New Investigator Clinical Trial
Section 2: Statistics for Clinical Trials
Part 1: Basics of Phase III Trial Design

Bingshu Chen, Ph. D.

Outline

- Phase III Randomized Clinical Trial
  - Design
  - Randomization
  - Blindness
- Endpoints
- Sample Size
- Non-inferiority Trials
- Summary

Example of statistical section in a Phase III protocol

PROTOCOL DATE: 2010-APR-09
NCIC CTG TRIAL: MA.32

14.0 STATISTICAL CONSIDERATIONS

14.1 Objectives and Design

The primary objective of this study is to compare invasive disease-free survival (IDFS) in pre- and post-menopausal women with T1-T3, N+/-, ER/PgR+/-, HER2 +/- invasive breast cancer. Eligible subjects will be randomized to one of the following two treatment groups: metformin (850 mg po bid for 5 years, experimental arm) or placebo (one caplet po bid for 5 years, control arm). Subjects will be stratified by: 1) ER and PgR status (both negative vs. either ER or PgR positive), 2) Body Mass Index (≤ 30 kg/m^2 versus > 30 kg/m^2), 3) HER2 (Positive = 3+ over-expression by IHC in > 30% of invasive tumour cells *OR* HER2 gene amplification by FISH/CISH> 6 HER2 gene copies per nucleus, *OR* a FISH/CISH ratio: HER2 gene copies to chromosome 17 signals of ≥ 2.2. All other results will be considered negative.) and 4) Chemotherapy administration (any versus none). We will also compare overall survival, distant relapse free survival, breast cancer free survival, health-related quality of life, body mass index between metformin and placebo arm. The intent-to-treat (ITT) population will comprise all randomized patients, will be based on the allocated treatment regardless of whether the patient received the assigned treatment, and will be based on the at-randomization values of the stratification factors. The adverse events in the two different treatment groups will also be compared. The embedded correlative science studies are described in Section 13 of this protocol.

A minimization procedure *[White,1978]* will be used to allocate patients with equal probabilities to one of the two treatment groups.

Example of statistical section in a Phase III protocol

14.2 Primary Endpoints and Analysis

The primary endpoint of this study is invasive disease-free survival. It is defined as the time from randomization to the time of documented ipsilateral and contralateral invasive breast tumour, local/regional invasive recurrence, distant recurrence, death from breast cancer, death from non-breast cause, death from unknown cause, second primary invasive cancer (non-breast, except for adequately treated BCC or SCC of the skin). If a subject has not had invasive disease or died at the time of data cut-off for final analysis, IDFS will be censored on the date of last follow-up. The survival experiences of subjects in both treatment arms will be described by the Kaplan-Meier method. Stratified two-sided log-rank tests adjusting for stratification factors as defined in the protocol will be the primary method to compare IDFS between metformin and placebo arm. As an exploratory analysis, a Cox proportional hazards model will be use to identify and adjust for factors significantly related to invasive disease-free survival.

Example of statistical section in a Phase III protocol

14.4 Safety Monitoring

Adverse events will be monitored on an ongoing basis by the central office and their frequencies reported annually at investigators' meetings.

14.5 Interim Analysis

Chris will cover details about interim analysis

Two interim analyses are planned for this study when 144 and 288 events are observed, to allow early termination of the study if the results are extreme. Lan-DeMets error spending function will be used to assess for superiority, and futility for superiority *[Lan 1983]*. The early stopping boundaries are based on a power family with power 3, which approximates the O'Brien-Flemming boundaries *[Jennison 2000]*. The actual p-values for superiority and futility will be calculated based on the number of events observed at the time of interim analysis, controlling the two-sided Type I error of 0.05 and the power of 80% at the end of the study.

14.6 Quality Of Life Analysis

QoL will be covered tomorrow.

The EORTC QLQ-30 Global Score will be used for our primary assessment of quality of life but subscales and specific symptoms (diarrhea, bloating, flatulence, dyspepsia, abdominal cramps, nausea and vomiting, taste alteration, limitation of activities because of gastrointestinal symptoms, joint/musculoskeletal symptoms) will be used for our secondary hypothesis.

Randomized Clinical Trial

- Objectives
  - Study efficacy of an intervention in a given ***study population***
  - Need an intervention group and a control group
  - Method of selection: Randomization
- Randomized Clinical Trials
  - Cross-over design
  - Factorial design
  - Large simple clinical trial
  - Superiority trial
  - Non-inferiority trial (Study of equivalency)

Trial Design – II

- Cross-over design
  - Each participant to serve as his or her own control
    - Receives either intervention or control in the first period and the alternative in the succeeding period
    - The order is randomized
  - Assumption: no carry over effect, which is inappropriate in many oncology clinical trials
    - e.g. cured patients may not return to the initial state

Trial Design – III

- Factorial design
  - Evaluate two or more interventions compared with control in a single trial, e.g. 2x2 design:

	Intervention A	Control
Intervention B	A and B	B only
Control	A only	Control

  - It is possible to leave some of the cells empty
  - Sample size depends on the *interaction*
  - Impact on recruitment and compliance
- Example: NCIC CTG: MA.27 (Original Design)

Trial Design – III

- Example: NCIC CTG: MA.27
  - Original Design: 2 by 2 factorial

	Exemestane	Anastrozole
Celecoxib	C & E	C & A
Placebo	P & E	P & A

  - Amendment (2005)
    - Remove Celecoxib arm

PROTOCOL DATE: 2003-MAY-26
NCIC CTG TRIAL: MA.27

AMENDMENT: 2003-NOV-06
Amendment #2 (Canada) / Amendment #1 (USA)

NATIONAL CANCER INSTITUTE OF CANADA
CLINICAL TRIALS GROUP (NCIC CTG)

A RANDOMIZED PHASE III TRIAL OF EXEMESTANE VERSUS ANASTROZOLE WITH OR WITHOUT CELECOXIB IN POSTMENOPAUSAL WOMEN WITH RECEPTOR POSITIVE PRIMARY BREAST CANCER

NCIC CTG Protocol Number MA.27
CTSU Protocol Number: MA.27

MA.27: 2003 Design

RANDOMIZE	ARM 1 Exemestane, 25 mg/day X 5 years + Celecoxib 200 mg, 2 capsules twice daily X 3 years
	ARM 2 Exemestane, 25 mg/day X 5 years + Placebo, two capsules twice daily X 3 years
	ARM 3 Anastrozole, 1mg/day X 5 years + Celecoxib, 200 mg, two capsules twice daily X 3 years
	ARM 4 Anastrozole, 1 mg/day X 5 years + Placebo, two capsules twice daily X 3 years

MA.27: 2005 Amendment

RANDOMIZE	~~ARM 1~~ ~~Exemestane, 25 mg/day X 5 years + Celecoxib 200 mg, 2 capsules twice daily X 3 years~~
	ARM 2 Exemestane, 25 mg/day X 5 years + Placebo, two capsules twice daily X 3 years
	~~ARM 3~~ ~~Anastrozole, 1mg/day X 5 years + Celecoxib, 200 mg, two capsules twice daily X 3 years~~
	ARM 4 Anastrozole, 1 mg/day X 5 years + Placebo, two capsules twice daily X 3 years

Trial Design – IV

- Large simple design
  - Uncover modest benefits of intervention
    - Short term
    - Easily to implement in a large population
  - Unbiased allocation of participants
  - Unbiased assessment of the outcomes
  - No for trials with
    - Complex interventions
    - Complex Outcomes

Randomization – Why?

- Randomized Clinical Trials (RCT)
  - Assign participants to control or treatment group using formal randomization procedure
- Advantages of RCT
  - Ensure ***balance*** for all baseline factors
  - Remove potential ***bias***
  - Produce ***comparable*** groups
  - Guarantee ***validity*** of statistical tests

Randomization – How?

- **Simple Randomization**
  - Toss a coin or use a computer-based algorithm
  - Not guarantee for balance of important factors
  - e.g. randomly assign 12 subjects to arm A and arm B: A A B A B B A A A A B A
- **Block Randomization**
  - Balance between *treatment* groups
  - Not guarantee for balance of *important factors*
  - E.g. Randomly assign treatment to 12 patients:
    - List all possible block size of four: AABB, ABAB, ABBA, BBAA BABA BAAB
    - Randomly select 3 of the 6 blocks: BBAA|ABAB|BBAA

Randomization – How?

- **Stratified Block Randomization**
  - Identify important stratification factors
    - eg. Age, Gender, Centre, etc
  - Ensure treatments are balanced for a few pre-selected stratification factors
  - Randomly assign treatment group (Block Randomization) within each combination (cell) of stratification factors
  - *Risk of not balanced if the number of cell is large*

Randomization – How?

- Example: stratify with age and gender,
  - Cell 1 (4 pts): Age > 60, Male: ABAB BA~~BA~~
  - Cell 2 (6 pts): Age > 60, Female: AABB AB~~AB~~
  - Cell 3 (6 pts): Age ≤ 60, Male: ABBA
  - Cell 4 (4 pts**):** Age ≤ 60, Female: BAAB

	Arm A	Arm B	Total
Age			
≤60	4 (50%)	4 (50%)	8
>60	6 (50%)	6 (50%)	12
Gender			
Female	5 (50%)	5 (50%)	10
Male	5 (50%)	5 (50%)	10

Randomization – How?

- **Dynamic Randomization / Minimization**
  - Large number of cells
    - Age (3 levels), Gender (2 levels), smoking history (3 levels), centre (5 levels), node status (3 levels): 270 cells
  - For a new subject in a give cell, total number of patients allocated in each treatment group is counted. The subject will be allocated to group with smallest number (coin tossing if tied)
  - Guarantee balance but treatment allocation may be predictable
  - ***Default procedure used in NCIC CTG for randomized phase III studies.***

Blindness

- Objective
  - Reduce ***bias***
- Type of Trials
  - Un-blinded trials (open trials)
    - Both the participant and the investigator are aware of the intervention assignment. e. g. lifestyle intervention
  - Single-blind trials
    - Only the investigators are aware of the intervention assignment
  - Double-blind trials

Blindness - II

- Double-blind trials
  - Neither the participants nor the investigators know the intervention assignment.
  - Usually restrict to trials of drug efficacy
  - ***Bias*** is reduced (but can't be completely eliminated)
  - An outside body to monitor the data for toxicity and benefit (e.g. DSMC)

Blindness - III

- Special problems in double-blind trials
  - Participants and investigators may try to unblind the medication
    - Consciously
    - Unconsciously
  - Matching of drugs: Tablets or capsules closely resembled one another
  - Coding of drugs: Labeling of individual drug bottles or vials so that the identity of the drug is not disclosed

Blindness - IV

- Ideally, a clinical trial should have double-blind design to avoid potential bias

- If a double-blind design is impossible, use a single-blind approach or other measures to reduce potential bias

Example: Double Blind Trial

NCIC CLINICAL TRIALS GROUP (NCIC CTG)

A PHASE III DOUBLE-BLIND STUDY OF DEXAMETHASONE VERSUS PLACEBO IN THE PROPHYLAXIS OF RADIATION-INDUCED PAIN FLARE FOLLOWING PALLIATIVE RADIOTHERAPY FOR BONE METASTASES

NCIC CTG Protocol Number: **SC.23**

Example: Un-blinded Trial

NCIC CLINICAL TRIALS GROUP (NCIC CTG)

A PHASE III STUDY OF THE IMPACT OF A PHYSICAL ACTIVITY PROGRAM ON DISEASE-FREE SURVIVAL IN PATIENTS WITH HIGH RISK STAGE II OR STAGE III COLON CANCER: A RANDOMIZED CONTROLLED TRIAL (CHALLENGE)

NCIC CTG Protocol Number: CO.21

Endpoints

- Primary Endpoints
  - Most clinically relevant and direct related to primary objective of the trial
  - Base for sample size calculation
  - Analysis to be adjusted for Type I error if there are multiple primary endpoints
- Secondary Endpoints
  - Supportive measurements of effects related to the secondary objective
  - Hypothesis generation
  - No need to adjust trial results for secondary endpoints

Endpoints - II

- Examples of Endpoints
  - Time to event endpoints
    - Overall survival
    - Event free survival
    - Progression free survival
    - Recurrent free survival
  - Binary endpoints
    - Overall response rate
    - Complete response rate
  - Continuous endpoints
    - Quality of life scores
    - Incremental cost-effectiveness ratios

Example: Time to Event Endpoint
NCIC CTG HN.6

- Progression free survival (PFS), the primary endpoint of this study, is defined as the time from randomization to the time when a failure defined in 10.3 is observed. If a patient has not developed a failure at the time of final analysis, PFS will be censored on the date of the last disease assessment.

Example: Binary Endpoint
NCIC CTG SC.23

- The primary endpoint is the per-patient incidence of radiation-induced pain flare that occurs from the time of radiotherapy to ten days after the completion of radiation treatment. Pain flare is defined in section 10.2.1.

Sample Size

- Objectives
  - Provide an estimate of the needed size of a study
  - Ensure sufficient statistical power to detect clinical meaningful difference between groups
  - Provide adequate levels of significance
  - Parameters for sample size estimation shall be as conservative as possible while still being realistic

Sample Size - II

- Parameters for Sample Size Calculations
  - Specify Type I error (or Significant level)
  - Specify Type II error (or Power)
  - Determine the minimum difference to be detected or of clinical interest (defined by δ)
- Sample size calculations for
  - Continuous endpoint
  - Binary endpoint
  - Time to event endpoint
- Details will be covered in this afternoon's Workshop.

Non-inferiority Trials

- Objectives
  - Study of equivalency
  - Test whether a new intervention is as good as an established one
  - Trials with positive control
- Requirements
  - Control or standard treatment must have been shown to be effective (i.e. better than placebo)
  - Similar populations, concomitant therapy and dosage
  - Trials that demonstrated efficacy of the standard shall be recent and properly designed and conducted

Non-inferiority Trials - II

- Other important factors to be considered
  - Frequency and severity of adverse effects
  - Changes in Quality of Life (QoL)
  - Ease of applying the new intervention
  - Cost of the new intervention

Non-inferiority Trials - III

- What is meant by equivalence?
  - Two therapies are identical? – Require infinity sample size to test $\delta = 0$
  - New intervention falls sufficiently close to the standard as defined by reasonable boundaries
  - ***Non-inferiority margin***
    - Specify some value, δ, such that interventions with differences that are less than this might be considered equally effective or equivalent

Non-inferiority Trials - IV

- Example
  - A recent trial shows that drug A has response rate of 60%, compared with 30% of the placebo
  - Drug B is less expensive and has fewer side effects
  - Drugs A and B are considered to be equivalent if the difference in response rate is less than δ=10%
    - Null hypothesis H_0: $P_A - P_B > \delta$ vs H_a: $P_A - P_B < \delta$
  - Calculate sample size such that one can reject H_0 with power 80% if the upper 5% confidence interval for the difference of the response rate does not exceed δ (e.g. $\alpha = 0.05$, power = 80%: N = 594)

Summary

- Define the term randomized clinical trials
- Randomization methods used in clinical trials
- The importance of blindness in clinical trials
- Different endpoints used in clinical trials
- Necessary parameters for sample size calculation
- Difference between superiority and non-inferiority trials