Epidemiology

Introduction

Objectives

At the end of this session 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

Epidemiological dataset

ID	Age	Sex	Disease	Smoking	…
1	54	1	0	0	…
2	65	2	0	1	…
3	47	1	1	1	…
4	53	1	0	0	…
…	…	…	…	…	…

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

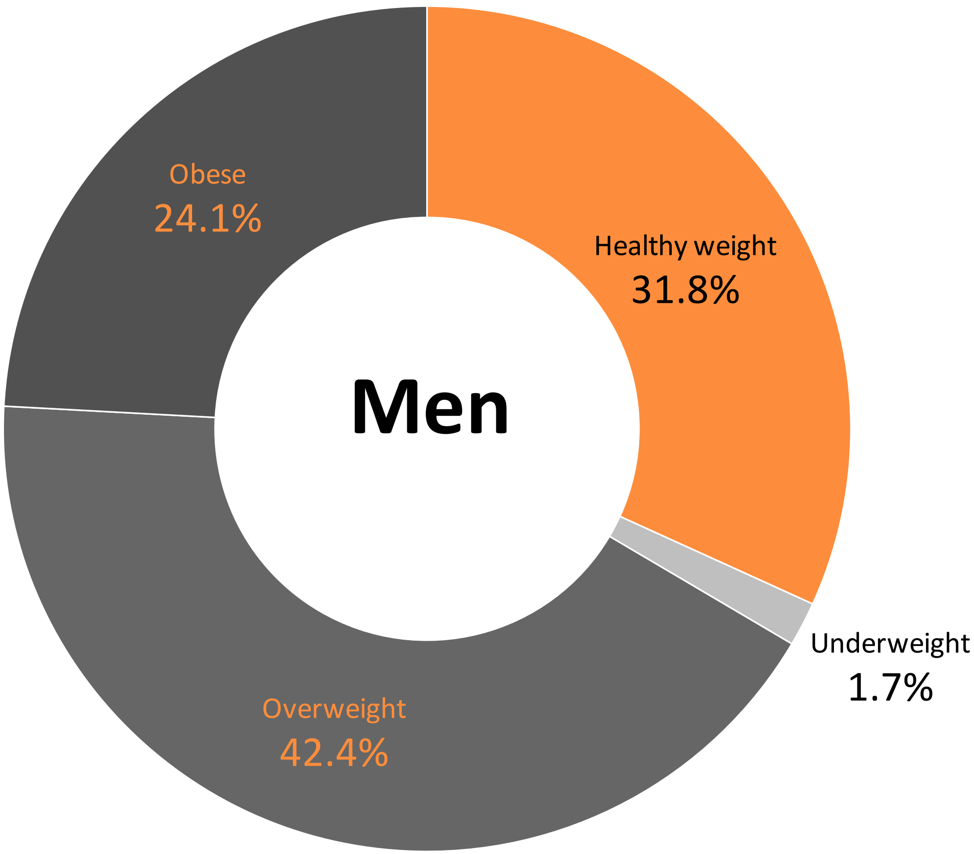
- number of **existing** cases / population of interest at a defined time
  - Unable to work now for health reasons
  - Injury ever in the past
  - Ever wheezing or whistling in the chest

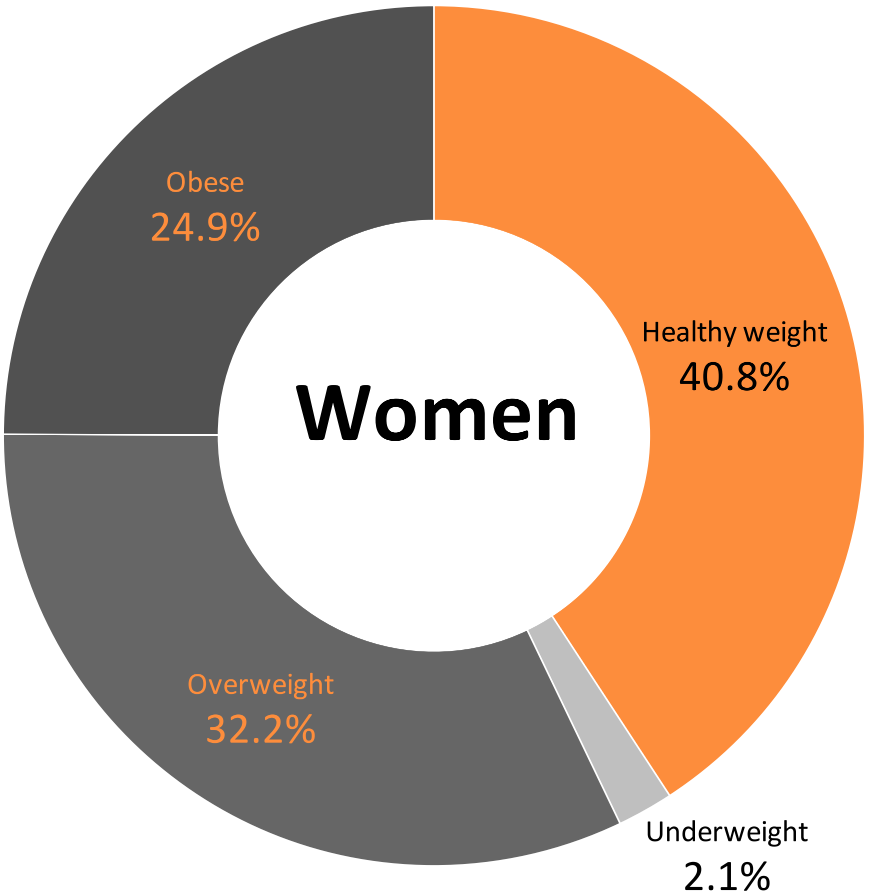
NOTE a **denominator** is needed for prevalence

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.

Other examples of specific mortality

- Infant mortality (first year)
- Neonatal (first 28 days)
- Early neonatal (first 7 days)
- Post-neonatal (29^{th} day to 1 year)
- Maternal (while pregnant or within 42 days of the end of pregnancy)
- Stillbirth (baby is born dead after 24 completed weeks of pregnancy)
- Child (usually under 5 years)

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%.

Life expectancy

LIFE EXPECTANCY THROUGH THE AGES

Early humans did not generally live long enough to develop heart disease, cancer or loss of mental function. A snapshot of how life expectancy has changed, and the big killers of each era:

AVERAGE LIFE EXPECTANCY

30 years

Neanderthals (30,000 years ago): Died of injuries caused by rock falls, hunting accidents and conflicts. Food scarcity led to malnutrition. These hunter-gatherer groups contracted diseases that spread from animals. Rabies, tuberculosis, brucellosis, yellow fever and encephalitis were widespread.

38

Neolithic (8500 BC to 3500 BC): Agriculture, irrigation and urbanization brought problems associated with settled populations, such as fecal contamination of water and diseases such as cholera, smallpox, typhoid, polio and influenza. Malaria and other diseases carried by mosquitoes and insects, which fed on domesticated animals, appeared.

35 **Classical Greece and Rome** (500 BC to 500 AD): Tuberculosis, typhoid fever, smallpox and scarlet fever spread among the denser urban populations. Malnutrition, gastroenteritis and violence were also big killers.

48 EARLY MEDIEVAL

38 LATE MEDIEVAL

Medieval period (500 AD to 1500 AD):

Life expectancy grew with urbanization, but famine caused by crop failures and bubonic plague were the big killers. The Black Death (1347-1351) wiped out 25 million people in Europe and 60 million in Asia, returning several times, culminating in the Great Plague of London (1664-1666). By 1500, life expectancy had dropped back to 38.

40

Victorian (1850s to 1900): Typhus, typhoid fever, rickets, diphtheria, tuberculosis, scarlet fever and cholera raged in crowded cities.

MEN **70** WOMEN **75**

1900s: Better health care, sanitation and living conditions boosted life expectancy to 70 for men and 75 for women by 1950.

CANADA: MEN **82** WOMEN **85**

Today: Cancer, heart disease and stroke are the biggest killers in the developed world. Our longer lifespan also comes with unprecedented loss of mental function and mobility problems.

RESEARCH BY RICK SZNAJDER/TORONTO STAR LIBRARY

SOURCES: JOURNAL OF POPULATION RESEARCH, PRINCETON UNIVERSITY, STANFORD UNIVERSITY, WORLD HEALTH ORGANIZATION

Numbers of women expected to die at each age, out of 100,000 born, assuming mortality rates stay the same as 2010-2012. The expectation is 83 (mean), median 86, the most likely value (mode) is 90.