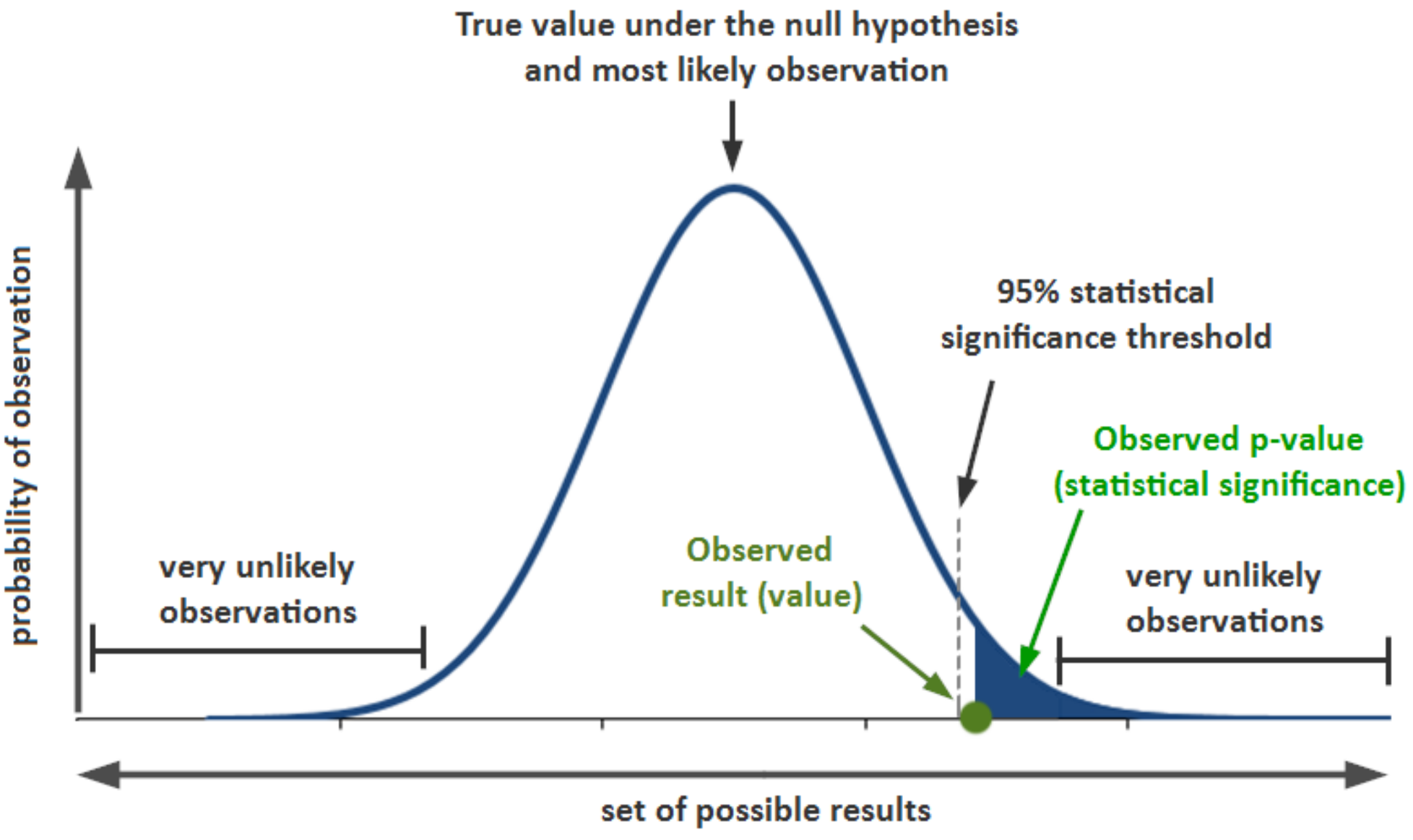
EVIE PIERCE

STATISTICAL FOUNDATIONS OF EPIDEMIOLOGY

OVERVIEW

- What is a p-value?
- What about Confidence Intervals?
- Statistical Foundations.

Probability & Statistical Significance Explained



ASA News
AMERICAN STATISTICAL ASSOCIATION
Promoting the Practice and Profession of Statistics®
732 North Washington Street, Alexandria, VA 22314 • (703) 684-1221 • Toll Free: (888) 231-3473 • www.amstat.org • www.twitter.com/AmstatNews

AMERICAN STATISTICAL ASSOCIATION RELEASES STATEMENT ON STATISTICAL SIGNIFICANCE AND *P*-VALUES

Provides Principles to Improve the Conduct and Interpretation of Quantitative Science

March 7, 2016

The American Statistical Association (ASA) has released a "Statement on Statistical Significance and *P*-Values" with six principles underlying the proper use and interpretation of the *p*-value [http://amstat.tandfonline.com/doi/abs/10.1080/00031305.2016.1154108#.Vt2XIOaE2MN]. The ASA releases this guidance on *p*-values to improve the conduct and interpretation of quantitative science and inform the growing emphasis on reproducibility of science research. The statement also notes that the increased quantification of scientific research and a proliferation of large, complex data sets has expanded the scope for statistics and the importance of appropriately chosen techniques, properly conducted analyses, and correct interpretation.

CONFIDENCE INTERVALS

- Interval estimate, calculated from the statistics of the observed data, that might contain the true value of an unknown population parameter

- If the analysis were done an infinite amount of times, 95% of the time, the true value would fall within the interval.

STATISTICAL FOUNDATIONS

EVALUATION OF DIAGNOSTIC TESTS

	Disease +	Disease -
Test +	TP	FP
Test -	FN	TN

SENSITIVITY (TRUE POSITIVE RATE)

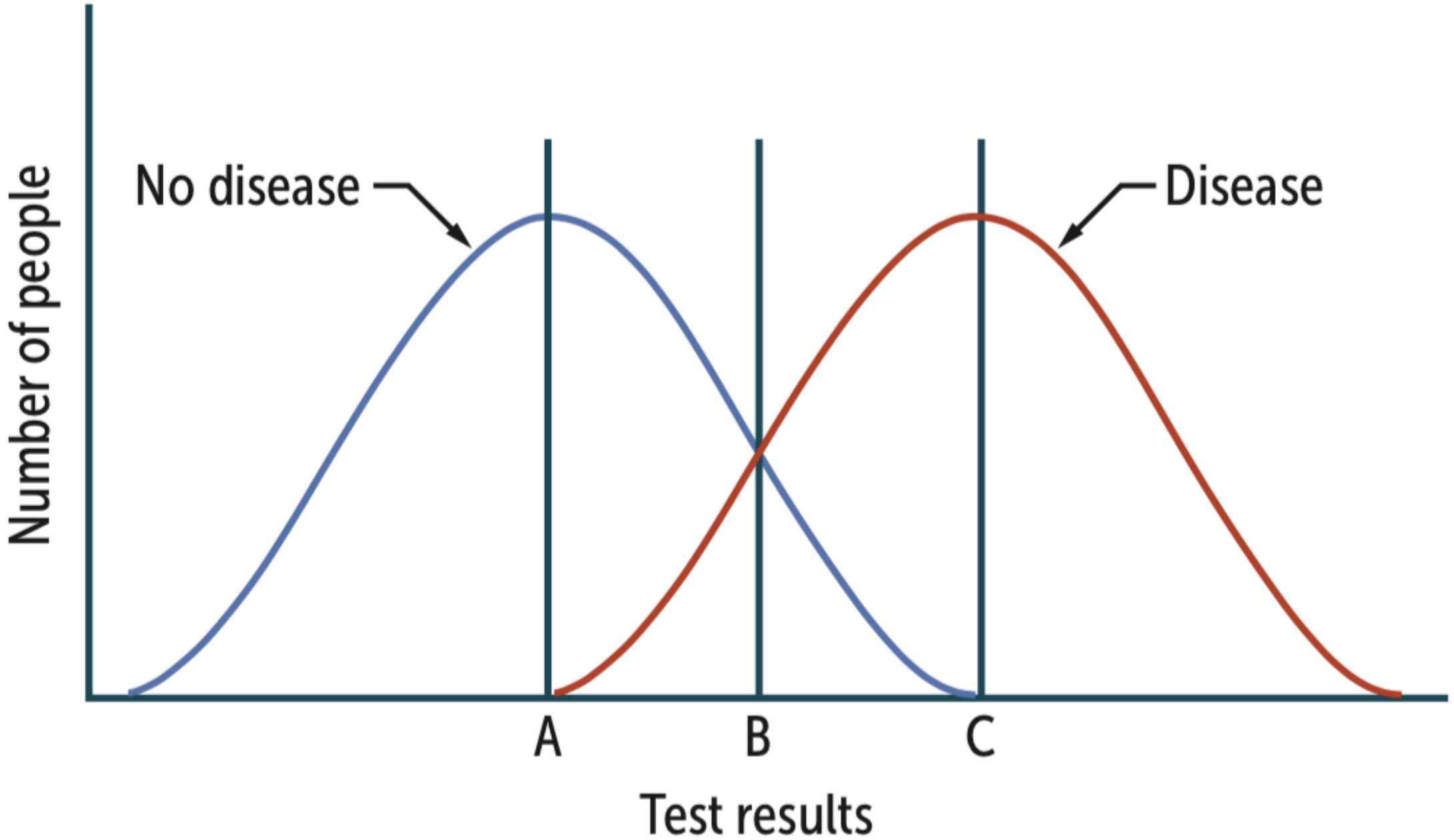
	Disease +	Disease -
Test +	TP	FP
Test -	FN	TN

$$\frac{TP}{(TP + FN)}$$

SPECIFICITY (TRUE NEGATIVE RATE)

	Disease +	Disease -
Test +	TP	FP
Test -	FN	TN

$$\frac{TN}{(TN + FP)}$$


No disease
Disease
Number of people
A
B
C
Test results

POSITIVE PREDICTIVE VALUE

	Disease +	Disease -
Test +	TP	FP
Test -	FN	TN

$$\frac{TP}{(TP + FP)}$$

NEGATIVE PREDICTIVE VALUE

	Disease +	Disease -
Test +	TP	FP
Test -	FN	TN

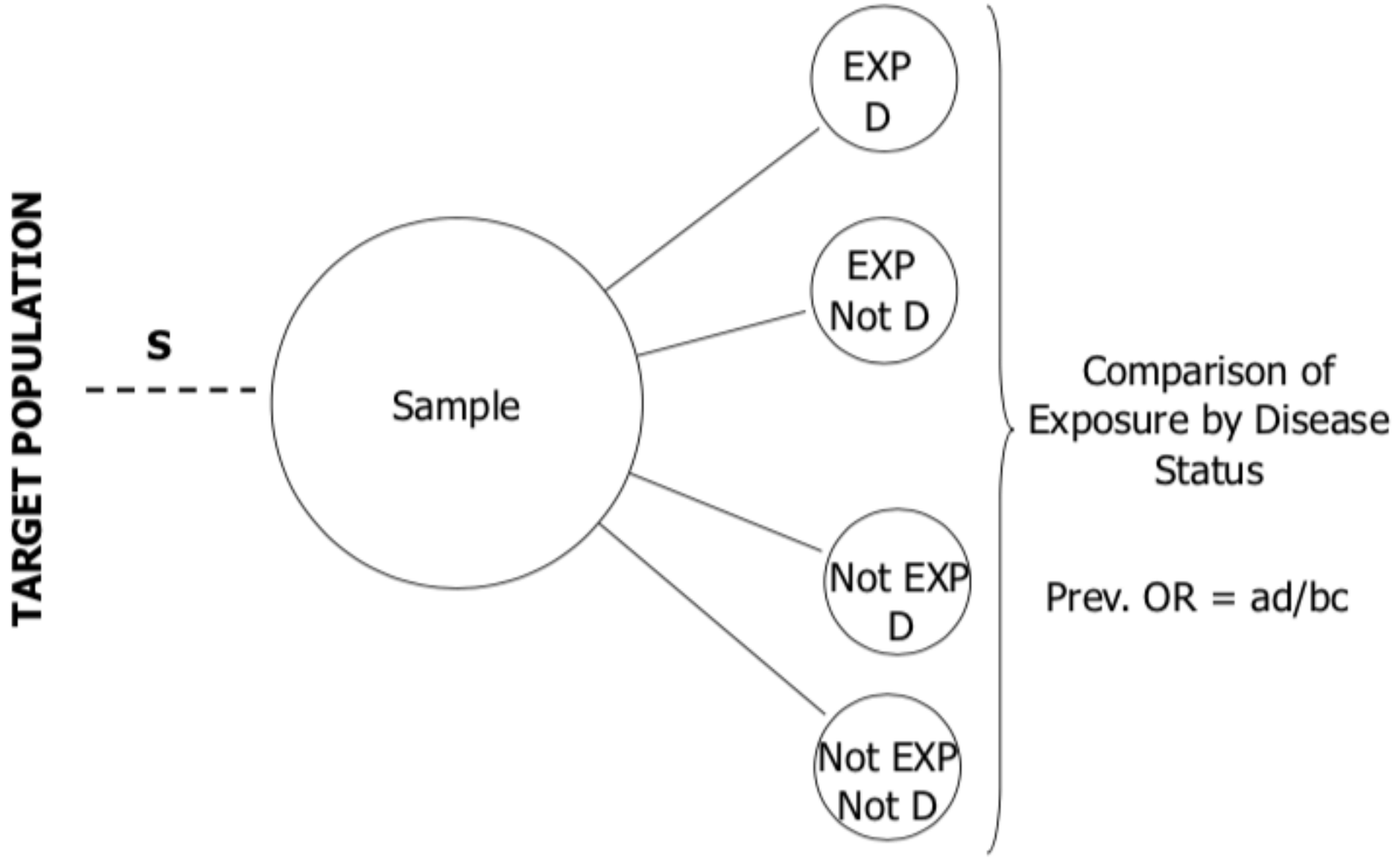
$$\frac{TN}{(TN + FN)}$$

CROSS-SECTIONAL

- Observational, all measurements are made on a single occasion.

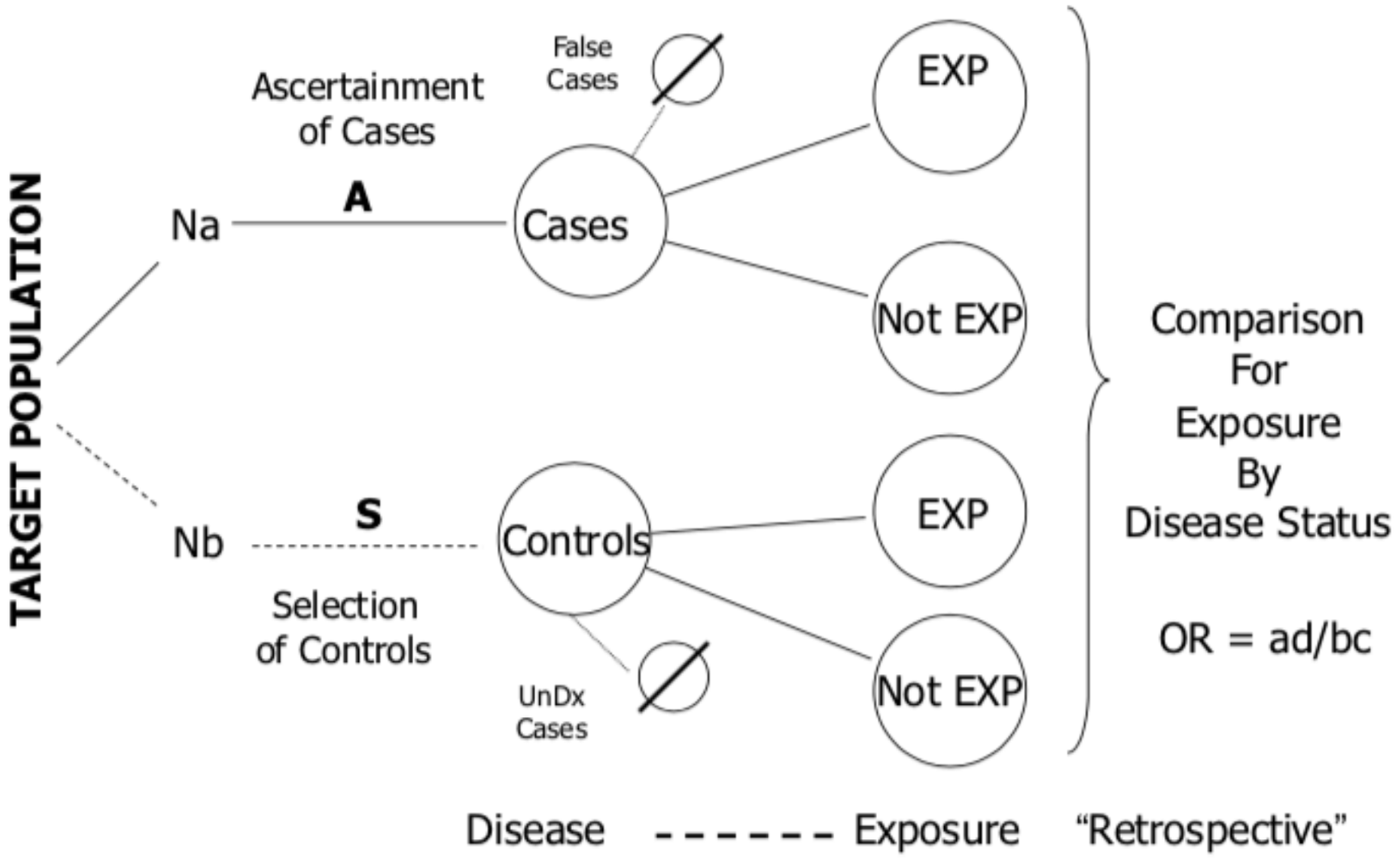
- Collects data from a group of people to assess frequency of disease (and related risk factors) at a particular point in time.

- Can show risk factor association with disease, but do not establish causality.


TARGET POPULATION
S
Sample
EXP
D
EXP
Not D
Not EXP
D
Not EXP
Not D
Comparison of
Exposure by Disease
Status
Prev. OR = ad/bc

CASE-CONTROL

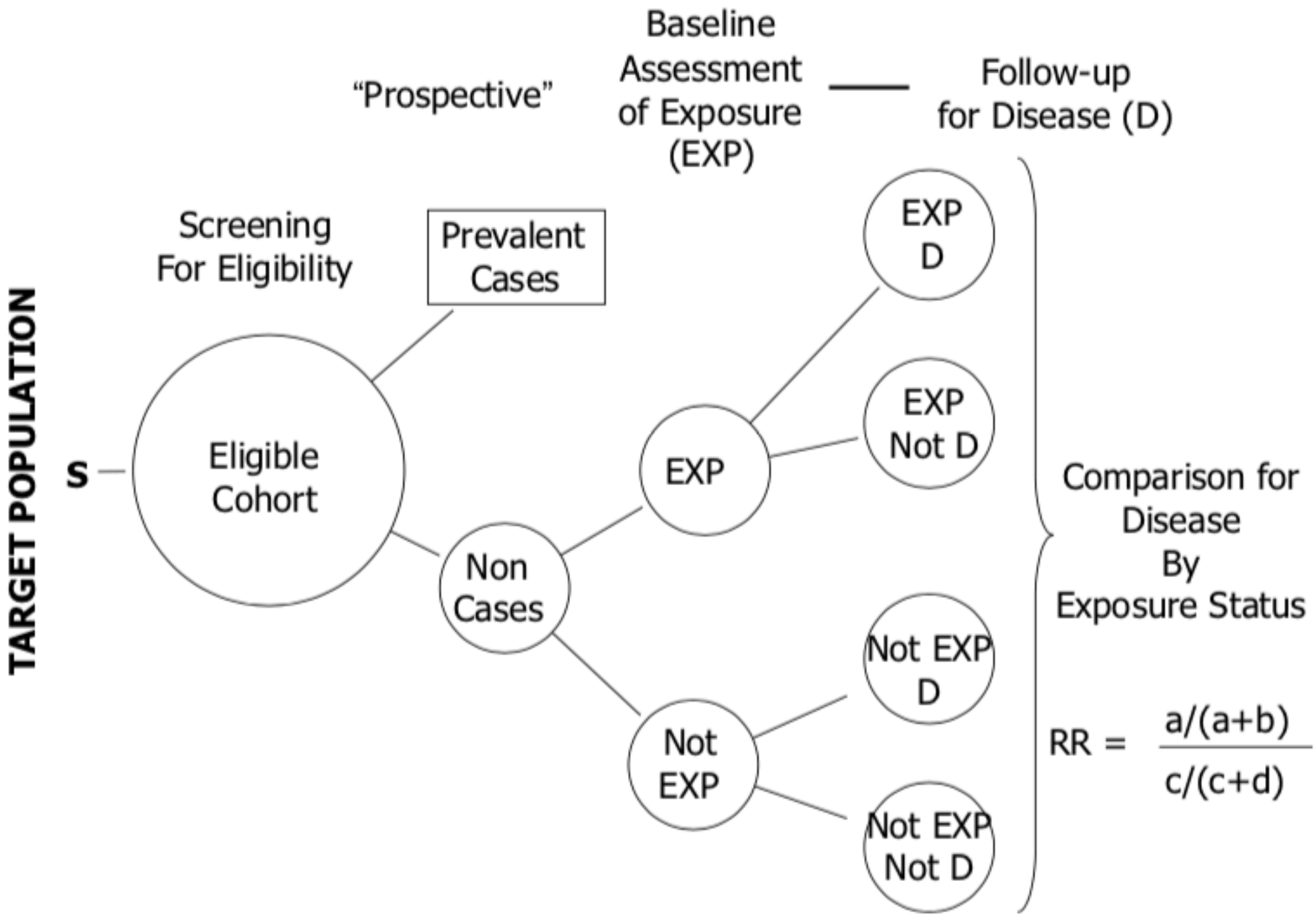
- Compares a group of people with disease to a group without disease.
- Selection of cases based on disease state


TARGET POPULATION
Na
A
Ascertainment of Cases
False Cases
Cases
EXP
Not EXP
Nb
S
Selection of Controls
Controls
UnDx Cases
EXP
Not EXP
Comparison For Exposure By Disease Status
OR = ad/bc
Disease ------ Exposure "Retrospective"

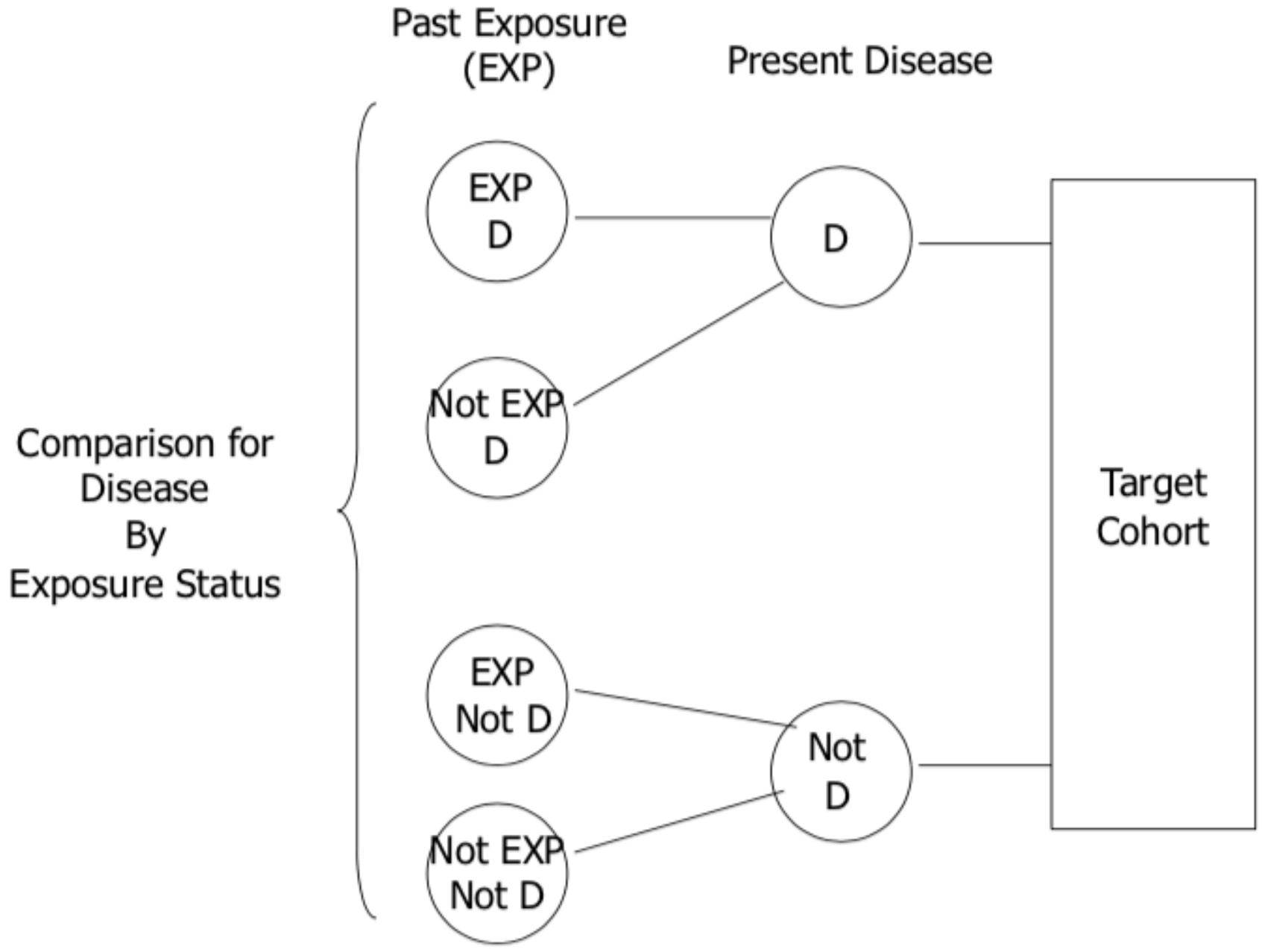
COHORT

- Compares a group with a given exposure or risk factor to a group without such exposure.
- Selection based on exposure to risk factor.
- Looks to see if exposure changes the likelihood of disease.
- Can be prospective or retrospective.

Cohort Study Design



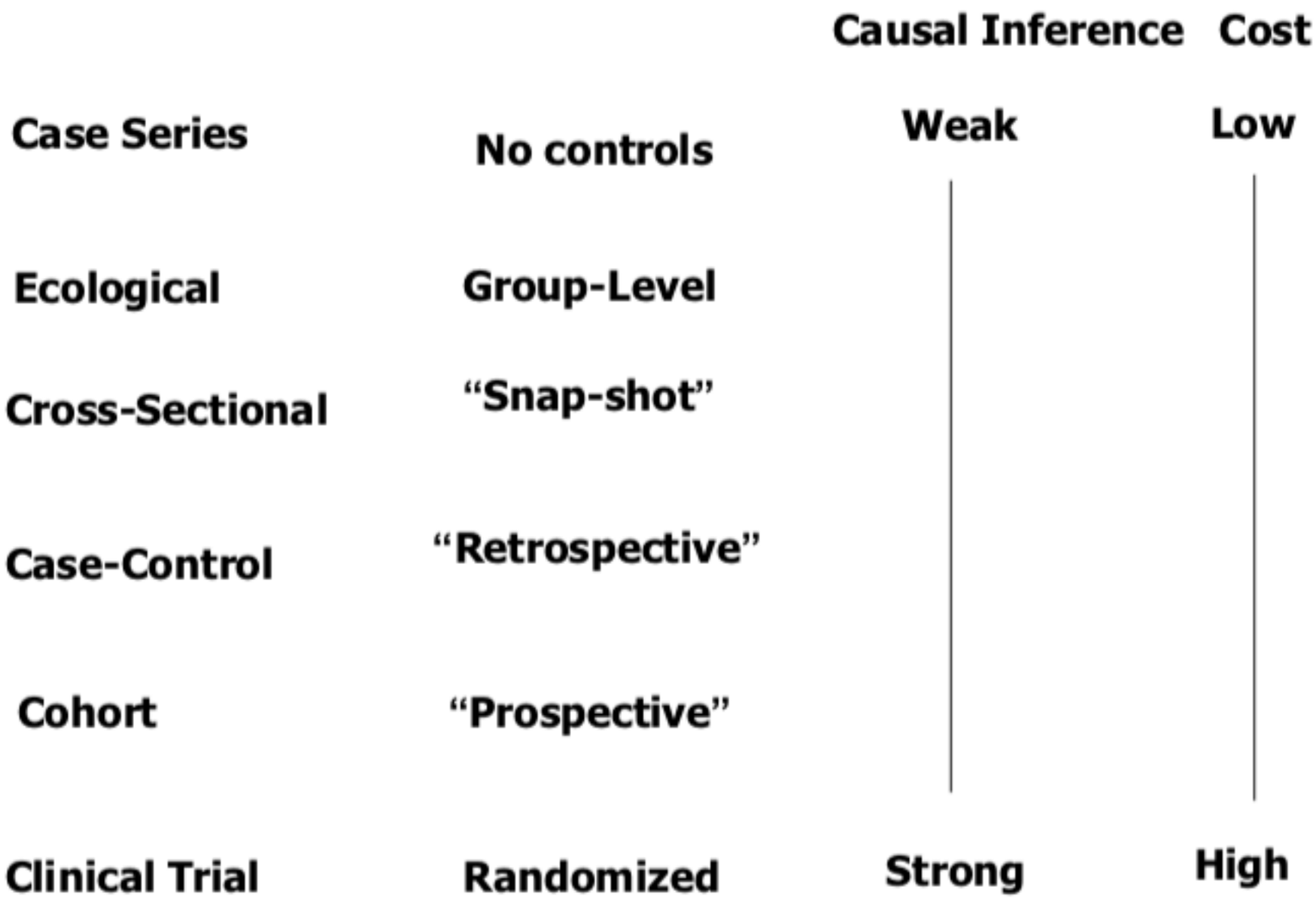
Retrospective Cohort Study Design



COMBINATION STUDIES

- Nested Case Control
- Case-Cohort
- Case Cross-Over

CAUSAL INFERENCE



QUANTIFIABLE RISK

	Disease +	Disease -
Exposure +	A	B
Exposure -	C	D

ODDS RATIO

	Disease +	Disease -
Exposure +	A	B
Exposure -	C	D

$$OR = \frac{A \times D}{B \times C}$$

RELATIVE RISK

	Disease +	Disease -
Exposure +	A	B
Exposure -	C	D

$$RR = \frac{A / (A + B)}{C / (C + D)}$$

ATTRIBUTABLE RISK

	Disease +	Disease -
Exposure +	A	B
Exposure -	C	D

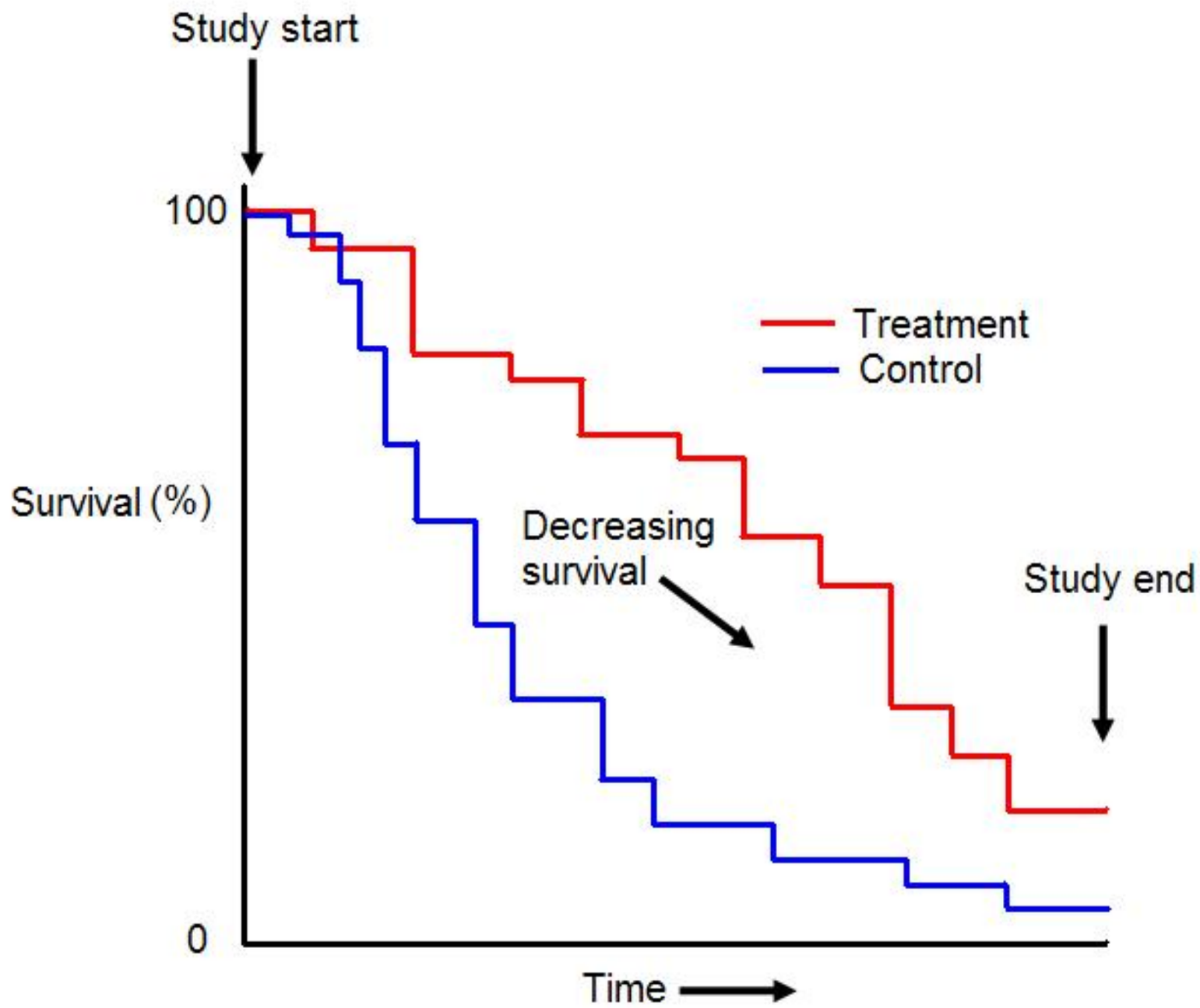
$$AR = \frac{A}{A + B} - \frac{C}{C + D}$$

NUMBER NEEDED TO TREAT AND NUMBER NEEDED TO HARM

$$\text{NNT} = \frac{1}{\text{ARR}}$$

$$\text{NNH} = \frac{1}{\text{AR}}$$

HAZARD RATIOS



RECRUITING PARTICIPANTS

- Selection bias
    - Berkson bias
    - Loss to follow-up
    - Healthy worker and volunteer biases

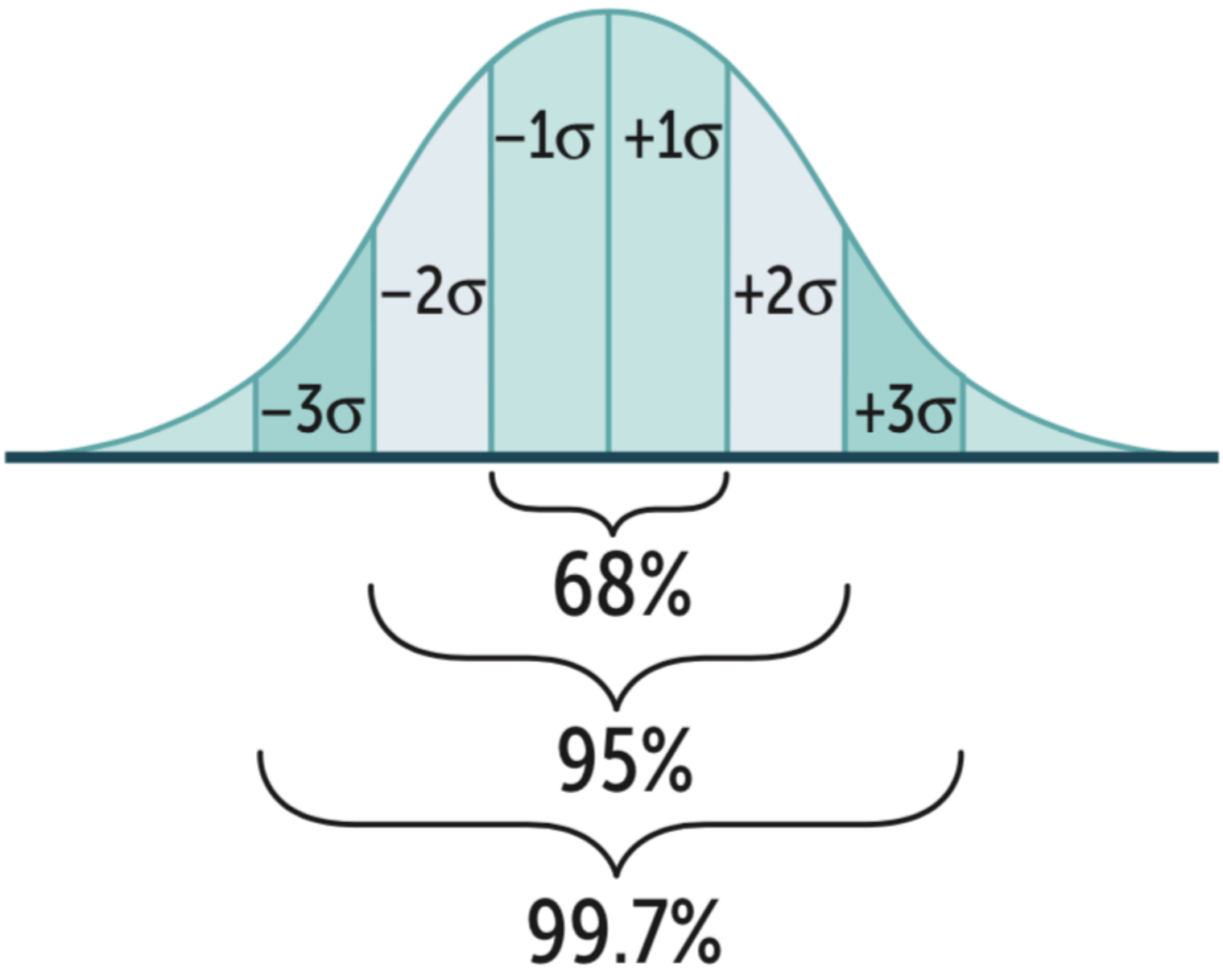
PERFORMING STUDY

- Recall bias
- Measurement bias
- Procedure bias
- Observer-expectancy bias

INTERPRETING RESULTS

- Confounding bias
- Lead-time bias

DISTRIBUTION



T-TEST VS. ANOVA VS. X^2

- T-Test: differences between means of 2 groups
- ANOVA: differences between means of 3 or more groups
- X^2: difference between 2 or more percentages or proportions of categorical outcomes