METHODOLOGY IN CLINICAL RESEARCH

By

Jen-pei Liu[1,2], Ph.D. and

Wei-Chu Chie[1], M.D. and Ph.D.

[1]National Taiwan University

and

[2]Division of Biostatistics and Bioinformatics

National Health Research Institutes

I. Introduction

II. Observational Studies

III. Experimental Studies

IV. Summary

I. INTRODUCTION

- **Bias and Precision**

Goal：To make a **unbiased** inference with the possibly **best precision** to scientifically answer clinical questions with respect to a **targeted patient population**.

(1) To minimize bias.

(2) To maximize precision.

Bias

- Source: systemic error
  - selection: two definitions!
  - information
- Prevention/avoidance
  - better design (RCT)
- Evaluation and analysis
  - additional data
  - check consistency

Chance (Variability)

- Source: random error
- Prevention/avoidance
  - increase sample size/power of test
  - more accurate measurement
- Analysis and evaluation
  - p value/ confidence intervals
  - meta-analysis

Confounding

- Source: other factors associated with both exposure and outcome
- Prevention/avoidance
  - better design (RCT) (*matching is not suitable)
- Analysis and evaluation restriction
  - restriction
  - stratification
  - modeling (adjusting)

Major designs

- Experimental:
  - » exposure (treatment) is manipulated
  - » analog to laboratory work
  - » gold standard: randomized controlled trials
- Observational:
  - » no any manipulation of exposure (treatment)
  - » natural observation
- Common purpose: causal inference

美國預防醫學特別委員會
判斷標準

- Review of Evidence:
- Literature retrieval and exclusion criteria
- Evaluating the quality of the evidence
  - » grade I: RCT (randomized controlled trials)
  - » grade II-1: CT without R
  - » grade II-2: well-designed cohort or case-control studies, multi-center preferable

美國預防醫學特別委員會判斷標準

- » grade II-3: multiple time-series with or without intervention, dramatic results of uncontrolled experiments
- » grade III: opinion of respected authorities, based on clinical experiences; descriptive studies and case reports; case reports of expert committees

– cost-benefit, utility and effectiveness analysis

– meta-analysis and synthesis of research results

– updating evidence

Ideal and reality

- Ideal: experimental
  - » good for causal inference/fewer bias and confounders but not always generalizable
  - » more ethical concerns and costly
  - » therapeutic efficacy evaluation
- Reality: observational
  - » easier to implement or data ready to use
  - » fewer ethical concerns but more bias or confounders
  - » prognostic factor identification

II. OBSERVATIONAL STUDIES

- Cohort, prospective
  - variations of prospective cohort
- Case-control
- Cross-sectional
- Other related designs

The major difference

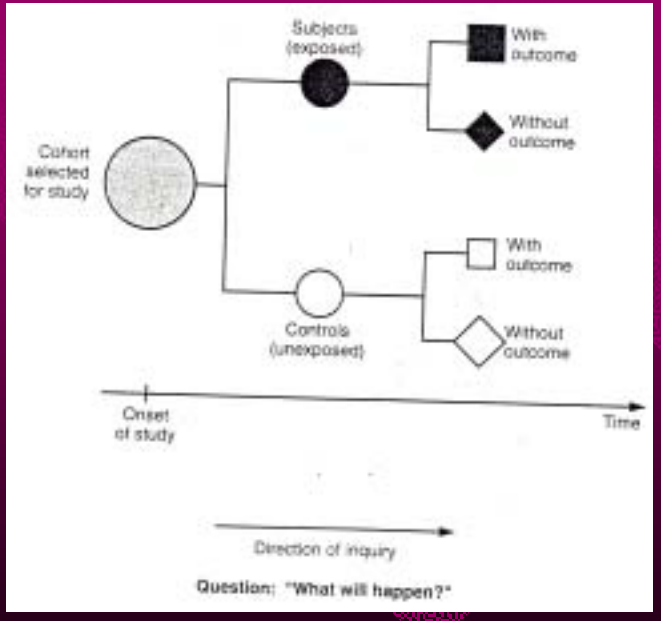
- Time/timing between measurement of exposure and outcome
- Strength in causal inference
- Efficiency of subjects recruitment

Cohort study

- Original definition of a cohort 羅馬軍團
- Prospective cohort study:
  - the most classical design and attractive nature of epidemiology
  - causal inference without experiment
  - the best in observational studies
- Variants:
  - retrospective cohort, ...

Prospective cohort study in outcome research

- Assemble the cohort
  - inception cohort: onset of disease/zero time
- Measure predictor variables (prognostic/predictive)
- Follow-up and measure outcomes
  - time to event (incidence): change of status
  - surrogate, qol, …: change of value

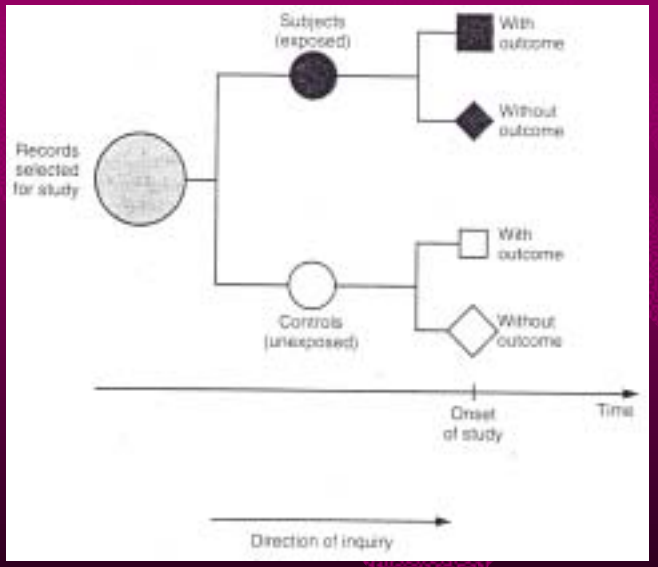


Strengths and weakness in outcome research

- Strengths:
  - proper time sequence: predictors (exposures measured before outcomes)
  - fewer bias: information and selection
  - time-dependent variables available if measured
  - binary: rates obtainable/ non-binary: value/change
- Weakness:
  - inefficient for rare outcomes
  - expensive, time consuming in maintenance/follow-up
  - confounders unavoidable

Variant : retrospective cohort study

- Identify a suitable cohort
- Collect data about predictor variables
- Collect data about outcomes at a later time
  - » basically also a cohort or follow-up study
  - » only difference: time of measurement
  - » common in clinical studies/data linkage
  - » not necessarily collecting outcomes “later” but at a later time than the occurrence of the exposure

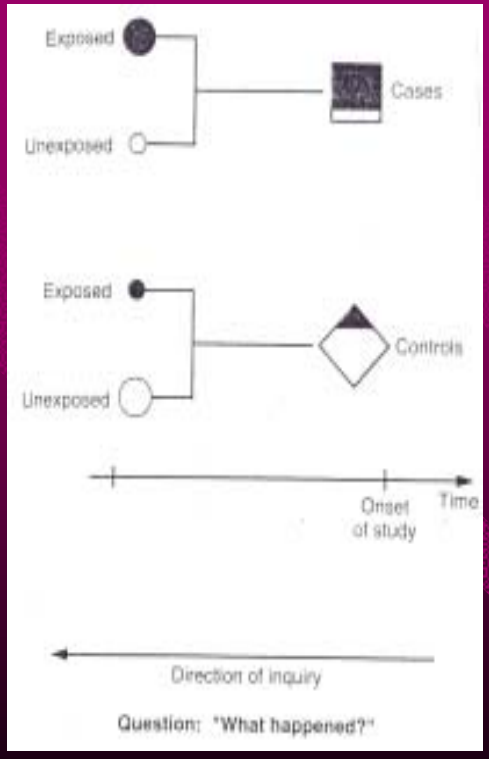


Strengths and weakness in outcome research

- Strengths:
  - same as prospective cohort
  - less costly and time consuming
- Weakness:
  - same as prospective cohort except for cost & time
  - no QA/QC for data collection
  - may not include information needed

Case-control (reference) study

- An important breakthrough of epidemiologic study
- classical definition
- new perspective
  - control as a sample of hypothetical population from which cases came from
  - can be seen as a variant of cohort study



Case-control study in outcome research

- Draw a sample of new (incident) cases (outcome +)
- Draw a sample of controls (outcome - at a certain time)
  - » a sample of hypothetical population from which cases came from
- Measure the predictor variables
  - » usually at the time when cases and controls are drawn

Strengths and weakness in outcome research

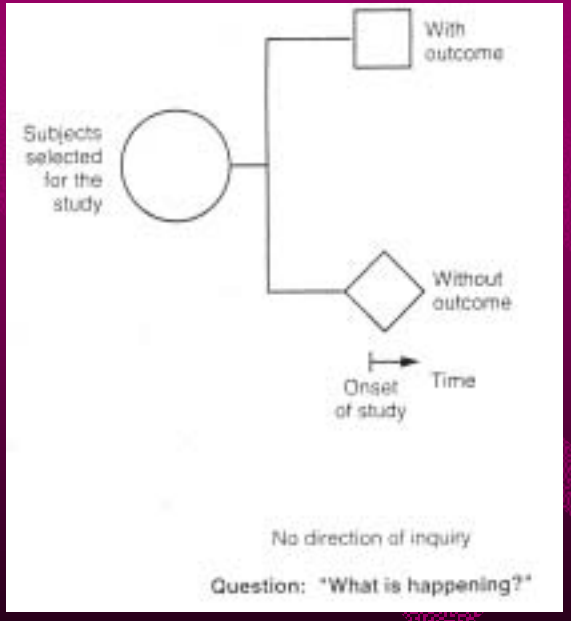
- Strengths
  - efficient for rare outcomes: time and cost
- Weakness
  - not always proper time sequence
  - bias: selection and information
  - confounding
  - non-binary outcomes not obtainable
  - binary outcomes: only odds ratio obtainable

Cross-sectional study

- The most easy type
- usually by surveys
- current status/prevalence and prevalence ratios only
- poor in causal inference

Cross-sectional study in outcome research

- Select a sample from population
- measure the predictor variables and the outcomes at the same time
  - case/non-case (not controls)
  - exposure/non-exposure



Strengths and weakness in outcome research

- Strengths
  - time saving
  - get prevalence/status data both binary and non-binary
- Weakness
  - no proper time sequence: poor in causal inference
  - inefficient in rare outcomes
  - bias: selection and information/confounding
  - binary outcomes: only prevalence ratio obtainable, no incidence or change of status

Serial surveys or panels

- Follow-up a single population
  - Serial surveys: like multiple cross-sectional studies
  - Panel: like cohort studies
- multiple measurements

III. EXPERIMENTAL STUDIES

- **Clinical Trials**
  - FDA (21 CFR 312.3, April 1994)
    A clinical trial is the clinical investigation of a drug which is administrated or dispensed or used involving one or more human subjects.
  - Chow and Liu (July 1998)
    A clinical trial is a clinical investigation in which treatments are administrated, dispensed or used involving one or more human subjects for evaluation of the treatments

Techniques to Avoid Bias

- Use of controls
- Blinding.
- Randomization.

Types of Controls

- **Concurrent Controls**
  - To provide internal validity
  - – Placebo concurrent control
    - » The standard concurrent control
  - – Active treatment concurrent control
    - » Ethical reasons
    - » Equivalence trials
  - – Dose-comparison control
  - – No treatment control
    - » Should avoid

Types of Controls

- **Historical Control**
  - Results of the controlled group were not obtained concurrently within the same trial.
  - Not recommended unless the drug is self-evident such as general anesthetics

Blinding (Masking)

- To limit the occurrence of conscious and unconscious bias in the conduct and interpretation of a clinical trial arising from the knowledge of treatment.

Goal ：to prevent identification of the treatments until all opportunities for bias have passed.

Levels of Blinding

- **Open-labeled**
- **Single-blind**
- **Double-blind**
  - Gold standard for most of clinical trials
- **Triple-blind**
  - Gold standard for the trials sponsored by the pharmaceutical industry.

***Principle*：** blindness thorough out the entire course of the study.

Randomization

Goals：

- To introduce a deliberate element of chance into assignment of treatments to patients
- To avoid bias in selection and allocation of subjects from the predictability of treatment assignments
- To minimize the differences in relevant characteristics of the treatment groups and to produce similar distributions of prognostic factors between groups
- To provide a sound statistical basis for the quantitative evaluation of the evidence relating to treatment effects

Randomization

Methods：

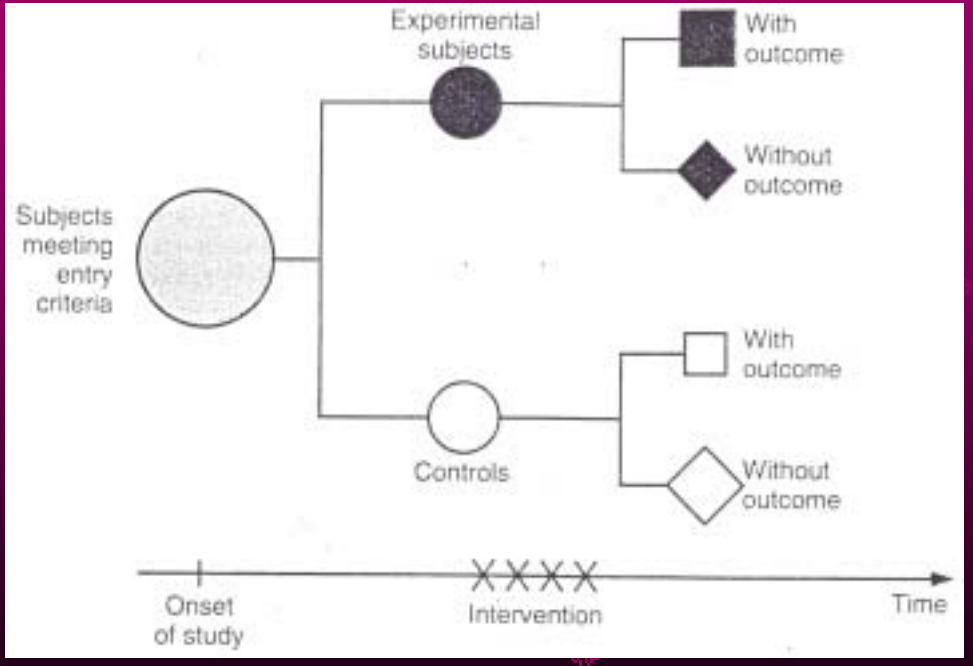
- Unrestricted randomization
- Permuted-block randomization
- Stratification
  - By important prognostic factors: center, gender, age, baseline characteristics
  - Separate randomization within strata
  - The number of stratified factors <= 2
  - Separate random scheme for each center

COMMON DESIGNS

- **Parallel Group Designs**

  The patients are randomized to one of two or more arms, each arm being allocated to a different treatment.

  - Advantages：
    - » Simple and easy to implement.
    - » Less complicated analysis and interpretation.
  - Drawbacks：
    - » Relative large variability
    - » Inter-patient + Intra-patient



- ***Example***

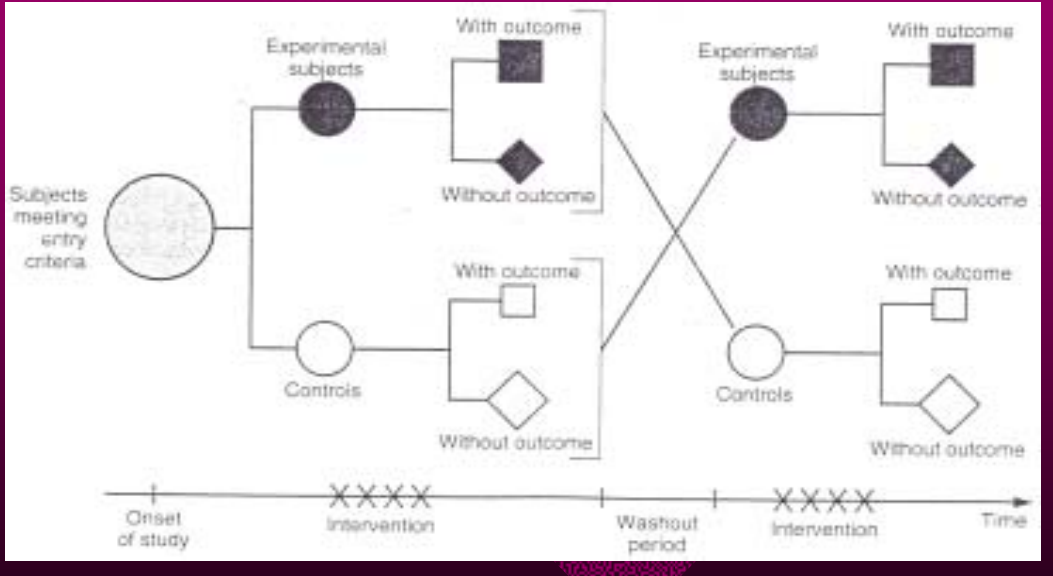
  NINDS rt-PA Stroke Study (NEJM 1995;333:1581-7) rt-PA in treatment of acute ischemic stroke

  - Patient population：624 patients with acute ischemic stroke.
  – Treatment: rt-PA and placebo
  – Design: randomized double-blind placebo-controlled
  – Stratified randomization: time to onset of stroke to initiation of treatment: 0-90 and 91-180 min; early and late improvement
  – Primary endpoint: a 4-point improvement from baseline in NIHSS

Crossover Design

Each subject is randomized to a sequence of two or more treatments.

- Advantages
  - » Subjects act their own control for treatment comparison
  - » Reduction of variability
  - » Fewer patients required



Crossover Design

- Drawbacks
  - » More difficult to implement (more dropouts)
  - » Treatment effect should fully develop within the treatment period
  - » For stable and chronic diseases only
  - » Biased inference due to carryover effects
  - » More complicated analysis and interpretation, e.g, adverse events in later treatment periods

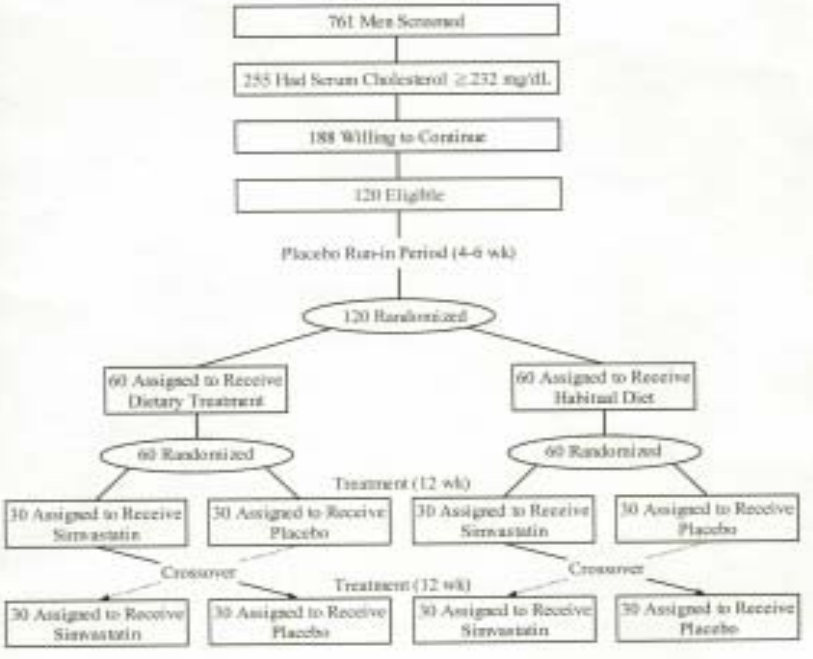


Figure 5.4.3 Joint Application of Parallel-group and Crossover Designs.

Source: Jula, et al. (2002)

Factorial Design (Combination Trials)

- Two or more treatments are evaluated simultaneously in the same sets of patients via various of combinations of two treatments.

- ***Example***：The Medical Therapy of Prostatic Symptom (MTOPS) Research Group (NEJM 2003;349:2387-98)

- ***Example：***
  - Long-term effect of doxazosin and finasteride on the clinical progression of BPH
  - Double-blind, randomized, parallel group
  - 3047 patients with a mean follow-up of 4.5 years

Treatment	Doxazosin	Finasteride
I	Placebo	Placebo
II	Placebo	5 mg/d
III	8 mg/d	Placebo
IV	8 mg/d	5 mg/d

Factorial Design (Combination Trials)

- Advantages
  - can efficiently use patients for evaluation of efficacy of both treatments if no interaction
  - can investigate the joint treatment effects
  - can establish dose-response relationship of simultaneous use of two treatments
- Drawbacks
  - difficult to implement because of large number of treatment groups
  - lack of power for interaction

Multicenter Trials

- A multicenter study is a single study conducted under a common protocol, involving several centers (e.g., clinics, practices, hospitals) where the data collected are intended to be analyzed as whole (as opposed to a post-hoc decision to combine data or results from separate studies)

Multicenter Trials

- ***Goals*：**
  - To accrue patients efficiently (all stages)
  - To provide a basis for generalization of its findings (later phases)
  - Generalizability：
    The extent to which the findings of a clinical trial can be reliable extrapolated from the subjects who participated in the trial to a broader patient population and a broader range of clinical settings.

- ***Examples*** **：**
  Tacrine in Alzheimer's disease
  (Farlow, et al. JAMA 1992;268:2523-2529)
  A multinational and multicenter trial
  - Targeted population：
    - 468 randomized patients
    - 50-89 years old
    - Criteria by NINCDS-ADRDA
    - Min-Mental State Examination
    - (MMSE) score: 1-26
  - No. Centers: 23 in 2 countries

- ***Issues :***
  - Variations in implementing protocol
    - » Common protocol
    - » Standardization of procedures
    - » Pre-study investigator's meeting
    - » Training of personnel
    - » Careful monitoring
  - Variation in the number of patients
    - » Few small centers vs. lots of large centers
    - » Few large centers vs. lots of small centers
    - » All small centers

Superiority Trials

The objective of the trial is to establish the efficacy by demonstrating that the test treatment is superior to

- a concurrent placebo control
- a concurrent active treatment control

  or

- a dose-response relationship

Equivalence or Non-inferiority Trials

The objective of the trial is to show that the efficacy of the test treatment is either

- similar (or equivalent) to or
- no worse than the concurrent active treatment control.

- ***Equivalence Trial***
  A trial with the primary objective of showing that the response to two or more treatments differs by an amount which is clinically unimportant

- ***Non-inferiority Trial***
  A trial with the primary objective of showing that the response to the investigational product is not clinically inferior to a comparative agent (active or placebo control)

- ***Example：***

  COBALT Investigators and Ware and Antman (NEJM 1997;337:1124-30 NEJM 1997;337:1159-61)

  – Objective:

    To show non-inferiority of double-bolus of alteplase (a bolus of 50 mg over 1-3 minutes followed 30 minutes later by a second bolus of 50 mg) to accelerated infusion of 100 mg of alteplase in 30-day mortality of patients with acute M.I.

- ***Joint Applications of Superiority and Equivalence Trials***
  - Moseley, et al. NEJM, 2002: 347:81-8
  - Patient: Osteoarthritis of the knee
  - **Design**: Randomized, parallel-group, Placebo-controlled, evaluator-blind
  - Treatments:
  - Arthroscopic debridement (n=59)
  - Arthroscopic lavage (n=61)
  - Placebo surgery (n=60)

- ***Joint Applications of Superiority and Equivalence Trials***
  - Objectives:
  - (1) Superiority of the arthroscopic procedures over placebo surgery
    - **Results: NO**
  - (2) If lack of evidence of superiority, equivalence of arthroscopic prcocedures to placebo surgery
    - **Results: YES**

Group Sequential Trials

A group sequential trial allows to evaluate the efficacy and safety of test treatment by means of interim analyses during the study for possible early termination based on convincing evidence of either benefit or harm before its scheduled completion. Example: WHI study.

- Description of statistical methods and pre-planned interim analyses in the protocol with adjustment of p-values
- Documentation of everything
- Independent data monitoring committee

Genomic Information in Designs

- Imatinib mesylate (Gleevec) for Chronic Myelogenous Leukemia (CML) and gastrointesitnal stromal tumors (GIST)
- Philadelphia (Ph+) chromosome from reciprocal translocation of long arms of 9 and 22 in 90% of patients with CML
- Formation of BCR-ABL fusion gene ⇒BCR-ABL tyrosine kinase ⇒ CML
- KIT proto-oncogene ⇒ transmembrane receptor KIT ⇒ GIGS
- Kantarjian et al. 2002; Demetri, et al., 2002

Genomic Information in Designs

- Other examples
- HER2 gene in metastatic breast cancer - Herceptin - requirement of screening the patients with over-expressed HER2 level (Slamon, 2001).
- Estrogen receptor ploymorphism - Estrogen Replacement Atherosclerosis trial (ERA, Herrington, et al, 2002): a total of 9 SNPs were identified and interaction between treatment of HRT and some of SNPs in elevation of lipid levels is suggested
- Sample size determination: Fijal, et al. (2000)

IV Summary

- **Basic Design Considerations**
  - Bias and Variability
- **Ethical Considerations**
  - Protection of human subjects
    - Prevention of physical injuries
    - Privacy of personal data
    - Sound methodology
    - Unethical for under-powered studies
- **No Free Lunch**