Summary: Cross-Sectional Studies

Cord Heuer
Prof. em. Veterinary Epidemiology and Infectious Diseases
EpiCentre, Massey University
## Schedule

<table>
<thead>
<tr>
<th>Workshop Contents</th>
<th>Facilitators</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>• <strong>W1</strong>: Cross-sectional</td>
<td>• Cord Heuer, Arata Hidano</td>
<td>- 11 Nov’21</td>
</tr>
<tr>
<td>• W2: Questionnaire design</td>
<td>• Naomi Cogger</td>
<td>- 18 Nov’21</td>
</tr>
<tr>
<td>• W3: Case-control</td>
<td>• Cord Heuer, Ashish Sutar</td>
<td>- 25 Nov’21</td>
</tr>
<tr>
<td>• W4: Scenario 1: X-sectional</td>
<td>• Cord Heuer, Art Subharat</td>
<td>- 19 Nov’21</td>
</tr>
<tr>
<td>• W5: Scenario 2: case-control</td>
<td>• Cord Heuer, Art Subharat</td>
<td>- 10 Dec’21</td>
</tr>
</tbody>
</table>
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

```
Total population (hypothetical) --> Reference population (in sampling frame) --> missing in sampling frame

intuitive inference

Reference population (in sampling frame) --> Study population (accessible participants) --> not accessible

statistical inference

Study population (accessible participants) --> Sampled by chance (random subset) --> not sampled by chance
```
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 $\rightarrow$ farms $\rightarrow$ 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)
1. If D+ were 20%, how would OR compare to PR?

   \[
   \text{OR} = \frac{80 \times 780}{720 \times 20} = 4.33 \quad \rightarrow \quad \text{OR} = \frac{256 \times 666}{614 \times 64} = 4.33
   \]

   \[
   \text{PR} = \frac{80}{800} / (20/780) = 4.00 \quad \rightarrow \quad \text{PR} = \frac{256}{870} / (64/730) = 3.35
   \]

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Outcome</th>
<th>E+</th>
<th>E-</th>
</tr>
</thead>
<tbody>
<tr>
<td>D+</td>
<td>80</td>
<td>720</td>
<td></td>
</tr>
<tr>
<td>D-</td>
<td>20</td>
<td>780</td>
<td></td>
</tr>
</tbody>
</table>

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:

$$\text{OR} = \frac{80 \times 780}{720 \times 20} = 4.33$$  \quad \rightarrow \quad \text{OR} = \frac{24 \times 1387}{56 \times 133} = 4.33$$

$$\text{PR} = \frac{80}{800} / \frac{20}{800} = 4.00$$  \quad \rightarrow \quad \text{PR} = \frac{8}{80} / \frac{38}{1520} = 4.00$$

Frequent exposure: $\frac{800}{1600} = 50\%$

Rare exposure: $\frac{80}{1600} = 5\%$
‘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: \((1 - P) \times \text{Incidence} = P \times \frac{1}{\text{Duration}}\)

\[
\text{Incidence} = \frac{P}{(1 - P) \times \text{Duration}}
\]

• 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

<table>
<thead>
<tr>
<th>Group</th>
<th>Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thailand, Laos, Cambodia</td>
<td>- What are the strengths and limitations of slaughterhouse sampling as opposed to sampling on farm?</td>
</tr>
<tr>
<td></td>
<td>- How can we minimise the limitation of slaughterhouse sampling?</td>
</tr>
<tr>
<td>Vietnam, PNG</td>
<td>- What was the target and what was the sampling population; was selection bias likely?</td>
</tr>
<tr>
<td></td>
<td>- Are samples collected at different time points likely describing the same population? If not, what information would you need to answer this question?</td>
</tr>
<tr>
<td>Malaysia, Indonesia, Philippines</td>
<td>- How do you interpret the result of nasal swab samples tested by qPCR as opposed to seroprevalence?</td>
</tr>
<tr>
<td></td>
<td>- How does this affect the sample size?</td>
</tr>
<tr>
<td>China, Mongolia</td>
<td>- What are important risk factors determining prevalence of either antibody (ELISA) or viral antigen (PCR)?</td>
</tr>
<tr>
<td></td>
<td>- What outcome measures would you use for this study?</td>
</tr>
</tbody>
</table>