Mini lecture:

Conducting and managing observational studies

Caroline Sabin

Conflict of interest

I have received funding for the membership of Data Safety and Monitoring Boards, Advisory Boards and for the preparation of educational materials from:

- Gilead Sciences
- ViiV Healthcare
- Janssen-Cilag

Outline of Session

- The limitations of RCTs
- Designing a cohort study
- Designing a case-control study

Outline of Session

- The limitations of RCTs
- Designing a cohort study
- Designing a case-control study

Limitations of RCTs

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

Limitations of RCTs

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

Limitations of RCTs

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

Limitations of RCTs

3. RCTs may be short (48/96 weeks) and may focus on two or three main treatment comparisons

Limitations of RCTs

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

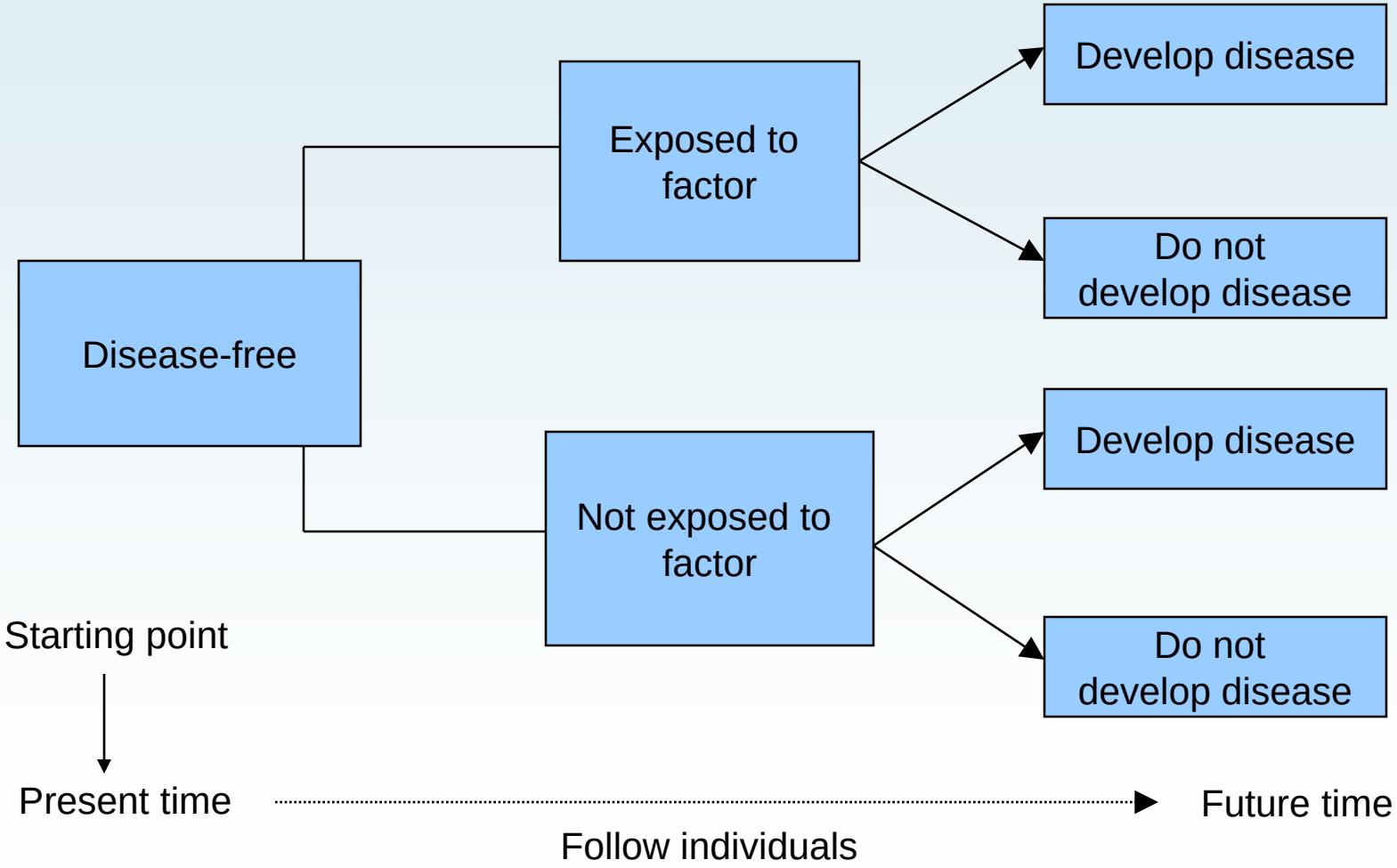
   Example…

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

Outline of Session

- The limitations of RCTs
- Designing a cohort study
- Designing a case-control study

Reminder – Cohort Studies



Basic study design issues

- Important to have a clear objective for the study and to design accordingly
- Ensure that sample size will be sufficient to address at least one key hypothesis
- Participants included in cohort should be representative of the population to which the results will be generalised

Ways of following individuals

- Failure to ascertain all disease events can result in under-estimation of event rates
- Can also lead to bias in comparisons between exposure levels
- Nationally recorded registers/databases
  - Death certificates
  - Disease registers, e.g. cancer registry
  - (In UK) NHS electronic information systems e.g. hospital episode statistics, GP databases
  - Office of national statistics (ONS)
- Other efforts to contact people (e.g. phone call, letter)

Key outcome variables

- Ideal outcome should address the primary aim of the study, have biological/clinical relevance and be appropriate for the population studied
- Should be ascertainable on all cohort participants (including those lost-to-follow-up)

Toxicity outcomes

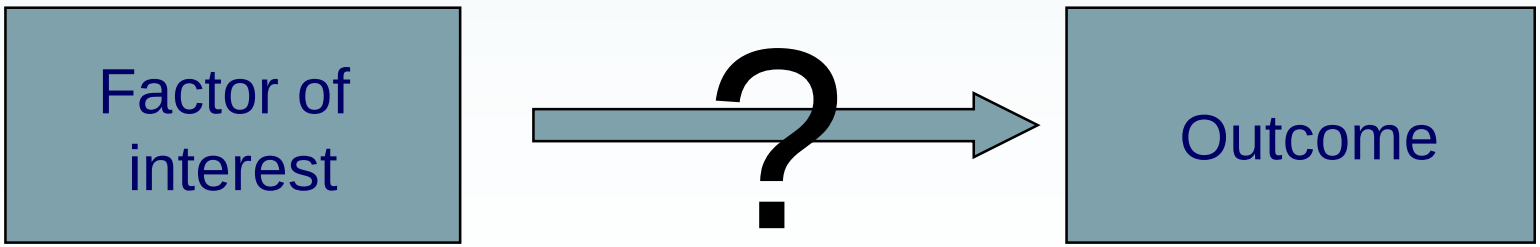
- Cohorts may be the only study design that is able to capture data on long-term toxicities of HAART
- Toxicity outcomes may be based on clinical symptoms and/or laboratory data
- Need to be aware of possible biases when interpreting results from such studies:
  - Irregular/infrequent laboratory monitoring
  - Selective laboratory monitoring
  - Between-clinic assay variability
  - Clinic differences in monitoring policies
  - Bias due to confounding

What other data should be captured?

- Potential confounders (traditional definition):
  - Factors that are associated both with the exposure and outcome of interest
  - Failure to adjust for confounders may introduce bias, as they may lead to a spurious association between the exposure and outcome
- Effect modifiers:
  - Factors that modify the size of the association in one group compared to another
  - Provide important clinical information
  - Often referred to as a ‘statistical interaction’

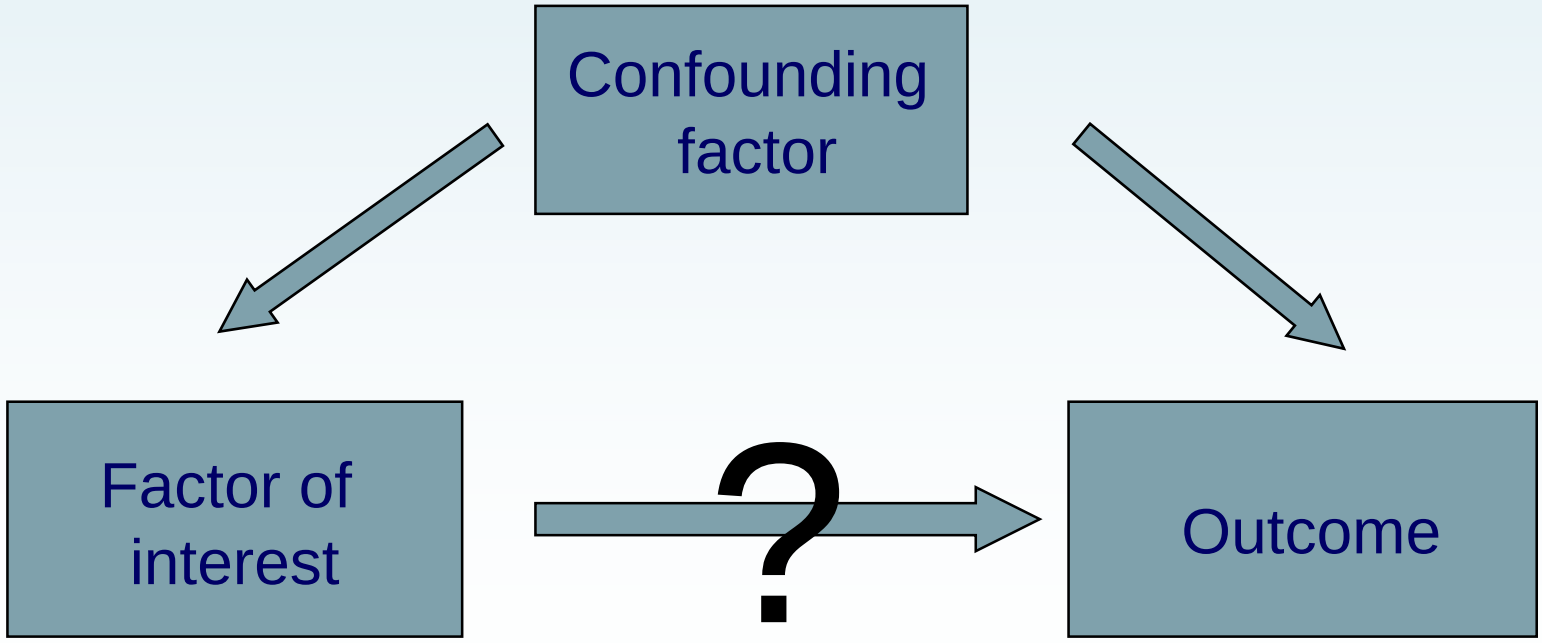
Bias due to confounding

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



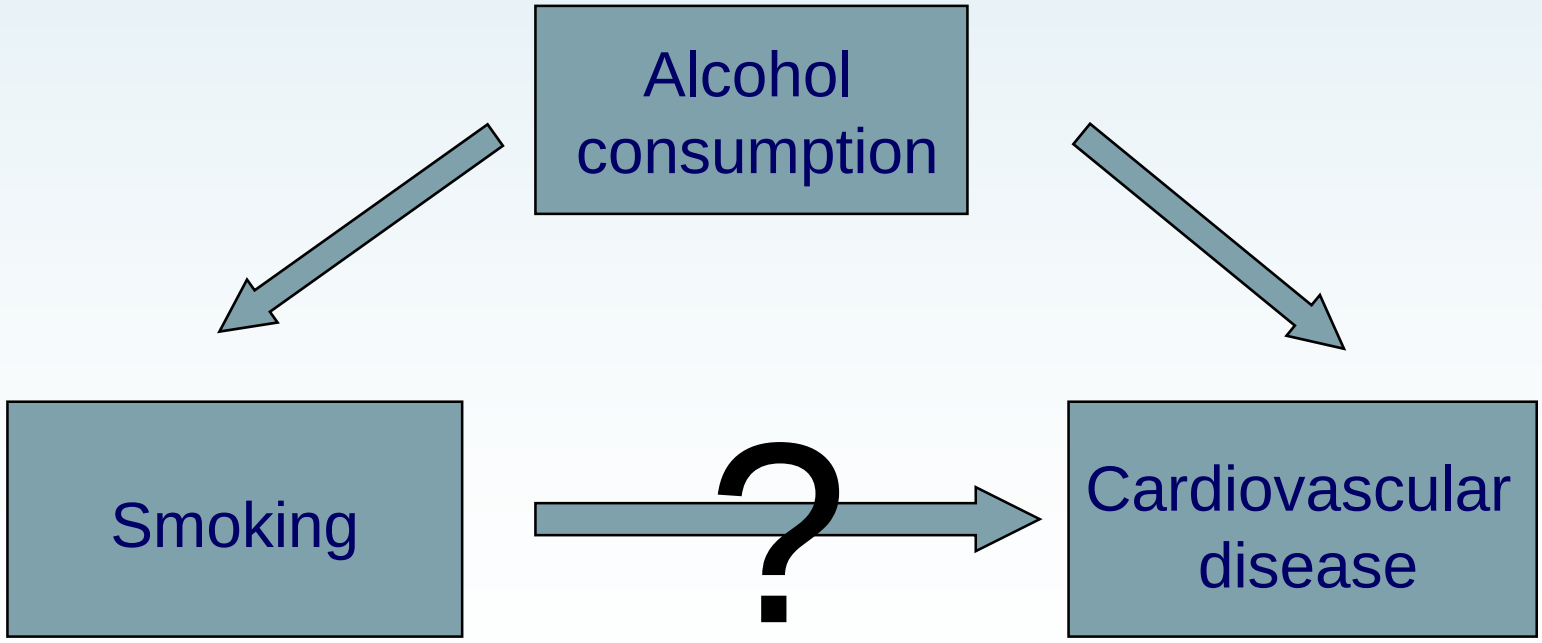
Bias due to confounding

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



Bias due to confounding

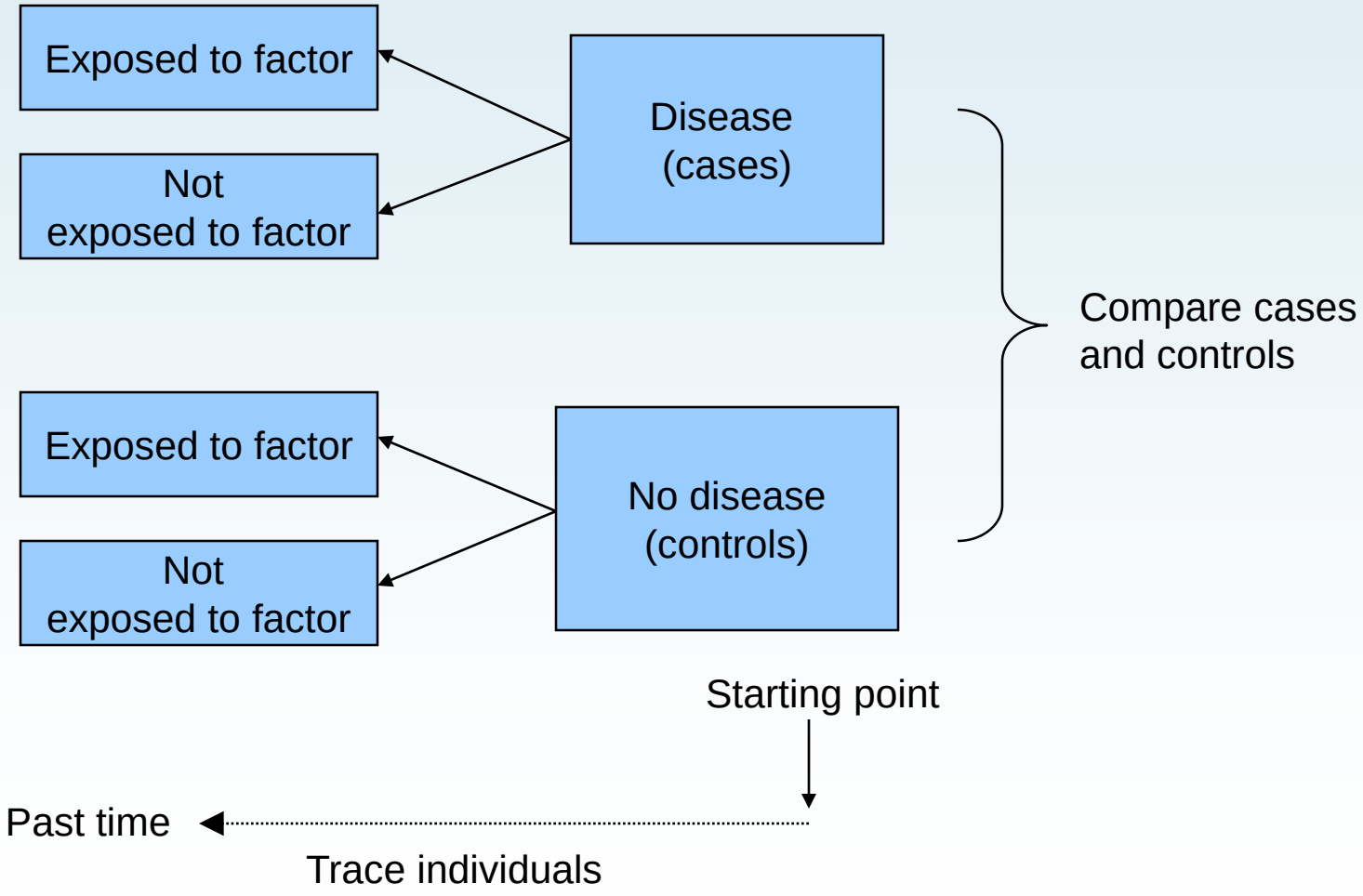
- Occurs when a spurious association arises due to a failure to fully adjust for factors related to both the risk factor and outcome



Outline of Session

- The limitations of RCTs
- Designing a cohort study
- Designing a case-control study

Reminder – Case-control studies



General points

- Retrospective, so reliant on recorded data (which may contain inaccuracies and be subject to missing data)
- Useful for rare diseases and diseases with long latency periods
- Care should always be taken to ensure that the timing of events (e.g. exposures, outcome) is captured accurately

Selection of cases

- Develop a **case definition** to identify cases that is precise, objective and unambiguous
- This could include:
  - Histologically or laboratory confirmed diagnosis
  - Clinical diagnosis
  - Stages of disease (standardised e.g. CDC AIDS definition)
- Source of cases needs to be carefully considered
  - Population based or clinic based cases?
  - How complete is your source of cases?
  - Time lag between diagnosis and notification/ identification?
  - What about patients who may have moved or died?

Selection of controls

- Controls should be selected to be as similar to cases as possible except for the outcome of interest
  - Drawn from the same population
  - Fulfil the same eligibility criteria
- Aim is to provide estimate of “level of exposure” in those without outcome
- Should represent same population as cases are
  - General population (voting registries, random digit dialling etc)
  - Hospital/clinic based controls – care should be taken to ensure they don’t have another condition also related to the exposure
  - People related to the case – i.e. friends, relatives, neighbours

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
- But, 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

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

Nested case-control studies

- Case-control study may often be nested within a larger cohort or RCT
- Provides a means of studying associations between novel biomarkers and disease outcome
- May also 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

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