Chapter 644

Meta-Analysis of Tests for Two Means using a Random-Effects Model in a Cluster-Randomized Design

Introduction

This procedure calculates power and number of studies for random-effects meta-analyses of studies that used cluster- randomized designs and that are summarized by the difference in two group means.

Meta-Analysis

Meta-analysis refers to the systematic assessment of a group of two or more individual studies (experiments or clinical trials). The studies must be similar so that their results may be numerically compared. For example, in this procedure all studies must have as their outcome two group means whose difference is used to compare their values.

By combining the results of several studies, their individual results may be summarized and assessed. The power of the combined test will be increased since it represents a larger sample.

Fixed-Effects versus Random-Effects Models

The *fixed-effects model*, which is assumed by another procedure, assumes that all studies have the same true effect size. That is, it assumes that all factors that might influence the effect size are the same so that the true effect size for all studies is the same. Therefore, only the *within study variability* needs to be considered since the effect size variation occurs because of random error.

The *random-effects model*, which is assumed by this procedure, assumes that the studies in the analysis are similar, but not to the extent that they all have the same true effect size. Instead, it is assumed that each study has its own true effect size and that the particular outcome of the study is a sample from a distribution with that mean. The random-effects model must not only account for the *within study variability*, but it must also account for the *between study variability*.

Technical Details

The Meta-Analysis of the Difference of Two Means

The power formulas used by this procedure are based on the following references: Hedges and Pigott (2001), Campbell and Walters (2014), Ahn, Heo, and Zhang (2015), Donner and Klar (2002), Donner, Piaggio, and Villar (2003), and Eldridge and Kerry (2012) page 204.

Cluster-randomized designs are those in which whole clusters of subjects (classes, hospitals, communities, etc.) are randomized into a treatment group or a control group. In this procedure, the analysis anticipates that a series of H studies will be meta-analyzed using the individual results of each study.

It is assumed that each study resulted in two group means, made up of K_i clusters of M_{ij} subjects each, was analyzed using summary statistics created from the individual subject data. The sample sizes for each study were adjusted to compensate for the cluster-randomization using a *design effect* (DE) factor.

A series of H studies are assumed to each have an effect size δ which is the standardized difference between two group means. The results of these individual studies may be combined into one overall statistical test that draws a more precise conclusion.

Test Statistic

Suppose the ith study yields the two group sample sizes $N1'$ and $N2'$ as well as a *standardized difference D* which is calculated as

$$D_i = \frac{\bar{X}_{1i} - \bar{X}_{2i}}{\hat{\sigma}_{D_i}}$$

Using a random-effects model, these results may be summarized and tested using the statistic

$$Z = \frac{\bar{D} - \delta_0}{\hat{\sigma}_{\bar{D}}}$$

where $\bar{D}$ is a weighted average of the individual D_i and $\hat{\sigma}_{\bar{D}}$ is a weighted estimate of its standard deviation that includes an estimate of both the within-study variation and the between-study variation provided by the random-effects model.

We assume that Z follows the standard normal distribution and so it can be used to test the null hypothesis $H_0: \delta = \delta_0$ at significance level of α.

Hypotheses

Three statistical hypotheses commonly used are

$$H_0: \delta \leq \delta_0 \quad \text{vs.} \quad H_1: \delta > \delta_0$$

$$H_0: \delta \geq \delta_0 \quad \text{vs.} \quad H_1: \delta < \delta_0$$

$$H_0: \delta = \delta_0 \quad \text{vs.} \quad H_1: \delta \neq \delta_0$$

Power Calculation

The power of a one-tailed test that $\delta = \delta_0$ is given by

$$Power = 1 - \Phi(c_\alpha - \lambda)$$

where δ is set by the alternative hypothesis, c_α is the $1 - \alpha$ quantile of the standard normal distribution, and

$$\lambda = \frac{\delta - \delta_0}{\sqrt{V_{\bar{D}}}}$$

$$V_F = \left(\frac{N_1' + N_2'}{N_1' N_2'} + \frac{\delta^2}{2(N_1' + N_2')} \right)$$

$$V_{\bar{D}} = \frac{1}{H}(V_F + RV_F)$$

Here V_F represents the within-study variance and R is the heterogeneity ratio. This ratio allows us to specify the between-study variance, V_B, as a multiple of V_F. Hedges and Pigott (2001) recommend using 0.333 for a *small* value, 0.667 for a *medium* value, and 1.0 for a *large* value. Setting R to 0 results in the fixed-effect result.

The power of a two-tailed test that $\delta = \delta_0$ is given by

$$Power = 1 - \Phi(c_{\alpha/2} - \lambda) + \Phi(-c_{\alpha/2} - \lambda)$$

These formulas can be used in search algorithms to determine the necessary number of studies required to achieve a desired level of power.

I^2 - Heterogeneity Index

Some authors state the amount of heterogeneity in the study using the heterogeneity index I^2 (also called the *degree of heterogeneity* or *inconsistency*) instead of R. This value is related to R using the relationship

$$R = \frac{I^2}{1 - I^2}$$

Effective Sample Sizes

To plan appropriate sample sizes for cluster-randomized designs, the regular sample sizes found from the number of clusters and the average cluster size must be adjusted. This adjusted sample size is named the effective sample size. The adjustment is made by dividing the original sample size by a *design effect* (DE). The *Design Effects DE_1 and DE_2* are adjustment factors that are used to compensate for the fact that cluster-randomization was used rather than simple random sampling. They are always greater than one.

The average *effective* sample size of each group across all studies is given by

$$N_1' = N_1/DE_1 \quad \text{and} \quad N_2' = N_2/DE_2$$

where

$$DE_1 = 1 + \big((COV^2 + 1)M_1 - 1\big)\rho$$

$$DE_2 = 1 + \big((COV^2 + 1)M_2 - 1\big)\rho$$

Here, M_1 and M_2 are the average cluster size across all studies of clusters in groups 1 and 2, respectively. *COV* is the coefficient of variation of the cluster sizes of all clusters, and ρ is the intracluster correlation coefficient (ICC) for all studies. This ICC may be interpreted as the correlation between any two observations within a particular cluster.

Example 1 – Calculating Sample Size

Researchers wish to determine a reasonable number of studies for a random-effects meta-analysis with the following characteristics: two-group, equal-allocation trials, two-sided significance level of 0.05, power of 0.9, average cluster size of all studies is 15, average number of clusters = 10, COV = 0.65, $\rho = 0.04$, $\delta_1 = 0.15$, and $\delta_0 = 0.0$. They want to compare the results of setting $I^2 = 0.25\ 0.5\ 0.75$.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For ....**Number of Studies**
Alternative Hypothesis ....**Two-Sided (H1: δ ≠ δ0)**
Power..**0.90**
Alpha...**0.05**
K1 (Average Number of Clusters).................**10**
M1 (Average Cluster Size)............................**15**
K2 (Average Number of Clusters).................**K1**
M2 (Average Cluster Size)............................**M1**
COV of Cluster Sizes....................................**0.65**
δ0 (Standardized Mean Difference|H0)..........**0.0**
δ1 (Standardized Mean Difference|H1)..........**0.15**
ρ (Intracluster Correlation, ICC)....................**0.04**
Calculate Between-Study Variance From ......**I² (Between-Study Var / Total-Study Var)**
I² (Between-Study Var / Total-Study Var).......**0.25 0.5 0.75**

Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Reports

Numeric Results
_____

Solve For: Number of Studies
Hypotheses: H0: δ = δ0   vs.   H1: δ ≠ δ0
Calculate Between-Study Variance From: I² (Between-Study Var / Total-Study Var)
_____

Power	Number of Studies H	Number of Clusters		Cluster Size		Sample Size		Standardized Mean Difference		Between-Study Variance Calculation		ICC	Alpha
		Average K	Total HK	Average M	COV	Average N	Total HN	δ|H0 δ0	δ|H1 δ1	R	I²	ρ	
0.91573	16	20	320		0.65	300	4800	0	0.15	0.33333	0.25	0.04	0.05
[Group 1]		10	160	15		150	2400						
[Group 2]		10	160	15		150	2400						
0.90434	23	20	460		0.65	300	6900	0	0.15	1.00000	0.50	0.04	0.05
[Group 1]		10	230	15		150	3450						
[Group 2]		10	230	15		150	3450						
0.90434	46	20	920		0.65	300	13800	0	0.15	3.00000	0.75	0.04	0.05
[Group 1]		10	460	15		150	6900						
[Group 2]		10	460	15		150	6900						
_____

Power   The probability of rejecting a false null hypothesis about the standardized mean difference (SMD). This power was
           calculated at the value of the SMD assumed by the alternative hypothesis, H1.
[Group]   The group to which the results on this line belong. Group 1 is the treatment group and group 2 is the control group.
           The first line shows the settings that are used for both groups.
H   The number of studies included in the meta-analysis.
K   The average number of clusters per study. These clusters are split between groups 1 and 2. The average cluster
           count in each group is shown below on the next two lines.
HK   The total number of clusters in the meta-analysis. These clusters are split between groups 1 and 2. The total cluster
           counts for each group are given below on the next two lines.
M   The average cluster size (i.e., number of subjects per cluster) in groups 1 and 2.
COV   The coefficient of variation of the cluster sizes of the clusters within the group. This value is used for both groups.
N   The average number of subjects per study. The total sample size is shown as the top row followed by N1 and N2.
HN   The total number of subjects represented in the meta-analysis. HN = H x N. The subtotals for groups 1 and 2 are
           shown below.
δ0   The value of the standardized mean difference (SMD) assumed by the null hypothesis, H0. Note that δ = [(μ1 - μ2) /
           σ(δ)] where σ(δ) is the standard deviation of the estimate of δ.
δ1   The value of the SMD assumed by the alternative hypothesis, H1. This is the value of δ at which the power is
           calculated.
R   The value of the heterogeneity ratio used to specify the between study variance. The formula is (Between Study
           Variance) = R x (Within Study Variance). The following values are often used: Small = 0.333, Medium = 0.667,
           Large = 1.000.
I²   The value of the heterogeneity index used to specify the Between Study variance. The formula is (Between Study
           Variance) = I² / (1 - I²) x (Within Study Variance). The following values are often used: Small = 0.25, Medium =
           0.5, Large = 0.75.
ρ   The intracluster correlation (ICC) between any two observations from within a cluster. This statistic has two
           common interpretations: (1) the proportion of variation related to the between-cluster variation or (2) the
           correlation between two subjects from the same cluster.
Alpha   The probability of rejecting a true null hypothesis.

Summary Statements

A collection of two-group cluster-randomized-design studies will be used to test whether the Group 1 mean is different from the Group 2 mean (H0: $\delta = 0$ versus H1: $\delta \neq 0$, $\delta = [\mu_1 - \mu_2] / \sigma[\delta]$). The meta-analysis comparison will be made using a two-sided random-effects-model Z-test, with a Type I error rate (α) of 0.05. In Group 1, the studies are assumed to have an average of 10 clusters with an average of 15 subjects per cluster (totaling 150 subjects), and in Group 2, an average of 10 clusters with an average of 15 subjects per cluster (totaling 150 subjects). The coefficient of variation of the cluster sizes within a group is assumed to be 0.65. The intracluster correlation coefficient of responses within a cluster is assumed to be 0.04. The between-to-total variance ratio (I^2), used to calculate the between-study variance, is assumed to be 0.25. To detect a standardized mean difference (δ) of 0.15 with 90% power, the number of needed studies will be 16.

References

Borenstein, M., Hedges, L.V., Higgins, J.P.T., and Rothstein, H.R. 2021. Introduction to Meta-Analysis, Second Edition. John Wiley and Sons.

Cooper, J., Hedges, L.V., Valentine, J.C. 2019. The Handbook of Research Synthesis and Meta-Analysis, Third Edition. Russell Sage Foundation.

Donner, A. and Klar, N. 2002. 'Issues in the meta-analysis of cluster randomized trials.' Statistics in Medicine, Vol. 21, pages 2971-2980.

Donner, A., Piaggio, G., and Villar, J. 2003. 'Meta-Analysis of Cluster Randomization Trials.' Evaluation & The Health Professions, Vol. 26, No. 3, pages 340-351.

Eldridge, S. and Kerry, S. 2012. A Practical Guide to Cluster-Randomised Trials in Health Services Research. John Wiley. NY, NY.

Harrer, M., Cuijpers, P., Furukawa, T.A., and Ebert, D.D. 2022. Doing Meta-Analysis with R: A Hands-On Guide. CRC Press.

Hedges, L.V. and Pigott, T.D. 2001. 'The Power of Statistical Tests in Meta-Analysis', Psychological Methods, Volume 6, No. 3, pages 203-217.

Higgins, J.P.T. and Thomas, J. 2019. Cochrane Handbook for Systematic Reviews of Interventions, Second Edition. Cochrane, Wiley, Blackwell.

Schmid, C. H., Stijnen, T., and White, I.R. 2021. Handbook of Meta-Analysis. CRC Press. New York.

This report shows the number of studies needed for the indicated scenarios.

Plots Section

Plots
_____

H vs I²

Pwr=0.9 A=0.05 K1=10 M1=15 K2=K1 M2=M1 COV=0.65 δ0=0 δ1=0.15 ρ=0.04 2-Sided Z Test
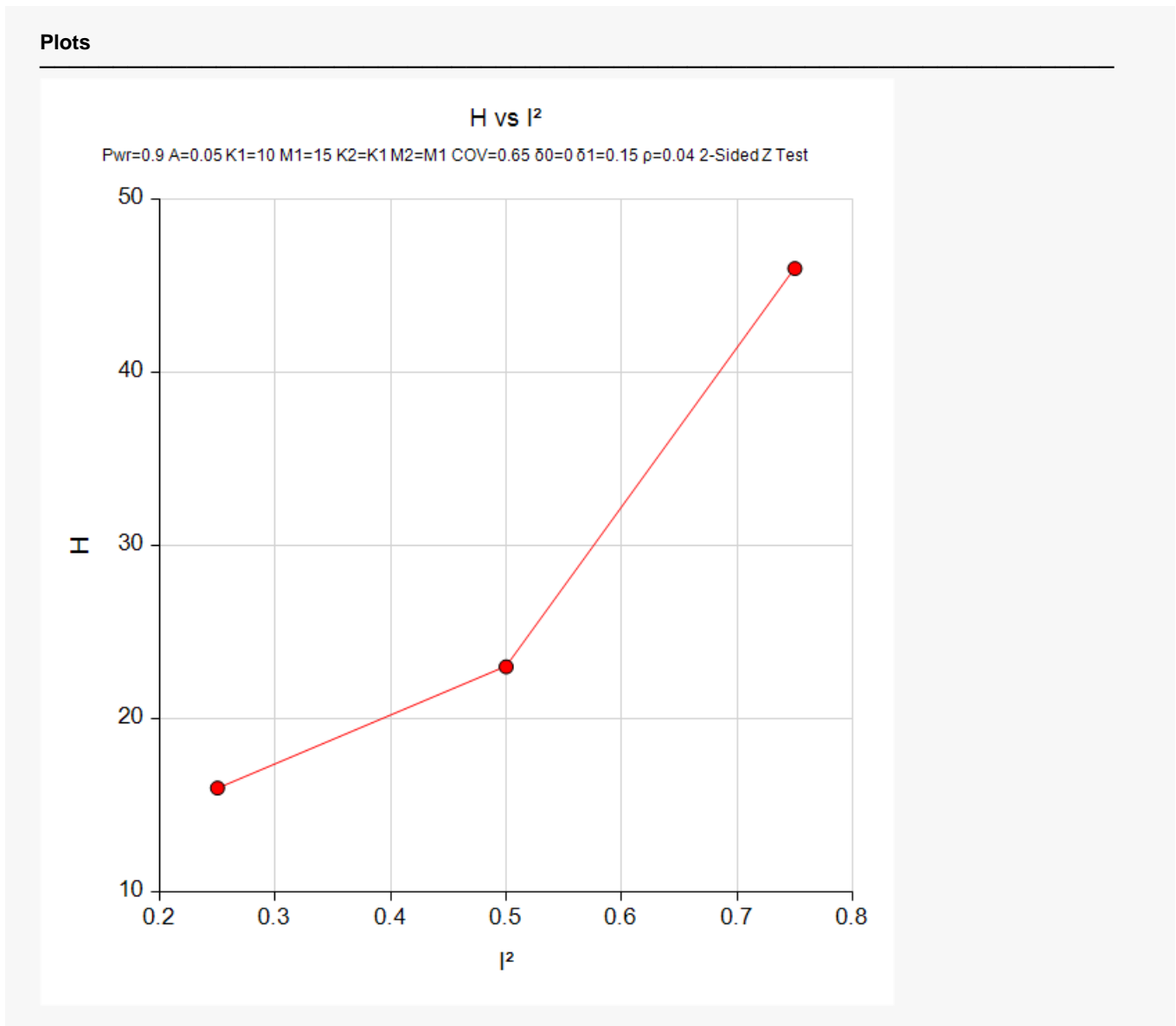
This plot represents the required number of studies for various values of the heterogeneity index.

Example 2 – Validation using Another Validated PASS Procedure

This procedure will be validated using results from a previous validated **PASS** procedure.

A fixed-effects meta-analysis is run with the following characteristics: two-group, equal-allocation trials, two-sided significance level of 0.05, number of studies = 23, average cluster size of all studies is 15, average number of clusters = 10, COV = 0.65, $\rho = 0.04$, $\delta_1 = 0.15$, $\delta_0 = 0.0$, and $I^2 = 0.5$. The power is calculated as 0.90434.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 2** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For ....**Power**
Alternative Hypothesis ....**Two-Sided (H1: δ ≠ δ0)**
Alpha..**0.05**
H (Number of Studies) **23**
K1 (Average Number of Clusters).................**10**
M1 (Average Cluster Size)............................**15**
K2 (Average Number of Clusters).................**K1**
M2 (Average Cluster Size)............................**M1**
COV of Cluster Sizes....................................**0.65**
δ0 (Standardized Mean Difference|H0)..........**0.0**
δ1 (Standardized Mean Difference|H1)..........**0.15**
ρ (Intracluster Correlation, ICC)...................**0.04**
Calculate Between-Study Variance From ......**I² (Between-Study Var / Total-Study Var)**
I² (Between-Study Var / Total-Study Var).......**0.5**

Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

Solve For: Power
Hypotheses: H0: δ = δ0   vs.   H1: δ ≠ δ0
Calculate Between-Study Variance From: I² (Between-Study Var / Total-Study Var)

| Power | Number of Studies H | Number of Clusters | | Cluster Size | | Sample Size | | Standardized Mean Difference | | Between-Study Variance Calculation | | ICC | |
		Average K	Total HK	Average M	COV	Average N	Total HN	δ|H0 δ0	δ|H1 δ1	R	I²	ρ	Alpha
0.90434	23	20	460		0.65	300	6900	0	0.15	1	0.5	0.04	0.05
[Group 1]		10	230	15		150	3450						
[Group 2]		10	230	15		150	3450						

PASS calculates a power of 0.90434.

This result can also be determined from the **PASS** procedure **Meta-Analysis of Tests for Two Means using a Random-Effects Model**. This procedure will be run with similar parameter settings: two-group, equal-allocation trials, two-sided significance level of 0.05, number of studies = 23, average cluster size of all studies is 15, average number of clusters = 10, COV = 0.65, $\rho = 0.04$, $\delta_1 = 0.15$, and $\delta_0 = 0.0$.

This procedure does not use cluster parameters, but it does calculate results for a meta-analysis of means using a random-effects model. Hence, we can use it to calculate the required power value by finding the appropriate sample size values. To accomplish this, we must calculate the value of DE by hand as follows.

The formula for DE is

$$DE = 1 + \{[COV^2 + 1]\bar{M} - 1\}\rho$$

Plugging in *COV* = 0.65, *M* = 15, and ρ = 0.05 gives

$$DE = 1 + \{[0.65^2 + 1]15 - 1\}(0.04) = 1.8135$$

Next, we use this value of DE to find adjusted sample sizes as follows

$$N_1' = \frac{N_1}{DE} = \frac{10(15)}{1.8135} = 82.7129859$$

Since the sample sizes are equal, this is the sample size used for group 2 as well.

The **Meta-Analysis of Tests for Two Means using a Random-Effects Model** procedure is set up as follows.

Design Tab
_____

Solve For ....**Power**

Alternative Hypothesis ....**Two-Sided (H1: δ ≠ δ0)**

Alpha...**0.05**

K (Number of Studies) **23**

N1 (Average Sample Size of Group 1)...........**82.7129859**

N2 (Average Sample Size of Group 2)...........**N1**

δ0 (Standardized Mean Difference|H0)..........**0.0**

δ1 (Standardized Mean Difference|H1)..........**0.15**

Calculate Between-Study Variance From ......**I² (Between-Study Var / Total-Study Var)**

I² (Between-Study Var / Total-Study Var).......**0.5**

This set of options generates the following report.

Numeric Results
_____

Solve For: Power
Hypotheses: H0: δ = δ0   vs.   H1: δ ≠ δ0
Between-Study Variance Input: I² (Between-Study / Total-Study Variances)
_____

	Average Sample Size Per Study			Number of		Standardized Mean Difference		Between-Study Variance Calculation		
Power	**Group 1 N1**	**Group 2 N2**	**Total N**	**Studies K**	**Subjects KN**	**δ|H0 δ0**	**δ|H1 δ1**	**R**	**I²**	**Alpha**
0.90434	82.7	82.7	165.4	23	3804.8	0	0.15	1	0.5	0.05
_____

The power of 0.90434 here matches the power given above which validates the procedure.