London School of Hygiene and Tropical Medicine

PLANNING OF INVESTIGATIONS

D.R. Cox
david.cox@nuffield.ox.ac.uk

LECTURE 3

- Formulation of questions
- Study design
- Measurement issues
- Analysis
- Interpretation

- Units of study
- Intrinsic variables
- Explanatory variables
- Outcomes

- Cross-sectional studies
- Prospective observational studies
- Retrospective observational studies
- Experiments

1. Implementation criteria
   - Avoidance of systematic error
   - Control of random error

1. Implementation criteria
   - Avoidance of systematic error
   - Control of random error
     1. Quality of measurement
     2. Appropriate scale of effort

1. Implementation criteria
   - Avoidance of systematic error
   - Control of random error
     1. Quality of measurement
     2. Appropriate scale of effort
   - Multiplicity of questions (factorial principle)

1. Implementation criteria
   - Avoidance of systematic error
   - Control of random error
     1. Quality of measurement
     2. Appropriate scale of effort
   - Multiplicity of questions (factorial principle)
   - Intermediate assessment and modification

1. Implementation criteria
   - Avoidance of systematic error
   - Control of random error
     1. Quality of measurement
     2. Appropriate scale of effort
   - Multiplicity of questions (factorial principle)
   - Intermediate assessment and modification
     1. Initial data audit; sampling principles
     2. Inadequate accrual
     3. More major intermediate changes; when are these desirable?
     4. Dealing with unanticipated complications

2. Factorial design

- Contrast with the large simple trial
- A traditional type example
    1. Rahman et al (2011); Ma et al (*J. Nat. Cancer Inst.* **104** (2012), 488-492)

2. Factorial design

- Contrast with the large simple trial
- A traditional type example
    1. Rahman et al (2011); Ma et al (*J. Nat. Cancer Inst.* **104** (2012), 488-492)
    2. Shandong intervention trial; 3 distinct treatments for reducing prevalence of precancerous gastric lesions and longer term outcomes: 7 years of garlic extract and oil; 7 years of vitamin and selenium treatment; 2 weeks of antibiotic treatment. 8 combinations. More than 3000 patients with follow up. Strong effect of antibiotic treatment. At most minor effects of others and no serious interaction.

- Fractional replication

- Fractional replication
- Key idea; three binary factors, A, B, C: 8 combinations.
  2^3 design
  $a_0 b_0 c_0, a_1 b_0 c_0, a_0 b_1 c_0, a_1 b_1 c_0, a_0 b_0 c_1, a_1 b_0 c_1, a_0 b_1 c_1, a_1 b_1 c_1,$
  Half replicate:
  $a_0 b_0 c_0, a_1 b_1 c_0, a_1 b_0 c_1, a_0 b_1 c_1$

- Fractional replication
- Key idea; three binary factors, A, B, C: 8 combinations.
  2^3 design
  $a_0 b_0 c_0, a_1 b_0 c_0, a_0 b_1 c_0, a_1 b_1 c_0, a_0 b_0 c_1, a_1 b_0 c_1, a_0 b_1 c_1, a_1 b_1 c_1,$
  Half replicate:
  $a_0 b_0 c_0, a_1 b_1 c_0, a_1 b_0 c_1, a_0 b_1 c_1$
  Main effect of A
  $a_0 b_0 c_0^-, a_1 b_1 c_0^+, a_1 b_0 c_1^+, a_0 b_1 c_1^-$

- Fractional replication
- Key idea; three binary factors, A, B, C: 8 combinations.
  2^3 design
  $a_0 b_0 c_0, a_1 b_0 c_0, a_0 b_1 c_0, a_1 b_1 c_0, a_0 b_0 c_1, a_1 b_0 c_1, a_0 b_1 c_1, a_1 b_1 c_1,$
  Half replicate:
  $a_0 b_0 c_0, a_1 b_1 c_0, a_1 b_0 c_1, a_0 b_1 c_1$
  Main effect of A
  $a_0 b_0 c_0^-, a_1 b_1 c_0^+, a_1 b_0 c_1^+, a_0 b_1 c_1^-$
  Interaction B times C

- Fractional replication
- Key idea; three binary factors, A, B, C: 8 combinations.
  2^3 design
  $a_0 b_0 c_0, a_1 b_0 c_0, a_0 b_1 c_0, a_1 b_1 c_0, a_0 b_0 c_1, a_1 b_0 c_1, a_0 b_1 c_1, a_1 b_1 c_1,$
  Half replicate:
  $a_0 b_0 c_0, a_1 b_1 c_0, a_1 b_0 c_1, a_0 b_1 c_1$
  Main effect of A
  $a_0 b_0 c_0^-, a_1 b_1 c_0^+, a_1 b_0 c_1^+, a_0 b_1 c_1^-$
  Interaction B times C
- Consequence: aliasing

- Fractional replication
- Key idea; three binary factors, A, B, C: 8 combinations.
  2^3 design
  $a_0 b_0 c_0, a_1 b_0 c_0, a_0 b_1 c_0, a_1 b_1 c_0, a_0 b_0 c_1, a_1 b_0 c_1, a_0 b_1 c_1, a_1 b_1 c_1,$
  Half replicate:
  $a_0 b_0 c_0, a_1 b_1 c_0, a_1 b_0 c_1, a_0 b_1 c_1$
  Main effect of A
  $a_0 b_0 c_0^-, a_1 b_1 c_0^+, a_1 b_0 c_1^+, a_0 b_1 c_1^-$
  Interaction B times C
- Consequence: aliasing
- Possible application: community trials

- Fractional replication
- Key idea; three binary factors, A, B, C: 8 combinations.
  2^3 design
  $a_0 b_0 c_0, a_1 b_0 c_0, a_0 b_1 c_0, a_1 b_1 c_0, a_0 b_0 c_1, a_1 b_0 c_1, a_0 b_1 c_1, a_1 b_1 c_1,$
  Half replicate:
  $a_0 b_0 c_0, a_1 b_1 c_0, a_1 b_0 c_1, a_0 b_1 c_1$
  Main effect of A
  $a_0 b_0 c_0^-, a_1 b_1 c_0^+, a_1 b_0 c_1^+, a_0 b_1 c_1^-$
  Interaction B times C
- Consequence: aliasing
- Possible application: community trials
- Blot et al (*J. Nat. Cancer Inst.* **85** (1993), 1483-1482) used one-half replicate of 2^4 with 30,000 patients. Justification

- Fractional replication
- Key idea; three binary factors, A, B, C: 8 combinations.
  2^3 design
  $a_0b_0c_0, a_1b_0c_0, a_0b_1c_0, a_1b_1c_0, a_0b_0c_1, a_1b_0c_1, a_0b_1c_1, a_1b_1c_1,$
  Half replicate:
  $a_0b_0c_0, a_1b_1c_0, a_1b_0c_1, a_0b_1c_1$
  Main effect of A
  $a_0b_0c_0^-, a_1b_1c_0^+, a_1b_0c_1^+, a_0b_1c_1^-$
  Interaction B times C
- Consequence: aliasing
- Possible application: community trials
- Blot et al (*J. Nat. Cancer Inst.* **85** (1993), 1483-1482) used one-half replicate of 2^4 with 30,000 patients. Justification
- 4 combinations of nutrients for lowering incidence and mortality from stomach cancer

- Fractional replication
- Key idea; three binary factors, A, B, C: 8 combinations.
  2^3 design
  $a_0b_0c_0, a_1b_0c_0, a_0b_1c_0, a_1b_1c_0, a_0b_0c_1, a_1b_0c_1, a_0b_1c_1, a_1b_1c_1,$
  Half replicate:
  $a_0b_0c_0, a_1b_1c_0, a_1b_0c_1, a_0b_1c_1$
  Main effect of A
  $a_0b_0c_0^-, a_1b_1c_0^+, a_1b_0c_1^+, a_0b_1c_1^-$
  Interaction B times C
- Consequence: aliasing
- Possible application: community trials
- Blot et al (*J. Nat. Cancer Inst.* **85** (1993), 1483-1482) used one-half replicate of 2^4 with 30,000 patients. Justification
- 4 combinations of nutrients for lowering incidence and mortality from stomach cancer
- design: 1/2 replicate of 2^4

3. Randomization

- Definition

3. Randomization

- Definition
- Justification both in design of experiments and in sampling
    1. Impersonality, with concealment, controls conscious or unconscious selection effect
    2. Public establishment of impartiality; badger trial

3. Randomization

- Definition
- Justification both in design of experiments and in sampling
    1. Impersonality, with concealment, controls conscious or unconscious selection effect
    2. Public establishment of impartiality; badger trial
    3. Estimation of error in standard designs under basic assumptions
    4. Availability of exact test of significance

3. Randomization

- Definition
- Justification both in design of experiments and in sampling
    1. Impersonality, with concealment, controls conscious or unconscious selection effect
    2. Public establishment of impartiality; badger trial
    3. Estimation of error in standard designs under basic assumptions
    4. Availability of exact test of significance
- Multi-stage

3. Randomization

- Definition
- Justification both in design of experiments and in sampling
    1. Impersonality, with concealment, controls conscious or unconscious selection effect
    2. Public establishment of impartiality; badger trial
    3. Estimation of error in standard designs under basic assumptions
    4. Availability of exact test of significance
- Multi-stage

- When is randomization a bad idea?

- When is randomization a bad idea?
  1. Too complicated organizationally
  2. Study too small

- When is randomization a bad idea?
  1. Too complicated organizationally
  2. Study too small
  3. Four units, two treatments
  4. 6 sequences

- When is randomization a bad idea?
  1. Too complicated organizationally
  2. Study too small
  3. Four units, two treatments
  4. 6 sequences
  5. Only possibilities

     $A \quad A \quad B \quad B$

     $A \quad B \quad A \quad B$ and three more by interchanging A and B

     $A \quad B \quad B \quad A$

- More complex schemes of randomization
- Data dependent randomization with unequal probabilities
- Outcome dependent randomization

- Formulation of questions
- Study design
- Measurement issues
- Analysis
- Interpretation

Additional References

Blot,W.J. and 17 others (1993). Nutrition intervention trials in Linxian, China: supplementation with specific vitamin/mineral combinations. Cancer incidence, and disease-specific mortality in the general population. *J. Nat. Cancer Inst.* **85**, 1483-1491.

Ma, J.-L and 15 others (2012). Fifteen-year effects of *Helicobacter pylori* and vitamin treatments on gastric cancer incidence and mortality. *J. Nat. Cancer Inst.* **104**, 488-492.

Rahman, N.M. and 22 others (2011). Intrapleural use of tissue plasminogen activator and DNase in pleural infection. *New England J. Med.* **365**, 518-526.

Rosenbaum, P.R. (2010). *Design of observational studies.* New York: Springer.