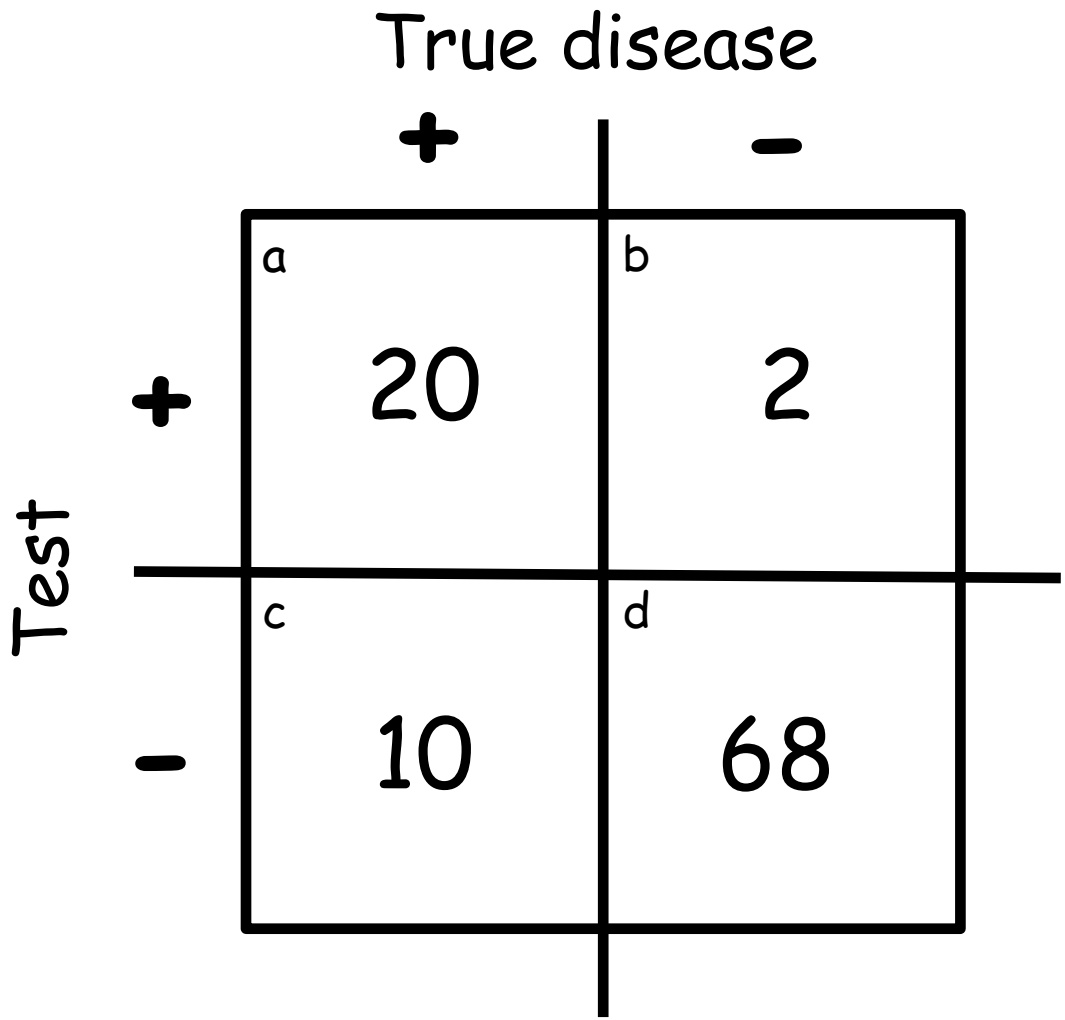


Positive Predictive Value = 20/22 ↑ 10/12

Negative Predictive Value = 68/78 ↓ 83/88

Sensitivity = 67% 98% = Specificity

Sensitivity and Specificity and Predictive values



Positive Predictive Value = 20/22 ↑ 10/12

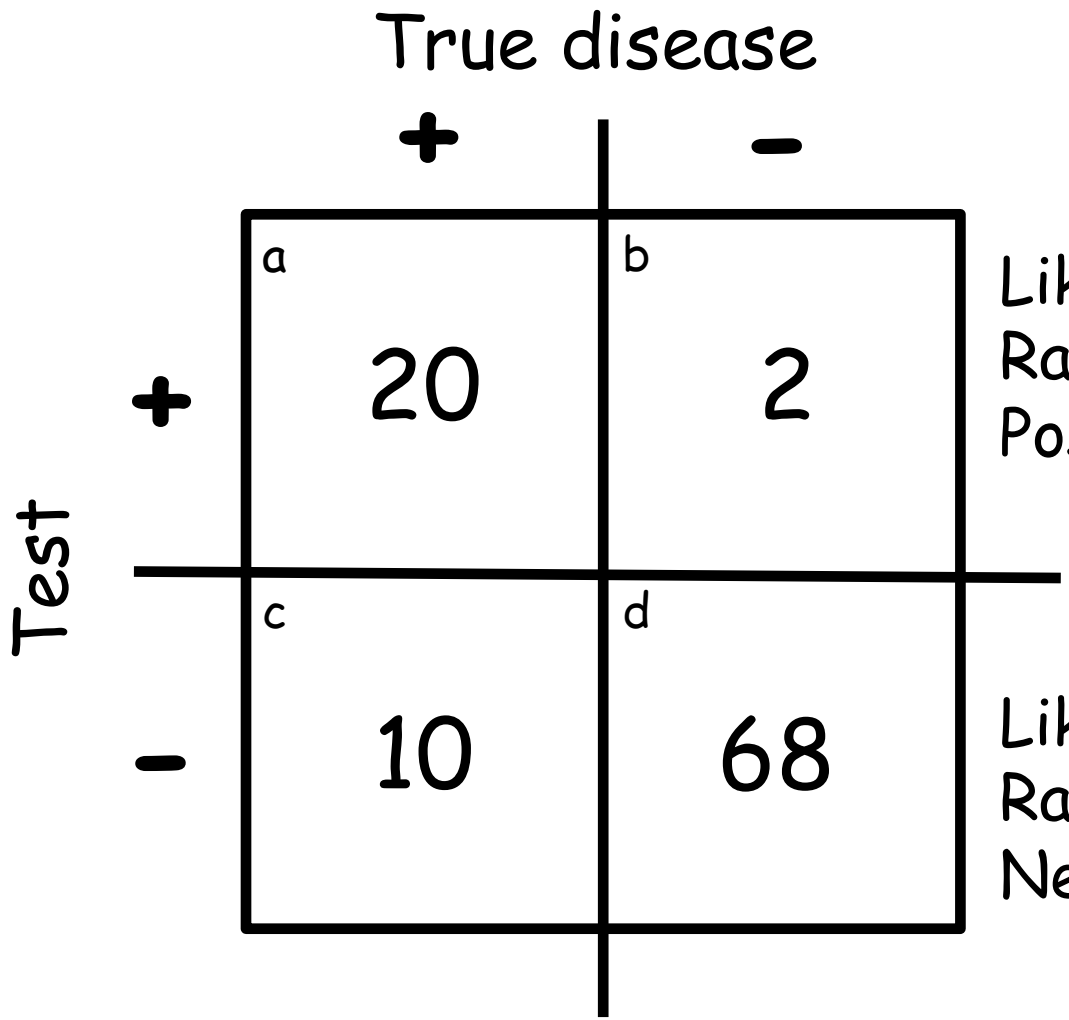
Negative Predictive Value = 68/78 ↓ 83/88

Sensitivity = 67% 98% = Specificity

Can sensitivity and specificity of tests differ indirectly?

- An example:
  - Disease burden (eg – heartworm)
  - Breed variation
  - High and low prevalence areas

Liklihood ratios – diagnostic utility of a test



Liklihood Ratio for a Positive Test = $\frac{a/a+c}{1-(d/b+d)}$

$33.5 = \frac{20/30}{1-(68/70)}$

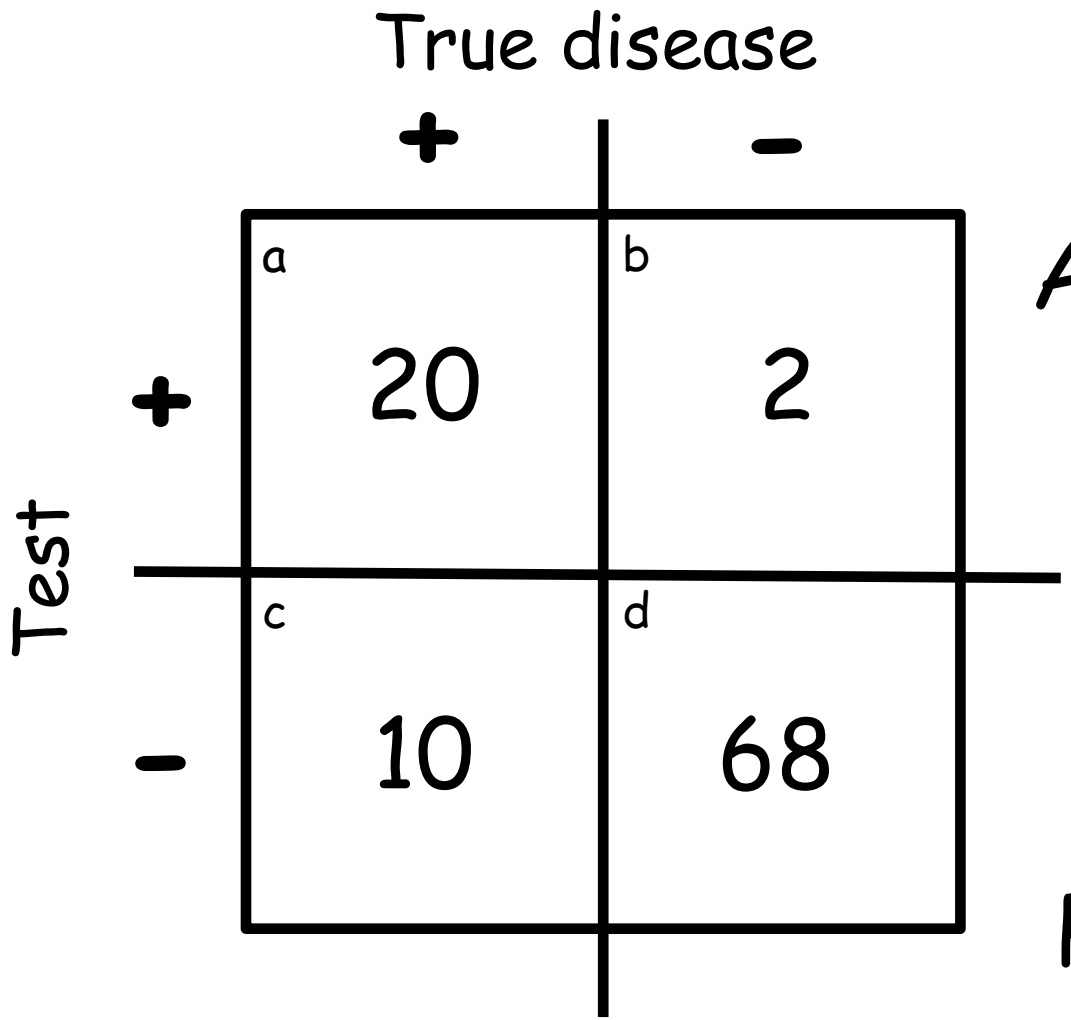
Liklihood Ratio for a Negative Test = $\frac{1-(a/a+c)}{d/b+d}$

$0.33 = \frac{1-(20/30)}{68/70}$

Sensitivity = 67%

98% = Specificity

Test Accuracy



Accuracy =

88/100 = 88%

Prevalence =

30/100 = 30%

Comparing tests?

- When is a test with high sensitivity most useful?

- When is a test with high specificity most useful?

Some sources of bias to consider in evaluating test performance

- Improper standards of validity
- The spectrum of test subjects

Some Basic Epidemiology

To add context to the infectious disease microbes

Epidemiology

Epidemiology is "the study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to the prevention and control of health problems"

Last JM. A dictionary of epidemiology, 4th ed. Oxford, Oxford University Press, 2001

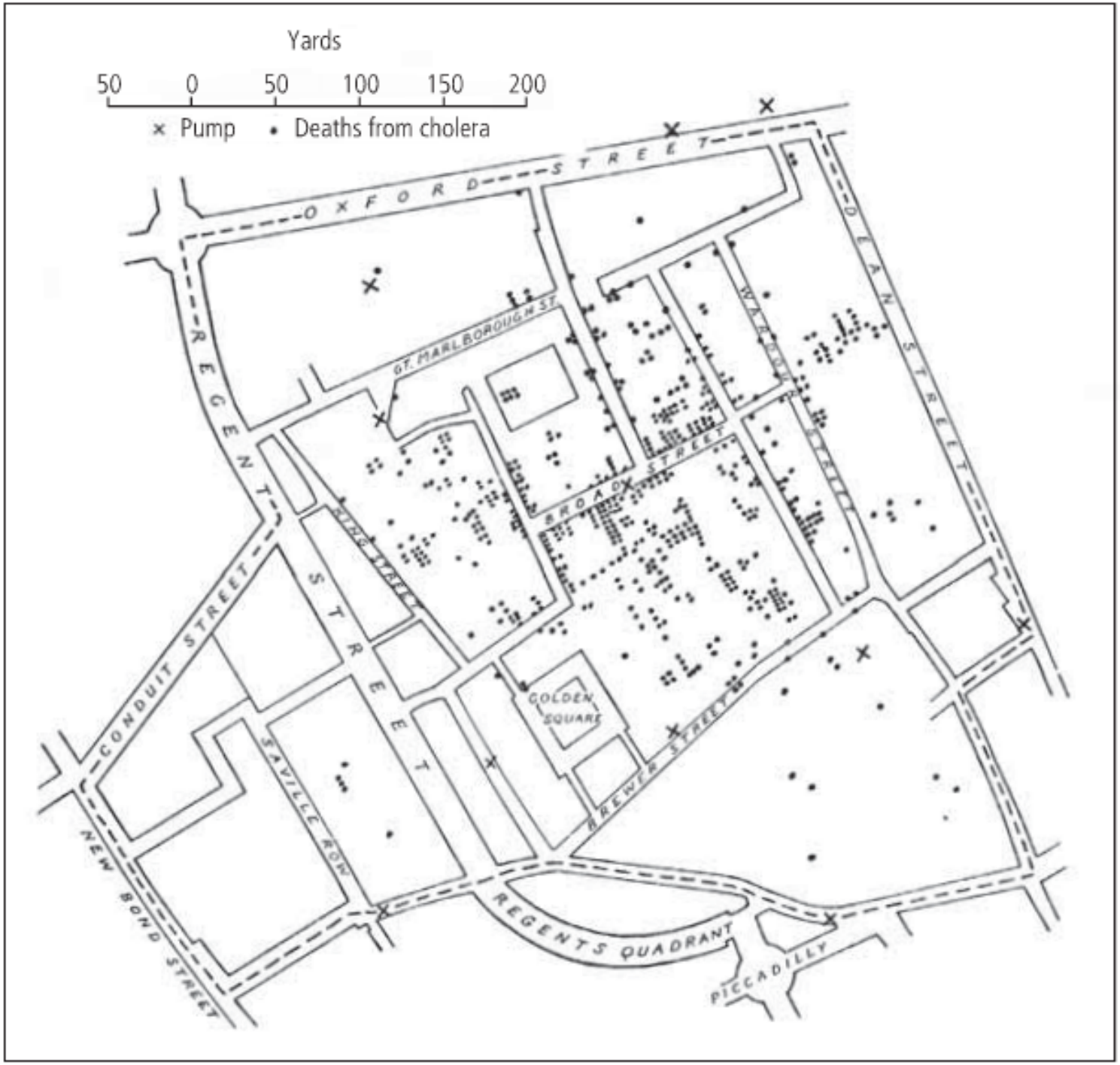

Yards
50 0 50 100 150 200
× Pump • Deaths from cholera
OXFORD STREET
ST. MARLBOROUGH ST.
BROAD STREET
WARWICK STREET
DEAN STREET
REGENT STREET
KING STREET
CONDUIT STREET
GOLDEN SQUARE
BREWER STREET
SAVILLE ROW
NEW BOND STREET
REGENTS QUADRANT
PICCADILLY

Table 1.1. Deaths from cholera in districts of London supplied by two water companies,[3] 8 July to 26 August 1854

Water supply company	Population 1851	Cholera deaths (n)	Cholera death rate (per 1000 population)
Southwark	167 654	844	5.0
Lambeth	19 133	18	0.9

Table 2.2. Differences between incidence and prevalence

	Incidence	Prevalence
Numerator	Number of **new** cases of disease during a specified period of time	Number of **existing** cases of disease at a given point of time
Denominator	Population at risk	Population at risk
Focus	Whether the event is a new case Time of onset of the disease	Presence or absence of a disease Time period is arbitrary; rather a "snapshot" in time
Uses	Expresses the risk of becoming ill The main measure of acute diseases or conditions, but also used for chronic diseases More useful for studies of causation	Estimates the probability of the population being ill at the period of time being studied. Useful in the study of the burden of chronic diseases and implication for health services

Note: If incident cases are not resolved, but continue over time, then they become existing (prevalent) cases. In this sense, prevalence = incidence × duration.

Figure 2.2. Factors influencing prevalence

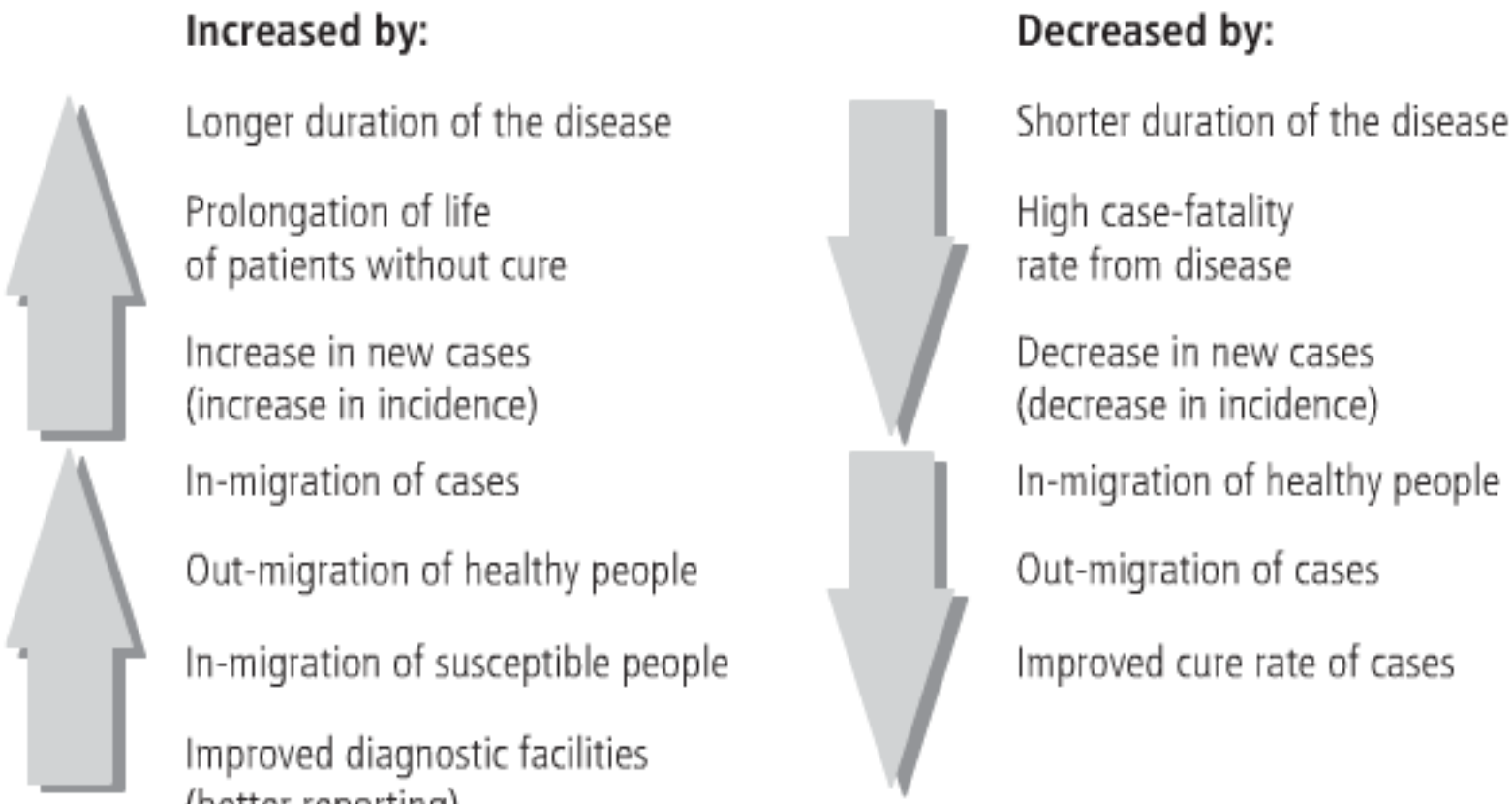


Figure 2.3. Calculation of disease occurrence

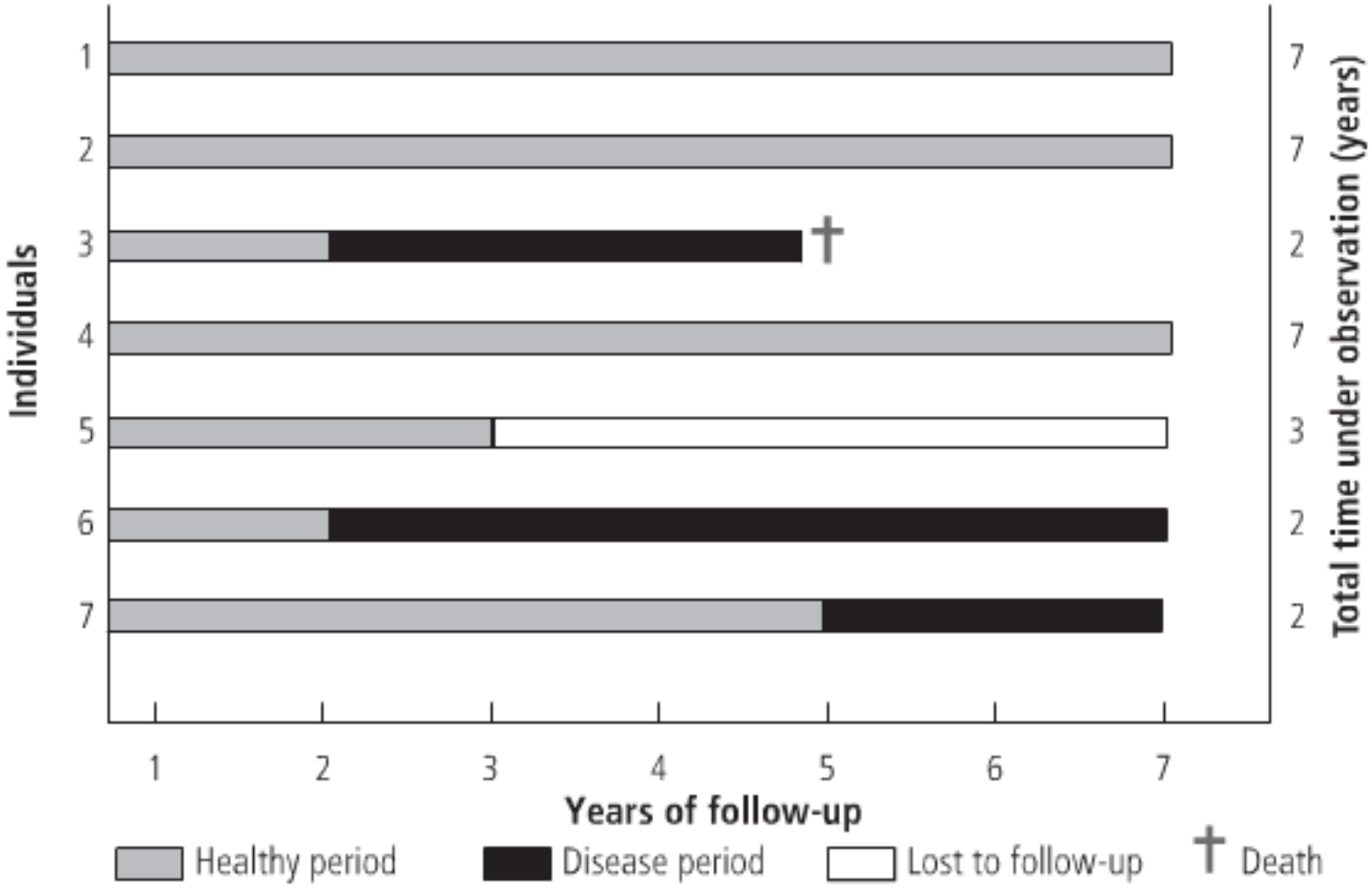


Table 3.1. Types of epidemiological study

Type of study	Alternative name	Unit of study
Observational studies		
Descriptive studies		
Analytical studies		
Ecological	Correlational	Populations
Cross-sectional	Prevalence	Individuals
Case-control	Case-reference	Individuals
Cohort	Follow-up	Individuals
Experimental studies	***Intervention studies***	
Randomized controlled trials	Clinical trials	Individuals
Cluster randomized controlled trials		Groups
Field trials		
Community trials	Community intervention studies	Healthy people Communities

Figure 3.10. Confounding: relationship between coffee drinking (exposure), heart disease (outcome), and a third variable (tobacco use)

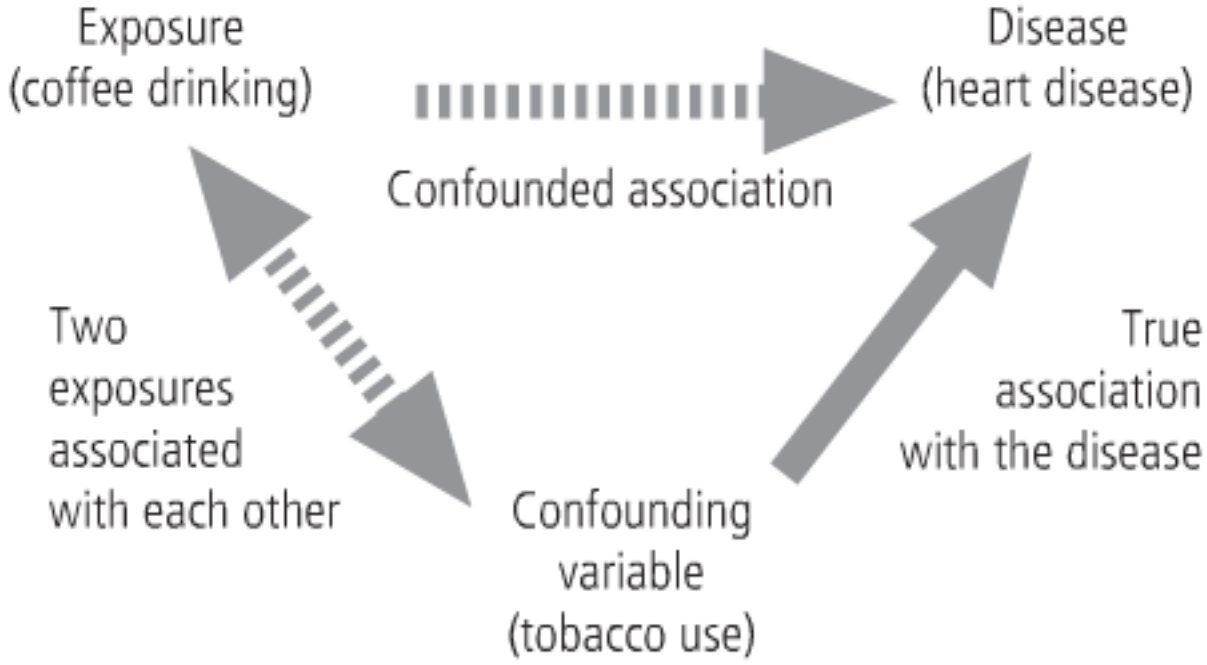
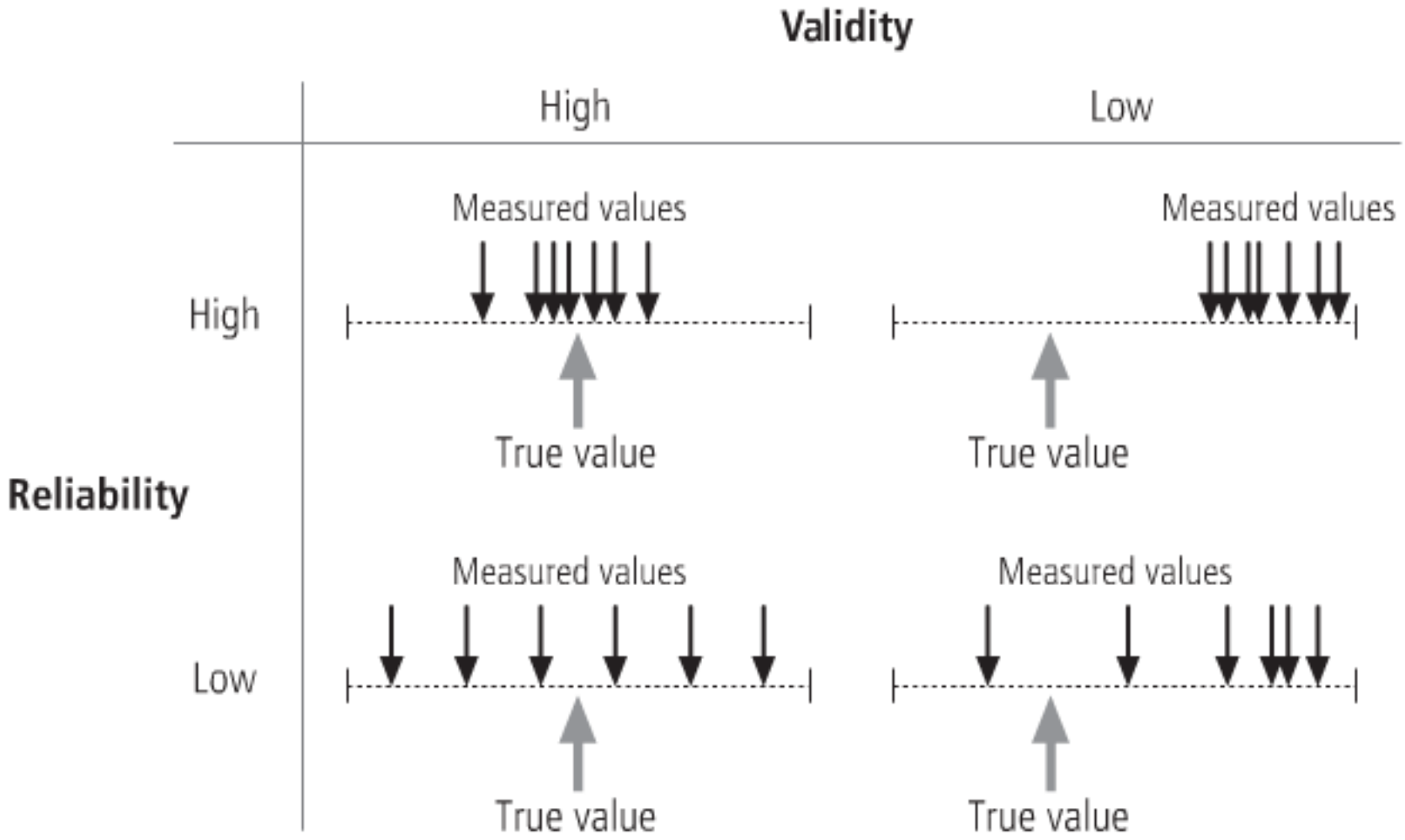


Figure 3.11. Validity and reliability



Comparing sequencing platforms in microbiome analysis

Platform	Method	Reads	16S	Metagenome	Notes
Sanger	Dideoxy terminator	750 bp	2-3 reads to cover	Good for database comparisons	Accurate, costly, slow
Pyroseq.	Light emission	400 bp			Good for 16S but not meta
Illumina	Flourescent step-by-step	100-150		More coverage makes up for short reads	High coverage, low cost
3rd generation	Electronic signal	10-100 kb		Great for assembly	Unknown error, usability

Bases to Bytes (Technology Review April 2012)

Cheap sequencing technology is flooding the world with genomic data.
Can we handle the deluge?