Revision of introduction to epidemiology

Dr David A McAllister

Wellcome Trust Intermediate Clinical Fellow

Aim

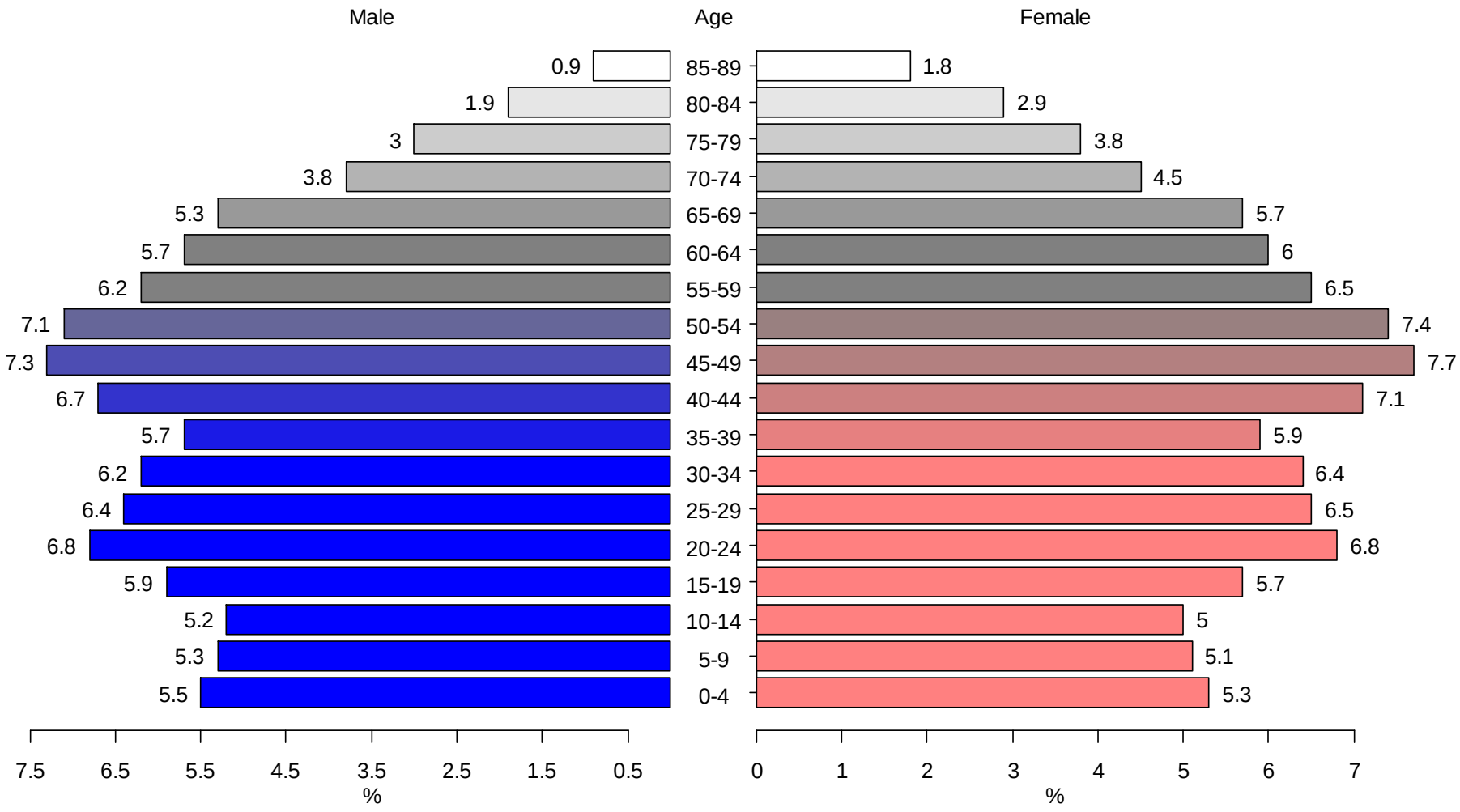
- To review the epidemiological concepts covered in previous courses

BASIC EPIDEMIOLOGICAL TERMS

Mortality rates

- Am I worried?
  - 50 people died from flu
  - 50 people in Glasgow died from flu
  - 50 people in Glasgow died from flu in past 100 years
- Meaningful statistics need
  - A denominator population
  - A time frame

Scotland’s population 2013



Other denominators

- Health board
- City
- Hospital
- Disease register
- Recruited to a study

THE DENOMINATOR MUST CORRESPOND TO THE NUMERATOR

Two sorts of time

- Person-time
  - 10 deaths per 10,000 person years
    - 10,000 people for 1 year
    - 5,000 people for 2 years
    - 2,000 people for 5 years
- n-year follow-up
  - 5-year mortality of 10 per 10,000 people
- **Without denominator population and time death rates are meaningless**

Incidence

- Number of new cases
  - person-time, eg 5 per 10,000 person years
  - n years of follow-up, 10-year cancer incidence rate is 1 per 10,000

Prevalence

- Proportion of population that has disease
  - at a specified time, eg 1% of population had COPD in 2010
  - over a specified period, eg lifetime prevalence of dementia 40%

Compare/contrast

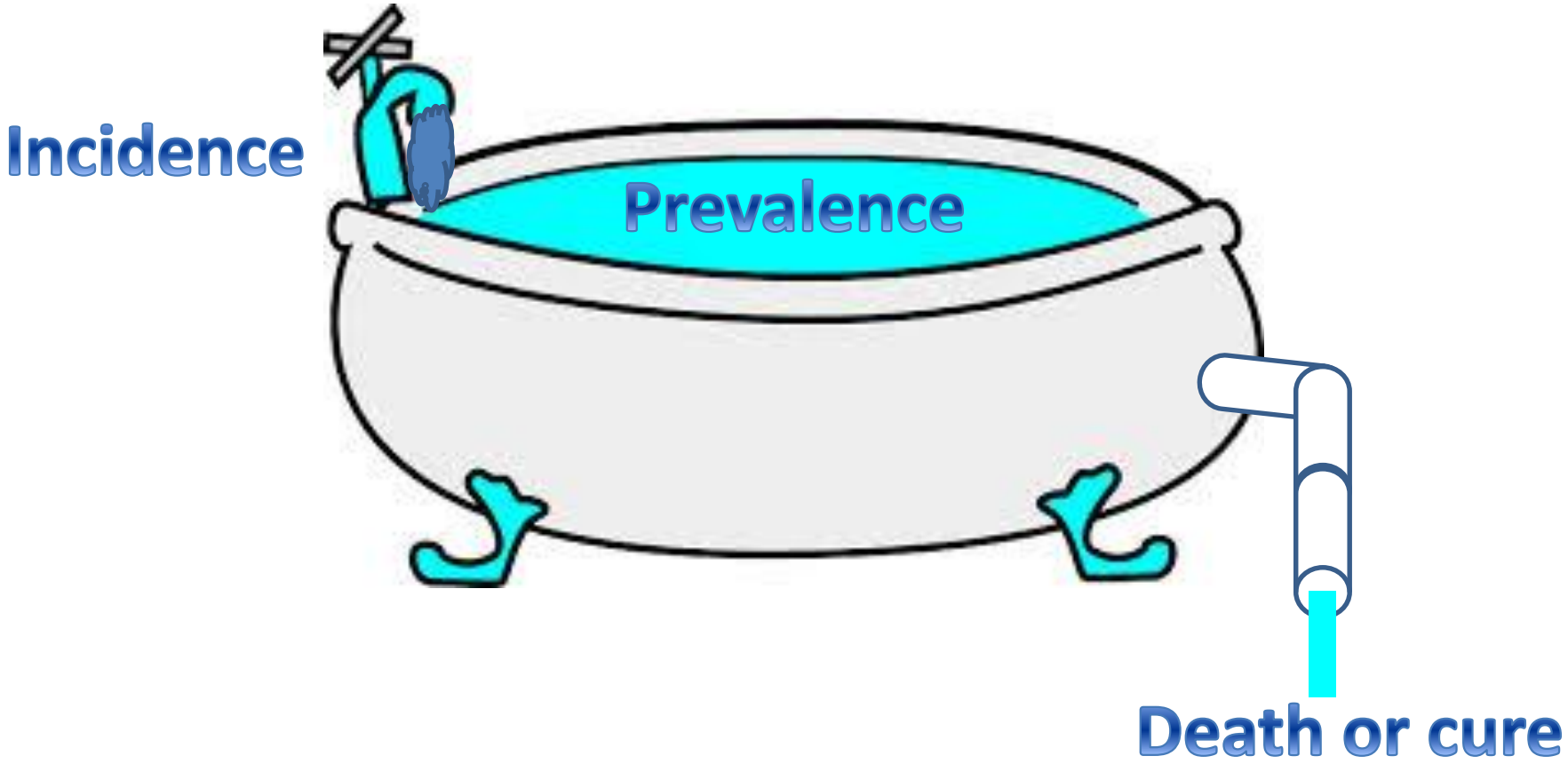
Prevalence

- A proportion
- Useful for planning services
- Depends partly on incidence

Incidence

- A rate or a proportion
- Useful for identifying causes of diseases
- Occurs, by definition, only in people without the disease

How incidence relates to prevalence

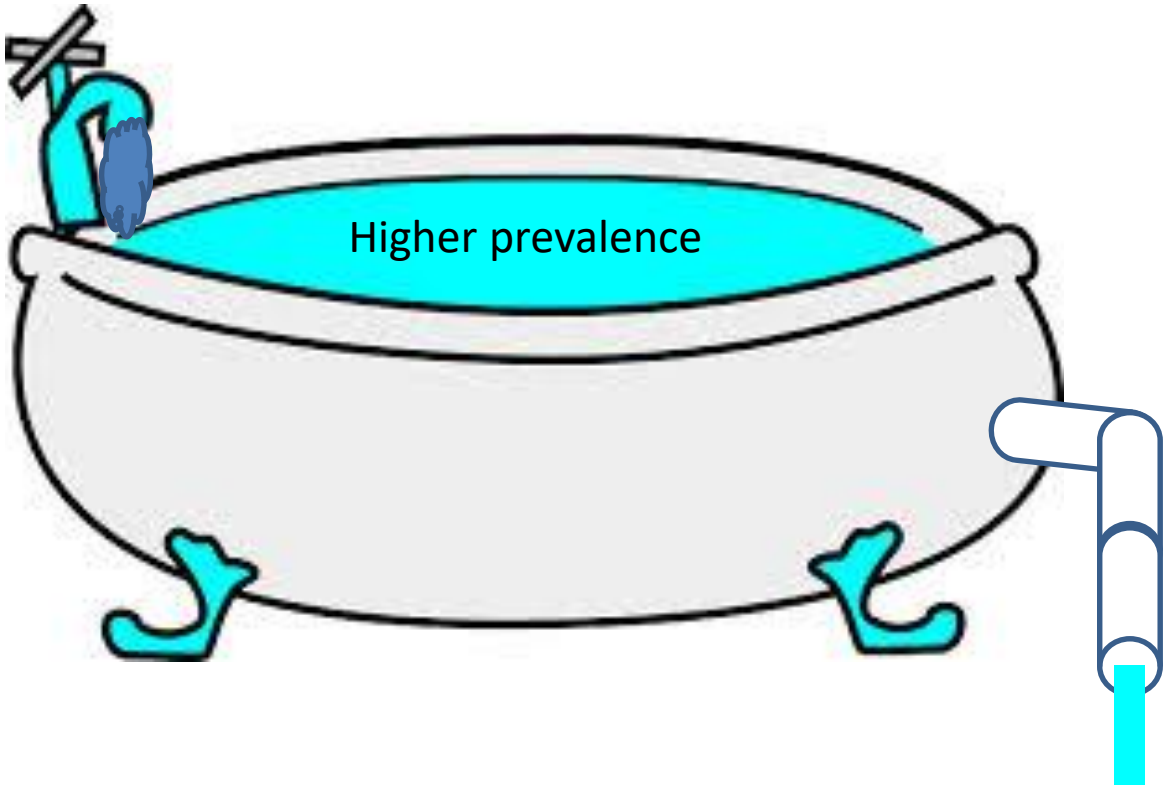


Example - heart failure

Increased myocardial infarction

↓

Increased heart failure incidence



Improved long-term management → Reduced death rate

ASSOCIATIONS BETWEEN EXPOSURES AND OUTCOMES

Outcomes

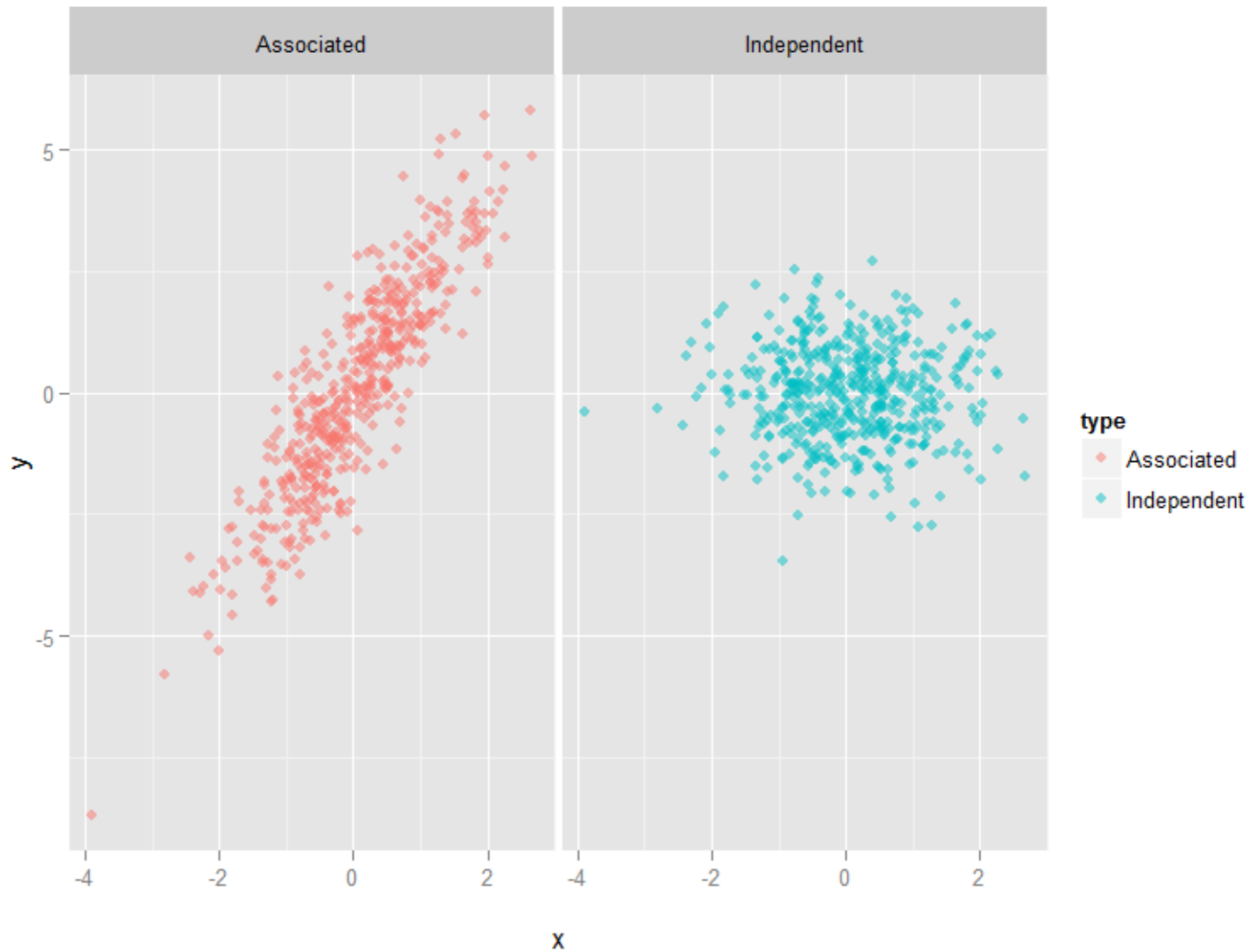
- Outcomes
  - death
  - hospitalisation
  - first diagnosis with a disease
  - recurrence, eg cancer
  - quality of life
  - surrogates – blood pressure, lung function

Exposures

- Non-modifiable
  - age, sex, genotype
- Modifiable
  - smoking, weight, diet, alcohol consumption
- Interventions – a special kind of exposure
  - drug therapy
  - surgery
  - lifestyle advice

Association

- a and b are associated when a and b are not independent
  - $P(a|b) = P(a)$
  - $P(b|a) = P(b)$
- Knowing something about a tells you nothing about b
- Knowing something about b tells you nothing about a


Associated
Independent
5
0
-5
y
-4
-2
0
2
x
type
Associated
Independent

Effect estimates

	Died	N	Proportion
Thrombolysis	100	1000	**10%**
No thrombolysis	130	1000	**13%**

$$RR = \frac{risk_{exposed}}{risk_{unexposed}}$$

$$ARR = risk_{unexposed} - risk_{exposed}$$

Measure	Result
Risk ratio (RR)	=10/13 = 0.77
Relative risk reduction (RRR)	100 x (1-0.77) = 23%
Risk difference (absolute risk reduction – ARR)	= 13-10 = 3%
Number needed to treat (NNT)	= 1/0.03 = 33

$$RRR = 100 \times (1 - RR)$$

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

Other commonly encountered measures

- Odds ratio – commonly used estimate of risk ratio
- Rate ratio – ratio between two mortality rates, hospitalisation rates etc
- Hazard ratio – a special kind of rate ratio

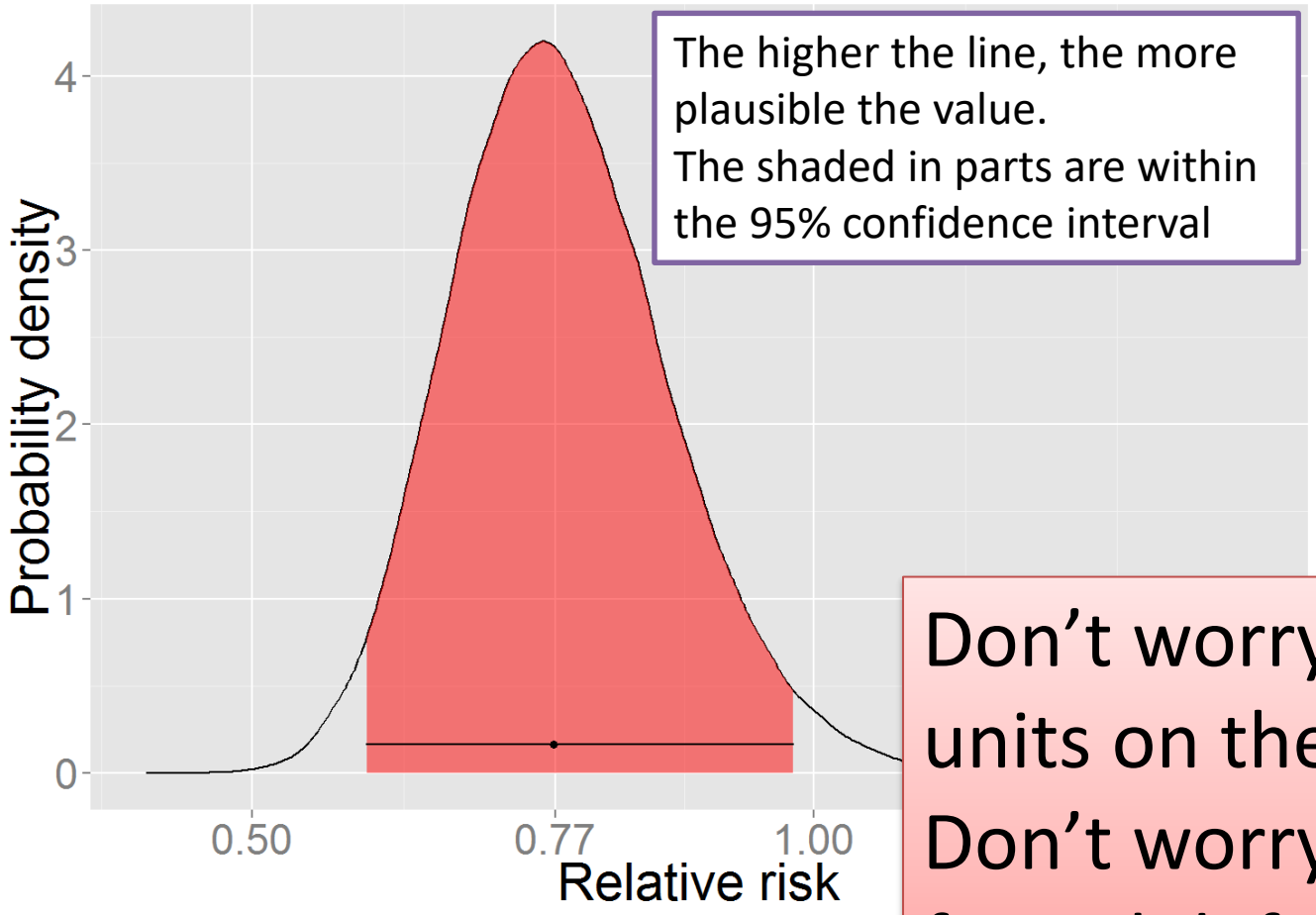
CONFIDENCE INTERVALS

Confidence intervals

- Vital to reading **ANY**! research paper
- Formal definitions are counterintuitive
- Pragmatically
  - "A confidence interval can be thought of as a range of plausible values"
  - Values near the limits are much less plausible than those in the middle

EXAMPLE – AS IN TABLE EFFECT ESTIMATES - THROMBOLYSIS

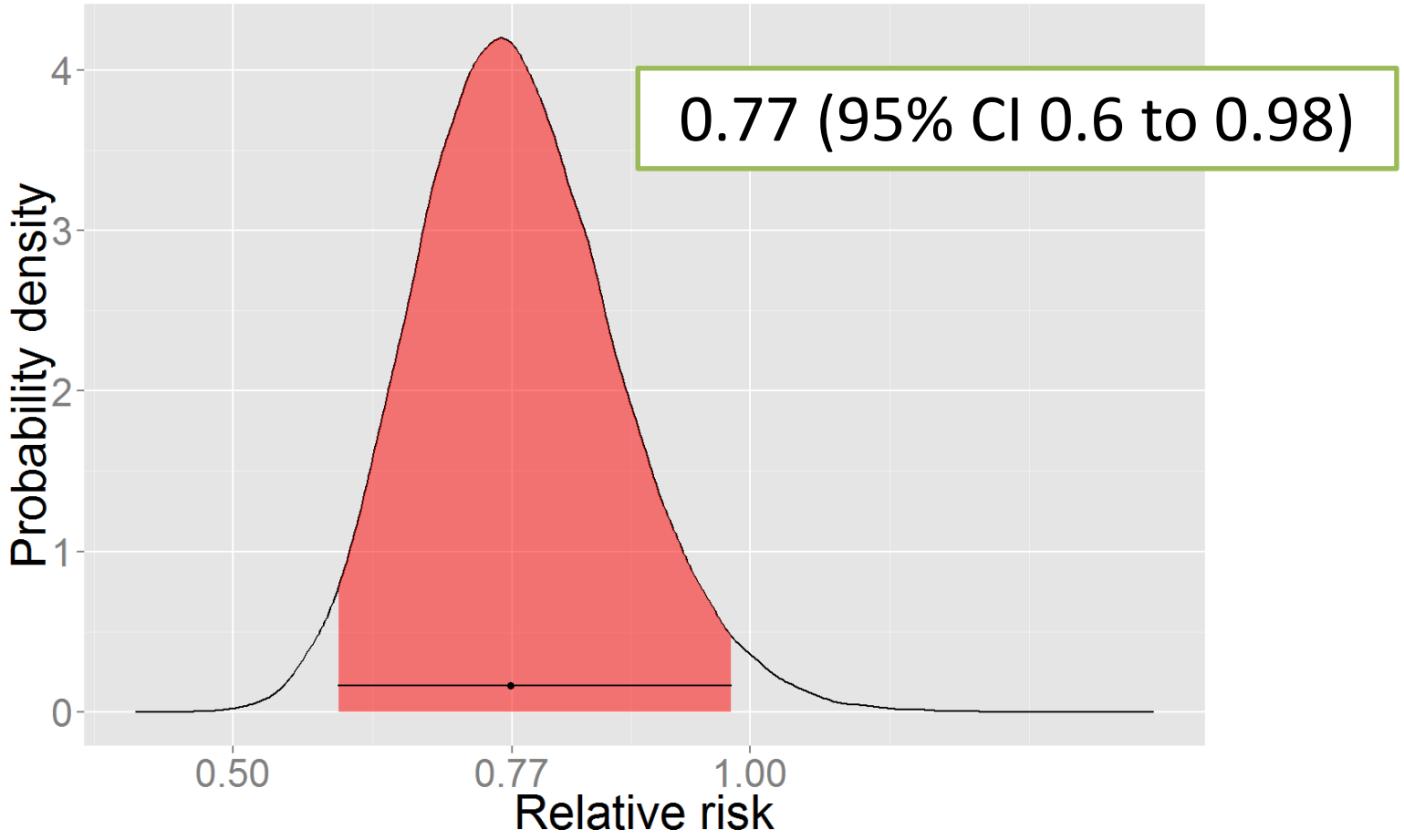
95% CI represented visually



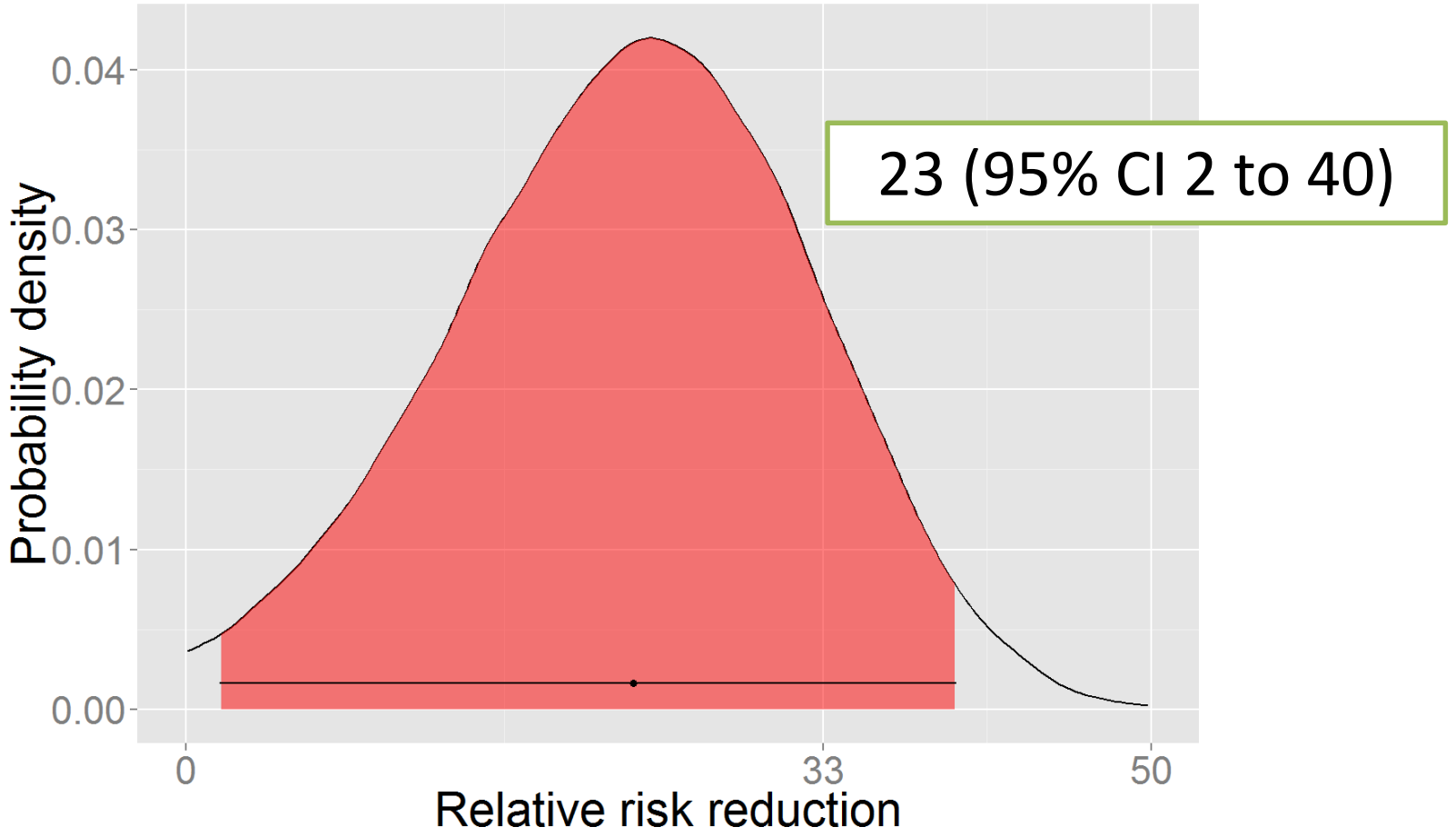
The higher the line, the more plausible the value.
The shaded in parts are within the 95% confidence interval

Don’t worry about the units on the y-axis
Don’t worry about formal definitions!!!

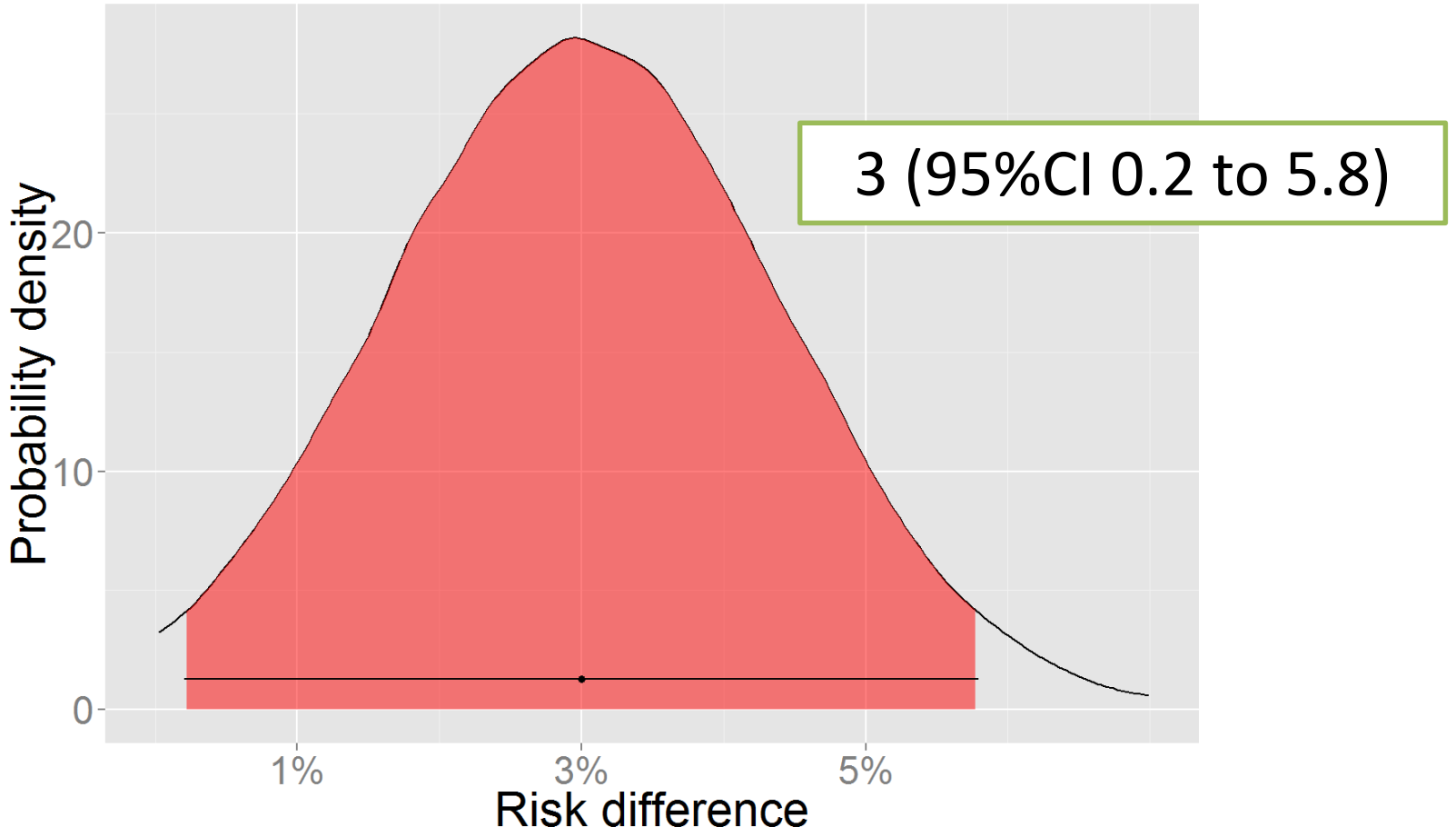
Relative risk



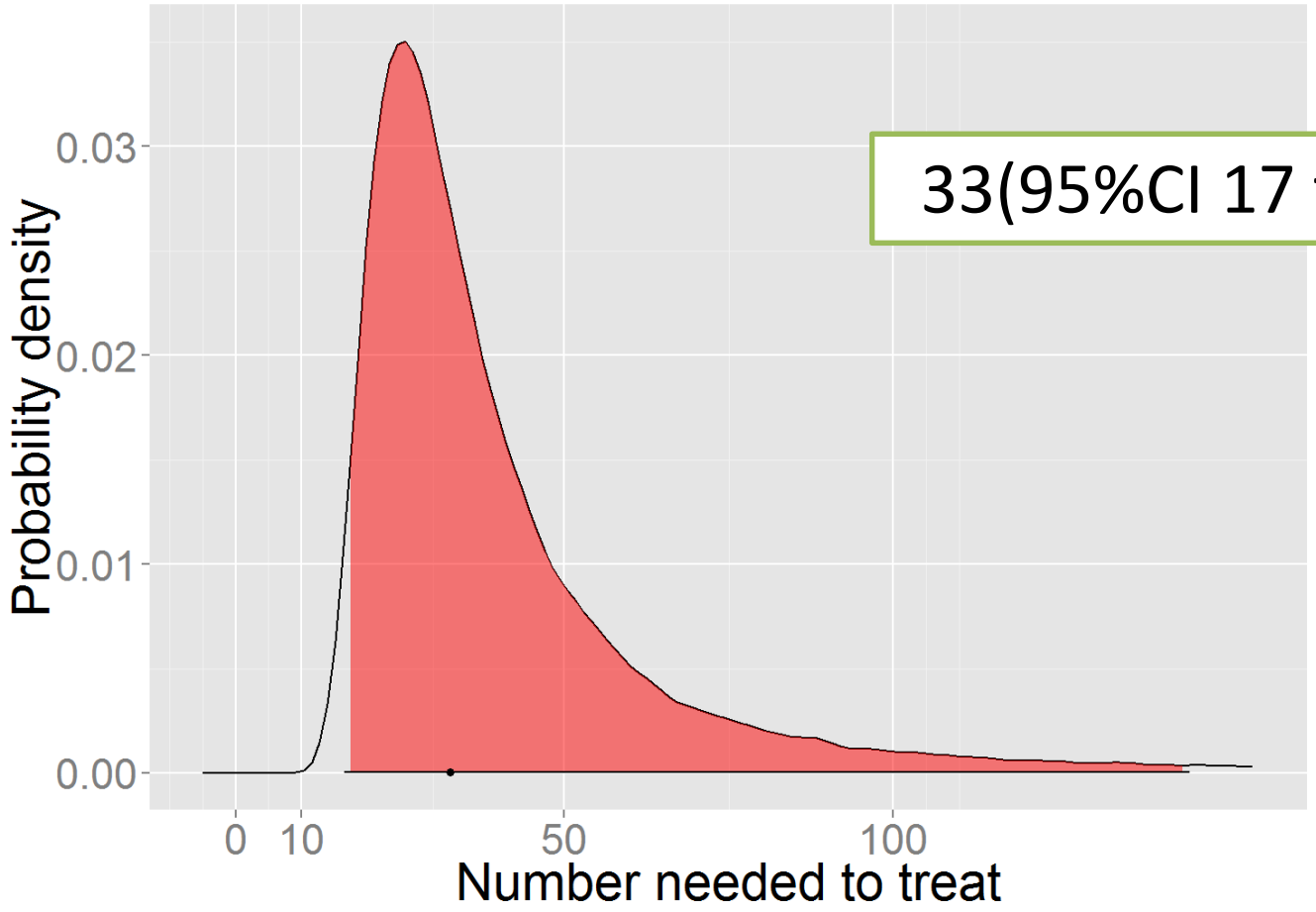
Relative risk reduction



Risk difference
(absolute risk reduction)



Number needed to treat (NNT)



33(95%CI 17 to 145)

Confidence intervals

- Can be presented for any statistic/effect measure
- Represents range of plausible values
- More extreme values less likely
- Very useful in appraising published research

DESCRIPTIVE EPIDEMIOLOGY

Crude mortality

State	Deaths in 2013	Population in 2013	Crude annual mortality
New York City	48,000	8,000,000	6 per 1,000
Florida	19,000	19,000,000	10 per 1000
		Rate ratio	**1.67**

WHY?

Standardised mortality ratio (SMR)

State	Deaths	Expected deaths	SMR
New York City	48,000	50,000	96
Florida	19,000	22,000	86

- Calculate expected deaths based on
  - age-sex specific mortality rates in whole of US
  - age and sex of people in NY and Florida
- Both places have lower than expected mortality
- New York has a higher SMR than Florida

CONFOUNDING

NY vs Florida was an example of confounding

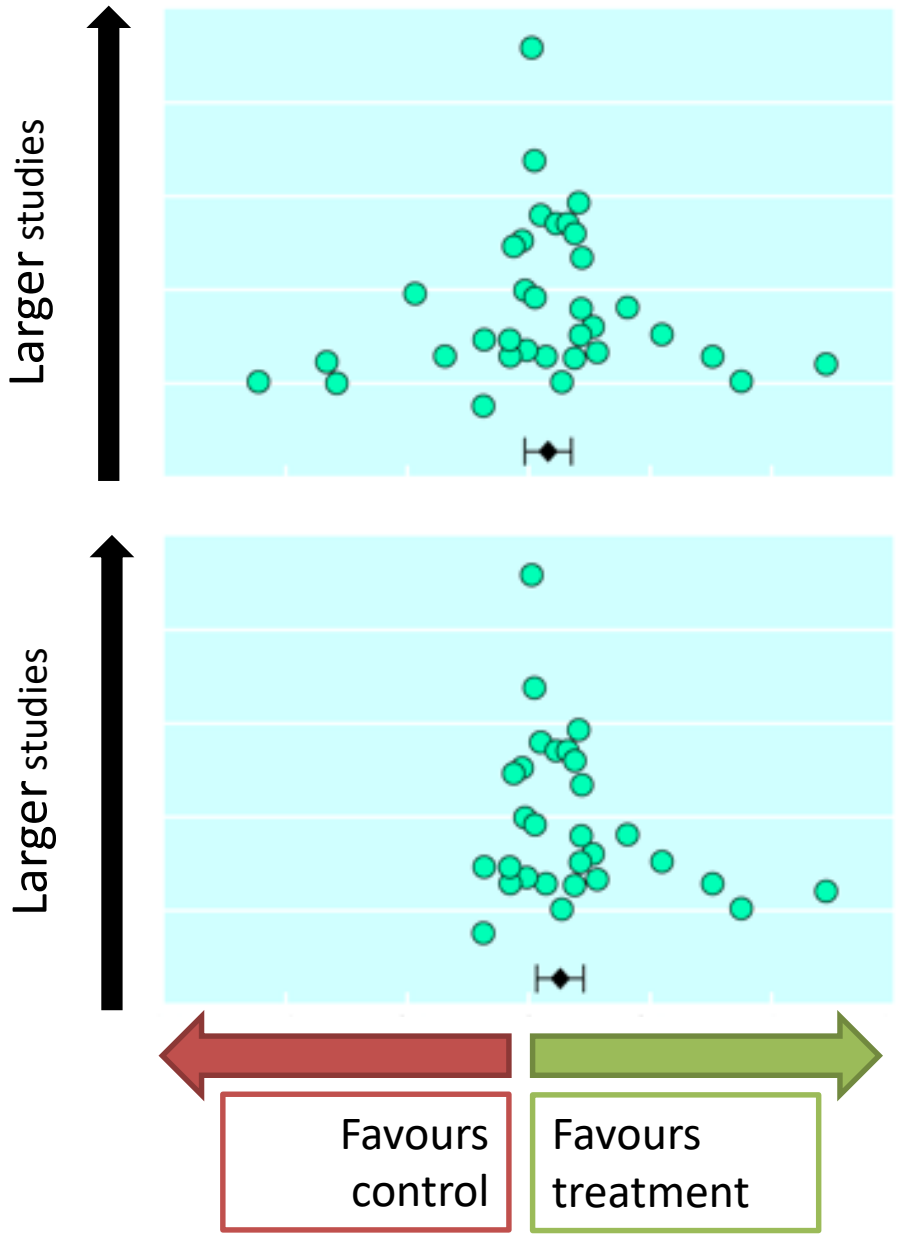
- True relationship confused by a third factor
- Can deal with confounding
  - study design
  - data analysis – eg via standardisation

Bias

- **Systematic** error in
  - what data are collected
  - how data are collected
  - how data are analysed
  - how data are interpreted
  - how data are reported
- Bias leads to wrong conclusions about
  - disease causation
  - treatment effectiveness

Bias - examples

- **Systematic** error in
  - what data are collected - novel treatments
  - how data are collected - different follow-up
  - how data are analysed - misconduct
  - how data are interpreted - highlight certain findings
  - how data are reported - publication bias
- Causes wrong conclusions about
  - disease causation
  - treatment effectiveness



All studies

Published studies

Empirical assessment of effect of publication bias on meta-analyses *BMJ* 2000;320:1574

DIFFERENT TYPES OF STUDY ARE MORE/LESS PRONE TO CONFOUNDING AND BIAS

Hierarchy of evidence

Less prone to confounding and bias

Case-control study

- Select CASES with a disease – eg lung cancer
- Select CONTROLS without that disease
- Find our what EXPOSURES the cases and controls had – eg smoking
- Compare exposures in cases and controls
- Identify if association

Cohort study

- Select people without a disease (eg heart disease)
- Classify them according to an exposure eg
  - high cholesterol
  - low cholesterol
- Follow-them, eg for 10 years
- Compare RISK of disease in exposed and unexposed

Randomised controlled trial (RCT)

- Like cohort study, except
- **ALLOCATE** PEOPLE TO INTERVENTION
  - cohort study ‘observe exposure’
  - randomised controlled trial ‘allocate intervention’
- Compare RISK of outcome in treatment and control groups

Cross-sectional study

- Sample a population
- Estimate the proportion of people with
  - different exposures (eg smoking, alcohol consumption, income)
  - different symptoms (eg breathlessness, cough)
  - different diseases (eg angina)
- Use these data
  - to describe disease burden
  - to explore associations

Hierarchy of evidence

Less prone to confounding and bias

Topics covered

- Basic epidemiological terms
- Confidence intervals
- Confounding
- Bias
- Case-control and cohort studies
- The hierarchy of evidence

Questions?