The minimum required number of clusters with unequal cluster sizes in cluster randomized trials

Zhiying You, MD, PhD
Michigan State University

May 19, 2015

SCT 36th Annual Meeting
Arlington, VA

Outline

- Introduction
- Two-level design
- Three-level design
- Summary

Introduction

- Clusters are randomized, NOT individuals
- All individuals within a cluster receive same intervention
- In a two-level design, there are two sample size components, number of clusters (g) at level-2 and number of individuals within clusters (m) at level-1
- Examples:
  - Residents (m) in communities (g)
  - Students (m) in classrooms(g) and classrooms(g) in schools (c)
- m usually varies across clusters, i.e. unequal sizes
- There are three sample size components in a three level design, denoted c, g, and m.

Important characteristics

- There are multiple sample size components, e.g. (g, m),; g or m alone CANNOT tell you the whole story; and total sample size (total subjects) can be very large
- Small degrees of freedom due to small number of clusters (i.e. small g) or highest/top level units → decreased power
- Individuals within the same cluster are positively correlated, measured as intraclass (intracluster) correlation coefficient (ICC, $\rho > 0$) → decreased power
- Unequal cluster sizes (coefficient of variation of cluster size, $cv > 0$) → decreased power

Important considerations

- Large number of small clusters (large g, small m) is generally preferred statistically over small number of large clusters (small g, large m)
- Available number of clusters is small in practice, i.e. small g
  - Qualification, and willing
- Same study power can be achieved by many different combinations of g and m
- You may prefer large m
  - Reduction in cost and logistical complexity
  - Smaller variation in cluster sizes
  - More concerns about contamination with small adjacent clusters
- Trade-offs between g and m
- Questions trade-offs between g and m

  (1) How small g can be? 6 or 10 per group? Why?

  (2) What $\bar{m}_i$ should be if $g = 6, 100 \text{ or } 1000$?

Two-level designs

- Nested model $y_{ijk} = \mu + A_i + u_{j(i)} + \varepsilon_{ijk}$

  $i = 1{,}2$, two intervention groups

  $j = 1, \ldots, g$, number of clusters per group

  $k = 1, \ldots, m_{ij}$, clusters sizes, mean size $\bar{m}_i$,

  coefficient of variation by cv_i

  μ is constant (grant mean)

  A_i is fixed intervention effect, $\sum_{i=1}^{2} A_i = 0$

  $u_{j(i)} \sim N(0, \sigma_u^2)$, random effect of cluster

  $\varepsilon_{ijk} \sim N(0, \sigma_\varepsilon^2)$, random error

  $u_{j(i)}$ and ε_{ijk} are independent

ICC and t-test

- ICC $\rho = \frac{\sigma_u^2}{\sigma_u^2+\sigma_\varepsilon^2} = \frac{\sigma_u^2}{\sigma_y^2}$

- $Var(\bar{y}_i) = \sigma_y^2 F_i,$ where

$$F_i = \frac{1}{g\bar{m}_i}\left\{1 + \left\{\left(\left(\frac{g-1}{g}\right) cv_i^2 + 1\right)\bar{m}_i - 1\right\}\rho\right\}$$

$$= \frac{1}{g\bar{m}_i}\left\{1 + (\bar{m}_i - 1)\rho + \left(\frac{g-1}{g}\right) cv_i^2 \bar{m}_i \rho\right\}$$

- The two-sample t-test is used to test H_0: $\mu_1 = \mu_2$

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{(F_1 + F_2)\hat{\sigma}_y^2}}$$

Power

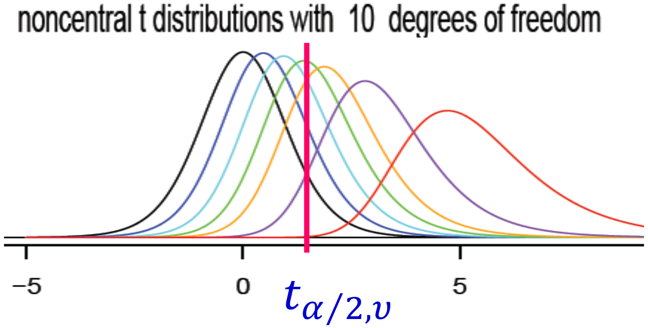
- Under H$_1$: $\mu_1 \neq \mu_2$, t has noncentral t distribution with degrees of freedom (df) $\upsilon = 2(g-1)$, and the noncentrality parameter can be calculated as

$$\tau = \frac{\delta}{\sqrt{(F_1+F_2)}}, \quad \delta = \frac{|\mu_1-\mu_2|}{\sigma_y}$$ is effect size (ES)

- Power of the test can be calculated as

$$\psi = P(noncentral \ |t| \geq t_{\alpha/2,\upsilon}|\tau)$$

From left to right, noncentrality parameters are: 0, 0.5, 1, 1.5, 2, 3, 5



- Power increases ONLY with τ, given $\upsilon = 2(g-1)$

Noncentrality parameter

- $\tau = \dfrac{\delta\sqrt{g}}{\sqrt{2\rho + \left(\frac{1}{\bar{m}_1} + \frac{1}{\bar{m}_2}\right)(1-\rho) + \left(\frac{g-1}{g}\right)\left(cv_1^2 + cv_2^2\right)\rho}}$

  the parts relevant to cluster sizes

- τ increases with cluster sizes and achieve its upper bound as $\tau_{\max(g)} = \dfrac{\delta\sqrt{g}}{\sqrt{2\rho}}$ for given g when both $\bar{m}_1$ and $\bar{m}_2$ tend to infinity.
- There is an upper bound in power if g is fixed and only cluster sizes can increase.
- g should NOT be too small !!
- Questions about trade-offs between g and $\bar{m}_i$:

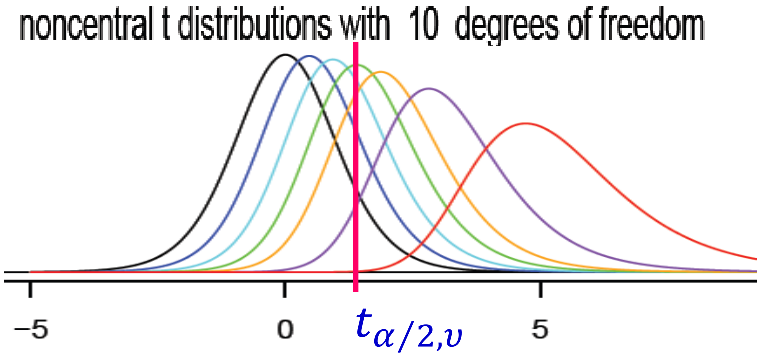
  (1) How small g can be? 10 or 8per group? Why?

  (2) What $\bar{m}_i$ should be if $g = 10, 100$ or 1000 ?

Minimum parameter for power

- Let $\tau_{\min(\upsilon,\psi)}$ be the minimum value of τ such that $\psi = P\left(noncentral\ \ |t| \geq t_{\alpha/2,\upsilon} \middle| \tau_{\min(\upsilon,\psi)}\right) = 1 - \beta$ with $\upsilon = 2(g-1)$, β is type II error rate.

From left to right, noncentrality parameters are: 0, 0.5, 1, 1.5, 2, 3, 5



- Then we want to have the sample size components such that

$$\tau = \frac{\delta\sqrt{g}}{\sqrt{2\rho + \left(\frac{1}{m_1} + \frac{1}{m_2}\right)(1-\rho) + \left(\frac{g-1}{g}\right)\left(cv_1^2 + cv_2^2\right)\rho}} = \tau_{\min(\upsilon,\psi)}$$

Minimum required # of clusters

- Assuming $\bar{m}_1 = \bar{m}_2 = \bar{m}$ and $cv_1 = cv_2 = cv$ gives

- $\bar{m} = \frac{2g(1-\rho)\tau^2_{\min(v,\psi)}}{\delta^2 g^2 - 2\rho\{g+(g-1)cv^2\}\tau^2_{\min(v,\psi)}}$ requires

$$g > g_0 = \frac{2\rho\left(1+cv^2\right)\tau^2_{\min(v,\psi)} + \sqrt{\{2\rho(1+cv^2)\}^2\tau^2_{\min(v,\psi)} - 8\delta^2\rho cv^2\tau^2_{\min(v,\psi)}}}{2\,\delta^2}$$

- g_{min} is the minimum integer greater than g_0

- Given $g \geq g_{min}$ → $\bar{m}$ can now be calculated
- Any trade-offs between g and $\bar{m}$ should have $g \geq g_{min}$. Waste of resource otherwise!

Examples

Table 1. $g_{min(\psi)}$ and $\bar{m}$ for power of 90% with 2-sided α = 0.05

δ	ICC	cv	g_{min}	$\bar{m}$ when g takes greater values			
				g_{min}	$g_{min}+1$	$g_{min}+2$	$g_{min}+3$
0.2	0.02	0.0	18	1157.1	354.4	148.6	79.5
0.2	0.02	0.5	20	26837.8	508.5	171.6	86.0
0.2	0.02	1.0	28	1128.1	358.1	151.2	80.9
0.2	0.03	0.0	23	2331.8	416.8	157.9	81.7
0.2	0.03	0.5	27	1067.6	346.4	147.3	79.0
0.2	0.03	1.0	39	528.3	261.3	129.7	73.9
0.3	0.02	0.0	12	1239.6	188.5	70.4	36.4
0.3	0.02	0.5	13	1074.7	188.6	71.2	36.8
0.3	0.02	1.0	16	3730.0	222.5	76.8	38.6
0.3	0.03	0.0	15	247.2	117.4	57.4	32.5
0.3	0.03	0.5	16	641.3	168.2	67.8	35.8
0.3	0.03	1.0	21	528.4	161.4	67.4	35.9

Three-level designs

- Nested model

$y_{ijkl} = \mu + A_i + u_{j(i)} + \tau_{k(ij)} + \varepsilon_{ijkl}$

$i = 1{,}2$, two treatment groups

$j = 1, .., c$, number of level-3 units per group

$k = 1, .., g$, number of level-2 units per level-3 unit

$l = 1, .., m_{ijk}$, clusters size, $\bar{m}_{ij}$, cv_{ij}

μ is constant (grant mean),

A_i is fixed effects of treatment, $\sum_{i=1}^{2} A_i = 0$,

$u_{j(i)} \sim N(0, \sigma_u^2)$, random effect of level-3 unit,

$\tau_{k(ij)} \sim N(0, \sigma_\tau^2)$, random effect of level-2 unit,

$\varepsilon_{ijkl} \sim N(0, \sigma_\varepsilon^2)$, random error,

$u_{j(i)}$ and $\tau_{k(ij)}$ and ε_{ijk} are independent one another.

ICCs and t-test

- Two ICCs: $\rho_3 = \frac{\sigma_u^2}{\sigma_y^2}$,

  $\rho_2 = \frac{\sigma_u^2+\sigma_\tau^2}{\sigma_y^2}$; note: $\rho_2 \geq \rho_3$

- $Var(\bar{y}_i) = \sigma_y^2 F_i$, where

$$F_i = \frac{1}{cg\bar{m}_i}\left\{1 + \bar{m}_i(g-1)\rho_3 + (\bar{m}_i - 1)\,\rho_2 + \left(\frac{g-1}{g}\right) cv_i^2 \bar{m}_i(\rho_2 - \rho_3)\right\}$$

$$= \frac{1}{c}\left\{\frac{1}{g\bar{m}_i} + \left(\frac{g-1}{g}\right)\rho_3 + \frac{1}{g}\left(\frac{\bar{m}_i - 1}{\bar{m}_i}\right)\rho_2 + \frac{1}{g}\left(\frac{g-1}{g}\right) cv_{m(i)}^2(\rho_2 - \rho_3)\right\}$$

- $t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{(F_1+F_2)\hat{\sigma}_y^2}}$

Power

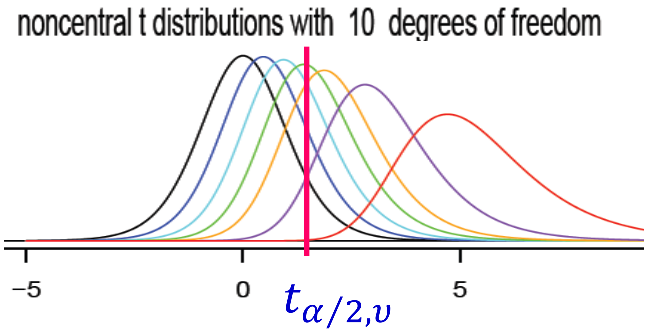
- Under H_1: $\mu_1 \neq \mu_2$, t has noncentral t distribution with degrees of freedom (df) $\upsilon = 2(c-1)$, and the noncentrality parameter can be calculated as

$$\tau = \frac{|\mu_1 - \mu_2|}{\sqrt{(F_1+F_2)\sigma_y^2}} = \frac{\delta}{\sqrt{(F_1+F_2)}}, \quad \delta = \frac{|\mu_1-\mu_2|}{\sigma_y}$$ is effect size (ES)

- Power of the test can be calculated as

$$\psi = P(noncentral\ \ |t| \geq t_{\alpha/2,\upsilon} | \tau)$$

From left to right, noncentrality parameters are: 0, 0.5, 1, 1.5, 2, 3, 5



- Power increases ONLY with τ, given $\upsilon = 2(c-1)$

Noncentrality parameter

- Assuming $\bar{m}_1 = \bar{m}_2 = \bar{m}$ and $cv_1 = cv_2 = cv$ gives

$$\tau = \frac{\delta\sqrt{c}}{\sqrt{2\left\{\frac{1}{g\bar{m}} + \left(\frac{g-1}{g}\right)\rho_3 + \frac{1}{g}\left(\frac{\bar{m}-1}{\bar{m}}\right)\rho_2 + \frac{1}{g}\left(\frac{g-1}{g}\right)cv^2(\rho_2-\rho_3)\right\}}}$$

- $\tau_{\max(c,g)} = \frac{\delta\sqrt{c}}{\sqrt{2\frac{1}{g}\{(g-1)\rho_3 + \rho_2\}}}$ if only $\bar{m}$ can increase
- $\tau_{\max(c)} = \frac{\delta\sqrt{c}}{\sqrt{2\rho_3}}$ if only g and $\bar{m}$ can increase
- There is an upper bound in power if c and g are fixed and only cluster sizes can increase
- Number level-2 and -3 units should NOT be too small
- Questions about trade-offs between c and g and $\bar{m}$:

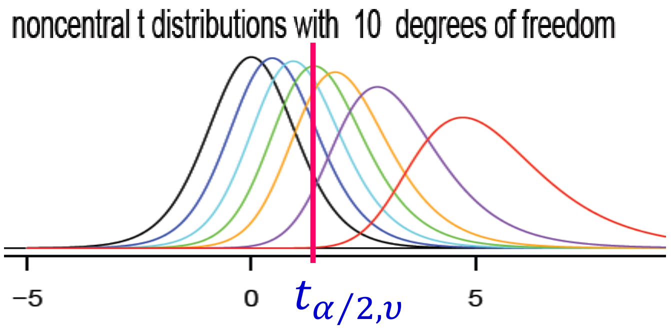
(1) How small c and g can be? $c = 10$ and $g = 12$? Why?

(2) What $\bar{m}$ should be if $c = 10$ and $g = 12,\ 100$ or 1000 ?

Minimum parameter for power

- Let $\tau_{\min(\upsilon,\psi)}$ is the minimum value of τ such that power $\psi = P\left(noncentral\ \ |t| \geq t_{\alpha/2,\upsilon} \middle| \tau_{\min(\upsilon,\psi)}\right) = 1 - \beta$ with $\upsilon = 2(c-1)$, β is type II error rate.

From left to right, noncentrality parameters are: 0, 0.5, 1, 1.5, 2, 3, 5



- Then we want to have the sample size components c, g, and $\overline{m}$ such that

$$\tau = \frac{\delta\sqrt{c}}{\sqrt{2\left\{\frac{1}{g\overline{m}}+\left(\frac{g-1}{g}\right)\rho_3+\frac{1}{g}\left(\frac{\overline{m}-1}{\overline{m}}\right)\rho_2+\frac{1}{g}\left(\frac{g-1}{g}\right)c\upsilon^2(\rho_2-\rho_3)\right\}}} = \tau_{\min(\upsilon,\psi)}$$

Minimum required # of clusters

$$\bar{m} = \frac{2g\tau^2_{\min(\upsilon,\psi)}\{1-\rho_2\}}{\left(\delta^2 c - 2\tau^2_{\min(\upsilon,\psi)}\rho_3\right)g^2 - 2\tau^2_{\min(\upsilon,\psi)}(1+cv^2)(\rho_2-\rho_3)g + 2\tau^2_{\min(\upsilon,\psi)}cv^2(\rho_2-\rho_3)}$$

Requires

$$g > \frac{2\tau^2_{\min(\upsilon,\psi)}(\rho_2-\rho_3)\left(1+cv^2\right) + \sqrt{4\tau^4_{\min(\upsilon,\psi)}(\rho_2-\rho_3)^2(1+cv^2)^2 - 8\left(\delta^2 c - 2\tau^2_{\min(\upsilon,\psi)}\rho_3\right)\tau^2_{\min(\upsilon,\psi)}cv^2(\rho_2-\rho_3)}}{2\left(\delta^2 c - 2\tau^2_{\min(\upsilon,\psi)}\rho_3\right)}$$

which requires

$$c > \frac{2\rho_3\tau^2_{\min(\upsilon,\psi)}}{\delta^2}$$

Take the minimum integer to be the minimum required number

Given c_{min} → g_{min} → $\bar{m}$

When consider trade-offs between c and g and $\bar{m}$, c and g should go first !!!

Examples

Table 1. $c_{min(\psi)}$, $g_{min(\psi)}$ and $\bar{m}$ for power of 80% with 2-sided α = 0.05

δ	ρ_3	ρ_2	cv	c_{min}	g_{min}	$\bar{m}$	c_{min} +1			g_{min} +1		c_{min} +1 and g_{min} +1		
							c	g	$\bar{m}$	g	$\bar{m}$	c	g	$\bar{m}$
0.2	0.02	0.025	0.0	12	6	11892.1	13	5	14044.5	7	1049.5	13	6	900.0
0.2	0.02	0.025	0.5	12	8	1429.1	13	6	23414.9	9	644.9	13	7	950.5
0.2	0.02	0.025	1.0	12	12	1679.2	13	10	1526.2	13	698.6	13	11	606.6
0.2	0.03	0.035	0.0	17	13	2364.5	18	10	3116.8	14	1170.9	18	11	1148.0
0.2	0.03	0.035	0.5	17	15	13132.5	18	12	4275.6	16	1992.6	18	13	1289.0
0.2	0.03	0.035	1.0	17	24	5011.1	18	19	2746.0	25	1607.7	18	20	1110.2
0.3	0.02	0.025	0.0	6	7	1825.8	7	4	700.5	8	736.1	7	5	326.1
0.3	0.02	0.025	0.5	6	8	4223.5	7	4	2145.8	9	971.1	7	5	490.0
0.3	0.02	0.025	1.0	6	13	1472.6	7	6	2315.0	14	684.1	7	7	513.1
0.3	0.03	0.035	0.0	9	2	338.8	10	2	249.2	3	142.5	10	3	116.1
0.3	0.03	0.035	0.5	9	2	434.0	10	2	297.1	3	162.5	10	3	129.1
0.3	0.03	0.035	1.0	9	2	2771.0	10	2	702.9	3	280.6	10	3	193.9

Summary

- Enough total sample size may not imply enough power
- Number of highest level units (clusters) dominate the power
- There is up bound in power for given number of clusters
- There exists a minimum required number of clusters at different levels for a specified power
- trade-offs between number of clusters and cluster sizes can work for your specified power ONLY if you have a minimum required numbers of clusters
- Consider the minimum required numbers of clusters, NOT ONLY total sample size!

Thank you!

Questions

?

zyyou@msu.edu