Session 2

Observational studies

Outline of Session

- The limitations of RCTs
- Observational studies – their benefits and limitations
- Designing a cohort study
- Designing a case-control study

Outline of Session

- The limitations of RCTs
- Observational studies – their benefits and limitations
- Designing a cohort study
- Designing a case-control study

The key features of RCTs

- The control group allows us to judge whether a new treatment provides additional benefits to those expected with standard care
- Randomisation and blinding (where used) allow us to conclude that any such benefits are due to the treatment itself, rather than to any other factor or to chance
- As a result, RCTs are perceived to be the gold-standard method for obtaining evidence of a treatment effect

However, RCTs may have some limitations

1. RCTs are only possible where there is an ‘intervention’ that people are willing to be randomised to

However, RCTs may have some limitations

1. RCTs are only possible where there is an 'intervention' that people are willing to be randomised to

   Examples…

   - Impact of smoking and/or alcohol consumption on response to HAART
   - Impact of co-infection with TB on HIV progression rates

However, RCTs may have some limitations

2. Patients in RCTs may not be representative of the clinic population, and follow-up of patients may differ to that in clinic – thus, outcomes may differ from what would normally be expected

However, RCTs may have some limitations

2. Patients in RCTs may not be representative of the clinic population, and follow-up of patients may differ to that in clinic – thus, outcomes may differ from what would normally be expected

Examples…

- Patients may be selected on the basis of their likely adherence to treatment
- Patients may attend clinic more frequently – outcomes may be detected sooner
- Monitoring may be more intensive

However, RCTs may have some limitations

3. RCTs may be short (48 weeks) and may focus on two or three main treatment comparisons

However, RCTs may have some limitations

4. RCTs may concentrate on short-term surrogate marker endpoints rather than long-term clinical events

However, RCTs may have some limitations

4. RCTs may concentrate on short-term surrogate marker endpoints rather than long-term clinical events

   Example…

   - Early studies of IL-2 treatment in HIV infection focussed on CD4 endpoints only

Outline of Session

- The limitations of RCTs
- Observational studies – their benefits and limitations
- Designing a cohort study
- Designing a case-control study

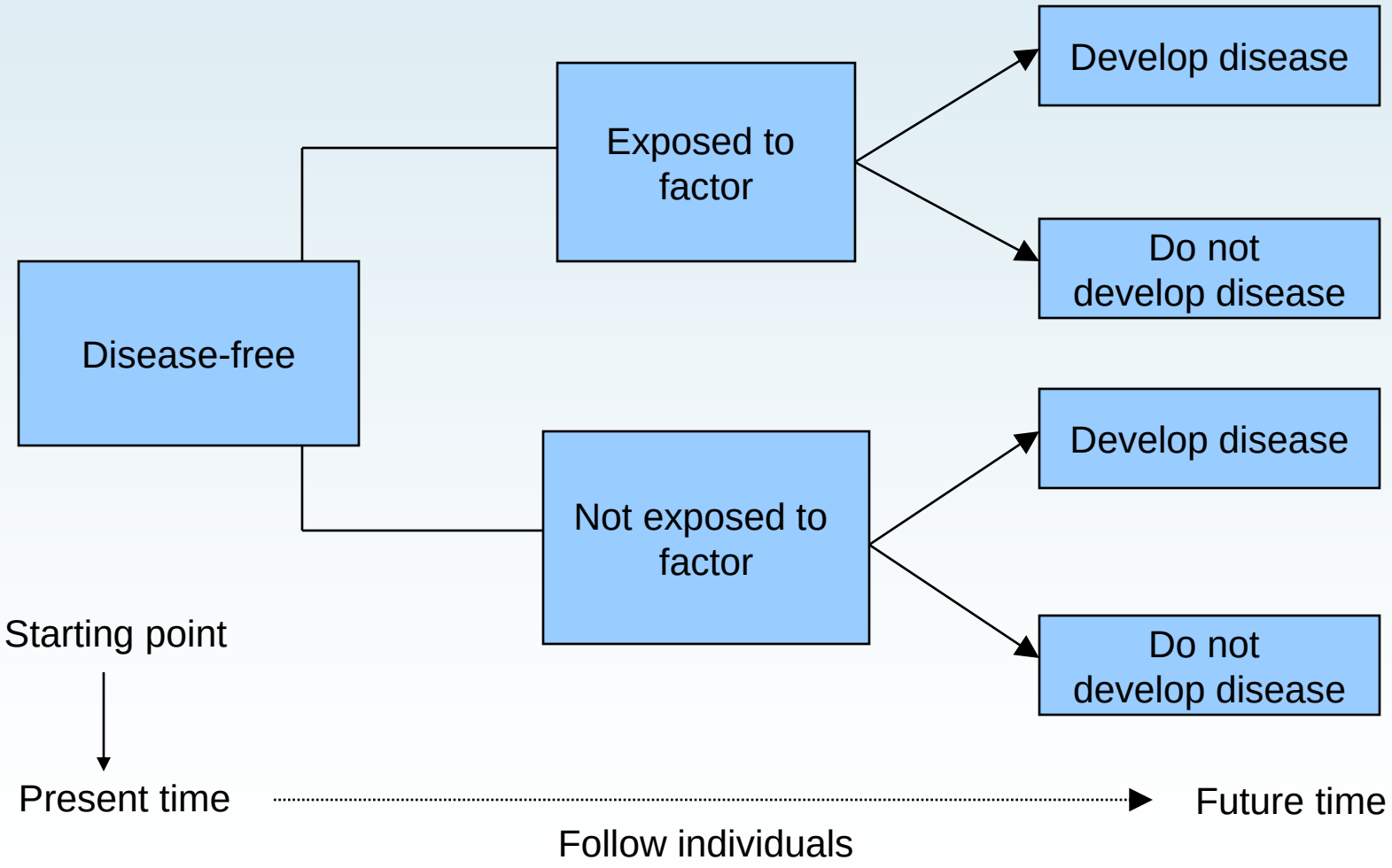
Experimental / observational studies

- In an experimental study (e.g. an RCT) the investigator intervenes in the care of the patient in a pre-planned way and records the outcome
- In an observational study, the investigator does not intervene in the care of the patient, but simply records outcomes when they occur
- Common observational studies: cohort studies, case-control studies, cross-sectional studies

Cohort studies

- Follow a group of individuals over time to assess the incidence of a disease (or some other outcome)
- Are used to describe the effect of exposure to one or more factors of interest (potential risk factors) on the incidence of the outcome
- Can be prospective or retrospective/historical

Cohort studies

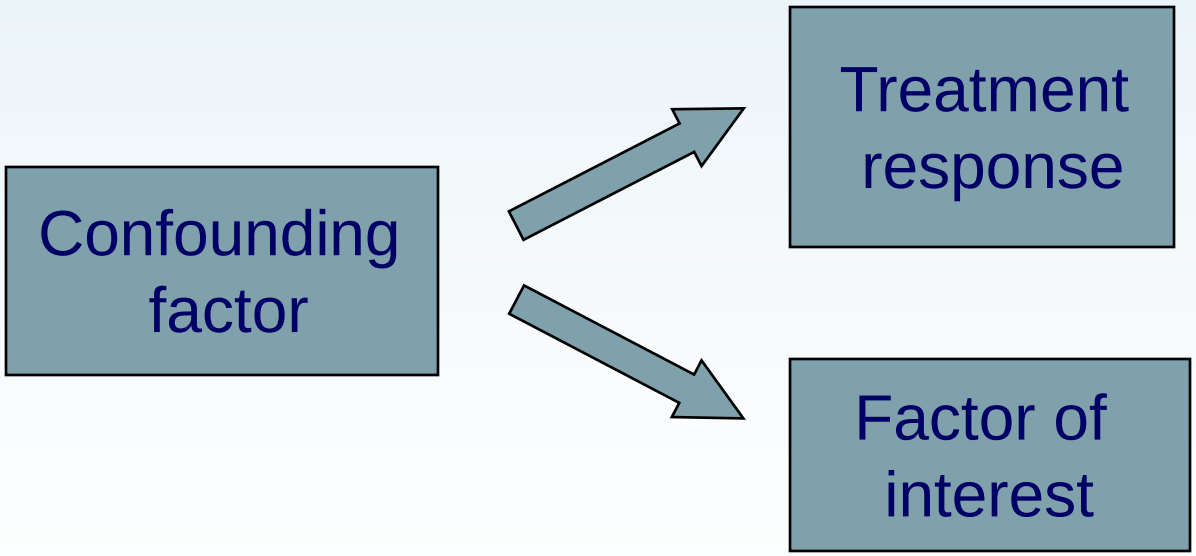


Cohort Studies

- Advantages
  - Can assess temporal relationships between exposure and disease
  - Can make some attempt to assess cause and effect (although RCTs are preferable)
  - Can sometimes be more representative of clinic population
- Disadvantages
  - If disease is rare then cohort may need to be large and follow-up long
  - May be problem with loss to follow up
  - Cannot ever rule out presence of unmeasured confounding

Bias due to confounding

- Occurs when a spurious association arises due to a failure to fully adjust for factors related to both the risk factor and outcome.

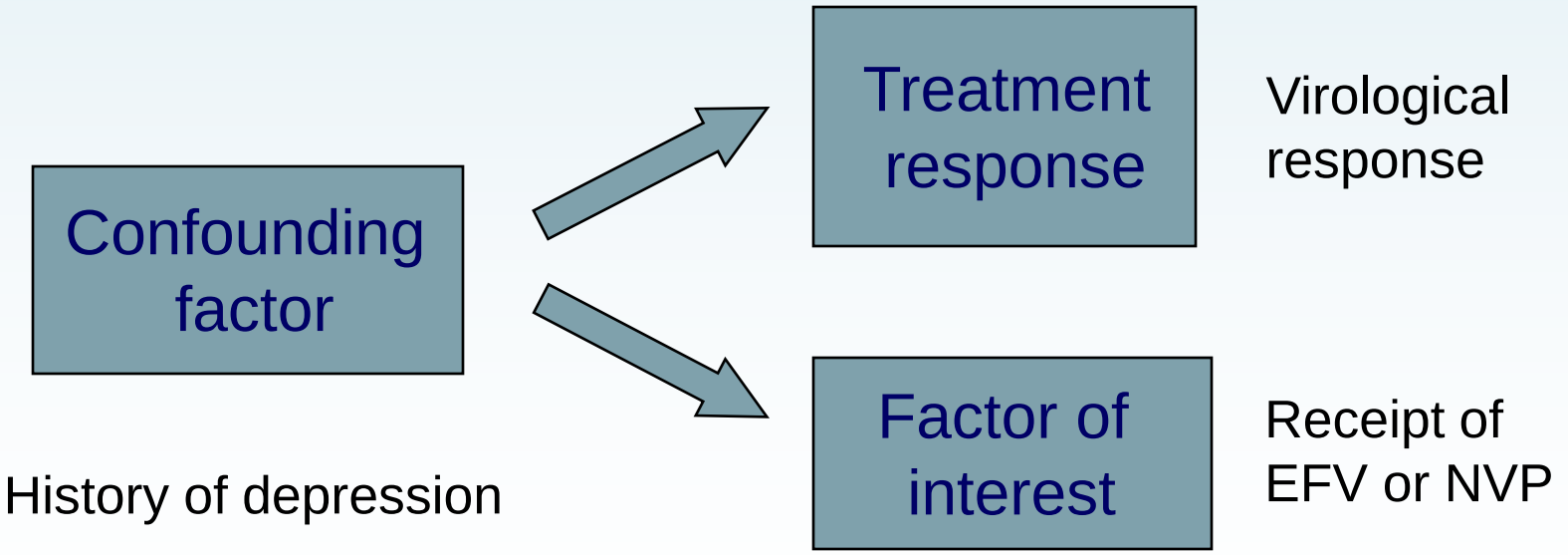


Confounding bias – treatment comparisons

- The reason why some patients were treated with one regimen and others with a different one is often unknown
- Whilst one treatment may appear to be associated with a better outcome, these other factors may explain the better outcome in these patients
- Can perform 'adjusted' analyses to reduce the impact of known confounders, but it is harder to reduce the impact of unknown or unmeasured confounders

Bias due to confounding

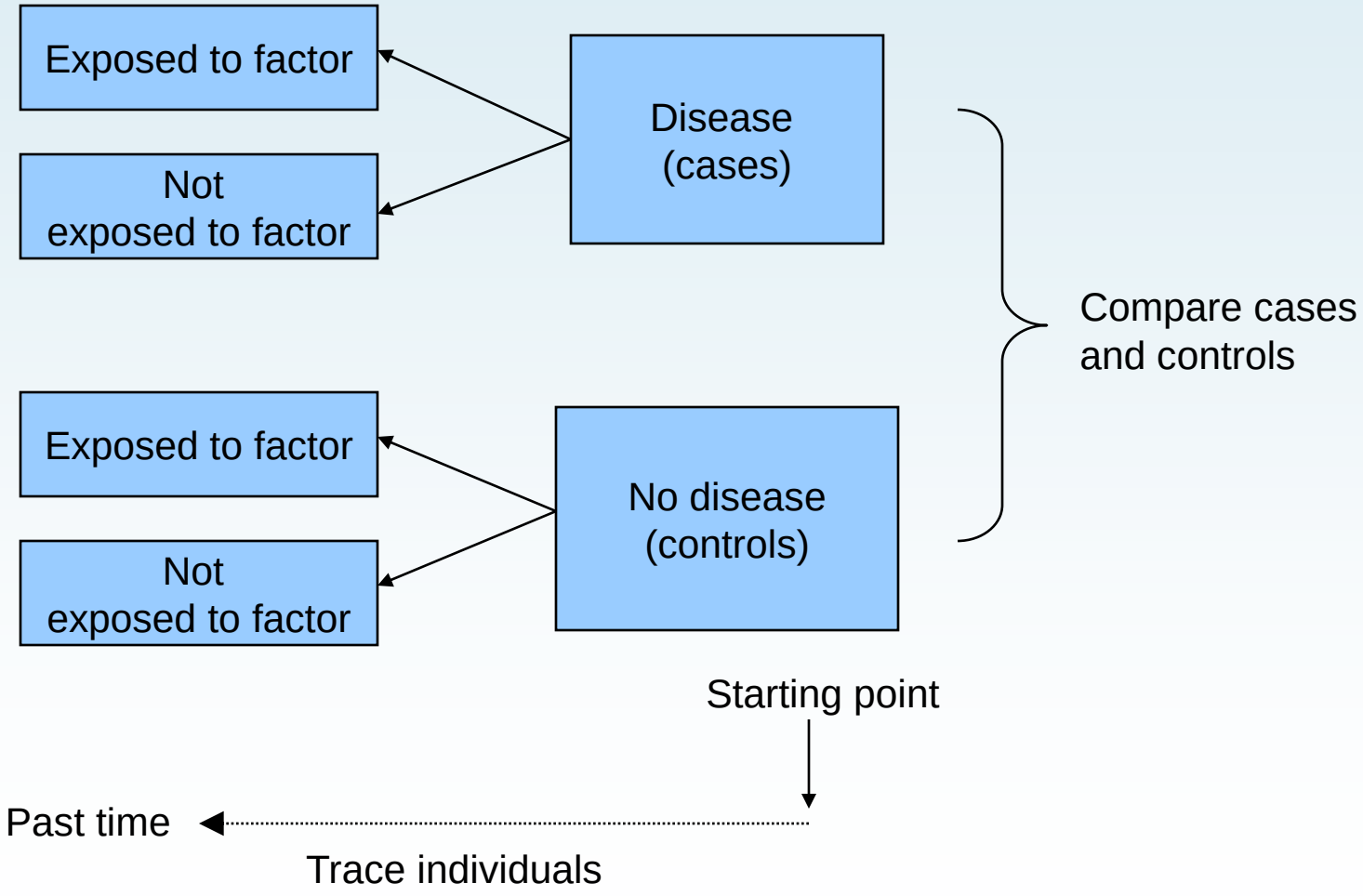
- Occurs when a spurious association arises due to a failure to fully adjust for factors related to both the risk factor and outcome.



Case-control studies

- Retrospective studies in which a group of patients with a disease (cases) are compared to a group of patients without the disease (controls)
- Aim is to see whether exposure to any factor has occurred more or less frequently *in the past* in cases than in controls
- Cases and controls may often be matched on basic demographic information (e.g. sex and age) to make the two groups as similar as possible

Case-control studies



Case control studies

- Advantages
  - Relatively cheap, quick and easy to carry out
  - No loss-to follow up
  - Particularly suitable for rare events
- Disadvantages
  - Potential for recall bias
  - Timing of events cannot be reliably established
  - Cannot assess incidence (proportion with disease is fixed as part of the study design)

Recall bias

- Tendency of cases to ‘recall’ information (particularly relating to exposure) differently to cases
- Can lead to apparent association between outcome and exposure, even if the association does not exist
- Example: cigarette smoking and lung cancer

Cross-sectional studies

- Carried out at a single point in time
- Often used to assess the prevalence of a condition, to describe the current situation or to assess attitudes and beliefs
- Whilst relatively cheap and quick to perform, it is not possible to estimate the incidence of disease, only the prevalence

Outline of Session

- The limitations of RCTs
- Observational studies – their benefits and limitations
- Designing a cohort study
- Designing a case-control study

General issues

- Cohorts may be fixed/closed (new patients are unable to join the study), dynamic/open (new patients are able to join the study) or a combination of both
- Patients may be seen at regular time intervals for specific ‘cohort’ visits or the cohort may make use of existing data collection systems within clinics

Traditional interval cohort

- Patients often seen at a study site (often different to their place of care) on regular occasions for 'study visits' (e.g. 6-monthly)
- Participants may complete questionnaire on their health since last visit, treatments received, etc.
- Laboratory tests performed at pre-defined time intervals – this information is unlikely to be available at intervening times or when an event occurs, unless this coincides with a study visit
- Patients must give consent to participate

Traditional interval cohort - example

- The Multicenter AIDS Cohort Study (MACS)
- HIV+ve and HIV-ve individuals from 4 centres in Baltimore, Chicago, Los Angeles and Pittsburgh
- Participants recruited from 1984-1985 (n=4954), 1987-1991 (n=668) and 2001-2003 (n=1351)
- Visits are bi-annual – at each visit, participants under go a detailed interview, physical examination, quality of life assessment and collection of blood for concomitant laboratory testing and storage

Observational databases

- Utilise data that are already routinely collected as part of patient's medical care
- Patient does not have to attend for a particular study visit or fill in any questionnaires
- Laboratory testing will be performed according to clinical needs – will be more frequently monitored if patient is ill or requires investigation
- Some data items may be difficult to collect if not part of routine care
- May or may not require patient consent, depending on local regulations in country

Observational databases - example

- The UK Collaborative HIV Cohort (CHIC) Study
- All HIV+ve individuals who have attended one of 11 UK clinical centres at least once between 1996 and 2007
- Median (inter-quartile range) time between consecutive CD4 counts: 95 (71-137) days
- Current dataset includes information on over 29,000 patients
- Median (range) duration of follow-up: around 4 years (1 day – 20+ years) years

Distinction between the two types of cohort

	Traditional cohort	Observational database
Study visits	At regular defined intervals	As and when patient attends for care
Data entry	Often form-based	Often electronic transfer of data
Representative?	May not be – patients must give consent	Often includes *all* patients – therefore representative
Loss to follow-up	May be substantial, but can be determined	May be difficult to assess as some patients attend infrequently
Data quality	Can introduce quality control measures	Difficult to regulate
Data items collected	Can determine at outset and change over time	May be difficult to influence

Key outcome variables

- An ideal outcome should address the primary aim of the study, should have biological/clinical relevance and should be appropriate for the population included in the cohort
- An ideal outcome should be ascertainable on *all* cohort participants (including those lost-to-follow-up), either through regular follow-up visits or through other sources, e.g.:
  - national death registries
  - cancer registries
  - death certificates/autopsy reports

Key outcome variables

- Aim: To investigate hepatocellular carcinoma risk factors in chronic hepatitis B patients
- Outcomes:
  - Ultrasound of the abdomen, computerised tomography, hepatic angiogram, and/or liver biopsy were performed if α fetoprotein levels were >50 mg/l or demonstrated a rising trend >20 mg/l to confirm the diagnosis of HCC
  - Clinical liver cirrhosis defined as ultrasonic features of liver cirrhosis plus evidence of hypersplenism (splenomegaly with platelet count <100 $x10^9$/l or white count <4 $x10^9$/l), clinical ascites, varices, and/or hepatic encephalopathy

H L-Y Chan et al; Gut; 2004; 53; 1494-1498

Exposures, confounders and effect modifiers

- Whilst most RCTs consider a single exposure, it is possible to consider more exposures in a cohort
- These exposures should be defined at the start of the study
- Potential confounders and effect modifiers (factors that change the strength of the association between the outcome and exposure) should also be defined in advance

Exposures, confounders and effect modifiers

- As measurements may change over the study period, a patient's status should be re-assessed at regular times during the study
- The frequency at which each measurement is assessed will depend on the likelihood of it changing over time, as well as the reliability of the data sources
- Example: dietary factors, smoking status, alcohol consumption

Key outcome variables

- Aim: To investigate hepatocellular carcinoma risk factors in chronic hepatitis B patients
- Ascertainment:
  - Patients were prospectively followed up every 6 months or more frequently if clinically indicated, with monitoring of liver biochemistry, HBeAg, and antibody to hepatitis B e antigen status, as well as α fetoprotein levels
  - For patients with normal α fetoprotein levels, ultrasound of the abdomen was performed every 1-2 years
  - HBV genotyping was performed and basal core promotor mutations were determined in the residual serum sample of each patient at the initial visit

H L-Y Chan et al; Gut; 2004; 53; 1494-1498

Choosing the optimal sample size

- The power of a cohort is largely determined by the number of events that occur; this can be increased either by increasing the size of the cohort or by lengthening the period of follow-up
- Whilst large cohorts may sometimes be desirable, the real value of many cohorts is provided by their length of follow-up and detailed data collection

Examples

- Royal Free Haemophilia HIV cohort
  - 111 men with haemophilia infected with HIV between 1979 and 1985; over the 25 years of the study, the cohort published >100 papers in peer-reviewed journals, including many that were highly influential at the time
- The D:A:D Study
  - Using existing data, it was estimated that 100 incident myocardial infarction (MI) cases would be required to detect a doubling in the risk of MI in those receiving cART; the investigators calculated that they would require 30,000 PYFU to achieve this

Outline of Session

- The limitations of RCTs
- Observational studies – their benefits and limitations
- Designing a cohort study
- Designing a case-control study

General points

- Retrospective, so reliant on recorded data (which may contain inaccuracies and be subject to missing data)
- Care should always be taken to ensure that the timing of events (e.g. exposures, outcome) is captured accurately
- As with cohort studies, criteria for selection of cases/controls (e.g. outcome) and exposures should be standardised, precise and unambiguous

Selection of cases

- Criteria for selection may include clinical definitions (e.g. CDC AIDS definition), laboratory or histological classifications
- Must decide whether the study will include only incident cases (i.e. new cases identified in population during study period) or prevalent cases (cases who had already been diagnosed at the study of the study period)

Matching in case-control studies

- Cases and controls may often be matched on a small number of factors associated with both the exposure and outcome (e.g. sex, age)
- Matching may help to minimise effects of confounding and may increase study power
- However, it may be impractical to match patients on many factors and special analytical methods may be required if matching is used
- If a factor has been used in matching, then it is not possible to evaluate its association with the outcome

Selection of controls

- Controls should be drawn from the same population as cases
- May include 1 (1:1) or >1 (n:1) control per case, but little to be gained from >4 controls per case
- Controls may be people using the same hospital as cases (hospital controls) – if so, care should be taken to ensure that they don't have another condition that is also related to the exposure
- Use of friends/relatives/neighbours of cases may give cases of similar socio-economic background

Example

- Risk factors for thrombocytopenia in HIV-infected persons attending one of two HIV clinics in New York
- Cases: patients with platelet count $<100x10^9/L$ persisting for >3 consecutive months; cases could be incident or prevalent (n=73)
- Controls: patients with at least one outpatient record but no thrombocytopenia (n=73)
- Cases were matched 1:1 by age ($\pm$5 years), sex and first appointment ($\pm$6 months)

Marks KM et al; JAIDS 2009; 52: 595-599.

Nested case-control studies

- A case-control study may often be nested within a larger cohort or RCT
- This provides a means of studying associations between novel biomarkers and disease outcome, particularly if these are expensive to measure
- Alternatively, may be useful if additional detailed information is required which cannot be collected through standard data collection mechanism
- Example: nested case-control study in SMART trial, measured lipoprotein particles in 248 patients with a CVD event (cases) and 480 matched controls

Which study design?

- Largely an issue of resources, although it is possible to use a combination of approaches
- Important to consider whether the data collected using a particular design will be able to answer the question of interest
- Can often be a compromise between costs and the amount and type of data that can be collected
- Ultimately it is important to be aware of, and transparent about, the limitations of each study design

Where to go for guidance?

- Similar to CONSORT but provides guidance on reporting of observational studies
- Provides a checklist for reporting studies, as well as educational material
- Recommendations limited to 3 main designs of observational studies
  - Cohort
  - Case-control
  - Cross-sectional studies
- www.strobe-statement.org

Summary

- Whilst RCTs are perceived to provide the highest quality evidence when assessing associations, they may sometimes suffer from limitations which make them inappropriate for use when addressing certain questions
- In these situations, observational studies may provide useful information
- However, observational studies are always subject to bias and must be designed, managed and interpreted with caution so as to minimise this

Your turn…