Epidemiological methods

Objectives

At the end of the week students should be able to:

- Differentiate between different types of data.
- Describe the structure of an epidemiological dataset
- Define and calculate measures of disease occurrence and measures of association
- Describe the basic features of the main types of epidemiological studies
- Explain the main features of bias, confounding, chance
- Be able to discuss causality of the association

Epidemiology

- The study of the **distribution** and **determinants** of the **frequency** of health-related outcomes in specified populations
- Quantitative discipline
- Measurement of disease / condition / risk factor frequency is central to epidemiology
- Comparisons require measurements

Much of epidemiological research is taken up trying

- to establish associations between exposures and disease rates
- to measure the extent to which risk changes as the level of exposure changes
- to establish whether the associations observed may be truly causal (rather than being just consequence of bias or chance)

- Epidemiology has a major role in developing appropriate strategies to improve public health through prevention
  - public health has wider meaning in this sense; it is about the health of the whole population.
  - it does not cover only classic areas, such as immunization or monitoring of diseases, it also covers factors such as poverty, smoking, nutrition
- In this sense, epidemiology has a crucial role in trying to put into perspective the effects on population health of different risk factors.

Variables (outcomes/risk factors)

- Binary
  - Deaths (y/n)
  - Disease (y/n)
- Categorical (ordinal or nominal)
  - Frequency of drinking (never, 1-3 times a month, 1-3 times a week, 4 times a week or more often)
  - Severity of pain (none, some, a lot)
- Continous
  - BMI, blood pressure etc

What type of variable is…

- Self-rated health
  - Very poor, poor, average, good, very good
- Total cholesterol concentration
- Economic activity
  - Employed, unemployed, housewife, pensioner
- Risk of CVD death in the next 10 years (SCORE)
- Ethnicity
- Quartile of income
- Sex
- Marital status (married, divorced, ever single, widowed)

Binary outcomes: "cases" vs. "non-cases"

- Persons with disease = "cases"
- Definition of case is crucial
- E.g.
  - Obesity: BMI≥30
  - Hypertension: SBP≥140 mm Hg or DBP≥90 mm Hg or treatment
  - High cholesterol: ≥6.2 mmol/L
- Must always be clearly specified

Measures of disease frequency

- Used for binary outcomes
- Require a numerator and denominator

$$\frac{\text{number of persons with disease}}{\text{number of persons examined}}$$

- expressed as X per 1000 persons (or per 100,000 etc)

Numerators and denominators

- The number of cancer cases in the UK is 247,667 whereas in Belgium it is 47,948.
- The UK has a bigger problem in numerical terms.
- But do Belgians have lower risk of getting cancer?
  - Numerators alone are meaningless
  - We need both **numerators** AND **denominators**

Numerators and denominators

- The number of cancer cases in the UK is 247,667 whereas in Belgium it is 47,948.
- The UK has a bigger problem in numerical terms.
- But do Belgians have lower risk of getting cancer?
  - Numerators alone are meaningless
  - We need both **numerators** AND **denominators**
- UK: 247 667 / 60 000 000 = 0.00413 = 413 per 100 000
- Belgium: 47 948 / 10 000 000 = 0.00479 = 479 per 100 000

Prevalence

- number of **existing** cases / population of interest at a defined time

Incidence

- number of **new** cases in a given time period / total population at risk

Prevalence

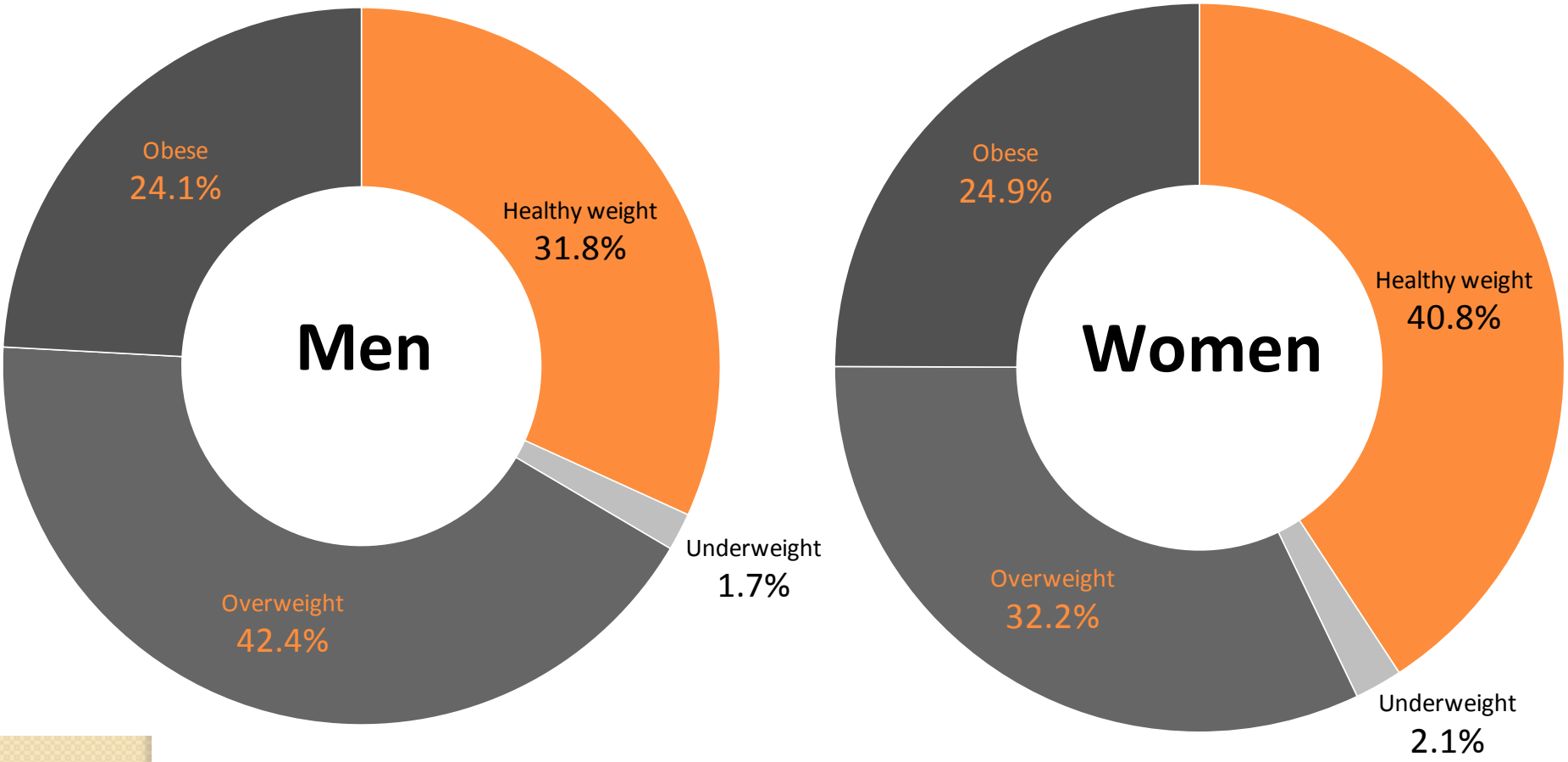
- number of **existing** cases / population of interest at a defined time
  - Unable to work now for health reasons
  - Injury ever in the past
  - Ever wheezing or whistling in the chest

NOTE a **denominator** is needed for prevalence

Adult prevalence by BMI status

Health Survey for England (2008-2010 average)

noo
National Obesity Observatory



Adult (aged 16+) BMI thresholds

Underweight: <18.5kg/m^2

Healthy weight: 18.5 to <25kg/m^2

Overweight: 25 to <30kg/m^2

Obese: ≥30kg/m^2

Incidence rates

- In 2014, 55,222 new cases of breast cancer were diagnosed in the UK.
- Approximately 65M people in the UK
- Most cases in women (only 389 cases in men)
- Population at risk?
- Cumulative incidence of breast cancer in the UK in 2014 in females was ?

$$\frac{???}{----------}$$

???

Incidence rates

- In 2014, 55,222 new cases of breast cancer were diagnosed in the UK.
- Approximately 65.5M people in the UK
- Most cases in women (only 389 cases in men)
- Population at risk?
- Incidence of breast cancer in the UK in 2014 in females was ?

$$\frac{55222-389}{65.5M/2} = \frac{54833}{32.75} = 0.001674 = 167.4/100{,}000$$

Incidence rate example:

3-year study with a sample size of 100, outcome of interest was fatal heart disease.

	year 1	*year 2*	*Study ends*
developed outcome	6	5	4
dropped out	4	10	-
sample at risk	90	75	71

- 10 participants were followed for 1 year
- 15 participants were followed for 2 years
- 75 participants were followed for 3 years

Total person-years:
Rate =

Incidence rate example:

3-year study with a sample size of 100, outcome of interest was fatal heart disease.

	year 1	*year 2*	*Study ends*
developed outcome	6	5	4
dropped out	4	10	-
sample at risk	90	75	71

- 10 participants were followed for 1 year
- 15 participants were followed for 2 years
- 75 participants were followed for 3 years

Total person-years of follow up = (10x1) + (15x2) + (75x3) = 265 person-years at risk

Incidence rate = 15 / 265 = 0.057 = 57 cases per 1000 person-years

Relationship between prevalence and incidence

- The prevalence of a health-related outcome depends both on the incidence rate and the time between onset and recovery or death.
- **Prevalence = Incidence x Average disease duration**
- E.g. volume of water in water tank depends on
  - Inflow
  - Outflow

Mortality

- **number of deaths / total population**
- Rate (or risk)
- the number of deaths in a specified population, divided by the number of that population, per unit time.
- If the mortality rate is to be calculated in a given year, the mid-year population is usually used as the denominator.
- Mortality rate is always expressed as deaths per X (e.g. 1000 persons per year). E.g.
  - A city has a population of 900,000, 30,000 deaths occur in a 3-year period.
  - Mortality rate for the period = 30 000 / 900 000 = 0.0033 or 33 deaths per 1000 per 3 years
  - = 11 deaths per 1000 per year.

Mortality rates can be:

- **All-cause mortality rates:** refers to the total number of deaths per 1000 people per year. This is also usually referred to just as all-cause mortality.
- **Cause-specific mortality rate** refers to total number of deaths due to a specific cause.

Mortality rates can be:

- **Crude mortality rates** – no care has been taken for age structure of the population
- **Standardised mortality rate** refers to a mortality rate which is age-standardised in order to permit comparisons between different countries, regions etc.

Case fatality

- **Case fatality rate** is the rate of death among people who already have a condition, usually in a defined period of time. usually measured as a decimal or as a percent.
- **Survival rate** is the proportion of people who remain alive for a given period of time after diagnosis of disease. E.g. breast cancer has 5-year survival rate around 70%.