Computing effect sizes in clustered trials

- In an experimental study, we are interested in the difference in performance between the treatment and control group
- In this case, we use the standardized mean difference, given by

\[ d = \frac{\bar{Y}_T - \bar{Y}_C}{S_p} \]

- Treatment group mean
- Control group mean
- Pooled sample standard deviation
Variance of the standardized mean difference

\[ S^2(d) = \frac{N^T + N^C}{N^T N^C} + \frac{d^2}{2(N^T + N^C)} \]

where \( N^T \) is the sample size for the treatment group, and \( N^C \) is the sample size for the control group.

**TREATMENT GROUP**

\( Y_1^T, Y_2^T, \ldots, Y_{N^T}^T \)

Overall Trt Mean

\[ \bar{Y}_g^T \]

\( S^2_{Trt} \)

**CONTROL GROUP**

\( Y_1^C, Y_2^C, \ldots, Y_{N^C}^C \)

Overall Cntl Mean

\[ \bar{Y}_g^C \]

\( S^2_{Cntl} \)

\( S^2_{pooled} \)
In cluster randomized trials, SMD more complex

- In cluster randomized trials, we have clusters such as schools or clinics randomized to treatment and control
- We have at least two means: mean performance for each cluster, and the overall group mean
- We also have several components of variance – the within-cluster variance, the variance between cluster means, and the total variance
- Next slide is an illustration

![Diagram showing cluster randomized trial analysis](image)
The problem in cluster randomized trials

• Which mean is the appropriate mean to represent the differences between the treatment and control groups?
  – The cluster means?
  – The overall treatment and control group means averaged across clusters?

• Which variance is appropriate to represent the differences between the treatment and control groups?
  – The within-cluster variance?
  – The variance among cluster means?
  – The total variance?

Three different components of variance

• The next slides outline the three components of variance that exist in a cluster randomized trial
• These components of variance correspond to the variances that we would estimate when analyzing a cluster randomized trial using hierarchical linear models
Pooled within-cluster sample variance

• One component of variance is within-cluster variation
• This is computed as the pooled differences between each observation and its cluster mean

\[
S^2_W = \frac{\sum_{i=1}^{m^T} \sum_{j=1}^{n} (\bar{Y}_{ij}^T - \bar{Y}_{ig}^T)^2 + \sum_{i=1}^{m^C} \sum_{j=1}^{n} (\bar{Y}_{ij}^C - \bar{Y}_{ig}^C)^2}{N - M}
\]

\(M\) is number of clusters: \(M = m^T + m^C\)

\(N\) is total sample size: \(N = n_{m^T} + n_{m^C} = N^T + N^C\) with equal cluster sample sizes, \(n\)
Pooled within-group (treatment and control) sample variance of the cluster means

- Another component of variation is the variance among the cluster means within the treatment and control groups.
- This is computed as the pooled difference between each cluster mean and its overall group (treatment and control) mean.

\[ S_B^2 = \frac{\sum_{i=1}^{m_T} (\bar{Y}_{ig}^T - \bar{Y}_{gg}^T)^2 + \sum_{i=1}^{m_C} (\bar{Y}_{ig}^C - \bar{Y}_{gg}^C)^2}{m_T + m_C - 2} \]
Total pooled within-group (treatment and control) variance

- The total variance is the variation among the individual observations and their group (treatment and control) overall means.

\[
S_T^2 = \frac{\sum_{i=1}^{m_T} \sum_{j=1}^{n} (\bar{Y}_{ij}^T - \bar{Y}_T^T)^2 + \sum_{i=1}^{m_C} \sum_{j=1}^{n} (\bar{Y}_{ij}^C - \bar{Y}_C^C)^2}{N - 2}
\]
Just a little distributional theory

• As usual, we will assume that each of our observations within the treatment and control group clusters are normally distributed about their cluster means, $\mu_T^i$ and $\mu_C^i$ with common within-cluster variance, $\sigma^2_W$, or

$$Y_{ij}^T \sim N(\mu_T^i, \sigma^2_W), \; i = 1,\ldots, m^T; \; j = 1,\ldots, n$$

$$Y_{ij}^C \sim N(\mu_C^i, \sigma^2_W), \; i = 1,\ldots, m^C; \; j = 1,\ldots, n$$

Just a little more distributional theory

• We also assume that our clusters are sampled from a population of clusters (making them random effects) so that the cluster means have a normal sampling distribution with means $\mu_T^*$ and $\mu_C^*$, and common variance, $\sigma^2_B$

$$\mu_T^i \sim N(\mu_T^*, \sigma^2_B), \; i = 1,\ldots, m^T$$

$$\mu_C^i \sim N(\mu_C^*, \sigma^2_B), \; i = 1,\ldots, m^C$$
The three components of variance are related:

\[ \sigma_T^2 = \sigma_W^2 + \sigma_B^2 \]

The intraclass correlation, \( \rho \)

- This parameter summarizes the relationship among the three variance components.
- We can think of \( \rho \) as the ratio of the between cluster variance and the total variance.
- \( \rho \) is given by

\[ \rho = \frac{\sigma_B^2}{\sigma_B^2 + \sigma_W^2} = \frac{\sigma_B^2}{\sigma_T^2} \]
Why is the intraclass correlation important?

• Later, we will see that we need the intraclass correlation to obtain the values of our standardized mean difference and its variance
• Also, if we know the intraclass correlation, we can obtain the value of one of the variances knowing the other two. For example:

\[
\sigma_W^2 = (1 - \rho) \sigma_B^2 / \rho
\]

\[
\sigma_W^2 = (1 - \rho) \sigma_T^2
\]

So far, we have

• Discussed the structure of the data
• Seen how to compute three different sampling variances: \(S^2_W, S^2_B,\) and \(S^2_T\)
• Discussed the underlying distributions of our clustered data
• Seen how the variance components in the distributions of our data are related to one another
• Next we will see how to estimate the variance components using our sampling variances
• (Yes, we are building up to the effect sizes…)
Estimator of $\sigma^2_W$

- As we might expect, we estimate $\sigma^2_W$ using our estimate of the within-cluster sampling variance, or,

\[
\hat{\sigma}^2_W = S^2_W
\]

Estimator of $\sigma^2_B$

- Unfortunately, the estimate of $\sigma^2_B$ is not as straightforward as for $\sigma^2_W$
- This is due to the fact that $S^2_B$ includes some of the within-cluster variance.
- The expected value for $S^2_B$ is

\[
\sigma^2_B + \frac{\sigma^2_W}{n}
\]
Thus, an estimator for $\sigma^2_B$ is

$$\hat{\sigma}^2_B = S^2_B - \frac{S^2_W}{n}$$

And, given prior information, an estimate of $\sigma^2_T$ is

$$\hat{\sigma}^2_T = S^2_B + \left( \frac{n-1}{n} \right) S^2_W$$
Now, we are ready for computing effect sizes

- There are 3 possibilities for computing an effect size in a clustered randomized trial
- These correspond to the 3 different components of variance we have been discussing
- The choice among them depends on the research design and the information reported in the other studies in our meta-analysis

The most common: $\delta_w$

- This effect size is comparable to the standardized mean difference computed from single site designs

$$\delta_w = \frac{\mu_g^T - \mu_g^C}{\sigma_w}$$
We estimate $\delta_w$ using $d_w$

$$d_w = \frac{\bar{Y}_T - \bar{Y}_C}{S_W}$$

Variance of $d_w$

- Variance of $d_w$ depends on the intraclass correlation, sample sizes, and $d_w$

$$V\{d_w\} = \left(\frac{N^T + N^C}{N^TN^T}\right) \left(\frac{1+(n-1)\rho}{1-\rho}\right) + \frac{d_w^2}{2(N-M)}$$

When $S_B$ is 0 (no between-group variation among the cluster means), the variance of $d_w$ is equal to the variance for the standardized mean difference in a single-site study.
A second estimator: $\delta_T$

- This estimator uses the total variation in the denominator
- We may compute this effect size when the other studies in our meta-analysis are also multi-site studies
- The general form is:

$$\delta_T = \frac{\mu^T - \mu^C}{\sigma_T}$$

There are two ways to compute $\delta_T$

- Method 1: $d_{T1}$
  - Used when the study reports $S^2_B$, the variance of the cluster means, and $S^2_W$, the pooled within-cluster variance
- Method 2: $d_{T2}$
  - Used when we the study reports the intraclass correlation, and $S^2_T$, the total variance
Estimator $d_{T1}$: When we have $S^2_B$ and $S^2_W$

$$d_{T1} = \frac{Y_T - Y_C}{\hat{\sigma}_T}, \text{ with}$$

$$\hat{\sigma}_T = \sqrt{S^2_B + \left(\frac{n-1}{n}\right)S^2_W}$$

The variance of $d_{T1}$ (a little messy)

$$V\{d_{T1}\} = \left(\frac{N^T + N^C}{N^T N^C}\right)(1 + (n-1)\rho) +$$

$$\left[\frac{1+(n-1)\rho}{2n^2(M-2)} + \frac{(n-1)^2(1-\rho)^2}{2n^2(N-M)}\right]d_{T1}^2$$
Estimator $d_{T2}$: When we have $S^2_T$ and $\rho$

$$d_{T2} = \left( \frac{\bar{Y}^T - \bar{Y}^C}{S_T} \right) \sqrt{1 - \frac{2(n-1)\rho}{N-2}}$$

Variance of $d_{T2}$ (also messy)

$$V\{d_{T2}\} = \left( \frac{N^T + N^C}{N^T N^C} \right) (1 + (n-1)\rho) +$$

$$d_{T2}^2 \left( \frac{(N-2)(1-\rho)^2 + n(N-2n)\rho^2 + 2(N-2n)\rho(1-\rho)}{2(N-2)[(N-2)-2(n-1)\rho]} \right)$$
A third estimator based on $S^2_B$

- We use this when the treatment effect is defined at the level of the clusters, or when the other studies in our meta-analysis have been analyzed using cluster means as the unit of analysis.
- Though it might not be of general interest, we can use this effect size to obtain other effect sizes of interest.
- There are two estimates of this effect.

Estimator $d_{B1}$: When we have $S^2_B$ and $S^2_W$

$$d_{B1} = \frac{\bar{Y}_T - \bar{Y}_C}{\hat{\sigma}_B}, \text{ with}$$

$$\hat{\sigma}_B = \sqrt{S^2_B - \frac{S^2_W}{n}}$$
The variance of $d_{B1}$ (a little messy)

$$V\{d_{B1}\} = \left(\frac{m^T + m^C}{m^T m^C}\right) \frac{1 + (n-1)\rho}{n\rho} + \left[\frac{\left[1 + (n-1)\rho\right]^2}{2(M-2)n^2\rho^2} + \frac{(1 - \rho)^2}{2(N-M)n^2\rho^2}\right]d_{B1}^2$$

Estimator $d_{B2}$: When we have $S_B$ and $\rho$

$$d_{B2} = \frac{\bar{Y}_T^T - \bar{Y}_C^T}{S_B} \sqrt{\frac{1 + (n-1)\rho}{n\rho}}$$
The variance of $d_{B2}$ (a little messy)

$$V \{d_{B2}\} = \left( \frac{m_T^T + m_C^T}{m_T m_C} \right) \frac{1 + (n-1)\rho}{n\rho} + \frac{(1 + (n-1)\rho) d_{B2}^2}{2(M-2)n\rho}$$

IMPORTANT: Can compute any $\delta$ from any other effect size and $\rho$

$$\delta_W = \delta_B \sqrt{\frac{\rho}{1 - \rho}} = \frac{\delta_T}{\sqrt{1 - \rho}}$$

$$\delta_T = \delta_B \sqrt{\rho} = \delta_W \sqrt{1 - \rho}$$
And, can compute the variance of the effect size using the same transformations

\[
\begin{align*}
Var \{ \delta_w \} &= Var \{ \delta_B \} \frac{\rho}{1 - \rho} = Var \{ \delta_T \} \\
Var \{ \delta_T \} &= Var \{ \delta_B \} \rho = Var \{ \delta_w \} (1 - \rho)
\end{align*}
\]

Where are we now?

- We have looked at the structure of the data
- We have examined the different components of variation
- We have defined estimators for the different components of variation
- We have outlined the computation of three different effect sizes, \( d_W \), \( d_T \), and \( d_B \) and their associated variances
But we all know it is not that simple: Some common problems

- We have unequal cluster sample sizes
  - There are formulas for unequal cluster sample sizes that are much more complex
  - We can assume equal cluster sizes since most studies attempt to use equal cluster sizes in the design
- We don’t know the intraclass correlation coefficient
  - We can estimate it from a number of sources listed in the References
  - Several researchers have provided estimates from other studies and from large-scale national samples

Let’s look at some examples

Impact of a Social-Emotional and Character Development Program on School-Level Indicators of Academic Achievement, Absenteeism, and Disciplinary Outcomes: A Matched-Pair, Cluster-Randomized, Controlled Trial

Frank Snyder and Brian Flay
Department of Public Health, Oregon State University, Corvallis, Oregon, USA

From the analysis section

School-Level Raw Means

Raw means for school-level academic achievement, absenteeism, suspensions, and retentions are presented in Figures 1 and 2, respectively. Overall, for the academic achievement outcomes, raw means for PA and control schools were statistically similar at baseline and demonstrated a clearly discernable divergence over time. State averages for academic achievement are shown for

Another clue that we have cluster means and sds reported

Matched Paired t Tests and Effect Sizes

The results of the matched paired t tests of difference scores and effect size calculations at posttest and 1-year post trial are presented in Table 4. At posttest,
The ICCs are also reported

Random-Intercept Growth Curve Models

The estimates for the intervention effect on academic achievement scores (random-intercept models) from baseline through posttest and 1-year post trial are presented in Table 5. At posttest, the intraclass correlation coefficient (ICC; expressed as the proportion of the total outcome variation that is attributable to differences among schools) for the unconditional means models (Singer & Willett, 2003) were .72, .67, .87, and .72 for math SAT and HCPS II and reading SAT and HCPS II, respectively. At 1-year post-trial, the

But I could not find sample sizes within clusters so looked for another report on the intervention

Use of a Social and Character Development Program to Prevent Substance Use, Violent Behaviors, and Sexual Activity Among Elementary-School Students in Hawaii

Michael W. Beets, PhD, MPH; Brian R. Ray, PhD; Samuel Vuchinich, PhD; Travis J. Snyder, MPH; Alan Acock, PhD; Him-Kul Li, MS; Kate Burns, PhD; Isaac J. Washburn, and Joseph Durkin, PhD

Am J Public Health | August 2009, Vol 99, No. 8

Sample

When students reached fifth grade (aged 10–11 years) they were asked to obtain active parental consent and to provide verbal assent to respond to 11 items asking about substance use (5 items), violent behavior (5 items), and sexual activity (1 item). This request garnered responses from 976 intervention students (50%) girls and 738 control students (50%).

Methods. We used a matched-pair, cluster-randomized, controlled design, with 10 intervention schools and 10 control schools. Fifth-graders (N = 1714) self-
So, what do we have in this example?

- We have $\rho$ for SAT Math of 0.72
- There are 10 matched-pairs of schools, with $N_T = 976$, and $N_C = 738$ (conservative common $n = 73$)
- We have the mean of the school (cluster) means and SDs for the treatment and control group at the level of school
We have all the information needed to compute $d_{B2}$

$$S_B^2 = \frac{(7.26)^2 + (7.48)^2}{2} = 54.33, \quad S_B = 7.37$$

$$d_{B2} = \frac{82.33 - 78.77}{7.37} \sqrt{\frac{1 + (73 - 1)0.72}{73(0.72)}}$$

$$= 0.48\sqrt{1.005} = 0.48$$

To compute $\text{Var}\{d_{B2}\}$

$$V\{d_{B2}\} = \left(\frac{10 + 10}{10 \times 10}\right) \frac{1 + (73 - 1)(0.72)}{(73)(0.72)}$$

$$+ \left(\frac{1 + (73 - 1)(0.72)}{2(20 - 2)(73)(0.72)}\right)(0.48)$$

$$= \left(\frac{20}{100}\right)\left(\frac{52.84}{52.56}\right) + \frac{(52.84)(0.48)}{1892.16} = 0.201 + 0.013$$

$$= 0.214$$
Say we want $d_w$ in our meta-analysis

$$d_w = d_B \sqrt{\frac{\rho}{1 - \rho}} = 0.48 \sqrt{\frac{0.72}{1 - 0.72}}$$

$$= 0.48(1.60) = 0.77$$

The variance of $d_w$ follows the same transformation

$$Var\{d_w\} = Var\{d_B\} \left( \frac{\rho}{\sqrt{1 - \rho}} \right)^2$$

$$= 0.214 \left( \frac{0.72}{(1 - 0.72)} \right)$$

$$= 0.214(2.57) = 0.55$$

$$SD\{d_w\} = 0.74$$
If instead, we want $d_T$

$$d_T = d_B \sqrt{\rho} = 0.48 \sqrt{0.72}$$

$$= 0.48(0.85) = 0.41$$

$$Var\{d_T\} = Var\{d_W\} \rho$$

$$= (0.214)(0.72) = 0.154$$

$$SD\{d_T\} = 0.392$$

Some observations on this example

- The intraclass correlation is large for this measure so we have a lot of between school variability
- Thus, it is not surprising that none of the effect sizes we calculated are significantly different from zero
- In this example, we have the intraclass correlation reported, but we may not be so lucky
Another example

Physical activity across the curriculum: year one process evaluation results
Cheryl A Gibson*,1, Bryan K Smith2, Katrina D DuBose3, J Leon Greene4, Bruce W Bailey5, Shannon L Williams6, Joseph J Ryan7, Kristin H Schmelzle5, Richard A Washburn2, Debra K Sullivan8, Matthew S Mayo8,10 and Joseph E Donnelly3


Physical activity across the curriculum: intervention and research design
PAAC is a cluster-randomized controlled, elementary school-based trial, involving 4905 children (2505 intervention, 2400 control) in 24 schools (14 intervention and 10 control). The study is being carried out in public ele-

We might have individual level info in Table 2

Table 2 displays the physical activity intensity levels reported from the SOFIT observations for intervention and control schools for Year 1. SOFIT observations were performed on a total of 4,515 students in the 2nd through 5th grades (intervention schools: 3,465 students; control schools: 1,050 students). Students in the intervention
What information do we have?

- \( m^T = 14 \) schools in the treatment group, \( m^C = 10 \) schools in the control group
- For treatment group, sample size average is \( n^T = 3429/14 = 244.9 \). For control group, sample size average is \( n^C = 1047/10 = 104.7 \)
- We have individual level means and standard deviations
- BUT
- We don’t have the intraclass correlation, or other estimates of the variance
To compute effect sizes, we need either 2 estimates of the variance, or 1 estimate and $\rho$

- We can only compute the estimate of $S^2_T$ from our 2 estimates of the individual level sds
- We will need to find an estimate for $\rho$ from another source
- The upper bound of the ICC for youth cohort studies with outcomes focused on diet is 0.0310

Since we have individual level stats:

\[
S^2_T = \frac{(3429 - 1)(0.52)^2 + (1047 - 1)(0.25)^2}{3429 + 1047 - 2} = 0.22
\]

\[
d_{T2} = \frac{3.40 - 2.16}{0.47} \sqrt{1 - \frac{2(104 - 1)0.031}{208 - 2}}
\]

\[
= 2.64 \sqrt{0.97} = 2.59
\]
The variance of $d_T$

$$V\{d_{r2}\} = \left( \frac{1456 + 1040}{1456 \times 1040} \right) \left( 1 + (104 - 1) \times 0.031 \right) +$$

$$2.59 \left( \frac{(2496 - 2)(1 - 0.031)^2 + 104(2496 - 2496 \times 2 \times 104) 	imes 0.031^2 + 2(2496 - 2) \times 104 \times 0.031 (1 - 0.031)}{2(2496 - 2)(2496 - 2) - 2(104 - 1) \times 0.031} \right)$$

$$= 0.007 + 2.59^2(0.0002) = 0.0083$$

$$SD\{d_{r2}\} = 0.091$$

If we want $d_w$

$$d_w = \frac{d_T}{\sqrt{1 - \rho}} = \frac{2.59}{\sqrt{1 - 0.031}} = 2.64$$

$$Var\{d_w\} = Var\{d_T\} \times \frac{1}{1 - \rho} = 0.0083 \times \frac{1}{1 - 0.031} = 0.0086$$

$$SD\{d_w\} = 0.093$$
What about other types of effect sizes, like odds-ratios?

- We do not yet have the same research for odds ratios
- What we can do is inflate the odds ratios from a clustered trial using recommendations from the Cochrane Handbook

Recommendations from the Cochrane Handbook

- We first compute the design effect given by
  \[ 1 + (n_{ave} - 1) \rho \]
  where \( n_{ave} \) is the average cluster size, and \( \rho \) is the intraclass correlation coefficient
- We divide the total sample sizes and number of events by the design effect to obtain the corrected numbers to compute the odds-ratio
- We adjust the variance by computing the variance of the log-odds ratio using the original sample sizes and dividing by the square root of the design effect
From the Positive Action program

<table>
<thead>
<tr>
<th>Sample size, no.</th>
<th>Ctrl Group, %</th>
<th>Intervention Group, %</th>
<th>OR^a (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substance use</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoked a cigarette</td>
<td>8.5</td>
<td>5.3</td>
<td>0.66 (0.30, 1.50)</td>
</tr>
<tr>
<td></td>
<td>6.7</td>
<td>2.7</td>
<td>0.38 (0.19, 0.76)</td>
</tr>
<tr>
<td></td>
<td>7.6</td>
<td>4.0</td>
<td></td>
</tr>
</tbody>
</table>

No smoking | Smoked
--- | ---
Intervention | 937 | 4.0% of 976 = 39
Control | 682 | 7.6% of 738 = 56

And the intraclass correlation coefficient

binary model) are presented in Table 2. The intraclass correlation coefficients for the unconditional models of student self-reports were 0.06, (0.05), and 0.28 for violent behaviors, substance use, and sexual activity, respec-
Adjusted values

- Average cluster size: \( \frac{(738+976)}{20} = 85.7 \)
- Design effect: \( 1 + (85.7 - 1) \times 0.05 = 5.235 \)

<table>
<thead>
<tr>
<th></th>
<th>No smoking</th>
<th>Smoked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention</td>
<td>937/5.235=178.99</td>
<td>4.0% of 976 = 39/5.235 = 7.45</td>
</tr>
<tr>
<td>Control</td>
<td>682/5.235=130.28</td>
<td>7.6% of 738 = 56/5.235 = 10.7</td>
</tr>
</tbody>
</table>

Variance for the log-odds ratio

- \( OR = \frac{(937 \times 56)}{(682 \times 39)} = 1.97 \)
- \( \ln OR = \ln(1.97) = 0.68 \)
- \( \text{SE}^2(\ln OR) = \frac{1}{937} + \frac{1}{56} + \frac{1}{682} + \frac{1}{39} = 0.046 \)
- \( \text{SE}(\ln OR) = 0.214 \)
- Adjusted \( \text{SE}(\ln OR) = \frac{0.214}{\sqrt{5.235}} = 0.094 \)
References


Cochrane Handbook:
http://www.cochrane-handbook.org/