Statistics – benefits and limitations

Valérie Decraene

PhD, EPIET Fellow

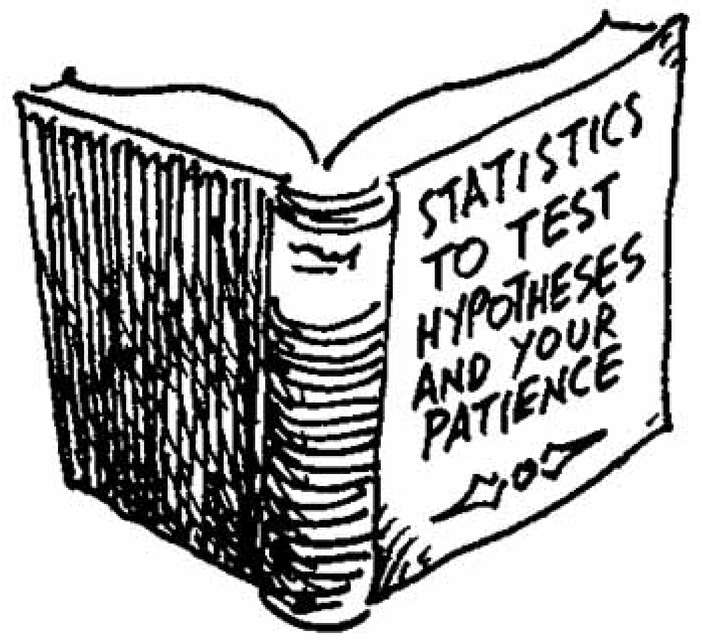
Swedish Institute for Infectious Disease Control

EpiTrain VI, Tallinn

Friday 10 December, 2010

First . . .

- What do you think about statistics?



Learning objectives

- To understand:
  - the importance of describing your data using confidence intervals
  - the problem with over-emphasising that ‘p<0.05’
  - the importance of reporting confidence intervals together with p-values in scientific communication

Study hypothesis

In our hospital, the prevalence of nosocomial infections is the same among men as among women (null hypothesis)

Conclusions of the study

Women have a higher prevalence of nosocomial infections ($p<0.05$)

→ the null hypothesis is rejected

EpiNorth

P-value

- P-value = probability, under the null hypothesis, to observe this result or more extreme results

- Large p-value: Data are consistent with the null hypothesis (of no association)
- Small p-value: Data are **<u>not</u> very consistent** with the null hypothesis

EpiNorth

The problem with such hypothesis testing

- It reduces epidemiology to a simple dichotomy: **yes or no**
  - Nature is much more interesting
  - The real picture is not usually just black or white

- Epidemiology is not about looking for "statistical significance"

EpiNorth

Epidemiology is about measurement (not hypothesis testing!)

- In epidemiology **we measure the size** of the occurrence, association, or effect
- Study purpose: define what to measure
- The result: an estimate of an epidemiological measure (of an epidemiological quantity)
  - Prevalence, incidence proportion, incidence rate
  - Prevalence ratio, risk ratio, incidence rate ratio, odds ratio

EpiNorth

Specific objective of the study

How much greater is the prevalence of nosocomial infections among women than among men?

Conclusion of the study

The relationship (ratio) between the prevalence of nosocomial infections among men and women is:

PR = **1,6**

(95% confidence interval: **1,1 – 2,1**)

EpiNorth

Which one is more informative?

Hypothesis testing

$p < 0,05$

- Tells us if the data is consistent with the null hypothesis
- Does **<u>not</u>** tell us about the strength of the association

Measurement (estimation)

1,6 (1,1 – 2,1)

- Gives the **strength** of the association (1,6)
- Gives the **precision** of the association (1,1 – 2,1)
- Gives the **direction** of the association

Which one is more informative?

Hypothesis testing

p < 0,05

- Tells us if the data is consistent with the null hypothesis
- Does **not** tell us about the strength of the association

Measurement (estimation)

1,6 (1,1 – 2,1)

- Gives the **strength** of the association (1,6)
- Gives the **precision** of the association (1,1 – 2,1)
- Gives the **direction** of the association

The p-value is not a good description of the information contained within the data

EpiNorth

Which question is the smartest?

Hypothesis testing

- Does the car cost more than 50 000 euros?
- Answer: **Yes** (p<0,05)

Measurement (estimation)

- How much does the car cost?
- Answer: **77 000 euros** (73 000 – 80 000)

EpiNorth

Terminology

Point estimate

Confidence interval

PR = 1,6 (1,1 – 2,1)

Lower confidence limit

Upper confidence limit

EpiNorth

Epidemiology and statistics

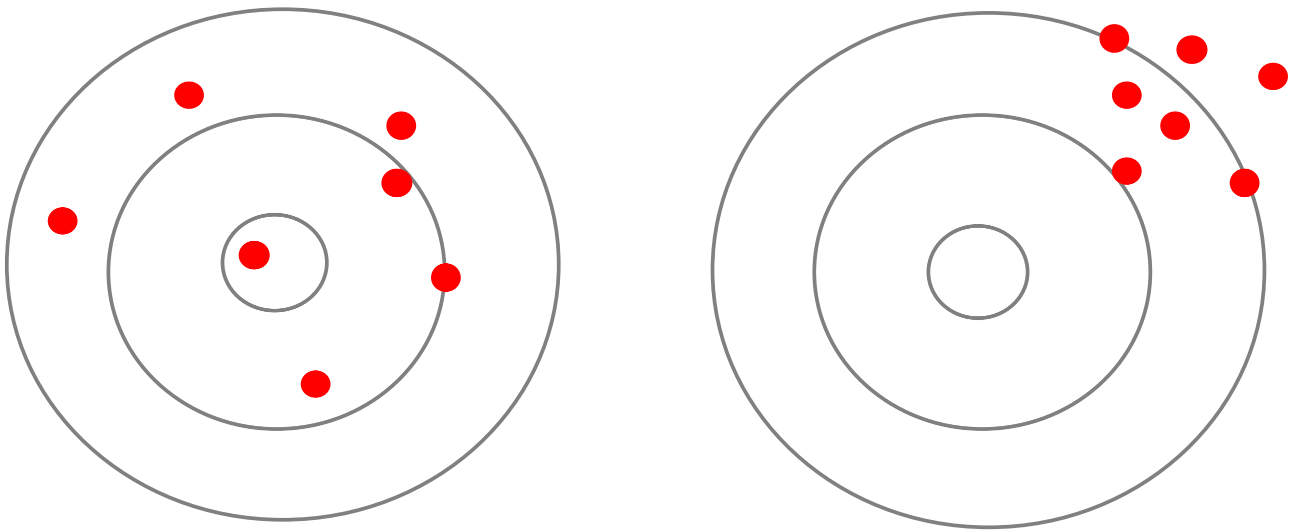
Epidemiology

- The study of the **occurrence, distribution** and **causes** of disease in the population and of the effect of **interventions** implemented to control the disease

Statistics

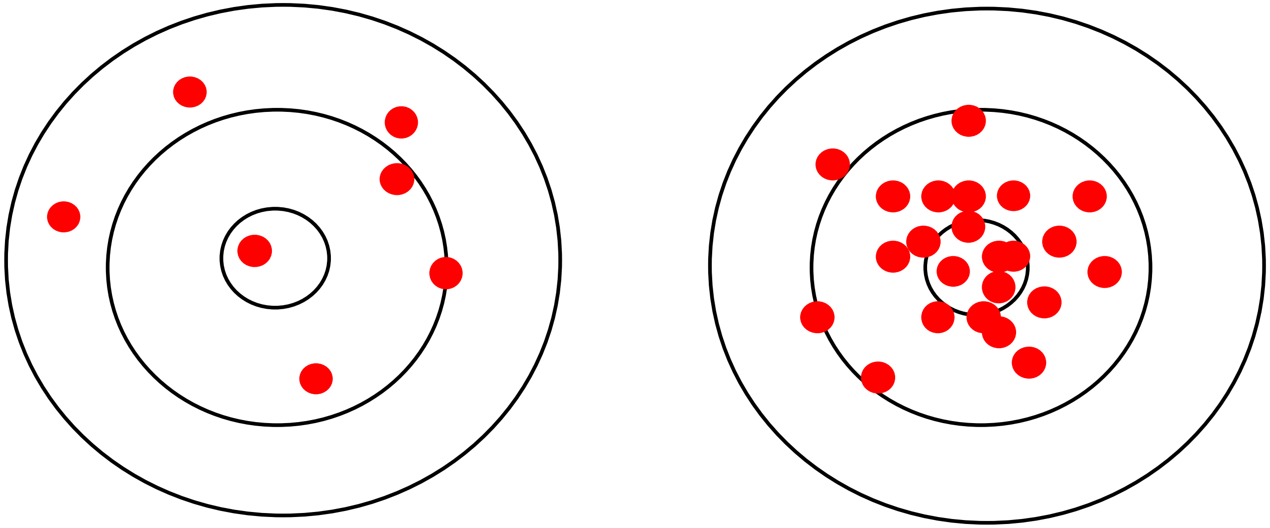
- A branch of mathematics
- Used in epidemiology for two purposes:
  - To consider variability
  - To measure effects after controlling for confounding

Random error versus systematic error (bias)



EpiNorth

Precision in the measurement



EpiNorth

Confidence Interval

- Indicates the amount of **random error** in the measurement
  - Requires that no bias be present
  - Requires a statistical model
- Formal definition: If the measurement could be repeated many times, the correct value would fall within the specified range 95% of the time

EpiNorth

Precision

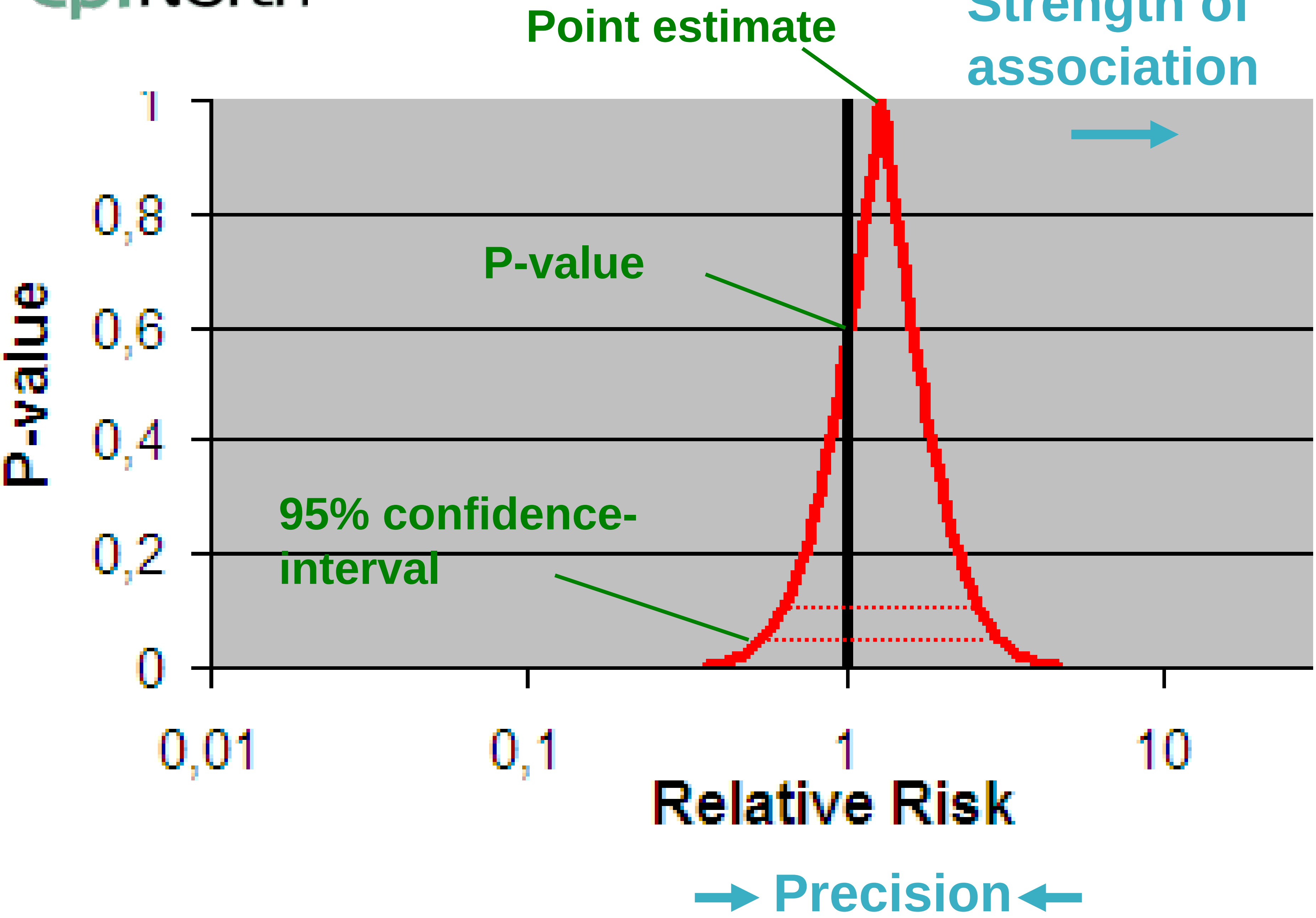
- Pragmatic definition: we can be 95% confident that the correct value lies within the confidence interval
- **Wide** confidence interval: **low** precision
  - Small study
- **Narrow** confidence interval: **high** precision
  - Big study
- Are all the values in the confidence interval equally probable?

EpiNorth

Confidence interval graphs

- A graph that shows all possible confidence intervals for the estimate
- The stronger the positive association, the further to the right the curve will be
- The width of the curve indicates the precision
- The graph can be easily drawn in Episheet based on the confidence interval

EpiNorth



Point estimate
Strength of association
P-value
95% confidence-interval
1
0,8
0,6
0,4
0,2
0
P-value
0,01
0,1
1
10
Relative Risk
Precision

EpiNorth

Hypothesis: Diabetics have the same prevalence of nosocomial infections

	Number with infection	Number of patients	Prevalence
Diabetics	12	76	15,8 %
Other	21	235	8,9 %

Prevalence ratio, PR = (12/76) / (21/235) = 1,8

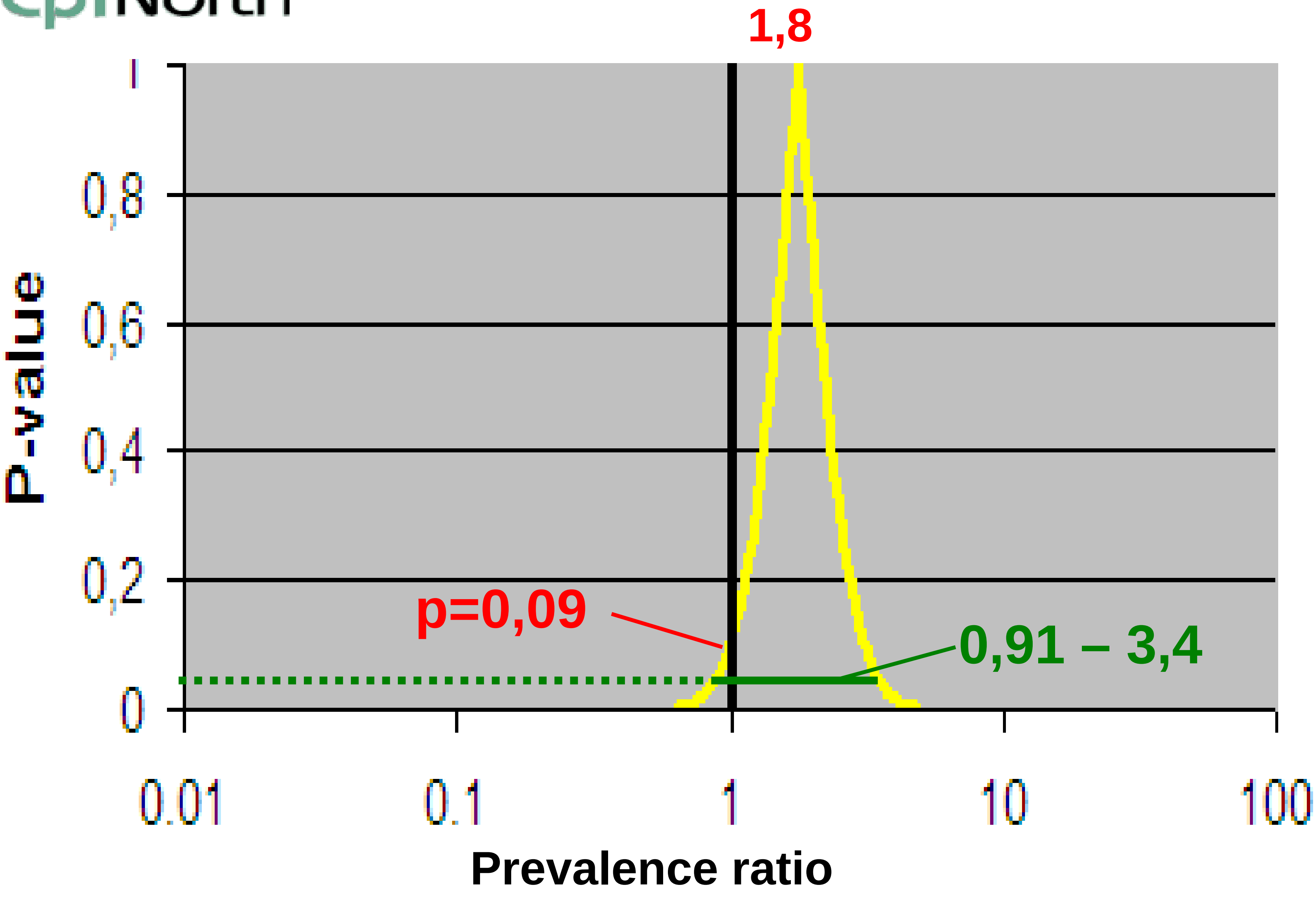
p = 0,09 (not significant)

So diabetics do *not* have a higher risk?

PR = 1,8 (0,91 – 3,4)

”The confidence interval includes 1 so the association is not significant” (?!)

EpiNorth



1,8
P-value
1
0,8
0,6
0,4
0,2
0
p=0,09
0,91 – 3,4
0.01
0.1
1
10
100
Prevalence ratio

EpiNorth

Hypothesis: Diabetics have the same prevalence of nosocomial infections. Investigated at all hospitals in the country

	Number with infection	Number of patients	Prevalence
Diabetics	140	890	15,7 %
Other	607	7101	8,5%

Prevalence ratio, PR = (140/890) / (607/1701) = 1,8

$p < 0.001$ (significant)

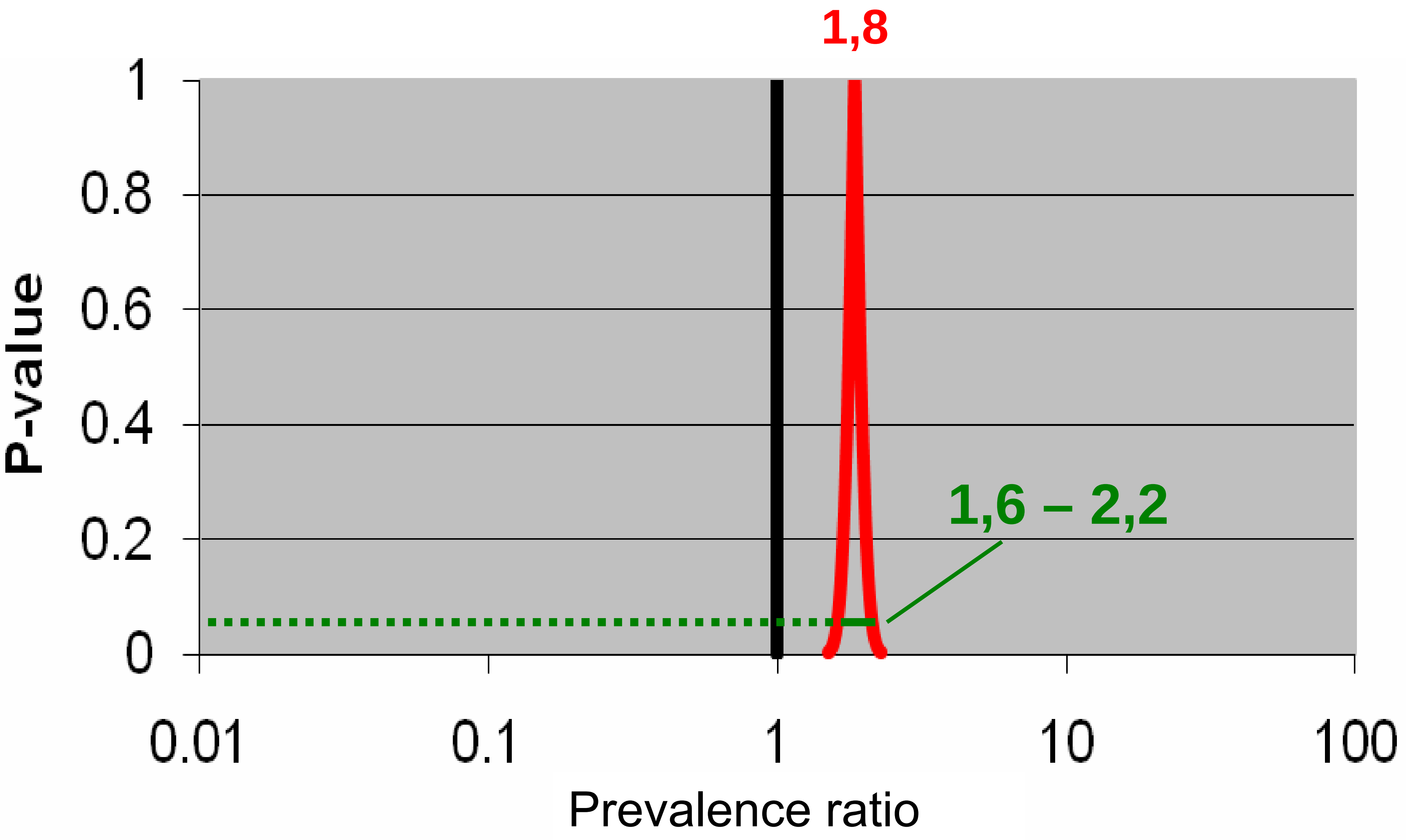
EpiNorth

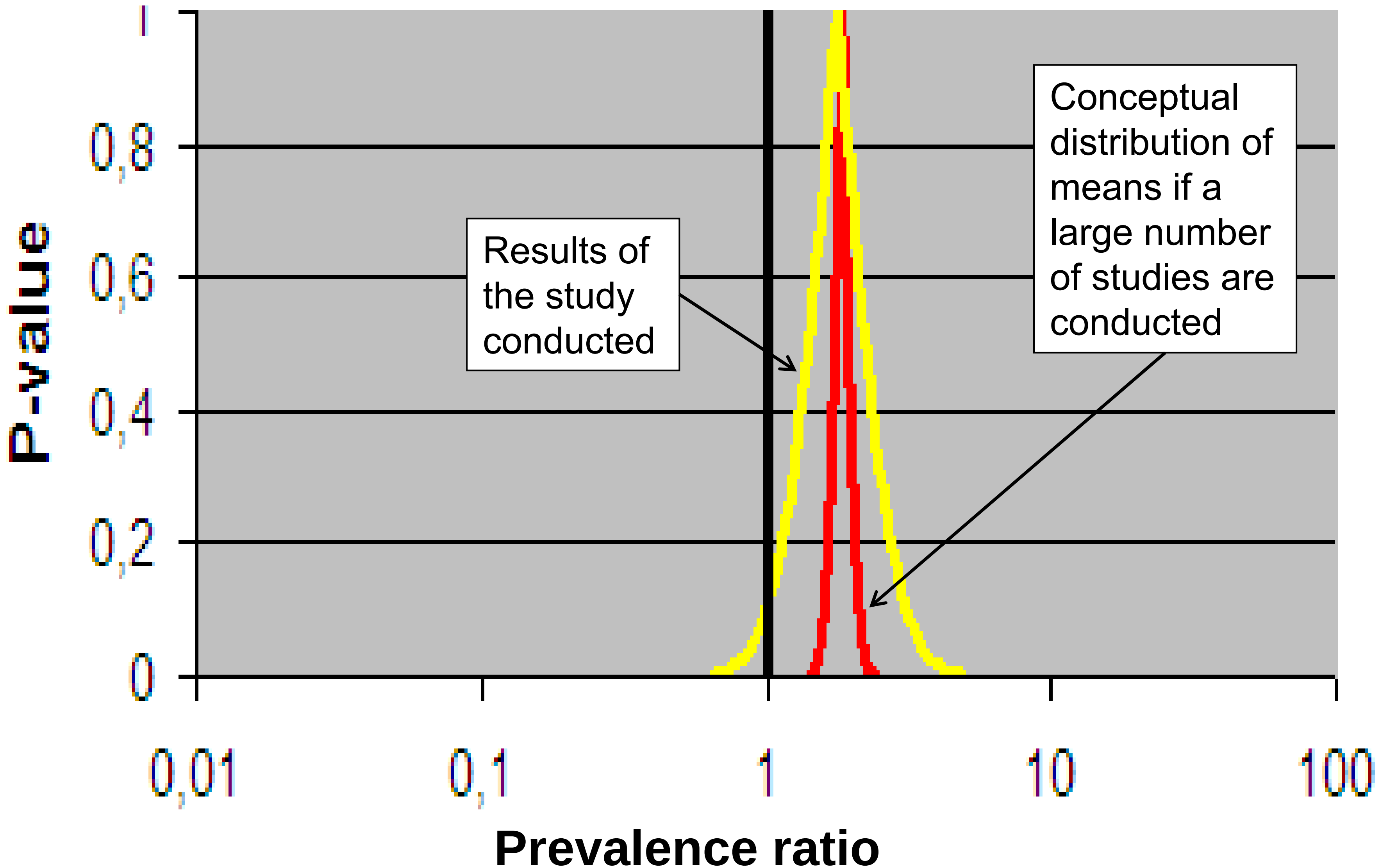
So diabetics do *not* have a higher risk?

PR = 1,8 (1,6 – 2,2)

”The confidence interval does not include 1 so the association is significant”(?!)

EpiNorth





P-value
1
0,8
0,6
0,4
0,2
0
Results of the study conducted
Conceptual distribution of means if a large number of studies are conducted
0,01
0,1
1
10
100
Prevalence ratio

Conclusions

- Significance testing evaluates only the role of chance as alternative explanation of observed difference or effect

- Confidence intervals are more informative than p-values

EpiNorth

General advice

- Do not rely only on p-values and hypothesis testing
- Calculate confidence intervals!
  - Study the **confidence interval graph**
- Always look at the raw data
  - Don’t forget about study size
- Interpret ’significant’ associations with caution

EpiNorth

Advice for scientific communication

- Concentrate on **measuring**
  - Use words like measure, estimate, determine the study objectives; not ”test”
- Always present the data together with BOTH p-values AND confidence intervals
- When reporting p-values, report them precisely ($p=0.027$ instead of $p<0.05$)

EpiNorth

Calculating confidence intervals for measures of disease occurrence

Proportions

- Prevalence
- Incidence proportion, cumulative incidence
- Case fatality

Episheet → Quickcalc → Proportion

Rates

- Incidence rate
- Mortality rate

Episheet → Quickcalc → Rate

EpiNorth

Calculating confidence intervals for measures of association

Cohort study

- If the measure of disease occurence is a *proportion*
- Risk ratio (RR)

Episheet → Risk data

- If the is measure of disease occurence is an incidence *rate*
- Incidence rate ratio (IRR)

Episheet → Rate data

Case-control study

- Odds ratio (OR)

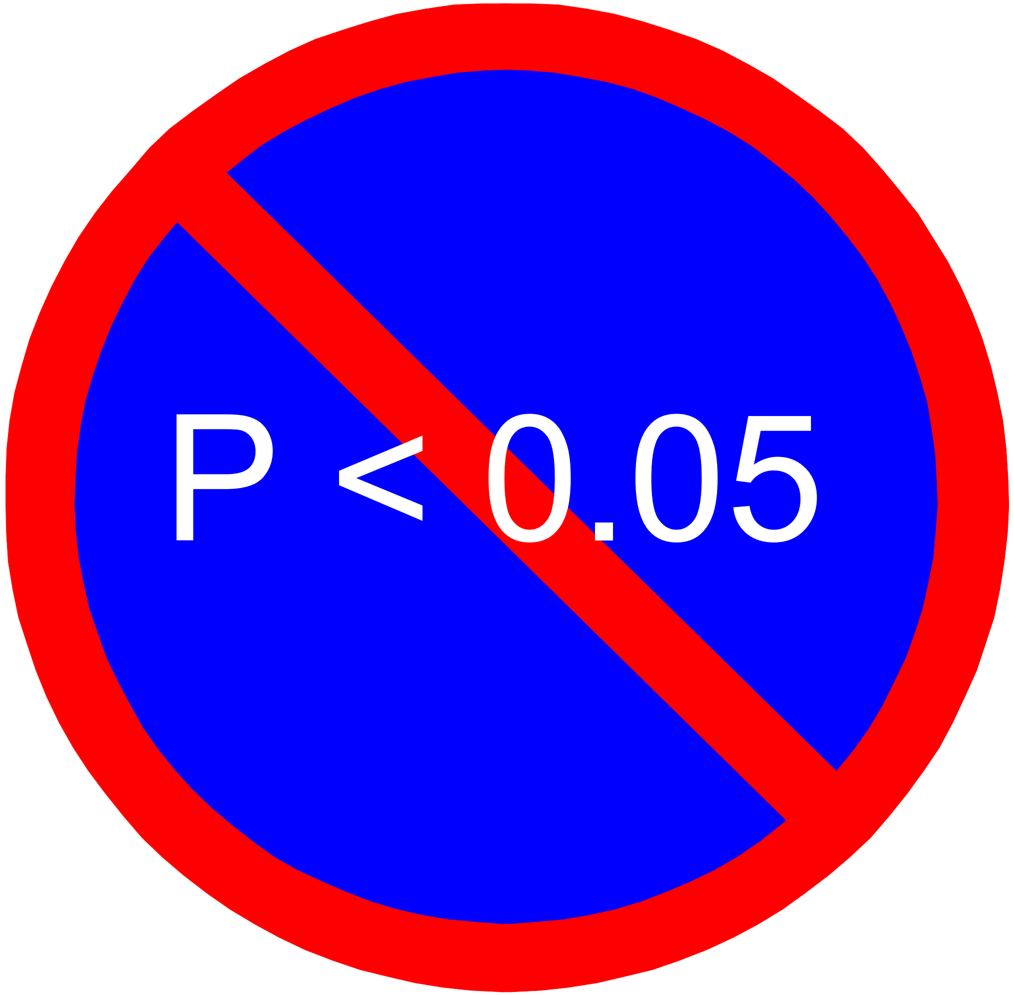
Episheet → Case-control data

If you have many potential causal factors to analyse, it is easier in Stata, SPSS, EpiInfo

“Epidemiological application needs more than a decision as to whether chance alone could have produced association.”

(KJ Rothman, 2002)

EpiNorth



Reference

- KJ Rothman, S Greenland, TL Lash. *Modern Epidemiology.* Lippincott, Williams & Wilkins, Philadelphia, USA, 2008.

Thank you for your attention!