Simplified power analyses for clustered sampling designs with compound symmetric covariance structure of *x* & y: A survey of sample size ratios (SSR)

UCSF CAPS Methods Core

May 11th, 2021

Steve Gregorich

Overview

Sample size ratios (SSR) provide convenient short-cuts for sample size calculations

Assumptions

. 2- and 3-level clustered sampling designs

. Limited coverage of 3-level designs in this talk

. Compound symmetric correlation structure of both *x* and *y*

Regression modeling contexts

. GLMM

. GEE

. Survey Sampling (SS)

Sample Size Ratios (SSR): Introduction

AKA design effect, misspecification effect, variance inflation/deflation factor.
I chose the SSR label because it is broadly applicable

Assume a simple random sample (SRS) of size *N* drawn from a population with population mean μ and variance σ^2

We choose the usual estimator $\hat{\mu}$ of the sample mean of *x*

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

The variance of the estimator $\hat{\mu}$ is

$$\sigma^2_{\hat{\mu}} = \sigma^2/N$$

The *precision* of the estimator is the inverse of the above quantity, i.e., $1/\sigma^2_{\hat{\mu}} = N/\sigma^2$, i.e., larger *N* obtains higher precision.

Sample Size Ratios (SSR): Introduction

Say we have an alternative estimator $\hat{\mu}_{\mathrm{a}}$ with variance equal to

$\sigma^2_{\hat{\mu}_{\mathrm{a}}} = \sigma^2 / N_{\mathrm{a}}$, and rearranging

$N_{\mathrm{a}} = \sigma^2 / \sigma^2_{\hat{\mu}_{\mathrm{a}}}$,

i.e., N_{a} equals population variance $(\sigma^2) \times$ estimator precision $(1/\sigma^2_{\hat{\mu}_a})$

Similarly, for estimator $\hat{\mu}$

$$N = \sigma^2 / \sigma^2_{\hat{\mu}}$$

SSR represents relative (effective) sample size and relative precision, i.e.,

$$\mathrm{SSR} = \frac{N}{N_a} = \frac{\frac{\sigma^2}{\sigma^2_{\hat{\mu}}}}{\frac{\sigma^2}{\sigma^2_{\hat{\mu}_{\mathrm{a}}}}} = \frac{\sigma^2_{\hat{\mu}_{\mathrm{a}}}}{\sigma^2_{\hat{\mu}}}$$

Sample Size Ratios (SSR): Introduction

$$\mathrm{SSR} = \frac{N}{N_a} = \frac{\sigma^2_{\hat{\mu}_\mathrm{a}}}{\sigma^2_{\hat{\mu}}}$$

Assume N=1000, estimator $\hat{\mu}_\mathrm{a}$ has $\sigma^2_{\hat{\mu}_\mathrm{a}}$=2, and estimator $\hat{\mu}$ has $\sigma^2_{\hat{\mu}}$=1

. SSR, as defined above, equals 2÷1=2

. I.e., $\hat{\mu}_\mathrm{a}$ has larger variance and lower precision than $\hat{\mu}$

Knowing N and SSR, we can calculate the effective sample size, Neff, for an application of $\hat{\mu}_\mathrm{a}$

For N=1000 and SSR=2, when applying $\hat{\mu}_\mathrm{a}$, Neff=N÷SSR= 500.

. I.e., when applying $\hat{\mu}_\mathrm{a}$ with N=1000, the Neff=500

. Or, wrt precision, $\hat{\mu}_\mathrm{a}$ with N=1000 is equivalent to $\hat{\mu}$ with N=500

At the same time, SSR=2 indicates the expectations that

. the variance of $\hat{\mu}_\mathrm{a}$ (i.e., $\sigma^2_{\hat{\mu}_\mathrm{a}}$) will equal 2× the variance of $\hat{\mu}$ (i.e., $\sigma^2_{\hat{\mu}}$)

. the std err of $\hat{\mu}_\mathrm{a}$ (i.e., $\sigma_{\hat{\mu}_\mathrm{a}}$) will equal $\sqrt{2}$ times the std err of $\hat{\mu}$ (i.e., $\sigma_{\hat{\mu}}$)

Ex #1a: Planning for a Cluster-Randomized Trial (CRT)

Context

. Clustered sampling: :Level1 participants nested w/in Level2 clusters

. Level2 clusters are randomized w/ 1:1 allocation to experimental groups

. N=1000: n_2=100 clusters, each of size n_1=10

. $y \sim N(0,1)$, $x \sim B(0.50)$, where x is the experimental group indicator

. Linear regression model

. Intra-cluster correlation (ICC) of y (ρ_y) equals 0.05

. 80% power with two-tailed $\alpha = .05$

Goal

. Solve for minimum detectable effect size, b_x

In this context, the familiar Design Effect (Deff) is a useful SSR.

. $\mathrm{SSR} = \mathrm{Deff} = 1 + r\rho_y, \quad \text{where } r = n_1 - 1$,

Deff was described by Kish

Kish (1965). *Survey Sampling.* New York: John Wiley & Sons, Inc

Ex #1a: Planning for a Cluster-Randomized Trial (CRT)

Application of SSR (Deff) to solve for b_x

Step 1. Calculate SSR & N_{eff} given r=9, ρ_y=.05, and N=1000

. $\text{SSR} = \text{Deff} = 1 + r\rho_y = 1 + (10 - 1) \times .05 = 1.45$

. $N_{\text{eff}} = N/\text{SSR} = 1000/1.45 = 689.7$

. Note. $N_{\text{eff}} < N$

Step 2. Calculate minimum detectable effect specifying N=689.7

Result: b_x = .21244 (from PASS Linear Regression routine, σ_y^2=1)

In this case, a GLMM/GEE/SS model fit to N=1000 clustered observations obtains the same power as plain linear regression model fit to $N \cong 690$ independent observations (SRS)

Ex #1b: Planning for a Cluster-Randomized Trial (CRT)

If we instead began w/ values of b_x and n_1, we could solve for n_2 and N

. b_x = 0.21244

. n_1 = 10

Step 1. Calculate N assuming b_x=.21244 & independent obs. (ρ_y=0)

. from PASS Linear Regression routine, N = 690, if ρ_y=0

(PASS only returns integer N values)

. In this case, N from PASS is our target effective sample size, Neff

Step 2. Calculate N assuming ρ_y=.05 and n_{L1}=10

. $\text{SSR} = 1 + r\rho_y = 1 + 9 \times .05 = 1.45$

. $N = N_{\text{eff}} \times \text{SSR} = 690 \times 1.45 = 1000.5$

. $n_2 = N/n_1 = 1000.5/10 \cong 100$

Ex #2: Planning for a RCT Randomizing Level1 Units

Context

. Clustered sampling: Level1 participants nested w/in Level2 clusters

. Level1 units are randomized with 1:1 allocation to experimental groups

. N=1000: n_2=100 clusters, each of size n_1=10

. $y \sim N(0,1)$, $x \sim B(0.50)$, where x is the experimental group indicator

. Linear regression model

. Intra-cluster correlation (ICC) of y (ρ_y) equals 0.05

. 80% power with two-tailed α = .05

Goal

. Solve for minimum detectable effect size, b_x

In this context, a familiar sample size ratio is

. $SSR = 1 - \rho_y$

This is the same SSR that applies to a paired *t*-test

Ex #2: Planning for a RCT Randomizing Level1 Units

Application of the SSR: solve for b_x

Step 1. Calculate SSR and Effective Sample Size (N_{eff})

. $\text{SSR} = 1 - \rho_y = 1 - .05 = 0.95$

. $N_{\text{eff}} = N/SSR = 1000/0.95 = 1052.6$

. Note. $N_{\text{eff}} > N$

Step 2. Calculate minimum detectable effect specifying N=1052.6

Result: $b_x = .17222$ (from PASS)

Here, a GLMM/GEE/SS model fit to N=1000 clustered observations obtains the same precision as a plain model fit to N=1053 independent observations

Sample Size Ratios (SSR): 2-Level Sampling Designs

So far, we've discussed two sample size ratios

$$\text{SSR} = \text{Deff} = 1 + r\rho_y$$

and

$$\text{SSR} = 1 - \rho_y$$

When does each apply?

The choice depends on whether the *x* has

. a between-cluster or

. a within-cluster effect

The intra-cluster correlation of *x* (ρ_x) is important

Sample Size Ratios (SSR): 2-Level Sampling Designs

We often think of ICC in terms of a variance component decomposition.

$$\rho = \frac{\sigma^2_{\text{between\_cluster}}}{\sigma^2_{\text{between\_cluster}} + \sigma^2_{\text{w/in\_cluster}}}$$

However, that formulation has positive bias.

When thinking about ρ_x, it is helpful to consider the unbiased formula (Harris 1913; Kish 1965; Wikipedia ICC page)

$$\rho_x = \frac{\sigma^2_{x.\text{between\_cluster}} - \sigma^2_{x.\text{w/in\_cluster}}/r}{\sigma^2_{x.\text{between\_cluster}} + \sigma^2_{x.\text{w/in\_cluster}}}$$

Harris JA (1913). On the calculation of intra-class and inter-class coefficients of correlation from class moments when the number of possible combinations is large. *Biometrika*, 9, 446–472.

Kish, Leslie (1965). *Survey Sampling.* New York: John Wiley & Sons, Inc (p. 170)

Sample Size Ratios (SSR): 2-Level Sampling Designs

$\mathrm{SSR} = \mathrm{Deff} = 1 + r\rho_y$

In a regression context,
this SSR applies when *x* has a fully <u>between-cluster</u> effect on *y*

When will *x* have a fully <u>between-cluster</u> effect on *y*?

. When *x* is a Level2 variable; it will have
positive between-cluster & <u>zero</u> (**0**) within-cluster variance

. In this case, the intra-cluster correlation of *x* (ρ_x) equals 1.0.

Consider the unbiased formula for intra-cluster correlation as applied to *x*

$$\rho_x = \frac{\sigma^2_{x.\text{between\_cluster}} - \mathbf{0}/r}{\sigma^2_{x.\text{between\_cluster}} + \mathbf{0}} = 1$$

Use of this SSR (i.e.,. $\mathrm{SSR} = \mathrm{Deff} = 1 + r\rho_y$) assumes
a fully between-cluster *x* effect, $\rho_x = 1$

Therefore, I label this SSR as SSR_b (i.e., sub-'b' for 'between')

Summary Implications: SSR_b

. SSR_b for a fully between-cluster effect, i.e., when $\rho_x = 1$

$$\mathrm{SSR_b} = 1 + r\rho_y,$$

Basic results for SSR_b

ρ_y	SSR_b result	N_{eff} vs N
$\rho_y = 0$	$\mathrm{SSR_b} = 1$	$N_{\mathrm{eff}} = N/1 \quad = N$
$\rho_y > 0$	$\mathrm{SSR_b} > 1$	$N_{\mathrm{eff}} = N/\mathrm{SSR_b} < N$

. I.e., when $\rho_x = 1$ and $\rho_y > 0$,

x has a between-cluster effect that will have <u>lower</u> precision versus within an alternative SRS design, all else being equal

Note. $\rho_y < 0$ is rare and not considered in this talk

Sample Size Ratios (SSR): 2-Level Sampling Designs

$\text{SSR} = 1 - \rho_y$

In a regression context…

. This SSR applies when *x* has a fully within-cluster effect on *y.*
. I.e., when *x* has exactly zero (**0**) between-cluster variation.
. In this circumstance, ρ_x takes its minimum value.

$$\rho_x = \frac{\mathbf{0} - \sigma^2_{\text{w/in\_cluster}}/r}{\mathbf{0} + \sigma^2_{\text{w/in\_cluster}}} = \frac{-1}{r}$$

When will *x* have a fully within-cluster effect?

. Often, a Level1 design variable w/ zero between-cluster variation, e.g.,
. RCT randomizing Level1 units w/ identical proportionate allocation across clusters
. *x* indicates scheduled assessment times (base, 6m, 12m)
. *x* indicates intra-cluster role/position, e.g., doctor versus patient
. Not always design vars. e.g., *x* holds deviations from cluster means

I label this SSR (i.e., $SSR = 1 - \rho_y$) as SSR_w (i.e., sub-'w' for 'within')

Summary Implications: SSR_w

. SSR_w for a fully within-cluster effect, i.e., when $\rho_x = -1/r$

$$\mathrm{SSR_w} = 1 - \rho_y$$

Basic results for SSR_w

ρ_y	SSR_w result	N_{eff} vs N
$\rho_y = 0$	$\mathrm{SSR_w} = 1$	$N_{\mathrm{eff}} = n_{\mathrm{L2}} \times n_{\mathrm{L1}} = N$
$\rho_y > 0$	$\mathrm{SSR_w} < 1$	$N_{\mathrm{eff}} = N/\mathrm{SSR_w} \;\; > N$

. I.e., when $\rho_x = -1/r$ and $\rho_y > 0$,

x has a within-cluster effect that will have <u>higher</u> precision vs SRS

SSR_w applies whenever $\rho_x = -1/r$

$\rho_x = -1/r$ when between-cluster *x* variation equals zero (exactly).

. Often, but not always, by design (e.g., assessment times) or analysis (e.g., deviation scores)

Sample Size Ratios (SSR): 2-Level Sampling Designs

So far, we've covered SSRs for

. fully between-cluster effects, i.e., when $\sigma^2_{x.\text{within}} = 0$, $\rho_x = 1$ and

. fully within-cluster effects, i.e., when $\sigma^2_{x.\text{between}} = 0$, $\rho_x = -1/r$

However, ρ_x values are not limited to $-1/r$ and 1

When $-1/r < \rho_x < 1$,

x can have both between- and within-cluster effects

You might expect $-1/r < \rho_x < 1$ when *x* is a(n)...

. design var. w/ some between-cluster variance (often $-1/r < \rho_x < 0$)

. observed random, eg, participant-reported, variable (often $0 \leq \rho_x < 1$)

Which SSR should be used when $-1/r < \rho_x < 1$?

The answer depends upon the regression modeling framework, i.e.,

. Survey Sampling (SS) versus

. Generalized Linear Mixed Models (GLMM) or GEE

Why is this the case?

Sample Size Ratios (SSR): 2-Level Sampling Designs

First, a side note. The formulation…

$$\rho_x = \frac{\sigma^2_{x.\text{between\_cluster}} - \sigma^2_{x.\text{w/in\_cluster}}/r}{\sigma^2_{x.\text{between\_cluster}} + \sigma^2_{x.\text{w/in\_cluster}}}$$

makes it clear that $\rho_x = 0$ when $\sigma^2_{x.\text{between\_cluster}} = \sigma^2_{x.\text{w/in\_cluster}}/r$

Thus, when $\rho_x = 0$, positive between-cluster variation is expected

Reminder…

In this talk, when $-1/r < \rho_x < 1$,
I assume equivalent between- and within-cluster effects of *x*

For reference, John Neuhaus describes modeling options that decompose between- and within-cluster effects.

Neuhaus, JM and Kalbfleisch, JD (1988). Between- and within-cluster covariate effects in the analysis of clustered data. *Biometrics*, 54, 638-645.

Neuhaus, JM (2001). Assessing change with longitudinal and clustered binary data. *Annual Review of Public Health*, 22, 115-118.

Sample Size Ratios (SSR): 2-Level Sampling Designs

SSR_{SS}: SSR for the Survey Sampling regression modeling framework

$$\text{SSR}_{\text{SS}} = 1 + r\rho_x\rho_y$$

Essentially, SSR_{SS} adds ρ_x to the product term of SSR_{b}

SSR_{SS} versus SSR_{b} and SSR_{w}

ρ_x	Result
1.0	$\text{SSR}_{\text{SS}} = \text{SSR}_{\text{b}} = 1 + r\rho_y$
$-1/r$	$1 - \rho_y = \text{SSR}_{\text{w}} = \text{SSR}_{\text{SS}}$
$-1/r < \rho_x < 1$	$1 - \rho_y < \text{SSR}_{\text{SS}} < 1 + r\rho_y$

SSR_{SS} was described by

Scott, AJ and Holt, D (1982). The Effect of Two-Stage Sampling on Ordinary Least Squares Methods. *Journal of the American Statistical Association*, 77(380), 848- 854.

(the SSR_{SS} label is mine)

Sample Size Ratios (SSR): 2-Level Sampling Designs

$\mathrm{SSR_{SS}} = 1 + r\rho_x\rho_y$ vs $\mathrm{SSR_b} = 1 + r\rho_y$ vs $\mathrm{SSR_w} = 1 - \rho_y$

Example $\mathrm{SSR_{SS}}$ results assuming SS model, $\rho_y = 0.10$, N=1000, r=10

ρ_x	ρ_y	x effect type	$N_{\mathrm{eff}\,SS}$= $1000/\mathrm{SSR_{SS}}$	$N_{\mathrm{eff}\,b}$= $1000/\mathrm{SSR_b}$	$N_{\mathrm{eff}\,w}$= $1000/\mathrm{SSR_w}$
1.00	0.10	btw-cluster	500	500	--
0.10	0.10	btw- & w/in-	909	--	--
0	0.10	btw- & w/in-	1000	--	--
-0.05	0.10	btw- & w/in-	1053	--	--
-0.10†	0.10	w/in-cluster	1111	--	1111

† $\rho_x = -1/r = -0.10$. '--': Inappropriate applications of $\mathrm{SSR_b}$ & $\mathrm{SSR_w}$

When applying the SS modeling framework

. Do not use $\mathrm{SSR_b}$ unless $\rho_x = 1$ (or, trivially, $\rho_y = 0$)
If $\rho_x < 1$, then use of $\mathrm{SSR_b}$ can underestimate N_{eff}, power

. Do not use $\mathrm{SSR_w}$ unless $\rho_x = -1/r$
if $-1/r < \rho_x$, then use of $\mathrm{SSR_w}$ can overestimate N_{eff}, power

Sample Size Ratios (SSR): 2-Level Sampling Designs

SSR$_{GE}$ for the GEE and GLMM modeling frameworks

$$\mathrm{SSR}_{\mathrm{GE}} = \frac{(1+r\rho_y)(1-\rho_y)}{1-\rho_y+r\rho_y(1-\rho_x)} = \frac{1-\rho_y+r\rho_y(1-\rho_y)}{1-\rho_y+r\rho_y(1-\rho_x)}$$

SSR$_{GE}$ is based upon the seminal (but underutilized) work of Basagaña, Liao, and Spiegelman (2011; BLS).

BLS were focused on power of longitudinal studies estimating the effects of a time-varying binary *x* variable.

BLS reported a SSR assuming compound symmetric (CS) ρ_x and ρ_y versus assuming $\rho_x = 1$ and CS ρ_y. See their Eq. 3.5.

SSR$_{GE}$ manipulates BLS Eq. 3.5 to reflect comparison of

. a clustered sampling design with CS ρ_x and ρ_y versus *N* independently sampled units (SRS).

. See Appendix A

Basagaña, X., Liao, X., and Spiegelman. D. (2011). Power and sample size calculations for longitudinal studies estimating a main effect of a time-varying exposure. *Statistical Methods in Medical Research*, 29, 181-192.

Sample Size Ratios (SSR): 2-Level Sampling Designs

SSR_{GE} for the GEE and GLMM modeling frameworks

$$\mathrm{SSR_{GE}} = \frac{1-\rho_y+r\rho_y(1-\rho_y)}{1-\rho_y+r\rho_y(1-\rho_x)}$$

SSR_{GE} versus SSR_b, SSR_w, and SSR_{SS}

ρ_x	Result
1.0	$\mathrm{SSR_{GE}} = \mathrm{SSR_{SS}} = \mathrm{SSR_b} = 1 + r\rho_y$
$-1/r$	$1-\rho_y = \mathrm{SSR_w} = \mathrm{SSR_{GE}} = \mathrm{SSR_{SS}}$
$-1/r < \rho_x < 1$	$1-\rho_y <$ $\mathrm{SSR_{GE}} < \mathrm{SSR_{SS}}$ $< 1 + r\rho_y$

Notes.

. $SSR_{GE} \leq SSR_{SS}$

. When $-1/r < \rho_x < 1$, $Neff_{GE} > Neff_{SS}$, i.e.,
superior power via the GEE/GLMM vs SS modeling framework

Sample Size Ratios (SSR): 2-Level Sampling Designs

$$\text{SSR}_{\text{GE}} = \frac{1-\rho_y+r\rho_y(1-\rho_y)}{1-\rho_y+r\rho_y(1-\rho_x)} \quad \text{vs} \quad \text{SSR}_{\text{SS}} = 1 + r\rho_x\rho_y$$

Results: *N*eff$_{\text{GE}}$ versus *N*eff$_{\text{SS}}$: *N*=1000, $\rho_y = 0.10,$ *r* =10

ρ_x	ρ_y	*x* effect type	*N*eff$_{\text{GE}}$= *1000/*SSR$_{\text{GE}}$	*N*eff$_{\text{SS}}$= 1000/SSR$_{\text{SS}}$	$\frac{Neff_{\text{GE}}}{Neff_{\text{SS}}}$
1.00	0.10	btw-cluster	500	500	=
0.10	0.10	btw- & w/in-	1000	909	+10% ‡
0	0.10	btw- & w/in-	1056	1000	+6% ‡
-0.05	0.10	btw- & w/in-	1083	1053	+3% ‡
-0.10†	0.10	w/in-cluster	1111	1111	=

note. † $\rho_x = -1/r = -0.10.$ ‡ GEE/GLMM has a power advantage over SS

The tabled results are not dramatic, but the GEE/GLMM advantage can be stark for some combinations of ρ_x, ρ_y, and r

When you have a choice, GEE/GLMM can be more efficient than SS

Sample Size Ratios (SSR): 2-Level Sampling Designs

$$\text{SSR}_{\text{GE}} = \frac{1-\rho_y+r\rho_y(1-\rho_y)}{1-\rho_y+r\rho_y(1-\rho_x)} \quad \text{vs} \quad \text{SSR}_{\text{SS}} = 1 + r\rho_x\rho_y$$

Results: *N*eff$_{\text{GE}}$ vs *N*eff$_{\text{SS}}$: *N*=1000, ρ_y=.50, *r* =1 (e.g., repeated measures)

ρ_x	ρ_y	*x* effect type	*N*eff$_{\text{GE}}$= *1000/*SSR$_{\text{GE}}$	*N*eff$_{\text{SS}}$= 1000/SSR$_{\text{SS}}$	$\frac{Neff_{\text{GE}}}{Neff_{\text{SS}}}$
1.00	0.50	btw-cluster	667	667	=
0.10	0.50	btw- & w/in-	1267	952	+33% ‡
0	0.50	btw- & w/in-	1333	1000	+33% ‡
-0.50	0.50	btw- & w/in-	1667	1333	+25% ‡
-1.00†	0.50	w/in-cluster	2000	2000	=

note. † $\rho_x = -1/r = -1.0$. ‡ GEE/GLMM has a power advantage over SS

The following slide compares SSR$_{\text{GE}}$ and SSR$_{\text{SS}}$ values for a range of ρ_x and ρ_y values and *r* =1

Expected $Neff_{GE}$ (**bold**) & $Neff_{SS}$ (regular): r=1 & N=1000

		ρ_y			
		0	0.3	0.6	0.9
ρ_x	$-1/r$	**1000** 1000	**1429** 1429	**2500** 2500	**10,000** 10,000
	0	**1000** 1000	**1099** 1000	**1563** 1000	**5263** 1000
	0.3	**1000** 1000	**1000** 917	**1281** 847	**3842** 787
	0.6	**1000** 1000	**901** 847	**1000** 735	**2421** 649
	0.9	**1000** 1000	**802** 787	**718** 649	**1000** 552
	1.0	**1000** 1000	**769** 769	**625** 625	**526** 526

- If ρ_y = 0 (green), $\rho_x = -1/r$ (purple), or ρ_x = 1 (pink), then $Neff_{GE}$=$Neff_{SS}$. Otherwise, $Neff_{GE}$ > $Neff_{SS}$
- If $\rho_x = \rho_y$, then $Neff_{GE}$ = N (yellow)
- If $\rho_x < \rho_y$, then $Neff_{GE}$ > N (purple & blue)
- If $\rho_x > \rho_y$, then $Neff_{GE}$ < N (orange & pink)

SSR inputs: linear versus logistic models

When planning for a logistic regression analysis there are some wrinkles

Population average (GEE, SS) versus unit-specific (GLMM)…

. estimates of *x* effects (b_x) as well as
model-predicted vs observed outcome probabilities, and

. estimates of ρ_x and ρ_y

Power analyses described in this talk require population average inputs

If inputting an effect estimate into power analysis for logistic regression,
choose a population average estimate
(e.g., based upon observed means or a GEE, ALR, or SS analysis)

Obtain ρ_x & ρ_y estimates from a GEE logistic model,
not a mixed logistic model

. GEE ICCs reflect intra-cluster homogeneity of observed values

. In contrast, mixed logistic model ICCs reflect underlying latent values

Sample Size Ratios (SSR): 2-Level Sampling Designs

Simulation comparing calculated versus simulated SSRs

Simulate 2-level data for *c*= 1 to 63 combinations of ρ_x and ρ_y values

. ρ_x ranging from $-1/r$ to 1.0

. ρ_y ranging from 0 to 0.9

Generate *i*= 1 to 10K replicate samples from each of 63 combinations

Fit regression model to each replicate sample assuming independent obs.

Save standard error estimates for fixed effect of *x*, $\hat{\sigma}_{\text{ind.c}i}$

Fit regression model to each replicate sample assuming clustered obs.

Save standard error estimates for fixed effect of *x*, $\hat{\sigma}_{\text{clus.c}i}$

A simulated SSR for combination *c* averages

$(\hat{\sigma}_{\text{clus.c}i}/\hat{\sigma}_{\text{ind.c}i})^2$ values across *i*=1 to 10K replicate samples

Compare simulated SSR to SSR_{GE} or SSR_{SS}, as appropriate

Simulated (bold) and Expected (plain) Values of SSR_{GE} assuming small clusters (r=1) and 63 Combinations of ρ_x and ρ_y: GEE Linear Regression Modeling Framework with exchangeable working correlation structure.

	ρ_y						
ρ_x	0	.05	.10	.25	.50	.75	.90
-1.0 $(-1/r)$	**1.000** 1.000	**0.950** 0.950	**0.900** 0.900	**0.749** 0.750	**0.499** 0.500	**0.249** 0.250	**0.099** 0.100
0	**0.998** 1.000	**0.996** 0.998	**0.988** 0.990	**0.936** 0.938	**0.749** 0.750	**0.437** 0.438	**0.188** 0.190
.05	**0.998** 1.000	**0.998** 1.000	**0.993** 0.995	**0.948** 0.949	**0.769** 0.769	**0.454** 0.455	**0.196** 0.199
.10	**0.998** 1.000	**1.000** 1.003	**0.998** 1.000	**0.960** 0.962	**0.789** 0.789	**0.471** 0.473	**0.206** 0.209
.25	**0.998** 1.000	**1.008** 1.010	**1.013** 1.015	**0.998** 1.000	**0.855** 0.857	**0.537** 0.538	**0.243** 0.245
.50	**0.998** 1.000	**1.021** 1.023	**1.040** 1.042	**1.069** 1.071	**0.998** 1.000	**0.699** 0.700	**0.341** 0.345
.75	**0.997** 1.000	**1.034** 1.036	**1.068** 1.070	**1.152** 1.154	**1.198** 1.200	**0.996** 1.000	**0.578** 0.585
.90	**0.997** 1.000	**1.043** 1.045	**1.086** 1.088	**1.207** 1.210	**1.362** 1.364	**1.343** 1.346	**0.991** 1.000
1.00	**0.997** 1.000	**1.048** 1.050	**1.098** 1.110	**1.249** 1.250	**1.499** 1.500	**1.750** 1.750	**1.901** 1.900

Note. Simulated data for each combination of ρ_x and ρ_y included m=500 clusters, n=2 units per cluster, N=1000, and 10K replicate samples. Simulated SSR_{GE} estimated from comparison of std errs estimated from GEE model versus cluster-naïve model. Cell shading codes for combinations of ρ_x and ρ_y values: Grey: E[SSR_{GE}]=1.0. Green: E[SSR_{GE}]<1.0; Orange: E[SSR_{GE}]>1.0

Simulated (bold) and Expected (plain) Values of SSR_{GE} assuming small clusters (r=1) and 63 Combinations of ρ_x and ρ_y: GLMM Linear Regression Modeling Framework with exchangeable working correlation structure.

	ρ_y						
ρ_x	0	.05	.10	.25	.50	.75	.90
-1.0 $(-1/r)$	**1.000** 1.000	**0.947** 0.950	**0.900** 0.900	**0.750** 0.750	**0.500** 0.500	**0.251** 0.250	**0.100** 0.100
0	**0.998** 1.000	**0.996** 0.998	**0.988** 0.990	**0.936** 0.938	**0.750** 0.750	**0.439** 0.438	**0.191** 0.190
.05	**0.998** 1.000	**0.998** 1.000	**0.993** 0.995	**0.948** 0.949	**0.769** 0.769	**0.456** 0.455	**0.200** 0.199
.10	**0.998** 1.000	**1.001** 1.003	**0.998** 1.000	**0.960** 0.962	**0.790** 0.789	**0.473** 0.473	**0.210** 0.209
.25	**0.998** 1.000	**1.009** 1.010	**1.013** 1.015	**0.998** 1.000	**0.856** 0.857	**0.539** 0.538	**0.247** 0.245
.50	**0.999** 1.000	**1.023** 1.023	**1.040** 1.042	**1.069** 1.071	**0.999** 1.000	**0.701** 0.700	**0.346** 0.345
.75	**0.999** 1.000	**1.037** 1.036	**1.069** 1.070	**1.153** 1.154	**1.199** 1.200	**0.998** 1.000	**0.586** 0.585
.90	**0.999** 1.000	**1.047** 1.045	**1.088** 1.088	**1.209** 1.210	**1.363** 1.364	**1.345** 1.346	**1.000** 1.000
1.00	**0.999** 1.000	**1.053** 1.050	**1.100** 1.100	**1.251** 1.250	**1.501** 1.500	**1.752** 1.750	**1.903** 1.900

Note. Simulated data for each combination of ρ_x and ρ_y included m=500 clusters, n=2 units per cluster, N=1000, and 10K replicate samples. Simulated SSR_{GE} estimated from comparison of std errs estimated from GLMM linear model versus cluster-naïve model. Cell shading codes for combinations of ρ_x and ρ_y values: Grey: E[SSR_{GE}]=1.0. Green: E[SSR_{GE}]<1.0; Orange: E[SSR_{GE}]>1.0

Simulated (bold) and Expected (plain) Values of SSR_{SS} assuming small clusters (r=1) and 63 Combinations of ρ_x and ρ_y: Survey Sampling Linear Regression Modeling Framework.

	ρ_y						
ρ_x	0	.05	.10	.25	.50	.75	.90
-1.0 $(-1/r)$	**1.001** 1.000	**0.951** 0.950	**0.901** 0.900	**0.750** 0.750	**0.501** 0.500	**0.251** 0.250	**0.100** 0.100
0	**1.000** 1.000	**0.999** 1.000	**1.001** 1.000	**0.999** 1.000	**0.999** 1.000	**1.001** 1.000	**1.000** 1.000
.05	**0.999** 1.000	**1.003** 1.003	**1.004** 1.005	**1.012** 1.013	**1.025** 1.025	**1.037** 1.038	**1.043** 1.045
.10	**1.001** 1.000	**1.004** 1.005	**1.010** 1.010	**1.025** 1.025	**1.050** 1.050	**1.073** 1.075	**1.088** 1.090
.25	**0.999** 1.000	**1.012** 1.013	**1.024** 1.025	**1.062** 1.063	**1.123** 1.125	**1.185** 1.188	**1.223** 1.225
.50	**0.999** 1.000	**1.023** 1.025	**1.049** 1.050	**1.123** 1.125	**1.248** 1.250	**1.372** 1.375	**1.447** 1.450
.75	**0.997** 1.000	**1.035** 1.038	**1.073** 1.075	**1.186** 1.188	**1.372** 1.375	**1.560** 1.563	**1.673** 1.675
.90	**0.996** 1.000	**1.043** 1.045	**1.088** 1.090	**1.222** 1.225	**1.447** 1.450	**1.672** 1.675	**1.809** 1.810
1.00	**0.997** 1.000	**1.047** 1.050	**1.096** 1.100	**1.246** 1.250	**1.496** 1.500	**1.748** 1.750	**1.900** 1.900

Note. Simulated data for each combination of ρ_x and ρ_y included m=500 clusters, n=2 units per cluster, N=1000, and 10K replicate samples. Simulated SSR_{SS} estimated from comparison of std errs estimated from SS linear model versus cluster-naïve model. Cell shading codes for combinations of ρ_x and ρ_y values: Grey: E[SSR_{GE}]=1.0. Green: E[SSR_{GE}]<1.0; Orange: E[SSR_{GE}]>1.0

Simulated (bold) and Expected (plain) Values of SSR$_{GE}$ assuming large clusters (r =50) and 63 Combinations of ρ_x and ρ_y: GEE Linear Regression Modeling Framework with exchangeable working correlation structure.

	ρ_y						
ρ_x	0	.05	.10	.25	.50	.75	.90
-.02 ($-1/r$)	**1.000** 1.000	**0.951** 0.950	**0.902** 0.900	**0.754** 0.750	**0.507** 0.500	**0.257** 0.250	**0.104** 0.100
0	**0.999** 1.000	**0.964** 0.964	**0.917** 0.915	**0.768** 0.764	**0.517** 0.510	**0.262** 0.255	**0.106** 0.102
.05	**0.998** 1.000	**0.999** 1.000	**0.956** 0.956	**0.806** 0.802	**0.544** 0.536	**0.275** 0.268	**0.112** 0.107
.10	**0.995** 1.000	**1.038** 1.039	**1.000** 1.000	**0.847** 0.844	**0.572** 0.565	**0.291** 0.283	**0.118** 0.113
.25	**0.990** 1.000	**1.171** 1.177	**1.158** 1.161	**1.001** 1.000	**0.684** 0.675	**0.348** 0.339	**0.141** 0.136
.50	**0.984** 1.000	**1.491** 4.511	**1.573** 1.588	**1.444** 1.446	**1.010** 1.000	**0.519** 0.507	**0.212** 0.204
.75	**0.978** 1.000	**2.065** 2.111	**2.464** 2.512	**2.590** 2.613	**1.936** 1.926	**1.025** 1.000	**0.421** 0.405
.90	**0.976** 1.000	**2.693** 2.771	**3.754** 3.857	**4.965** 5.063	**4.330** 4.333	**2.457** 2.406	**1.038** 1.000
1.00	**0.974** 1.000	**3.411** 3.500	**5.884** 6.000	**13.146** 13.500	**25.451** 26.000	**38.051** 38.500	**45.734** 46.000

Note. Simulated data for each combination of ρ_x and ρ_y included m=75 clusters, n=51 units per cluster, N=3825, and 10K replicate samples. Simulated SSR$_{GE}$ estimated from comparison of std errs estimated from GEE linear model versus cluster-naïve model. Cell shading codes for combinations of ρ_x and ρ_y values: Grey: E[SSR_{GE}]=1.0. Green: E[SSR_{GE}]<1.0; Orange: E[SSR_{GE}]>1.0

Simulation Results: Multivariate Logistic Models. Expected vs Simulated Sample Size Ratios and Statistical Power across the SS & GEE Modeling Frameworks: r=2, m=500, ρ_y=0.349 (population average), & 50K replicate samples.

x variables [a]		Survey Sampling Modeling Framework [b]					GEE Modeling Framework [c]					$\frac{Neff_{GE}}{Neff_{SS}}$
	ρ_x	SSR_{SS} expected [d]	SSR_{SS} simulated [e]	$Neff_{SS}$ expected [f]	Power expected [g]	Power simulated [h]	SSR_{GE} expected [d]	SSR_{GE} simulated [e]	$Neff_{GE}$ expected [f]	Power expected [g]	Power simulated [h]	
x_1	-1.00	0.651	0.653	1535	.715	.701	0.651	.647	1535	.715	.709	1.00
x_2	0	1.000	1.000	1000	.532	.511	0.878	.872	1138	.586	.572	1.14
x_3	0.25	1.087	1.086	920	.499	.477	0.962	.956	1039	.547	.530	1.13
x_4	0.50	1.174	1.173	852	.469	.453	1.064	1.059	934	.505	.496	1.10
x_5	0.75	1.261	1.259	793	.443	.427	1.189	1.187	841	.464	.452	1.06
x_6	1.00	1.349	1.346	742	.420	.400	1.349	1.351	742	.420	.403	1.00

[a] x_1-x_6 jointly uncorrelated & each unit-standardized; column 'ρ_x' reports the population intra-cluster correlation value of each *x* variable.

[b] Survey sampling modeling framework: fixed effect parameters estimated assuming independent observations and standard errors estimated via the Taylor series method ($\hat{\sigma}_{\text{Taylor}}$) assuming a compound-symmetric covariance structure.

[c] GEE Modeling framework: fixed effect parameters and model-based standard errors ($\hat{\sigma}_{\text{Mod}}$) estimated by a GEE linear model with compound symmetric working correlation structure.

[d] $SSR_{SS} = 1 + r\rho_x\rho_y$ and $SSR_{GE} = [1-\rho_y + r\rho_y(1-\rho_y)]/[1-\rho_y + r\rho_y(1-\rho_x)]$, where r=1, ρ_x values as tabled, and ρ_y=0.349.

[e] the quantity $\left(\hat{\sigma}_{\text{Tay}_i}/\hat{\sigma}_{\text{ind}_i}\right)^2$ or $\left(\hat{\sigma}_{\text{Mod}_i}/\hat{\sigma}_{\text{ind}_i}\right)^2$, as appropriate, averaged over i=1 to 50K replicate samples, where $\hat{\sigma}_{\text{ind}_i}$ denotes the corresponding standard error estimate assuming independent observations.

[f] $Neff_{SS}$=N÷SSR_{SS} and $Neff_{GE}$=N÷SSR_{GE}, where N=1000.

[g] Statistical power calculated by PASS assuming $Neff_{SS}$ or $Neff_{GE}$, as appropriate, two-tailed α=.05, fixed effect of *x* equal b_{pa}≈0.130 (population average), $P[y=1|x=0]$≈0.533 (population average), $P[y=1|\text{any } x=1]$≈0.565 (population average), and σ_x=1.0.

[h] Simulated statistical power represents the proportion of corresponding replicate-sample fixed effect parameter estimates with test *p*-value <.05.

<u>3-Level</u> Clustered Sampling Design: Simple Example

Example: Multisite RCT w/ sites (s), people (p), measures (m)

Levels and Sample sizes

. Sites (s) @ Level3: *n*3 = 30 sites

. People (p) @ Level2: *n*2 = 10 people per site. Units of randomization

. Measures (m) @ Level1: *n*1 = 2 assessment times per person

x variables

*x*3 is a site-level (L3) continuous covariate

. *x*3 has positive between site variation and zero within site variation

. $\rho_{x3} = 1.0.$ The intra-site correlation of *x*3

*x*2 is the person-level (L2) binary experimental group indicator

. Assume *x*2 has zero between-site variation

. $\rho_{x2} = -1/(10-1) = -.\overline{111}.$ The intra-site correlation of *x*2

*x*1 is the binary assessment time indicator at L1

. *x*1 has zero between-person variation

. $\rho_{x1} = -1/(2-1) = -1.$ The intra-person correlation of *x*1

3-Level Clustered Sampling Design: Simple Example

Example: Multisite RCT w/ sites (s), people (p), measures (m)

Common types of ρ_y estimates reported from a 3-level model

$$\rho_{y.\mathrm{s}} = \frac{\sigma^2_{y.\mathrm{s}}}{\sigma^2_{y.\mathrm{s}}+\sigma^2_{y.\mathrm{p}}+\sigma^2_{y.\mathrm{m}}}$$ Proportion of *y* var. attributable to sites

$$\rho_{y.\mathrm{s\&p}} = \frac{\sigma^2_{y.\mathrm{s}}+\sigma^2_{y.\mathrm{p}}}{\sigma^2_{y.\mathrm{s}}+\sigma^2_{y.\mathrm{p}}+\sigma^2_{y.\mathrm{m}}}$$ Prop. of *y* var attributable to sites & pts

$$\rho_{y.\mathrm{p}} = \frac{\sigma^2_{y.\mathrm{p}}}{\sigma^2_{y.\mathrm{s}}+\sigma^2_{y.\mathrm{p}}+\sigma^2_{y.\mathrm{m}}}$$ Prop. of *y* var. attributable to patients

Both $\rho_{y.\mathrm{s\&p}}$ and $\rho_{y.\mathrm{s}}$ may be described as 'ICC at Level2'

. When reading the literature, be clear whether $\rho_{y.\mathrm{s\&p}}$ or $\rho_{y.\mathrm{s}}$ is reported

. When reporting, be clear whether you are reporting $\rho_{y.\mathrm{s\&p}}$ or $\rho_{y.\mathrm{s}}$

3-Level Clustered Sampling Design: Simple Example

Example: Multisite RCT w/ sites (s), people (p), measures (m)

1. SSR for *x*3. A site variable w/ ρ_{x3}=1 has a fully a between-site effect

My initial, incorrect conjecture

$$\mathrm{SSR}_{x3} = \mathrm{SSR}_{\mathrm{b(s)}} = 1 + (10 \times 2 - 1)\rho_{y.s}$$

Correction

$$\mathrm{SSR}_{x3} = \mathrm{SSR}_{\mathrm{b(s)}} = 1 + (10 \times 2 - 1)\rho_{y.s②},$$

where $\rho_{y.s②}$ is the intra-site correlation of *y* estimated from a
2-level model that excludes Level2 cluster indicators (persons).
I.e., only top-level (site) clusters are modeled, i.e.,

$$\rho_{y.s②} = \frac{\sigma^2_{y.s②}}{\sigma^2_{y.s②} + \sigma^2_{y.m②}}$$

Proportion of *y* var. attrib. to sites: from 2-level model

3-Level Clustered Sampling Design: Simple Example

Example: Multisite RCT w/ sites (s), people (p), measures (m)

1. SSR for *x*3, which has a fully a between-site effect

. Given a 3-level data structure, when a model ignores the 2nd level,
the Level2 variation is distributed to both Level3 & Level1 (Moerbeek).

From a 3-level model: obtain prop. of variance explained at Levels 2 & 3

$\rho_{y.\mathrm{s}} = .05$ Proportion of *y* variance attributable to sites

$\rho_{y.\mathrm{p}} = .10$ Proportion of *y* variance attributable to people

$\rho_{y.\mathrm{p}} = .85$ Proportion of *y* variance attributable to measures

Given $\rho_{y.\mathrm{s}} = .05$, $\rho_{y.\mathrm{p}} = .10$, *n*2=10, and *n*1=2, estimate $\rho_{y.\mathrm{s}②}$…

$$\rho_{y.\mathrm{s}②} = \rho_{y.\mathrm{s}} + \rho_{y.\mathrm{p}}(n1 - 1)/(n2 \times n1 - 1) = .05 + .10/19 = .055263$$

$$\mathrm{SSR}_{x3} = 1 + (10 \times 2 - 1) \times 055263 = 2.05$$

Moerbeek, M (2004). The Consequence of Ignoring a Level of Nesting in Multilevel Analysis. *Multivariate Behavioral Research*, 39. 129-149.

3-Level Clustered Sampling Design: Simple Example

Example: Multisite RCT w/ sites (s), people (p), measures (m)

2. SSR for *x*2. A Level 2 variable w/ $\rho_{x2} = -1/r$
has a fully within-site/fully between-people effect

My initial, incorrect conjecture

$$\mathrm{SSR}_{x2} = \mathrm{SSR}_{\mathrm{w(s)}} \times \mathrm{SSR}_{\mathrm{b(p)}} = \left(1 - \rho_{y.\mathrm{s}}\right) \times \left[1 + (2-1)\rho_{y.\mathrm{p}}\right]$$

Correction

$$\mathrm{SSR}_{x2} = \mathrm{SSR}_{\mathrm{w(s)}} \times \mathrm{SSR}_{\mathrm{b(p)}} = \left(1 - \rho_{y.\mathrm{s}}\right) \times \left[1 + \left(\frac{2-1}{1-\rho_{y.\mathrm{s}}}\right)\rho_{y.\mathrm{p}}\right]$$

$$= \left(1 - \rho_{y.\mathrm{s}}\right) \times [1 + (2-1)]\rho_{y.\mathrm{p}\cancel{\mathrm{s}}} \; ,$$

where $\rho_{y.\mathrm{p}\cancel{\mathrm{s}}}$ is estimated via var. components from a 3-level model, i.e.,

$$\rho_{y.\mathrm{p}\cancel{\mathrm{s}}} = \frac{\sigma^2_{y.\mathrm{p}}}{\sigma^2_{y.\mathrm{p}} + \sigma^2_{y.\mathrm{m}}}$$

Prop. *y* var. attrib. to people, removing site variation

Given $\rho_{y.\mathrm{s}} = .05$, $\rho_{y.\mathrm{p}} = .10$, $\rho_{y.\mathrm{m}} = .85$, & *n*1=2.

$$\rho_{y.\mathrm{p}\cancel{\mathrm{s}}} = .10/(.10 + .85) = .10526$$

$$\mathrm{SSR}_{x2} = (1 - .05) \times [1 + (2-1)].10526 = 1.05$$

3-Level Clustered Sampling Design: Simple Example

Example: Multisite RCT w/ sites (s), people (p), measures (m)

3. <u>SSR for *x*1</u>, which has a fully within-site/fully within-people effect

<u>My initial, incorrect conjecture</u>

$$\mathrm{SSR}_{x1} = \mathrm{SSR}_{\mathrm{w(s)}} \times \mathrm{SSR}_{\mathrm{w(p)}} = \left(1 - \rho_{y.\mathrm{s}}\right)\left(1 - \rho_{y.\mathrm{p}}\right)$$

<u>Correction</u>

$$\mathrm{SSR}_{x1} = \mathrm{SSR}_{\mathrm{w(s)}} \times \mathrm{SSR}_{\mathrm{w(p)}} = \left(1 - \rho_{y.\mathrm{s}}\right)\left(1 - \rho_{y.\mathrm{p}\cancel{\mathrm{s}}}\right),$$

where $\rho_{y.\mathrm{p}\cancel{\mathrm{s}}}$ is estimated as described above

Given $\rho_{y.\mathrm{s}} = .05$ and $\rho_{y.\mathrm{p}\cancel{\mathrm{s}}} = .10526$

$$\mathrm{SSR}_{x1} = (1 - .05)(1 - .10526) = 0.85$$

3-Level Clustered Sampling Design: Simple Example

Simulated data from 3-Level Linear Mixed Model w/ 2K replicate samples

. *n*3=30 sites (L3), *n*2=10 subjects/site (L2), *n*1=2 assessments/subject
. *x*3: a normal random variate at Level3
. *x*2: a binary randomized group indicator at Level2
*x*1: a binary assessment time indicator at Level1

x(level)	ρ_x	*x* effect @		ρ_y	ρ_y adjustment		SSR_{GE}	
		L3	L2		$\rho_{y.s②}$	$\rho_{y.p\not{s}}$	expected	simulated
*x*3	1.0	btw	--	0.05	.05526	--	2.05	2.056
*x*2	$-1/r$	w/in	btw	0.10	--	.10526	1.05	1.053
*x*1	$-1/r$	w/in	w/in	0.85	--	--	0.85	0.854

SSR expected value calculations

$$\mathrm{SSR}_{\mathrm{GE}.x3} = \mathrm{SSR}_{\mathrm{b(s)}} = 1 + (10 \times 2 - 1) \times 0.05526 = 2.05$$

$$\mathrm{SSR}_{\mathrm{GE}.x2} = \mathrm{SSR}_{\mathrm{w(s)}} \cdot \mathrm{SSR}_{\mathrm{b(p)}} = (1 - 0.05) \times [1 + (2 - 1) \times 0.10526] = 1.50$$

$$\mathrm{SSR}_{\mathrm{GE}.x1} = \mathrm{SSR}_{\mathrm{w(s)}} \times \mathrm{SSR}_{\mathrm{w(p)}} = (1 - 0.05) \times (1 - 0.10526) = 0.85$$

SSR simulated values are relative size of std errs from LMM & Independence models, averaged over *i*=1 to 2K replicates: $\left(\hat{\sigma}_{\mathrm{LMM}_i} / \hat{\sigma}_{\mathrm{Ind}_i}\right)^2$

3-Level Clustered Sampling Design: Alternative Design A

Simulated data from 3-Level Linear Mixed Model w/ 2K replicate samples

. *n*3=30 sites (L3), *n*2=10 subjects/site (L2), *n*1=2 assessments/subject
. *x*3: a normal random variate at Level3
. *x*2: a binary randomized group indicator at Level2
*x*1: a binary assessment time indicator at Level1

x(level)	ρ_x	*x* effect @		ρ_y	ρ_y adjustment		$\mathrm{SSR_{GE}}$	
		L3	L2		$\rho_{y.s②}$	$\rho_{y.ps}$	expected	simulated
*x*3	1.0	btw	--	0.2	.2368	--	5.50	5.564
*x*2	$-1/r$	w/in	btw	0.7	--	.875	1.50	1.523
*x*1	$-1/r$	w/in	w/in	0.1	--	--	0.10	0.102

SSR expected value calculations

$$\mathrm{SSR_{GE.}}_{x3} = \mathrm{SSR_{b(s)}} = 1 + (10 \times 2 - 1) \times 0.2368 = 5.50$$

$$\mathrm{SSR_{GE.}}_{x2} = \mathrm{SSR_{w(s)}} \cdot \mathrm{SSR_{b(p)}} = (1 - 0.20) \times [1 + (2 - 1) \times 0.875] = 1.50$$

$$\mathrm{SSR_{GE.}}_{x1} = \mathrm{SSR_{w(s)}} \cdot \mathrm{SSR_{w(p)}} = (1 - 0.20) \times (1 - 0.875) = 0.10$$

SSR simulated values are relative size of std errs from LMM & Independence models, averaged over *i*=1 to 2K replicates: $\left(\hat{\sigma}_{\mathrm{LMM}_i} / \hat{\sigma}_{\mathrm{Ind}_i}\right)^2$

3-Level Clustered Sampling Design: Alternative Design B

Simulated data from 3-Level Linear Mixed Model w/ 2K replicate samples

. *n*3=30 sites (L3), *n*2=10 subjects/site (L2), *n*1=5 assessments/subject
. *x*3: a normal random variate at Level3
. *x*2: a binary randomized group indicator at Level2
*x*1: a uniform categorical assessment time indicator at Level1

x(level)	ρ_x	*x* effect @		ρ_y	ρ_y adjustment		SSR$_{GE}$	
		L3	L2		$\rho_{y.s②}$	$\rho_{y.p\text{s}}$	expected	simulated
*x*3	1.0	btw	--	0.05	.0908	--	5.45	5.479
*x*2	$-1/r$	w/in	btw	0.50	--	.5263	2.95	2.964
*x*1	$-1/r$	w/in	w/in	0.45	--	--	0.45	0.455

SSR expected value calculations

$$\text{SSR}_{\text{GE}.x3} = \text{SSR}_{\text{b(s)}} = 1 + (10 \times 5 - 1) \times 0.0908 = 5.45$$

$$\text{SSR}_{\text{GE}.x2} = \text{SSR}_{\text{w(s)}} \cdot \text{SSR}_{\text{b(p)}} = (1 - 0.05) \times [1 + (5 - 1) \times 0.5263] = 2.95$$

$$\text{SSR}_{\text{GE}.x1} = \text{SSR}_{\text{w(s)}} \cdot \text{SSR}_{\text{w(p)}} = (1 - 0.05) \times (1 - 0.5263) = 0.450$$

SSR simulated values are relative size of std errs from LMM & Independence models, averaged over *i*=1 to 2K replicates: $\left(\hat{\sigma}_{\text{LMM}_i} / \hat{\sigma}_{\text{Ind}_i}\right)^2$

Summary

Proper use of SSRs requires consideration of ρ_x

SSR$_b$ [$SSR_b = 1 + r\rho_y$] is well known, but perhaps over-applied.
Improper application of SSR$_b$ can lead to
substantially under-estimated power, w/ cost and ethical implications

When $-1/r < p_x < 1$, choose GEE/GLMM over SS modeling framework

All results reported here assumed
compound symmetric correlation structure of *x* and *y*

Power analysis for 3-level logistic models entails a few more wrinkles,
mostly regarding estimation of population average ρ_y values
(a future talk)

Some additional details and a quiz w/ answers included in the Appendix

Thank you

2-Level Clustered Sampling Designs: ρ_x: Quiz Yourself

Pre-post design

Goal: test pre-post mean *y* difference in a one-arm longitudinal trial

Will the pre-post comparison have a between- or within-cluster effect?

What is the value of ρ_x?

Clustered sample of teachers and their current students

Goal: regress students' SAT (*y*) onto teacher's years of experience (*x*)

Will teacher experience have a between- or within-cluster effect?

What is the value of ρ_x?

Multisite RCT. Randomization of patients within each site

. Is the intervention group effect a between- or within-cluster effect?

. What can be said about the expected ρ_x value?

Observational study with geographic cluster sampling

Goal: regress smoking status (*y*) onto respondent income (*x*)

Is income expected to have between- and/or within-cluster effects?

Appendix A: Derivation of SSR_{GE} from BLS (2011) Eq. 3.5.

For applications of the GLMM or GEE modeling frameworks,
BLS Eq. 3.5 relates effective sample sizes under assumptions of

. (i) CS correlation structures of *x* and *y* versus

. (ii) CS correlation structure of *y* with $\rho_x = 1$

$$\mathrm{SSR}_{\mathrm{BLS}} = \frac{Neff_{\rho_x,\rho_y}}{Neff_{\rho_{x=1},\rho_y}} = \frac{1-\rho_y+r\rho_y-r\rho_y\rho_x}{1-\rho_y}$$ [BLS Eq. 3.5]

A SSR that relates observed *N* to *N*eff assuming CS correlation structures of *x* and *y* can be derived from BLS Eq. 3.5, as follows.

$$\mathrm{SSR}_{\mathrm{GE}} = \frac{N}{Neff_{\rho x,\rho y}} = \frac{\frac{N}{Neff_{\rho x=1,\rho y}}}{\frac{Neff_{\rho x,\rho y}}{Neff_{\rho x=1,\rho y}}} = \frac{\mathrm{SSR_b}}{\mathrm{SSR_{BLS}}} =$$

$$= \frac{1+r\rho_y}{\frac{1-\rho_y+r\rho_y-r\rho_y\rho_x}{1-\rho_y}} = \frac{(1-\rho_y)(1+r\rho_y)}{1-\rho_y+r\rho_y-r\rho_y\rho_x}$$

$$= \frac{1-\rho_y+r\rho_y(1-\rho_y)}{1-\rho_y+r\rho_y(1-\rho_x)}$$

Appendix B. SSR$_{GE}$ results for selected values of ρ_x (2-level model)

$$SSR_{GE} = \frac{(1-\rho_y)(1+r\rho_y)}{1-\rho_y[1-r(1-\rho_x)]} = \frac{1-\rho_y+r\rho_y(1-\rho_y)}{1-\rho_y+r\rho_y(1-\rho_x)}$$

if $\rho_x = -1/r$, then

$$SSR_{GE} = \frac{(1-\rho_y)(1+r\rho_y)}{1-\rho_y+r\rho_y(1-\rho_x)} = \frac{SSR_w \times SSR_b}{SSR_b} = SSR_w$$

if $\rho_x = 1$, then

$$SSR_{GE} = \frac{(1-\rho_y)(1+r\rho_y)}{1-\rho_y+r\rho_y(1-\rho_x)} = \frac{SSR_w \times SSR_b}{SSR_w} = SSR_b$$

if $\rho_x = 0$, then

$$SSR_{GE} = \frac{(1-\rho_y)(1+r\rho_y)}{1-\rho_y+r\rho_y(1-\rho_x)} = \frac{SSR_w \times SSR_b}{1+(r-1)\rho_y} \cong 1 - {\rho_y}^{(n_1/r)}, \quad \text{(where } n_1 = r+1\text{)}$$

I.e., as $r \to \infty$, $\frac{SSR_b}{1+(r-1)\rho_y} \to 1$, and $SSR_{GE} \to SSR_w$

Additionally, if r=1 then $SSR_{GE} = \frac{(1-\rho_y)(1+\rho_y)}{1} = 1 - \rho_y^2$

2-Level Clustered Sampling Designs: ρ_x: Quiz Answers

Paired t-test

Goal: test pre-post mean *y* difference in a one-arm longitudinal trial

Here, respondents define the clusters and repeated measures (pre and post) are nested within respondents.

Will the pre-post comparison have a between- or within-cluster effect?

Pre-post indicator (*x*) is defined at Level1. Each cluster (person) has 2 assessments: one pre (*x*=0) and one post (*x*=1). There is zero between-cluster variation of *x* and positive within-cluster variation of *x*. Therefore, the pre-post comparison is a fully within-cluster (within-person) effect.

What is the value of ρ_x?

In this case $\rho_x = -1/(2-1) = -1.0$

2-Level Clustered Sampling Designs: ρ_x: Quiz Answers

Clustered sample of teachers and their current students

Goal: regress students' SAT (*y*) onto teacher's years of experience (*x*)

Teachers are clusters (Level2) and students (Level1) are nested within teachers

Will teacher experience have a between- or within-cluster effect?

Teacher experience is a Level2 variable. Therefore, teacher experience will have a fully between-cluster (between-teacher) effect.

What is the value of ρ_x?

Teacher experience will have positive between-cluster variation and zero within-cluster variation. Therefore, $\rho_x = 1$

2-Level Clustered Sampling Designs: ρ_x: Quiz Answers

Multisite RCT. Randomization of patients within each site
Sites are clusters (Level2) and patients (Level1) are nested in sites

. Is the intervention group effect a between- or within-cluster effect?
Intervention group indicator is a Level1 variable. Therefore, if the proportionate representation of Trt vs Ctrl assignment is identical across clusters, then the group effect will be a fully within-cluster effect. If the proportionate representation of group assignment varies slightly across site clusters, then a small of amount of between-cluster *x* variation will exist and the group comparison will be *mostly* a within-cluster effect.

. What can be said about the expected $\rho_x$value?
$\rho_x = -1/r$ if the proportionate allocation to Trt v Ctrl is identical across clusters.

If proportionate treatment assignment varies across clusters, then $\rho_x > -1/r$. Given sufficient cluster size and between-site variation, ρ_x could become positive.

2-Level Clustered Sampling Designs: ρ_x: Quiz Answers

Observational study with geographic cluster sampling

Goal: regress smoking status (*y*) onto respondent income (*x*)

Geographic areas are clusters (Level2) and respondents (Level1) are nested within clusters.

Is income expected to have between- and/or within-cluster effects?

We expect that respondent income (*x*) will have both between- and within-cluster variation. Therefore, we expect $0 < \rho_x < 1$, which means that income (*x*) can have both between- and within-cluster effects on smoking status.