

Summary: Cross-Sectional Studies

Cord Heuer

Prof. em. Veterinary Epidemiology and Infectious Diseases
EpiCentre, Massey University



Collaborating Centre for Veterinary
Epidemiology & Public Health

Ministry for Primary Industries
Manatū Ahu Matua



Schedule

Workshop Contents

- **W1: Cross-sectional**
- W2: Questionnaire design
- W3: Case-control
- W4: Scenario 1: X-sectional
- W5: Scenario 2: case-control

Facilitators

- **Cord Heuer, Arata Hidano**
- Naomi Cogger
- Cord Heuer, Ashish Sutar
- Cord Heuer, Art Subharat
- Cord Heuer, Art Subharat

Date

- **11 Nov'21**
- 18 Nov'21
- 25 Nov'21
- 19 Nov'21
- 10 Dec'21

Structure

- Listening: introductory video-lectures online
- Reading: material provided online
- Online learning modules (3), each structured as:
 - Participants have watched videos ahead of online session
 - **20min summary of online video lecture**
 - 20min presentation of a case study
 - 50min Q&A, discussion of specific design features/questions

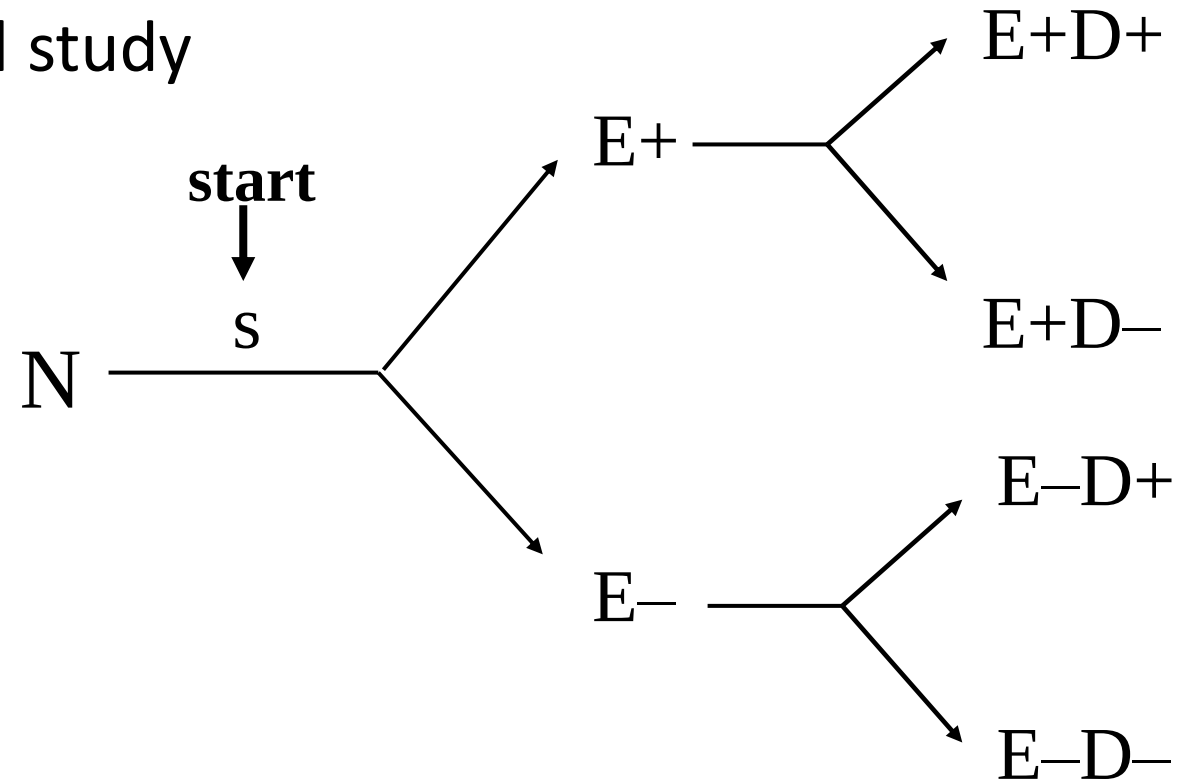
Learning outcomes

- Understand critical features of the cross-sectional design
- Critically review study design features of published reports
- Appropriately select study participants
- Determine the study size
- Choose appropriate analytical methods and adjust for bias
- Draw valid inferences

Summary

Cross-sectional design ('survey')

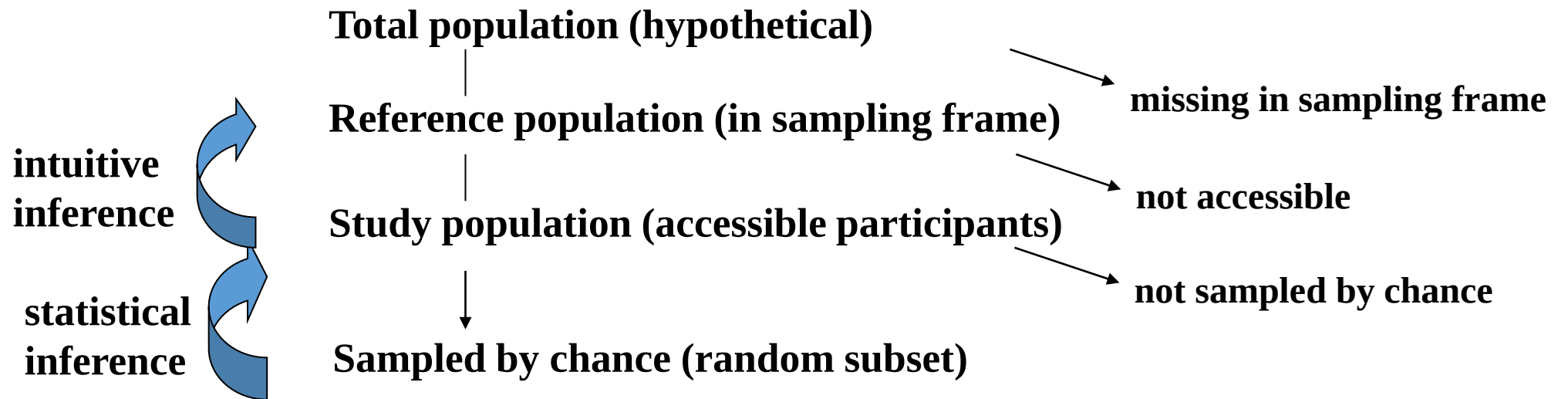
- Cross-sectional study



Q: What is the status?

Selection of study participants

- Random selection of sampling units from a **sampling frame**.
 - sampling independent of exposure and disease
 - sample $\approx$ population



Key features of cross-sectional studies

- To quantify the occurrence of 'disease' in the population
 - When 'disease' leaves a lasting response (e.g. antibody)
 - When 'disease' is reasonably frequent (e.g. >5%)
- To quantify the frequency of exposure in a population
- To evaluate associations between exposure and 'disease'
- Often: multi-level selection of subjects: villages → farms → animals
- *Not an efficient design for rare or short disease*
- *Not suitable to evaluate causal hypotheses*

Outcome measures: prevalence

- Prevalence (of exposure AND disease) and associations (OR, PR)
 - Answer of the homework question → next slide (1)
- Incidence
 - *“depending on the recall of participants, past disease events may be inquired and used to estimate incidence (assuming that population size has not changed over time)”*
 - **Recall** of disease events over time (e.g. 1 year, last season) → next slide (2)
 - Attack rate when TAR is short and as long as the population hasn't changed (e.g. outbreaks)
 - IF 'disease' is endemic at constant rates AND its **duration is known**, one can approximate incidence → next slide (3)

Outcome measures for association (1)

1. If D+ were 20%, how would OR compare to PR?

$$\text{OR} = 80 \cdot 780 / (720 \cdot 20) = 4.33 \rightarrow \text{OR} = 256 \cdot 666 / (614 \cdot 64) = \underline{4.33}$$

$$\text{PR} = 80/800 / (20/780) = 4.00 \rightarrow \text{PR} = 256/870 / (64/730) = \underline{3.35}$$

		<i>Outcome</i>	
		D+	D-
<i>Exposure</i>	E+	80	720
	E-	20	780



		<i>Outcome</i>	
		D+	D-
<i>Exposure</i>	E+	256	614
	E-	64	666

Rare disease: $100 / 1600 = 6.25\%$

Frequent disease: $320 / 1600 = 20\%$

Outcome measures for association (1)

2. Does the OR/PR relationship also depend on the frequency of exposure? No:

$$OR = 80 \cdot 780 / (720 \cdot 20) = 4.33 \rightarrow OR = 24 \cdot 1387 / (56 \cdot 133) = \underline{4.33}$$

$$PR = 80/800 / (20/800) = 4.00 \rightarrow PR = 8/80 / (38/1520) = \underline{4.00}$$

Outcome

D+

D-

E+

80

720

E-

20

780

Exposure



Exposure

Outcome

D+

D-

E+

8

72

E-

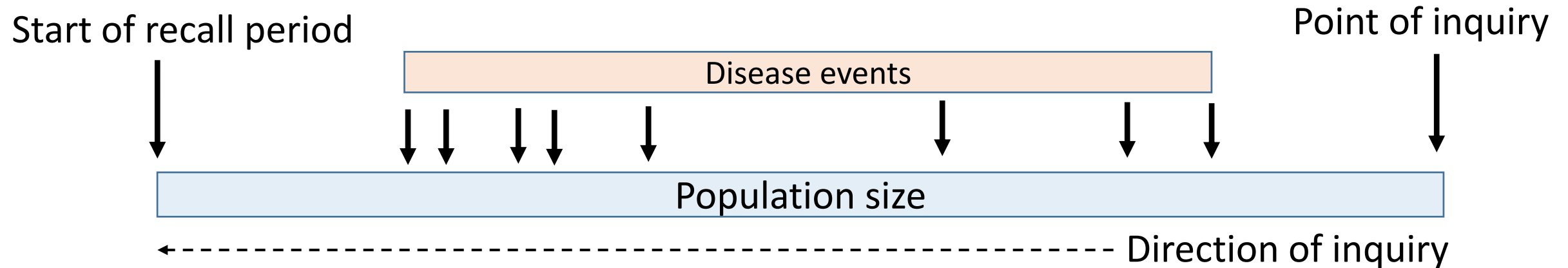
38

1482

Frequent exposure: $800 / 1600 = 50\%$

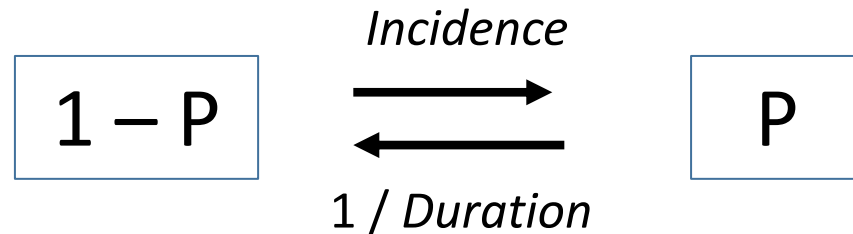
Rare exposure: $80 / 1600 = 5\%$

'Incidence' by Recall (2)



- Conditions:
 - Disease is endemic
 - Population size and composition are constant (e.g. large surveys of 1000's of households) or changes were recorded or can be recalled
 - Recall is 'perfect' (e.g. mortality, clear clinical signs like ASF)

'Incidence' when Duration is known (3)



At equilibrium: $(1 - P) * \textit{Incidence} = P * 1 / \textit{Duration}$

$$\textit{Incidence} = \frac{P}{(1 - P) * \textit{Duration}}$$

- Conditions:
 - Disease is endemic
 - Population size and composition are constant
 - Example: serological prevalence where antibody duration is known

Group exercise

- 4 groups

- | | |
|-------------------------------------|------------|
| 1. Thailand, Laos, Cambodia | Art |
| 2. PNG, Vietnam | Ash, Arata |
| 3. Malaysia, Indonesia, Philippines | Cord |
| 4. China, Mongolia | Boloru |

Discussion questions

Group

- Thailand, Laos, Cambodia
- Vietnam, PNG
- Malaysia, Indonesia, Philippines
- China, Mongolia

Questions

- What are the strengths and limitations of slaughterhouse sampling as opposed to sampling on farm?
- How can we minimise the limitation of slaughterhouse sampling?
- What was the target and what was the sampling population; was selection bias likely?
- Are samples collected at different time points likely describing the same population: if not, what information would you need to answer this question?
<https://statulator.com/SampleSize/ss1P.html>
- How do you interpret the result of nasal swab samples tested by qPCR as opposed to seroprevalence?
- How does this affect the sample size?
- What are important risk factors determining prevalence of either antibody (ELISA) or viral antigen (PCR)?
- What outcome measures would you use for this study?