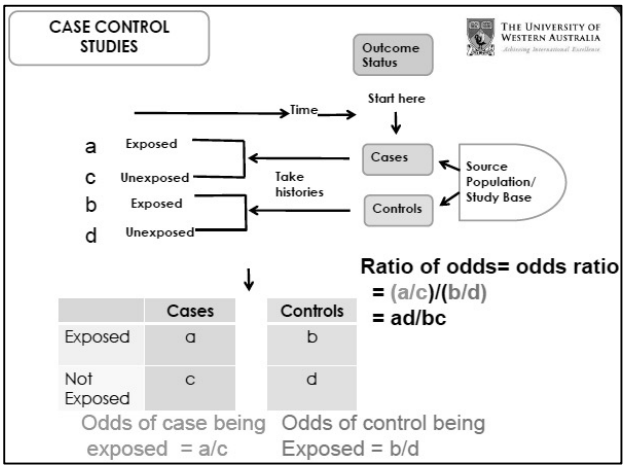
Case-control:

Start with OUCTOME & look back at EXPOSURE



***Q:** What are the **odds** that a **case** was **exposed?**
***A **rare OUTCOME** and an **EXPOSURE** that can recalled
***Measuring **EXPOSURE** but **Selection of subjects** is **INDEPENDENT** of it
***To sufficiently match cases and non-cases**
***Selection and recall bias**
***Retrospective**

Design:

1) Study **base**/source: the source of all exposed unexposed cases & controls
    a. All people **at risk** in the population
    b. All **OUTCOMES** must come from this person-time source
2) Selection strategies
    a. On the basis of **OUTCOME** which is a **rare DISEASE**
    b. Chosen Independently of **EXPOSURE** (unknown); otherwise there would be **selection bias**
    c. Must be **representative** of the study base
    d. **Cases**: all **incident** cases selected
        i. Person, place, time, **exclusion** criteria
        ii. **Not prevalent cases;** only used if disease onset is difficult to identify
        iii. Is there selection bias? Is **EXPOSURE** influencing selection into the study?
    e. **Controls**: should represent probability of **EXPOSURE** in population **AT RISK**
        i. Person, place, time, **exclusion** criteria
        ii. **Chosen independently** of **EXPOSURE**
        iii. One of the **trickiest** parts in this study design; provide estimate of **EXPOSURE** level expected to occur if there was no association found between **DISEASE** and **EXPOSURE**
3) Design questions:
    a. Explicit research question
    b. Define cases and controls – how many?
    c. Match / not-match?
    d. Bias and confounding
    e. Analysis
4) Types:
    a. **Matching:** selecting controls so that they are similar to the cases in certain factors thought to be associated with the outcome i.e. age, sex, race, SES, occupation
        i. **Group** matching: **frequency** matching
        ii. **Individual** matching: matched **pairs**
            1. **Conditional logistic regression** rather than OR
        iii. **CANNOT** study the factor that was matched on
        iv. Risk of **over-matching**
    b. **Nested**
    c. **Case-control within cohort study**
5) Measuring Association (**MoA**) = **OR**
    a. **OR = 1** no association between **EXPOSURE** and **OUTCOME/DISEASE**
    b. **OR > 1** positive association between **OUTCOME/DISEASE** and **EXPOSURE**
    c. **OR < 1** negative association between **OUTCOME/DISEASE** and **EXPOSURE**
    ***Sensitivity analysis** checks if the **OR** is free of **bias**

Item	**Case-control**
Population	Study base or source
Allocation	Based on outcome/disease
Outcomes	Rare, basis of subject allocation
Follow-up	Retrospective
Analysis	Deal with bias & confounding
MoA	OR, no incidence

Sources of error:

1) **Misclassification** of **OUTCOME** and **EXPOSURE**
    d. Minimised by:
        i. Using existing records
        ii. Using well-designed questionnaires
        iii. Use trained interviewers
        iv. Blind subjects to question; ethical?
        v. Memory triggers
        vi. Validate against records

2) Effect of **differential misclassification** on **OR**:
    a. Misclassification of **EXPOSURE** is different according to **OUTCOME** status
    b. Non differential misclassification **DECREASES** the **OR**
    c. Differential misclassification:
        i. With **DISEASE (cases)** deny **EXPOSURE = DECREASE OR**
        ii. Without **DISEASE (controls)** deny **EXPOSURE = INCREASE OR =** Causes havoc; better non-dif

****In **cohort**, misclassification of **OUTCOME** is different according to **EXPOSURE** status

3) **Confounding**: where the relationship between an **EXPOSURE** and an **OUTCOME** is due in part (or wholly) to a **3rd factor** that is differently distributed between the groups.
    a. Must be a **known risk factor**
    b. Be **associated** with the **EXPOSURE** but not a result of it

4) Bias
    a. **Selection bias** of both cases and controls
    b. Information/observation bias i.e. recall

Case control study within cohort study

1) Controls are not individually matched to cases
2) Advantages:
    a. Reduced recall bias
    b. No selection bias as cases and controls are from the same population
    c. Temporality of data; **EXPOSURE** data available predating **OUTCOME**
    d. Economical

Advantages	Disadvantages
-Choosing all cases is efficient -Efficient for rare diseases with long latency periods -Can study several exposures -Cheaper and takes less time	-Inefficient for rare exposures -Selection bias if based on exposure -Exposure is subject to observation bias -Recall and info bias -No incidence calculations

	Nested case-control	Case-control **within cohort**
Cases	Selected as they arise in the cohort	All cases identified in the cohort
Controls	Selected from full cohort at time of specific case identification; those at risk AT THE TIME of case identification	Randomly selected from whole cohort
Matching	Cases are matched on calendar time and length of follow up	Pro: can use same control series for case series of different diseases