Research Methodology

Assoc. Prof. Dr. Jamalludin Ab Rahman
B.Med.Sc., MD, MPH
Department of Community Medicine
Kulliyyah of Medicine
International Islamic University Malaysia

Kulliyyah of Medicine
International Islamic University Malaysia

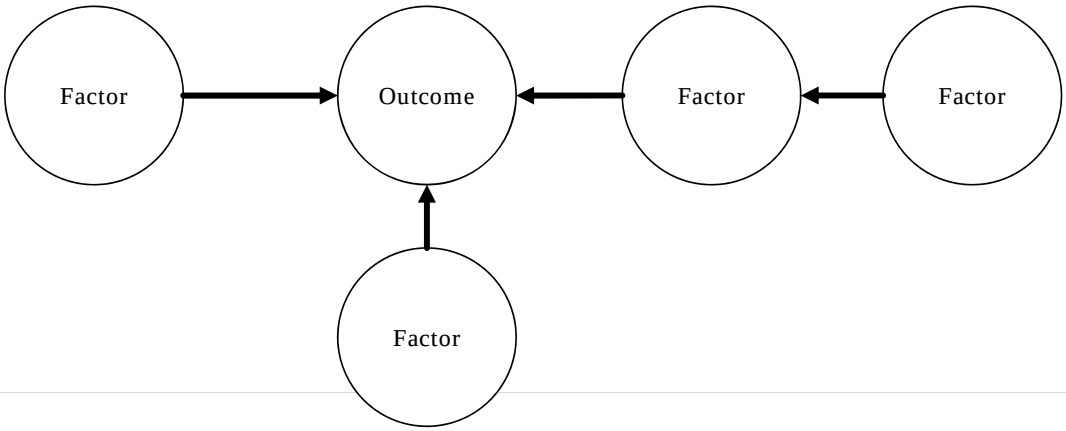
Why we do research?

- To fulfil academic requirements
- To get promotion
- To answer curiosity
- To solve problem
- To find alternative

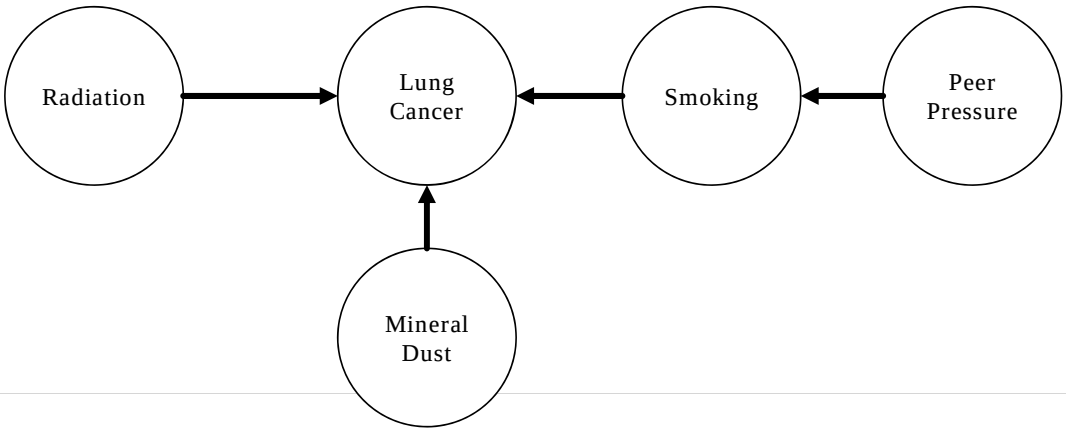
How to plan a research?

- Starts with the general broad idea
- Build the conceptual idea
- Identify main outcomes
- Identify explanatory (exposures, factors) variables
- Identify confounding variables

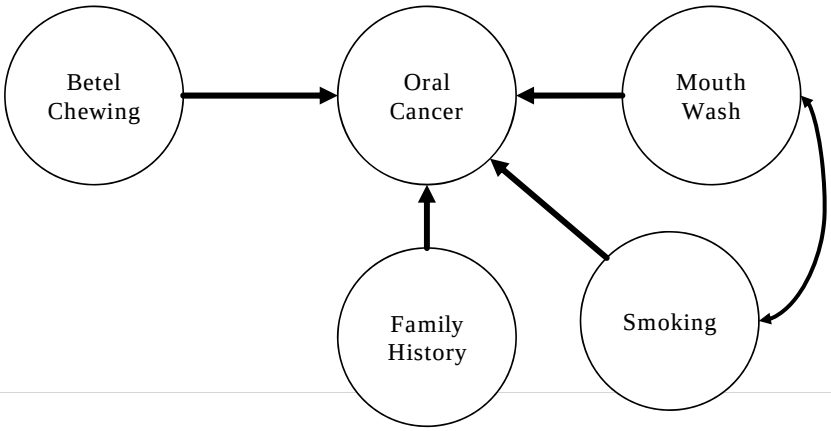
Conceptual idea

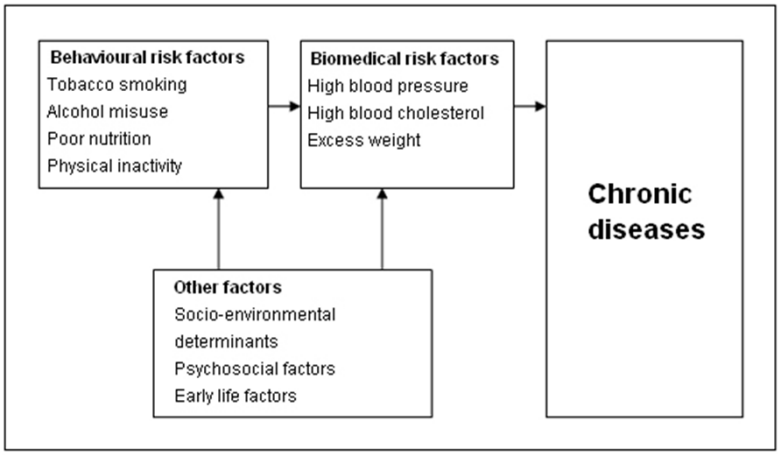


Example 1



Example 2





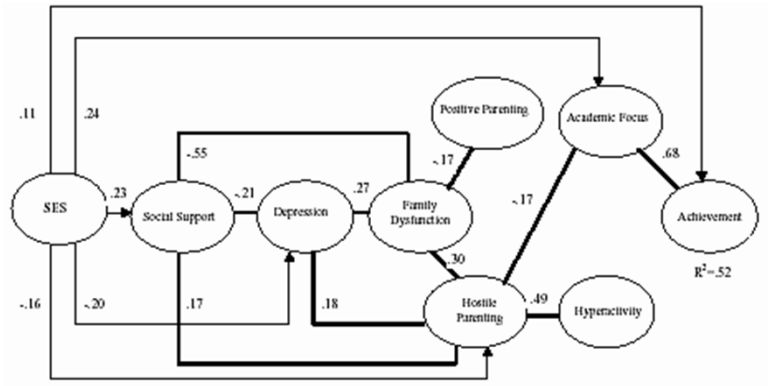


Figure 3 Structural Equation Model Showing the Relationship Between Family Processes, Child Characteristics, and Achievement for Girls Aged 6 to 11 years

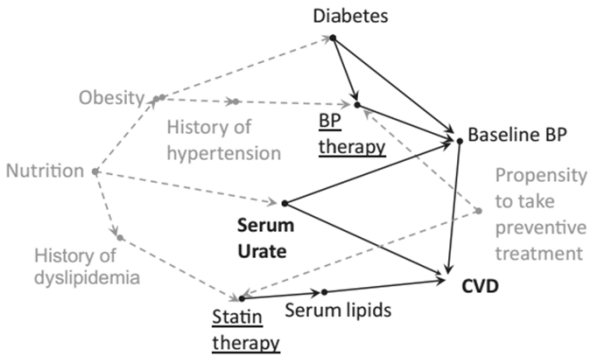


Figure 1: A directed acyclic graph (DAG), characterising the direction of selected influences* on risk of developing cardiovascular disease.
BP- Blood pressure; CVD-Cardiovascular disease.
*Dark font variables are those which are observed, whereas grey font represent unobserved. Drug treatment variables, which we consider colliders, which introduce the influence of unobserved variables not directly linked with the outcome, are underlined.

Thornley S (2012) Causation and Statistical Prediction: Perfect Strangers or Bedfellows? J Biom Biostat 3:e115. doi:10.4172/2155-6180.1000e115

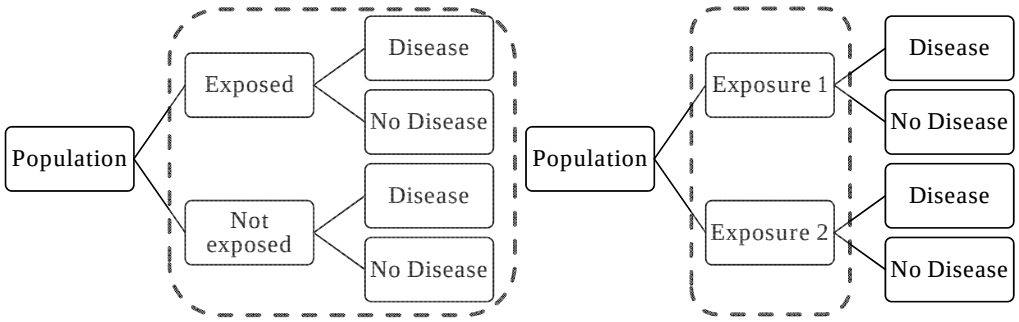
Exposure & Outcome

Smoking	Lung cancer
High carbo diet	Diabetes mellitus
Oral mouth wash	Oral cancer

Population → Exposure → Outcome

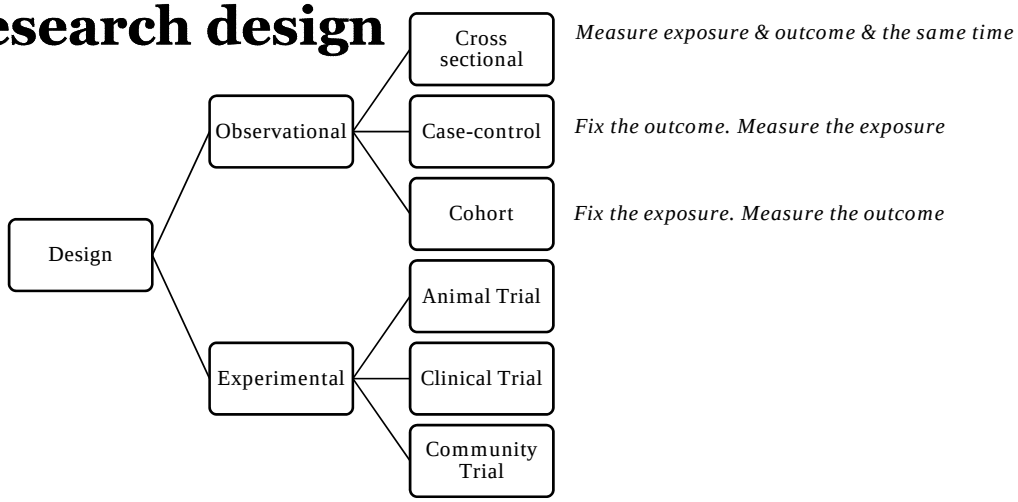
History — *To proof causation, exposure must precedes outcome* — Future

Observation vs. Experimental

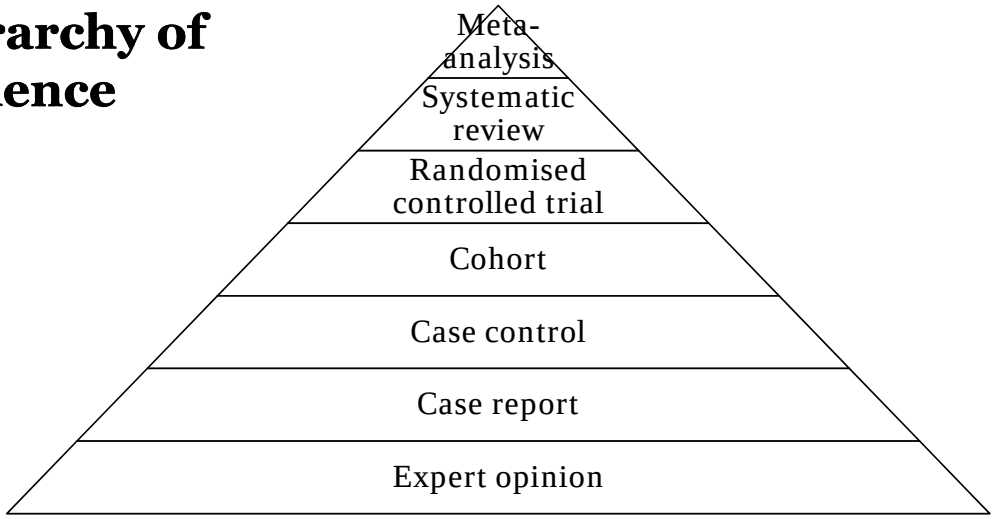


No intervention Intervention done

Research design



Hierarchy of Evidence



Sampling & sample size

- Type of sampling – *Random vs. Non random*
- Sample size – *Based on objective & research design*

Study place, duration & subjects

- Describe study place – especially if plan to represent a population
- State time & duration
- Who or what are the subjects – population, people, animal etc.

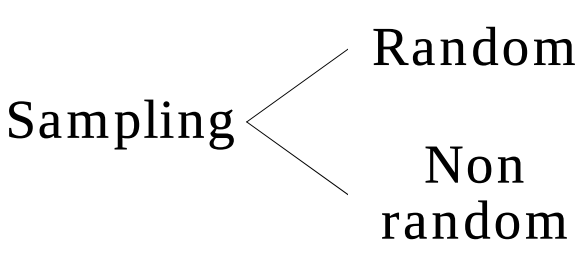
Subjects

- Target population
- Study population
- Sampling frame
- Sampling unit
- Observation unit

Example – NHMS III 2006

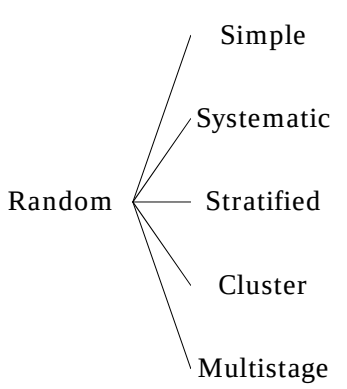
- Target population — *All Malaysian*
- Study population — *Household up to strata 6*
- Sampling frame — *List of Enumeration Block & Living Quarters*
- Sampling unit — *Enumeration Block & Living Quarters*
- Observation unit — *All household in the selected Living Quarters*

Sampling method



Random – equal chance to be selected

Random sampling



The ideal method. Randomly sample 10 students from a class of 50 students.

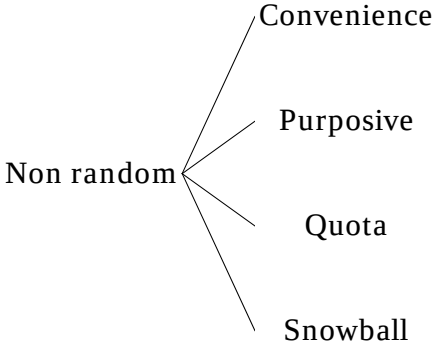
Follow certain pattern, order. Only first sample is random.

Study population divided intro strata. All strata selected. Portion of sample in each strata sampled.

Study population divided intro clusters. Assume all clusters are the same. Not all clusters selected. Only some will be sampled.

Several sampling techniques applied at different stage.

Non random sampling



Haphazard. Grab sampling. No rules. Just get the sample.

Non representative. Convenience but with certain purpose. E.g. diabetic patients on single insulin therapy.

When the sampling stop after achieving certain size.

Sample by reference or recommendation.

Example – NHMS III 2006

Two stage stratified random sampling

- Target population — *All Malaysian*
- Study population — *Household up to strata 6*
- Strata — *State & location (urban or rural)*
- Clusters — *Enumeration Block & Living Quarters*
- Sampling frame — *List of Enumeration Block & Living Quarters*
- Sampling unit — *Enumeration Block & Living Quarters*
- Observation unit — *All household in the selected Living Quarters*
- Sample distribution — *Proportionate to size*

Sample size

- How many sample required
- An estimate
- Adequate size to
  1. Represent population
  2. Test hypothesis
- Size affects duration & budget of the research
- Do not sample more than required

Sample size

- Sample size depends on
  1. Objective of the study
  2. Study design (design effect)
  3. Sampling method
  4. Expected (& precision) effect size
  5. Variability of sample
  6. Non-response rate
- Even the expected outcomes/effect size are estimates

Sample size – formula

- No single formula for all
- Depending on purpose – single proportion, compare two proportions, compare 3 proportions etc.
- Use of software or calculators e.g. PS Power and Sample Size, PASS etc.
- Anticipate non response, drop-out, loss to follow up, death (esp. animal study)

Example – Single proportion

- $N = \frac{z^2 p(1-p)}{d^2}$, where
- N is the sample size,
- z = z value for intended confidence interval,
- p is the estimated proportion (in decimal) &
- d is the precision of p (deviation from p) (in decimal)

Example – Single proportion

- Calculate sample size if you wish to do a study estimating the prevalence of DM as 20% with 5% variation at 95%CI (z for 95%CI is 1.96) anticipating 20% non response
- Answer: 246, ~ 250 then add 20% ~ 300 samples required

Data quality

- Valid value
  e.g. age > 200 years, weight > 500 kg, pregnant male etc
- No missing value
- Relevant skip response
  e.g. Not Applicable response for number of pregnancy for male respondent
- Declare method to ensure good data quality – e.g. double data entry

Statistical analysis plan

- Based on objective, especially the specific objectives – orderly manner
- Descriptive vs. analytical analysis
- The product is dummy table
- State all statistical tests planned to be used
- State significant level
- State software used
- Engage a statistician from the beginning!