Design and Conduct of Randomised Controlled Trials
(Part I)

STATS 773: DESIGN AND ANALYSIS OF CLINICAL TRIALS

Dr Yannan Jiang
Department of Statistics

May 7th 2012, Monday, 09:00-10:30
Randomised Controlled Trial (RCT)

A prospective study to compare one or more intervention techniques against a control group, with the assignment of the participant to a group determined by the formal procedure of randomisation

- Groups must be sufficiently similar at baseline so that differences in outcome may reasonably be attributed to the action of the intervention compared to the control, e.g.
  - Current standard therapy
  - Placebo (non-active intervention)
  - No treatment at all

- Randomisation removes the potential of allocation bias and tends to produce comparable groups
  - Not always feasible
  - Ethical aspects
Standard Design of RCT

Individual subject is randomly allocated to receive either the intervention or the control

- Participants (2N)
  - Randomisation (1:1)
  - Intervention (N)
  - Control (N)
  - Follow up
Other Types of Trial Design

• Cluster Randomised Trial
  – A group of subjects (“cluster”) are randomly allocated to receive either the intervention or the control
  – Example of clusters: School, GP clinic, Hospital, District
  – Logistic constraints with reduced efficiency

• Factorial Design
  – Evaluate two (or more) interventions vs. the control in a single experiment
  – Examples: chemotherapy and radiation therapy in cancel patients
  – Appeal of getting more experiments done at once, given the cost and effort in trial recruitment and conduct

• Cross-over Design
  – Each subject receives consecutively two or more treatments during the course of the study, i.e. serve as his/her own control
  – Smaller sample size with a prolonged trial period
  – Fairly strict assumption on carrying-over effect
Cluster Randomised Trial

A group of individuals is randomly allocated to receive either the intervention or the control.
# 2x2 Factorial Design

Individual subject is randomly allocated to receive one of the four combinations of two interventions (A and B, A only, B only, neither)

<table>
<thead>
<tr>
<th>Intervention B</th>
<th>Intervention A</th>
<th></th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>$n_{AB}$</td>
<td>$n_{BO}$</td>
<td>$N_B$</td>
</tr>
<tr>
<td>Yes</td>
<td>No</td>
<td>$n_{AO}$</td>
<td>$n_{CT}$</td>
<td>$N_B$</td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
<td>$n_{AB}$</td>
<td>$n_{BO}$</td>
<td>$N_B$</td>
</tr>
<tr>
<td>No</td>
<td>No</td>
<td>$n_{AO}$</td>
<td>$n_{CT}$</td>
<td>$N_B$</td>
</tr>
<tr>
<td>Total</td>
<td>$N_A$</td>
<td>$N_A$</td>
<td></td>
<td>$N$</td>
</tr>
</tbody>
</table>
2x2 Cross-Over Design

Individual patient is randomly allocated to receive either Intervention/Control [Group A] or Control/Intervention [Group B].
Randomisation

• A process by which each participant has the same chance of being assigned to either intervention or control
  – Tends to produce comparable groups with respect to known and unknown risk factors
  – Removes potential allocation bias by investigators
  – Produce valid statistical tests and significance levels

Fundamentals of Clinical Trials (3rd Ed), Friedman et al.
Randomisation

• Unit of Randomisation
  – Individuals
  – Clusters

• Method of Randomisation
  – Simple randomisation
  – Constrained randomisation
  – Dynamic randomisation

• Stratification
  – Important baseline prognostic factors to be considered in randomisation, e.g. study center, gender, ethnicity
  – Ensure allocation balance within each stratum
Simple randomisation

• This form of randomisation makes each new treatment assignment without regard to those already made
  – E.g. flipping a coin or drawing straws with 50:50 chance

• This approach, however, can produce imbalance in treatment allocations with small sample size

```r
> summary(factor(rbinom(10,1,0.5)))
  0 1
  7 3
> summary(factor(rbinom(10,1,0.5)))
  0 1
  6 4
> summary(factor(rbinom(10,1,0.5)))
  0 1
  5 5
> summary(factor(rbinom(10,1,0.5)))
  0 1
  2 8
```

• One way to prevent such times of imbalance is to use a constrained randomisation scheme
Constrained Randomisation

A general form of constrained randomisation is the *Blocked Randomisation*

- A block contains a pre-specified number and proportion of treatment assignments
- The size of each block must be an exact integer multiple of the number of treatment groups
- Within each block, the order of treatments is a random permutation that are exactly balanced at the end of the block
- A sequence of blocks makes up the randomisation list
An Example

• There are two treatment groups in a drug trial
  – Active (A) and Placebo (P)

• With a block size of 4, there are six possible permutations with balanced group allocation:
  – (1) AAPP
  – (2) APAP
  – (3) APPA
  – (4) PPAA
  – (5) PAPA
  – (6) PAAP

<table>
<thead>
<tr>
<th>Block</th>
<th>ID</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>101</td>
<td>ACTIVE</td>
</tr>
<tr>
<td></td>
<td>102</td>
<td>PLACEBO</td>
</tr>
<tr>
<td></td>
<td>103</td>
<td>PLACEBO</td>
</tr>
<tr>
<td></td>
<td>104</td>
<td>ACTIVE</td>
</tr>
<tr>
<td>2</td>
<td>105</td>
<td>ACTIVE</td>
</tr>
<tr>
<td></td>
<td>106</td>
<td>PLACEBO</td>
</tr>
<tr>
<td></td>
<td>107</td>
<td>ACTIVE</td>
</tr>
<tr>
<td></td>
<td>108</td>
<td>PLACEBO</td>
</tr>
</tbody>
</table>
Blocked Randomisation

- Usually it is best to use relatively small block sizes to balance the treatments at frequent intervals.
  - Maximum imbalance equals to block size / number of arms
- Varying the length of each block (randomly) makes the sequences of assignments appear more random and can prevent unblinding.
  - E.g. with two treatment groups we can randomly choose blocks of size 2, 4, 6, or any multiple of two
- Stratifications are often considered using prognostic factors to prevent large or significant differences in baseline characteristics between the treatment groups
  - This is also called ‘stratified’ blocked randomisation
  - Blocking within each stratum
Dynamic Randomization

• A process in which the probability of assignment to the treatments in a clinical trial does not remain constant, but is determined by the current balance and/or composition of the groups

“Any method of treatment assignment in which the treatment assignment ratio changes as a function of previous assignments, baseline data or observed outcomes”

• Many such schemes exist including simple use of a biased coin, urn designs, play the winner, and minimization
Minimization

• When a new participant is to be randomised, the overall imbalance (with all stratification factors) that will result is calculated in two ways:
  – Assuming the next assignment is made to Group A
  – Assuming the next assignment is made to Group B

• The treatment assignment that will produce the smallest imbalance is chosen ($p=1$) or with a higher probability (e.g. $p=2/3$) in random allocation

• If the imbalance is the same either way, the next treatment is assigned randomly (i.e. 50:50)
An Example

- Consider two stratification factors in a two-arm trial:
  - Gender (Male and Female)
  - Age (Young, Middle, Old)
- Assume 50 participants have already been randomized:

Table 1: Fifty randomised participants by group and factor

<table>
<thead>
<tr>
<th>Factor Level</th>
<th>1</th>
<th>2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>Male</td>
<td>Female</td>
<td>Young</td>
</tr>
<tr>
<td>1</td>
<td>16</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>14</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>20</td>
<td>25</td>
</tr>
</tbody>
</table>

- What if the next participant to be randomized is an old male?
An Example

Define $B(t)$ a function that measures the lack of balance over both stratification factors, if the next participant is assigned to Group $t$ ($t=1,2$)

$$B(t) = \sum_{i=1}^{f} w_i |x^t_{i1} - x^t_{i2}|$$

Where $w_i$ indicates the relative importance of factor $i$

- $B(1) = w_1|17 - 14| + w_2|5 - 6|$
- $B(2) = w_1|16 - 15| + w_2|4 - 7|$

Let $w_1=3$ and $w_2=2$, we have $B(1)=11 > B(2)=9$.

Therefore, the next participant is randomised to Group 2 with a higher probability, e.g. $p=2/3$
Why Minimization

• Tends to produce better balance than blocked randomization when
  – a number of stratification factors are used
  – the recruitment stops early

• Requires a good track of current status of imbalance which relies more on software functionality
  – Note: cross-classification of factors is not considered

• Treatment allocation may not be considered random (except for the first participant, or when there is a balance between groups)
Stratified Permuted Block vs Adaptive Minimization
Secrecy of Randomization

• Benefits of randomization can be undone if an investigator discovers the treatment allocation
  – Can arise if clinical investigators who are responsible for allocating subjects have access to the randomization list
  – Administration of the randomization process should be physically separated

• In multi-center trials it is commonly accomplished by designating an off-site location to serve as study coordinating center
  – Treatment allocations are either made by telephone or secure computer access
  – Sealed envelopes are sometimes used
Bias in Clinical Trials

• The systematic tendency of any factors associated with the design, conduct, analysis and evaluation of the results to make the estimate of a treatment effect deviate from its true value
  – Operational bias: deviations in conduct of the trial
  – Statistical bias: other sources of bias listed above

• Most important design techniques for avoiding bias in clinical trials are randomisation and blinding
Level of Blindness

“A clinical trial should, ideally, have a double-blind design to avoid potential problems of bias during data collection and assessment. In studies where such a design is impossible, a single-blind approach and other measures to reduce potential bias are favored”

- Type of design
  - Open-label
  - Single-blind
  - Double-blind
  - Triple-blind

- Parties involved
  - Participant
  - Study Investigator
  - Monitoring committee

Definitions vary in textbooks

Fundamentals of Clinical Trials (3rd Ed), Friedman et al.
Physician Interpretations and Textbook Definitions of Blinding Terminology in Randomized Controlled Trials

<table>
<thead>
<tr>
<th>Interpretations and Definitions</th>
<th>Physicians</th>
<th>Textbook Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Single blinding</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants</td>
<td>75 (66/91)</td>
<td>74 (17/23)</td>
</tr>
<tr>
<td>Participants and health care providers</td>
<td>9 (8/91)</td>
<td>0</td>
</tr>
<tr>
<td>Health care providers</td>
<td>4 (4/91)</td>
<td>9 (2/23)</td>
</tr>
<tr>
<td>Investigators*</td>
<td>NA†</td>
<td>9 (2/23)</td>
</tr>
<tr>
<td>Judicial assessors‡</td>
<td>3 (3/91)</td>
<td>4 (1/23)</td>
</tr>
<tr>
<td>Data collectors</td>
<td>2 (2/91)</td>
<td>4 (1/23)</td>
</tr>
<tr>
<td>Other groups</td>
<td>7 (6/91)</td>
<td>0</td>
</tr>
<tr>
<td><strong>Double blinding</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants and health care providers</td>
<td>38 (35/91)</td>
<td>43 (12/28)</td>
</tr>
<tr>
<td>Participants and investigators*</td>
<td>NA†</td>
<td>21 (8/28)</td>
</tr>
<tr>
<td>Participants and judicial assessors‡</td>
<td>5 (5/91)</td>
<td>14 (4/28)</td>
</tr>
<tr>
<td>Participants, health care providers, and judicial assessors‡</td>
<td>13 (12/91)</td>
<td>0</td>
</tr>
<tr>
<td>Participants, health care providers, data collectors, and judicial assessors‡</td>
<td>10 (9/91)</td>
<td>0</td>
</tr>
<tr>
<td>Participants, health care providers, data collectors, and data analysts</td>
<td>7 (6/91)</td>
<td>0</td>
</tr>
<tr>
<td>Participants, health care providers, and data analysts</td>
<td>7 (6/91)</td>
<td>0</td>
</tr>
<tr>
<td>Participants and data collectors</td>
<td>5 (5/91)</td>
<td>4 (1/28)</td>
</tr>
<tr>
<td>Judicial assessors‡ and assignment to treatment or control</td>
<td>NA†</td>
<td>4 (1/28)</td>
</tr>
<tr>
<td>Participants, health care providers, and judicial assessors‡</td>
<td>1 (1/91)</td>
<td>4 (1/28)</td>
</tr>
<tr>
<td>Participants, health care providers, investigators,§ and health care personnel§</td>
<td>NA†</td>
<td>4 (1/28)</td>
</tr>
<tr>
<td>Other groups</td>
<td>13 (12/91)</td>
<td>7 (2/28)</td>
</tr>
</tbody>
</table>

| **Triple blinding**                                                |            |                     |
| Participants, health care providers, data collectors, judicial assessors,‡ and data analysts | 18 (15/83) | 0                   |
| Participants, health care providers, data collectors, judicial assessors,‡ data analysts, and authors | 18 (15/83) | 0                   |
| Participants, health care providers, and data collectors           | 16 (13/83) | 0                   |
| Participants, health care providers, and data analysts             | 10 (8/83)  | 14 (1/7)            |
| Participants, health care providers, and judicial assessors‡       | 8 (7/83)   | 14 (1/7)            |
| Participants, data collectors, and data analysts                   | 1 (1/83)   | 14 (1/7)            |
| Participants, data analysts, and investigators*                    | NA†        | 14 (1/7)            |
| Participants, judicial assessors,‡ and data analysts               | 0          | 14 (1/7)            |
| Participants, data collectors, and judicial assessors‡             | 7 (6/83)   | 0                   |
| Participants, investigators,§ and monitors/sponsors§               | NA†        | 14 (1/7)            |
| Participants, health care providers, data collectors, and judicial assessors‡ | 7 (6/83)   | 0                   |
| Other groups                                                       | 14 (12/83) | 14 (1/7)            |

*Not defined further.
† Includes a term that was not an option available for physicians; NA indicates not applicable.
‡ Judicial assessors of outcomes.
§ Laboratory technician, pharmacist, or others.
Open-Label Trials

• No party is blinded to which intervention the participant has been assigned
  – Decision should be made carefully even with unconventional treatments (e.g. surgical procedures)

• Pros and Cons
  – Simpler and less expensive to execute, if blinding is not an issue
  – Subject to operational and statistical bias
  – Every effort should be made to minimize the known sources of bias
    • Use of centralized randomisation method and the list is concealed till time of randomisation
    • Blind clinical assessment by independent medical staff
    • Primary endpoints as objective as possible
Single-Blind Trials

• Only one party is aware of which intervention each participant is receiving, i.e. the investigators
  – Bias is partially reduced by keeping participants blinded
  – Participants’ health and safety may be best served if the investigators themselves are not blinded

• Pros and Cons
  – Subject to bias in administration of non-study therapy, data collection and assessment
  – Same effort should be made to minimize the known sources of bias as in open-label trials
Double-Blind Trials

• Neither the participants nor the investigators responsible for following the participants know the identity of intervention assignment
  – Most commonly used in clinical trials if possible
  – Can minimize the bias if well designed and conducted

• Difficult and Expensive to achieve the ideal
  – Different drug formulations and administrations
    ➢ “Double-dummy” technique

A technique for retaining the blind when administering supplies in a clinical trial, when the two treatments cannot be made identical. Supplies are prepared for Treatment A (active and indistinguishable placebo) and for Treatment B (active and indistinguishable placebo). Subjects then take two sets of treatment; either A (active) and B (placebo), or A (placebo) and B (active).
Triple-Blind Trials

• An extension of double-blind design, with the committee monitoring response variables is also unaware of the group identity
  – Theoretical advantage to objectively evaluate the responses

• Potential ethical issues on participant safety
  – Data monitoring committee, if blinded, cannot carry out its responsibility to minimize potential harm to the participants
    ➢ Introduce flexibility: the committee can ask to be unblinded at any time it chooses
Scope of Trials

• Selection of questions (research assumptions)
  – Primary objective
  – Secondary objectives

• Selection of outcome measures
  – Primary variable
  – Secondary variables

• Target population and sample size
  – Eligibility criteria
  – Sample size and power calculation
Study Objectives

• Primary objective
  – The most interesting question to answer
  – Can be adequately and objectively evaluated

• Secondary objectives
  – Secondary questions related to the primary
  – Subgroup hypotheses
    • Pre-specified (before data collection)
    • Limited in number (avoid chance finding)
    • With reasonable expectations
      – often too small to adequately test the hypothesis!
2. **STUDY OBJECTIVES**

The main objective of this study is to determine the effects of active video games over 6 months on: body mass index (BMI), percent body fat, waist circumference, cardio-respiratory fitness, and physical activity levels in children.

2.1. **Primary Objective**

The primary objective of this study is to evaluate the potential effect of playing active video games on change in children’s BMI from baseline to end of 6-month intervention.

2.2. **Secondary Objectives**

The secondary objectives are to evaluate the potential effect of playing active video games on change in children’s body composition (% body fat, waist circumference), cardio-respiratory fitness ($\text{VO}_{2\text{max}}$), physical activity levels (PAQ-C, Accelerometer), time spent on active/non-active video games, from baseline to end of 6-month intervention.

Additional interests are to evaluate the potential effect of playing active video games on change in children’s attitudes to physical activity and food behaviour, from baseline to end of 6-month intervention.
Study Endpoints

• Type of outcome measures
  – Efficacy (to evaluate the intervention effect)
  – Safety (to assess the adverse events)

• Level of importance
  – Primary variable
    • Providing most clinically relevant and convincing evidence directly related to the primary objective of the trial
    • Generally limited to one efficacy variable
    • Used for sample size and power calculation
  – Secondary variables
    • Supportive measurements related to the primary objective or measurements of effects related to the secondary objectives
    • Number of variables should be limited
Study Endpoints

• Level of importance (cont.)
  – Composite Variables
    • Integrate or combine the multiple measurements into a single or “composite” variable using a pre-defined algorithm
    • Used when a single primary variable cannot address multiple measurements associated with the primary objective
  – Exploratory Variables
    • Considered for exploratory trend analyses
    • Least important but potentially useful in future trials
Efficacy of an mHealth exercise-based cardiac rehabilitation programme (HEART)

5 STUDY ENDPOINTS

5.1 Efficacy

The following efficacy endpoints will be measured during the intervention period to evaluate the potential treatment effect.

5.1.1 Primary outcome

The primary outcome measure is change in maximal oxygen uptake ($\text{VO}_{2\text{max}}$) from baseline to 24 weeks, measured as Peak $\text{VO}_2$ (ml.kg$^{-1}$.min$^{-1}$).

5.1.2 Secondary outcomes

The following secondary outcome measures will be evaluated at 24 weeks, after the baseline assessment:

- Self-efficacy.
- Physical function measured with the 6-minute walk test.
- Total physical activity using the International Physical Activity Questionnaire (IPAQ).
- Meeting physical activity recommendations of 150 min/week moderate intensity activity.
- Cardiovascular risk factors (systolic blood pressure, weight, and waist hip ratio).
- Health related quality of life (SF-36) subscales.
- Cost-effectiveness – programme and direct medical cost, Quality Adjusted Life Year (QALY).

5.2 Safety

Information regarding any adverse events will be collected at 24 weeks.
Sample Size in Clinical Trials

• Clinical trials should be LARGE ENOUGH to detect reliably the smallest possible differences in the primary outcome with treatment that are considered clinically worthwhile

  ➢ “It is not uncommon for studies to be underpowered, failing to detect even large treatment effects because of inadequate sample size.”

• Sample size MUST BE PLANNED carefully to ensure research time, patient effort and support costs invested in any clinical trial are not wasted

  ➢ “… unethical to recruit patients into a study that does not have a large enough sample size for the trial to deliver meaningful information on the tested intervention.”

Determining the sample size in a clinical trial, Kirby et al. MJA 2002, 177
Type I and II Errors

<table>
<thead>
<tr>
<th>Test effect</th>
<th>True effect</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>✔</td>
</tr>
<tr>
<td>No</td>
<td>No</td>
<td>✔</td>
</tr>
</tbody>
</table>

- Type I error – “False Positive” (Level of Significance $\alpha$)
- Type II error – “False Negative” ($\beta$)

- *We want to reduce the probability of both (i.e. small $\alpha$ and $\beta$)*...
Sample Size Calculation

• Sample size estimation/determination
  – Calculate sample size to achieve desired accuracy and study power

• Sample size justification
  – Provide statistical justification for a selected sample size

• Sample size adjustment
  – Adjust for certain factors to yield sufficient number of evaluable subjects, e.g. loss to follow up
  – Control for overall type I error rate with planned and unplanned interim analyses

• Sample size re-estimation
  – Pre-planned during the trial, in a blinded or unblinded fashion

Sample size calculations in clinical research, Chow et al. 2003
Basic Considerations

• The following questions should be clarified before carrying out a sample size calculation:
  – Study design
    • Parallel, cross-over, factorial, cluster?
  – Primary objective and outcome
    • Variable definition: post-measure, or change from baseline?
    • Type of measurement: continuous, binary, or time to event?
  – Data collection and planned analysis
    • Repeated measures?
    • Simple test, or covariate-adjusted?
    • Test for equality, non-inferiority, superiority, or equivalence?
Components of Sample Size Calculation

• Minimum information required:
  - Power (1-\(\beta\))
  - Level of significance (\(\alpha\))
  - Size of the treatment effect sought (\(\varepsilon\))
    - E.g. 25% therapeutic improvement

• Other contributing factors:
  - Multiple comparisons (multiplicity)
    - Controlling the overall type I error rate with an adjusted \(\alpha\) level
    - E.g. Bonferroni adjustment or similar
  - Compliance rate
    - Loss to follow up (introduce missing data)
    - Drop in/out (switch treatments that are randomly allocated to)
Study Power

• Probability of detecting a true treatment effect between the groups
  – A power of 80% indicates a likelihood of 20% risk of missing a true effect (i.e. Type II error)
  – 90% power is often recommended for large confirmatory trials to reduce the risk of having a “false-negative” result

The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial - Survey of 71 "negative" trials.


Seventy-one "negative" randomized control trials were re-examined to determine if the investigators had studied large enough samples to give a high probability (greater than 0.90) of detecting a 25% and 50% therapeutic improvement in the response. Sixty-seven of the trials had a greater than 10 percent risk of missing a true 25% therapeutic improvement, and with the same risk, 50 of the trials could have missed a 50% improvement…
Level of Significance (\(\alpha\))

- Indicates the likelihood of detecting a treatment effect when no effect exists
  - So-called “false-positive” (Type I error)
  - Most commonly set to 5%, or 1%

- Defines the threshold “\(P\) value”
  - \(P\)-value<\(\alpha\) indicates an evidence of significant treatment effect between the groups
  - \(P\)-value\(\geq\alpha\) shows no/weak evidence of a treatment effect
    - Any observed difference may be due to chance alone
Size of Treatment Effect

• Defined based on the nature of primary variable
  ▪ Continuous → difference in means
  ▪ Categorical → difference in proportions / risk ratio
  ▪ Time to event → difference in number of events / hazard ratio

• Crucial to trial design
  – Minimum clinically meaningful difference to detect between groups
  – A question to the investigator rather than statistician
    • Earlier phase studies
    • Published related studies
    • Clinical experience
  – An implausibly large treatment effect will inevitably render statistical non-significance with a smaller real difference
An Generic Expression

Required sample size will increase with

- Higher study power (reduced Type II error)
- Lower level of significance (reduced Type I error)
- Smaller size of treatment effect to detect

2: Generic expression for calculating sample size

Sample size \( \propto \frac{\text{power, inverse function of significance level}^*}{(\text{absolute difference})^2} \)

* As the \( P \) value becomes smaller, the function of the significance level increases.

Determining the sample size in a clinical trial, Kirby et al. MJA 2002, 177
Non-Compliance

• Participants may not be 100% compliant to their allocated treatments during the study
  ➢ Drop in/out - Patients “switch” to another treatment group
    • Fail to take investigational products (drop out)
    • Cross-over to treatment from the control group (drop in)
    • Often problematic and need to be kept to an absolute minimum

• Another closely related issue is loss-to-follow-up (LTFU)
  • Patients withdrawal from the study or cannot be contacted
  • Common in longitudinal studies with a relatively long study period
  • Introduce missing data

• Both situations will affect the size of achievable treatment difference in statistical analysis and should therefore be adjusted in sample size calculation
Sample Size Adjustment

• LTFU can result in missing data and subsequently reduced sample size and study power
  – N per arm / (1 - %LTFU)
    • e.g. 100 / (1 - 20%) = 125 per arm

• Drop in/out may potentially dilute the underlying true effect of the intervention
  – Let R_o and R_i indicate the rate of drop-out and drop-in
    – N per arm / (1 – R_o – R_i)^2
      • e.g. 100 / (1 - 10% - 5%)^2 = 139 per arm

• Assume that the data are MCAR...