Chapter 489

Tests for Two Means in a Stepped-Wedge Cluster-Randomized Design

Introduction

A *parallel cluster (group) randomized design* is one in which whole units, or clusters, of subjects are randomized to the treatment or control group rather than the individual subjects in those clusters. However, the conclusions of the study concern individual subjects rather than the clusters. Examples of clusters are families, school classes, neighborhoods, hospitals, etc.

Cluster-randomized designs are often adopted when there is a high risk of contamination if cluster members were randomized individually. For example, it may be difficult for an instructor to use two methods of teaching individuals in the same class. The price of randomizing by clusters is a loss of efficiency--the number of subjects needed to obtain a certain level of precision in a cluster-randomized trial is usually much larger than the number needed when the subjects are randomized individually. Hence, the standard methods of sample size estimation cannot be used.

A *2×2 cluster-randomized cross-over design* as one in which each cluster receives both the treatment and control. The objective is to study difference between the two. Each cluster crosses over from the treatment group to the control group (or vice-versa). It is assumed that there is a washout period between applications during which the response returns back to its baseline value. If this does not occur, there is said to be a carry-over effect.

A *stepped-wedge cluster-randomized design* is similar to a cross-over design in that each cluster receives both the treatment and control over time. In a stepped-wedge design, however, the clusters switch or cross-over in one direction only (usually from the control group to the treatment group). Once a cluster is randomized to the treatment group, it continues to receive the treatment for the duration of the study. In a typical stepped-wedge design, all of the clusters are assigned to the control group at the first time point and then individual clusters are progressively randomized to the treatment group over time. The stepped-wedge design is particularly useful for cases where it is logistically impractical to apply a particular treatment to half of the clusters at the same time.

This procedure computes power and sample size for tests for the difference between two means in cross-sectional stepped-wedge cluster-randomized designs. In cross-sectional designs, different subjects are measured within each cluster at each point in time. No one subject is measured more than once. (This is not to be confused with cohort studies (i.e., repeated measures) where individuals are measured at each point in time. The methods in this procedure should not be used for cohort or repeated measures designs.)

Three Types of Cluster-Randomized Designs

In each design pattern matrix below, "0" represents the control and "1" represents the treatment.

Parallel		
		Time
		1
Cluster	1	1
	2	1
	3	1
	4	0
	5	0
	6	0

2×2 Cross-Over			
		Time	
		1	2
Cluster	1	1	0
	2	1	0
	3	1	0
	4	0	1
	5	0	1
	6	0	1

Stepped-Wedge								
		Time						
		1	2	3	4	5	6	7
Cluster	1	0	1	1	1	1	1	1
	2	0	0	1	1	1	1	1
	3	0	0	0	1	1	1	1
	4	0	0	0	0	1	1	1
	5	0	0	0	0	0	1	1
	6	0	0	0	0	0	0	1

Stepped-Wedge Cluster-Randomized Designs

There are two types of stepped-wedge designs that can be analyzed by this procedure: complete and incomplete (or custom).

Complete

In complete stepped-wedge designs, clusters sequentially switch from control to treatment in a balanced fashion over a fixed number of time periods (*T*). Once a cluster switches from control to treatment, the cluster continues to receive the treatment at each time point for the duration of the study. The number of clusters (*K*) is equal to number steps (*S*) multiplied by the number of clusters switching at each step (*R*), that is

$$K = S \times R.$$

The following stepped-wedge design pattern matrices illustrate complete designs (0 = Control, 1 = Treatment):

Complete Design K = 6 Clusters, S = 6 Steps, T = 7 Time Periods, R = 1 Switch per Step								
		Time						
		1	2	3	4	5	6	7
Cluster	1	0	1	1	1	1	1	1
	2	0	0	1	1	1	1	1
	3	0	0	0	1	1	1	1
	4	0	0	0	0	1	1	1
	5	0	0	0	0	0	1	1
	6	0	0	0	0	0	0	1

Complete Design K = 6 Clusters, S = 3 Steps, T = 4 Time Periods, R = 2 Switches per Step					
		Time			
		1	2	3	4
Cluster	1	0	1	1	1
	2	0	1	1	1
	3	0	0	1	1
	4	0	0	1	1
	5	0	0	0	1
	6	0	0	0	1

Incomplete (Custom)

In incomplete (or custom) stepped-wedge designs, there is no balance required. Furthermore, incomplete designs also allow for time periods at which no observations are made. The only requirement is that once a cluster switches from control to treatment, the cluster continues to receive the treatment if and when an observation is made.

When specifying incomplete designs using the number of clusters (*K*) and the number steps (*S*) (or the number of time periods (*T*)), **PASS** searches among all possible design configurations that satisfy the design constraints (*K*, *S*, *T*) to find an optimal design that achieves the highest power. This search is controlled by the Incomplete (Custom) Design Pattern Matrix Search Options on the Options tab.

The following stepped-wedge design pattern matrices illustrate incomplete designs (0 = Control, 1 = Treatment, • = No Observation):

Incomplete Design
K = 6 Clusters, S = 4 Steps, T = 5 Time Periods

Cluster	Time 1	2	3	4	5
1	0	1	1	1	1
2	0	1	1	1	1
3	0	0	1	1	1
4	0	0	0	1	1
5	0	0	0	1	1
6	0	0	0	0	1

Incomplete Design
K = 5 Clusters, S = 6 Steps, T = 7 Time Periods

Cluster	Time 1	2	3	4	5	6	7
1	0	•	1	1	1	1	1
2	0	0	•	1	1	1	1
3	0	0	0	•	1	1	1
4	0	0	0	0	•	1	1
5	0	0	0	0	0	•	1

Incomplete Design
K = 4 Clusters, S = 5 Steps, T = 6 Time Periods

Cluster	Time 1	2	3	4	5	6
1	0	1	1	•	•	•
2	•	0	1	1	•	•
3	•	•	0	1	1	•
4	•	•	•	0	1	1

Technical Details

This procedure is based on the results outlined in Hussey and Hughes (2007). In the technical discussions that follow, we will adopt the following notation:

K	Number of clusters
S	Number of steps in the stepped-wedge design
T	Number of time periods in the stepped-wedge design
M	Number of subjects (or items) per cluster
m	Number of subjects (or items) per cluster per time period
N	Total number of subjects (or items) from all clusters and all time periods combined
μ_1	The treatment group mean, assuming the alternative hypothesis
μ_2	The control, standard, reference, or baseline group mean
σ	The subject-to-subject standard deviation for both groups

The Linear Model

Linear mixed models are often used to model stepped-wedge cluster-randomized designs with time as a fixed factor at *T* levels and inter-cluster variation modeled as a random effect. As described in Hussey and Hughes (2007) and Hemming, Lilford, and Girling (2015) but using our notation and for a cross-sectional stepped-wedge design, let Y_{ikt} represent the response from individual *i* (*i* = 1, ..., *m*) in cluster *k* (*k* = 1, ..., *K*) at time *t* (*t* = 1, ..., *T*). The linear mixed model can then be written as

$$Y_{ikt} = X_{kt}\theta + \alpha_k + \beta_t + e_{ikt}$$

$$\alpha_k \sim N(0, \tau^2)$$

$$e_{ikt} \sim N(0, \sigma_w^2)$$

with

$$\mathrm{Var}(Y_{ikt}) = \sigma_y^2 = \tau^2 + \sigma_w^2,$$

where

X_{kt}	Indicator variable of the group assignment of cluster *k* at time *t*, with 0 = control and 1 = treatment
θ	Fixed treatment effect
α_k	Random effect for cluster *k*
β_t	Fixed effect for time *t*
e_{ikt}	Within-cluster error

The Treatment Effect

θ is the treatment effect and is equal to $D1$, the difference between the treatment mean, μ_1, and the control group mean, μ_2, such that

$$\theta = D1 = \mu_1 - \mu_2.$$

Between- and Within-Cluster Variation

From the model above, τ^2 is the between-cluster variance and σ_w^2 is the within-cluster variance. If τ^2 and σ_w^2 are known (as is generally assumed in power and sample size calculations), the model for the cell mean of cluster *k* at time *t* can then be written as

$$\bar{Y}_{kt} = X_{kt}\theta + \alpha_k + \beta_t + e_{kt}$$

$$\alpha_k \sim N(0, \tau^2)$$

$$e_{kt} \sim N\left(0, \frac{\sigma_w^2}{m}\right)$$

with

$$\mathrm{Var}(\bar{Y}_{kt}) = \tau^2 + \frac{\sigma_w^2}{m}.$$

For a complete stepped-wedge design, the *K*×*T* block-diagonal variance-covariance matrix, $\mathbf{V}$, of cell means has the form

$$\mathbf{V} = \begin{bmatrix} \mathbf{V_1} & 0 & \cdots & 0 \\ 0 & \mathbf{V}_{\boldsymbol{k}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{V}_{\boldsymbol{K}} \end{bmatrix},$$

where each *T*×*T* block matrix, $\mathbf{V}_{\boldsymbol{k}}$, represents a single cluster and has the form

$$\mathbf{V}_{\boldsymbol{k}} = \begin{bmatrix} \tau^2 + \frac{\sigma_w^2}{m} & \tau^2 & \cdots & \tau^2 \\ \tau^2 & \tau^2 + \frac{\sigma_w^2}{m} & \cdots & \tau^2 \\ \vdots & \vdots & \ddots & \vdots \\ \tau^2 & \tau^2 & \dots & \tau^2 + \frac{\sigma_w^2}{m} \end{bmatrix}.$$

An incomplete stepped-wedge design has a similar variance-covariance matrix structure with different dimensions depending on the incomplete design pattern matrix.

ICC (Intracluster Correlation Coefficient)

The correlation structure for the stepped-wedge cluster-randomized design is often characterized by the intracluster correlation coefficient, or ICC. The ICC may be interpreted as the correlation between any two observations in the same cluster. It may also be thought of as the proportion of the variation in the response that can be accounted for by the between-cluster variation. The ICC is calculated as

$$ICC = \frac{\tau^2}{\sigma_y^2} = \frac{\tau^2}{\tau^2 + \sigma_w^2}$$

and can be used along with the within-cluster variance, σ_w^2, to calculate the between-cluster variance, τ^2, as

$$\tau^2 = \frac{ICC \times \sigma_w^2}{1 - ICC}$$

COV (Coefficient of Variation of Outcomes)

The correlation structure for the stepped-wedge cluster-randomized design can also characterized by the coefficient of variation, or COV, of outcomes. (Note that this is not the COV of cluster sizes as is often referenced in conjunction with cluster-randomized designs.) If μ_2 is the control group mean, then the COV of outcomes for the control group is calculated as

$$COV = \frac{\tau}{\mu_2}$$

and can be used along with μ_2 to calculate τ^2 as

$$\tau^2 = COV^2 \times \mu_2^2.$$

Specifying the Total and Within-Cluster Variance

The standard deviation that is entered by the user, σ, may be considered to be either the total standard deviation, σ_y, such that the variance $\sigma^2 = \sigma_y^2$, or the within-cluster standard deviation, σ_w, such that $\sigma^2 = \sigma_w^2$. If σ is considered to be the total standard deviation, σ_y, then using the ICC, the between-cluster variance, τ^2, can be computed as

$$\tau^2 = ICC \times \sigma_y^2 = ICC \times \sigma^2,$$

and using the COV the between-cluster variance, τ^2, can be computed as

$$\tau^2 = COV^2 \times \mu_2^2.$$

and, finally, the within-cluster variance can be calculated as

$$\sigma_w^2 = \sigma_y^2 - \tau^2 = \sigma^2 - \tau^2.$$

Otherwise, if σ is considered to be the within-cluster standard deviation, σ_w, then

$$\sigma_w^2 = \sigma^2$$

and using the ICC, the between-cluster variance, τ^2, can be computed as

$$\tau^2 = \frac{ICC \times \sigma_w^2}{1 - ICC} = \frac{ICC \times \sigma^2}{1 - ICC}$$

and using the COV, the between-cluster variance, τ^2, can be computed as

$$\tau^2 = COV^2 \times \mu_2^2.$$

Test Statistic

Hussey and Hughes (2007) suggest that a statistical hypothesis test of H0: $\theta = \mu_1 - \mu_2 = 0$ vs. H1: $\theta = \mu_1 - \mu_2 \neq 0$ can be conducted using an asymptotic Wald test. The test statistic is

$$Z = \frac{\hat{\theta}}{\sqrt{\mathrm{Var}(\hat{\theta})}} = \frac{\hat{\mu}_1 - \hat{\mu}_2}{\sqrt{\mathrm{Var}(\hat{\theta})}}.$$

$\hat{\theta}$ is the estimated treatment effect from a weighted least-squares analysis. $\mathrm{Var}(\hat{\theta})$ is the appropriate element of $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$ from the weighted least-squares analysis. $\mathbf{X}$ is the design matrix (recall that X_{kt} is an indicator variable of the group assignment of cluster *k* at time *t*, e.g., 0 = control and 1 = treatment).

If X_{kt} contains only 0's and 1's, no missing cells, and *m* is equal for all clusters at all time points then $\mathrm{Var}(\hat{\theta})$ can be written in closed form as

$$\mathrm{Var}(\hat{\theta}) = \frac{K\left(\frac{\sigma_w^2}{m}\right)\left(\frac{\sigma_w^2}{m} + T\tau^2\right)}{\left(\frac{\sigma_w^2}{m}\right)(KU - W) + \tau^2(U^2 + KTU - TW - KV)}$$

where

$$U = \textstyle\sum_{kt} X_{kt}$$

$$V = \textstyle\sum_k \left(\sum_t X_{kt}\right)^2$$

$$W = \textstyle\sum_t \left(\sum_k X_{kt}\right)^2$$

Power Calculations

The power calculations available in this procedure for both complete and incomplete designs are based on the results outlined in Hussey and Hughes (2007). With τ^2 and σ_w^2 assumed to be known, the power for a two-sided Wald test is computed as

$$Power = \Phi\left(\frac{\theta_A}{\sqrt{\mathrm{Var}(\hat{\theta})}} - Z_{1-\alpha/2}\right)$$

$$= \Phi\left(\frac{\mu_1 - \mu_2}{\sqrt{\{(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\}}} - Z_{1-\alpha/2}\right)$$

where Φ is the cumulative standard Normal distribution, $\left\{\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\right\}$ is the appropriate element of the matrix, and $Z_{1-\alpha/2}$ is the $(1-\alpha/2)$ quantile of the standard Normal distribution function. θ_A is the value for θ under the alternative hypothesis, i.e., $\theta_A = \mu_1 - \mu_2$.

When solving for sample size (number of clusters), sample size (cluster size), or effect size, a search is conducted based on this power formula to find a solution that fits all required conditions.

A Note on Power Calculations for Incomplete Designs

When specifying an incomplete stepped-wedge design pattern using the number of clusters (*K*) and the number of steps (*S*) or the number of time periods (*T*), there are a variety of ways that the clusters can be arranged in an actual design and still meet the design criteria. **PASS** utilizes the following logic when creating a design pattern matrix using *K* and *S* or *T* and calculating power:

1. Sequentially assign *R* complete cluster sets to a design pattern sub-matrix, where *R* is the largest integer that satisfies (*R* × *S*) < *K* and ((*R* + 1) × *S*) > *K*. (Note: If (*R* × *S*) = *K*, then the design is complete and there are no “Extra” clusters no optimal design pattern search is required.)

2. Assign $J = K - (R \times S)$ extra clusters using the following options for the Design Pattern Assignment Type for Extra Clusters:

- **Balanced**

  The J "Extra" clusters are assigned in a balanced fashion. When one extra cluster is given a particular design pattern, it cannot be repeated.

Balanced Assignment

		Time				
		1	2	3	4	5
Cluster	1	0	1	1	1	1
	2	0	1	1	1	1
	3	0	0	1	1	1
	4	0	0	1	1	1
	5	0	0	0	1	1
	6	0	0	0	0	1
	7	0	0	0	0	1

- **Unbalanced**

  The J "Extra" clusters are assigned in a (potentially) unbalanced manner. When one extra cluster is given a particular design pattern, it may be repeated. This option results in maximum power but may result in a very unbalanced design.

Unbalanced Assignment

		Time				
		1	2	3	4	5
Cluster	1	0	1	1	1	1
	2	0	1	1	1	1
	3	0	1	1	1	1
	4	0	0	1	1	1
	5	0	0	0	1	1
	6	0	0	0	0	1
	7	0	0	0	0	1

- **Sequential**

  The *J* "Extra" clusters are assigned sequentially. The first extra cluster is given the same design pattern as the first already-assigned complete cluster, the second extra cluster is given the same design pattern as the second already-assigned complete cluster, and so on. When one extra cluster is given a particular design pattern, it cannot be repeated.

Sequential Assignment						
		Time				
		1	2	3	4	5
Cluster	1	0	1	1	1	1
	2	0	1	1	1	1
	3	0	0	1	1	1
	4	0	0	1	1	1
	5	0	0	0	1	1
	6	0	0	0	1	1
	7	0	0	0	0	1

3. Compute the power for all possible assignment combinations.
4. Return the design pattern matrix with the highest power.

Delayed Treatment Effect

Everything that has been discussed so far assumes that the full effect of the treatment occurs in the same time period in which the treatment is administered. This, however, might not always be the case. The full effect of the treatment may not be realized until several time periods after the treatment is applied. Hussey and Hughes (2007) propose that this situation can be modelled by simply altering the stepped-wedge design pattern matrix to include fractional numbers instead of 0's and 1's, where a fractional value indicates that the treatment is not fully efficacious at a particular time period.

The following is an example of a design pattern matrix exhibiting a delayed treatment effect where the treatment is 50% effective in the first time period, 80% effective in the second time period, and 100% effective by the third time period.

Design Pattern Matrix with a Delayed Treatment Effect								
		Time						
		1	2	3	4	5	6	7
Cluster	1	0	0.5	0.8	1	1	1	1
	2	0	0	0.5	0.8	1	1	1
	3	0	0	0	0.5	0.8	1	1
	4	0	0	0	0	0.5	0.8	1

The test statistic and power calculations are the same as for the case where the design pattern matrix contains all 0's and 1's, except that the closed form representation of $\mathrm{Var}(\hat{\theta})$ cannot be used. In this case, $\mathrm{Var}(\hat{\theta})$ must be calculated using matrix operations. The delay has the overall effect of reducing the power.

Example 1 – Finding Power of a Complete Design (Validation using Hemming and Taljaard (2016))

Hemming and Taljaard (2016) presents an example in section 3.2.3 where they calculate power for a complete stepped-wedge design when the number of clusters and the cluster sizes are both fixed. Their example computes power for a standardized effect size of 0.2, which equates to D1 = 0.2 and σ = 1 in **PASS**. What is the power with alpha = 0.05 for *K* = 10 clusters, *S* = 5 steps, *m* = 17 or 50 subjects per cluster per time period, and ICC values of 0.01 or 0.1?

In Table 5 on page 144, they present the following calculation results for the four ICC/*m* combinations:

ICC	m	Power
0.01	17	55%
0.1	17	49%
0.01	50	91%
0.1	50	90%

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 .... **Power**
Alternative Hypothesis .... **Two-Sided**
Alpha... **0.05**
Design Type... **Complete**
Design Parameter Entry Type................................ **Clusters (K) & Steps (S)**
K (Number of Clusters) **10**
S (Number of Steps) **5**
Cluster Size Entry Type .... **Subjects per Cluster per Time Period (m)**
m (Ave. Subjects per Cluster per Time Period)........ **17 50**
Input Type.. **Differences**
D1 (Difference = μ1 - μ2).. **0.2**
μ2 (Control Mean).. **0**
σ (Standard Deviation).. **1**
Use σ (Standard Deviation) as................................ **Total σ ($\sigma = \sigma y = \sqrt{[\tau^2 + \sigma w^2]}$)**
Between-Cluster Variability Input Type.................... **ICC**
ICC (Intracluster Correlation Coefficient) **0.01 0.1**

Output

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

Numeric Reports

Numeric Results

Solve For: Power
Design Type: Complete
Groups: 1 = Treatment, 2 = Control
Test Statistic: Wald Z-Test
Hypotheses: H0: μ1 - μ2 = 0 vs. H1: μ1 - μ2 ≠ 0

	Design Parameters			Number of Clusters	Cluster Size		Sample Size	Mean			Standard Deviation	Intracluster Correlation Coefficient	
Power	S	T	R	K	M	m	N	Treatment μ1	Control μ2	Difference D1	σ*	ICC	Alpha
0.54844	5	6	2	10	102	17	1020	0.2	0	0.2	1	0.01	0.05
0.48864	5	6	2	10	102	17	1020	0.2	0	0.2	1	0.10	0.05
0.91489	5	6	2	10	300	50	3000	0.2	0	0.2	1	0.01	0.05
0.90211	5	6	2	10	300	50	3000	0.2	0	0.2	1	0.10	0.05

* σ is considered to be the total standard deviation (σ = σy) in power computations.

Power	The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
S	The number of steps in the study design. S = T - 1.
T	The number of time periods in the study, including the baseline. T = S + 1
R	The number clusters switching from control to treatment at each step.
K	The total number of clusters to be randomized. K = S × R.
M	The average number of subjects per cluster. M = m × T.
m	The average number of subjects per cluster per time period. M = m × T.
N	The total sample size from all clusters and time periods combined. m = M / T.
μ1	The treatment mean assuming the alternative hypothesis.
μ2	The control, standard, reference, or baseline mean.
D1	The difference assuming the alternative hypothesis (H1). D1 = μ1 - μ2.
σ	The subject-to-subject standard deviation.
ICC	The intracluster correlation coefficient.
Alpha	The probability of rejecting a true null hypothesis.

Summary Statements

A complete stepped-wedge cluster-randomized design with 6 time periods (including the baseline) and 5 steps (with 2 clusters switching from control to treatment at each step) will be used to test whether the Group 1 (treatment) mean (μ1) is different from the Group 2 (control) mean (μ2) (H0: μ1 - μ2 = 0 versus H1: μ1 - μ2 ≠ 0). The comparison will be made using a two-sided Wald Z-test based on the mean difference (considered to be the total standard deviation) with a Type I error rate (α) of 0.05. The control group mean (μ2) is assumed to be 0. The subject-to-subject standard deviation (considered to be the total standard deviation) is assumed to be 1. The intracluster correlation coefficient is assumed to be 0.01. To detect a mean difference (μ1 - μ2) of 0.2 (or μ1 of 0.2) with 10 clusters with average cluster sizes of 102 subjects per cluster and an average of 17 subjects per cluster per time period (for a total sample size of 1020 subjects), the power is 0.54844.

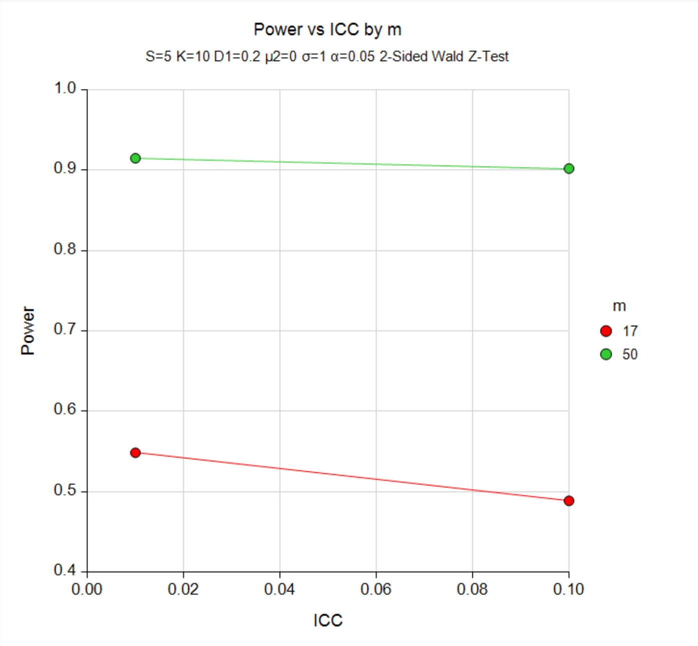
References

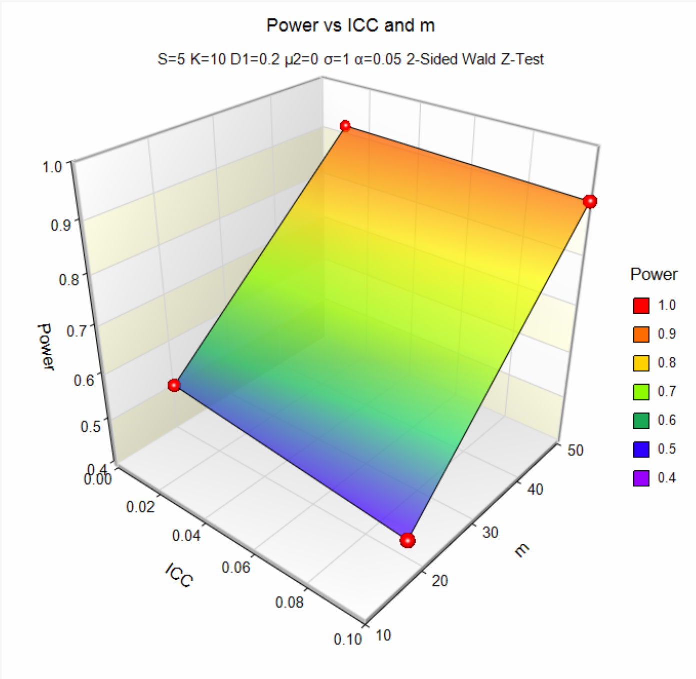
Hussey, M.A., and Hughes, J.P. 2007. 'Design and analysis of stepped wedge cluster randomized trials'. Contemporary Clinical Trials, Volume 28, pages 182-191.
Hemming, K., and Girling, A. 2014. 'A menu-driven facility for power and detectable-difference calculations in stepped-wedge cluster-randomized trials'. The Stata Journal, Volume 14, pages 363-380.
Hemming, K., Lilford, R., and Girling A.J. 2015. 'Stepped-wedge cluster randomised controlled trials: a generic framework including parallel and multiple-level designs'. Statistics in Medicine, Volume 34, pages 181-196.
Baio G., et al. 2015. 'Sample size calculation for a stepped wedge trial'. Trials, 16: 354.
Hemming, K., and Taljaard, M. 2016. 'Sample size calculations for stepped wedge and cluster randomised trials: a unified approach'. Journal of Clinical Epidemiology, Volume 69, pages 137-146.

This report shows the values of each of the parameters, one scenario per row. The power values calculated by **PASS** match those from Hemming and Taljaard (2016) exactly. The values from this table are shown in the plot below.

Plots Section

Plots





The values from the table are displayed on the plots. These plots give a quick look at the power that is achieved for the combinations of the *ICC* and *m*.

Design Details

Design 1 Details

Design Type: Complete

Power	Design Parameters S	T	R	Number of Clusters K	Cluster Size M	m	Mean Treatment μ1	Control μ2	Variance Total σy²	Between-Cluster τ²	Within-Cluster σw²	Intracluster Correlation Coefficient ICC	Coefficient of Variation COV
0.54844	5	6	2	10	102	50	0.2	0	1	0.01	0.99	0.01	

Design Pattern (Complete)

Cluster	T1	T2	T3	T4	T5	T6
1	0	1	1	1	1	1
2	0	1	1	1	1	1
3	0	0	1	1	1	1
4	0	0	1	1	1	1
5	0	0	0	1	1	1
6	0	0	0	1	1	1
7	0	0	0	0	1	1
8	0	0	0	0	1	1
9	0	0	0	0	0	1
10	0	0	0	0	0	1

(A Design Details report is given for each additional scenario)

This report gives the details about each design for which power was calculated. The design pattern matrix is also printed, showing exactly what design is being analyzed.

Procedure Input Settings

Procedure Input Settings

Autosave Inactive

Design Tab	
Solve For	Power
Alternative Hypothesis	Two-Sided
Alpha	0.05
Design Type	Complete
Design Parameter Entry Type	Clusters (K) & Steps (S)
K (Number of Clusters)	10
S (Number of Steps)	5
Cluster Size Entry Type	Subjects per Cluster per Time Period (m)
m (Ave. Subjects per Cluster per Time Period)	17 50
Input Type	Differences
D1 (Difference = μ1 - μ2)	0.2
μ2 (Control Mean)	0
σ (Standard Deviation)	1
Use σ (Standard Deviation) as	Total σ ($\sigma = \sigma y = \sqrt{[\tau^2 + \sigma w^2]}$)
Between-Cluster Variability Input Type	ICC
ICC (Intracluster Correlation Coefficient)	0.01 0.1
Options Tab	
Design Pattern Assignment Type for Extra Clusters	Balanced
Maximum Combinations for Incomplete Design Pattern Search	10000

This report displays the critical procedure input settings that were used to generate the report. This report is given primarily for the purpose of record-keeping.

Example 2 – Finding Power of an Incomplete Design (Validation using Hemming, Lilford, and Girling (2014))

Hemming, Lilford, and Girling (2014) presents an example on page 193-194 that calculates power for a staggered parallel cluster randomized design that can be analyzed as an incomplete stepped-wedge design. The study design aims to evaluate the effectiveness of a nutritional intervention program for children. The study involves nine control centers and nine treatment centers, each with m = 15 children measured per time interval for a total of 540 observations. The treatments are applied according to the following design pattern matrix:

	Time								
Cluster	0	1	2	3	4	5	6	7	8
1	0	•	•	•	•	•	0	•	•
2	0	•	•	•	•	•	0	•	•
3	0	•	•	•	•	•	0	•	•
4	0	•	•	•	•	•	1	•	•
5	0	•	•	•	•	•	1	•	•
6	0	•	•	•	•	•	1	•	•
7	•	0	•	•	•	•	•	0	•
8	•	0	•	•	•	•	•	0	•
9	•	0	•	•	•	•	•	0	•
10	•	0	•	•	•	•	•	1	•
11	•	0	•	•	•	•	•	1	•
12	•	0	•	•	•	•	•	1	•
13	•	•	0	•	•	•	•	•	0
14	•	•	0	•	•	•	•	•	0
15	•	•	0	•	•	•	•	•	0
16	•	•	0	•	•	•	•	•	1
17	•	•	0	•	•	•	•	•	1
18	•	•	0	•	•	•	•	•	1

The study aims to detect an increase in fruit and vegetable intake of one portion. The standard deviation is assumed to be 2.2. What is the power at an alpha level of 0.05 for ICC values from 0.05 to 0.5?

In Figure 10 on page 193, they present the following power calculation results for 7 different ICC values:

ICC	Power
0.05	0.891
0.1	0.870
0.15	0.869
0.2	0.877
0.3	0.905
0.4	0.937
0.5	0.967

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**
Alpha.. **0.05**
Design Type.. **Incomplete (Custom)**
Design Parameter Entry Type................................. **Custom Design Pattern Matrix**
Design Pattern Matrix Columns .... **ALL**
R (Number of Design Pattern Replicates)............... **1**
Cluster Size Entry Type .... **Subjects per Cluster per Time Period (m)**
m (Ave. Subjects per Cluster per Time Period)........ **15**
Input Type... **Differences**
D1 (Difference = μ1 - μ2) **1**
μ2 (Control Mean)... **1**
σ (Standard Deviation).. **2.2**
Use σ (Standard Deviation) as................................ **Total σ (σ = σy = √[τ² + σw²])**
Between-Cluster Variability Input Type................... **ICC**
ICC (Intracluster Correlation Coefficient) **0.05 0.1 0.15 0.2 0.3 0.4 0.5**

Input Spreadsheet Data

Row	C1	C2	C3	C4	C5	C6	C7	C8
1	0	.	.	.	.	0	.	.
2	0	.	.	.	.	0	.	.
3	0	.	.	.	.	0	.	.
4	0	.	.	.	.	1	.	.
5	0	.	.	.	.	1	.	.
6	0	.	.	.	.	1	.	.
7	.	0	.	.	.	.	0	.
8	.	0	.	.	.	.	0	.
9	.	0	.	.	.	.	0	.
10	.	0	.	.	.	.	1	.
11	.	0	.	.	.	.	1	.
12	.	0	.	.	.	.	1	.
13	.	.	0	.	.	.	.	0
14	.	.	0	.	.	.	.	0
15	.	.	0	.	.	.	.	0
16	.	.	0	.	.	.	.	1
17	.	.	0	.	.	.	.	1
18	.	.	0	.	.	.	.	1

Output

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

Numeric Results

Solve For: Power
Design Type: Incomplete (Custom)
Groups: 1 = Treatment, 2 = Control
Test Statistic: Wald Z-Test
Hypotheses: H0: μ1 - μ2 = 0 vs. H1: μ1 - μ2 ≠ 0

	Design Parameters			Number of Clusters	Cluster Size		Sample Size	Mean		Difference	Standard Deviation	Intracluster Correlation Coefficient	
Power	S	T	R	K	M	m	N	Treatment μ1	Control μ2	D1	σ*	ICC	Alpha
0.89096	7	8	1	18	30	15	540	2	1	1	2.2	0.05	0.05
0.87035	7	8	1	18	30	15	540	2	1	1	2.2	0.10	0.05
0.86936	7	8	1	18	30	15	540	2	1	1	2.2	0.15	0.05
0.87723	7	8	1	18	30	15	540	2	1	1	2.2	0.20	0.05
0.90459	7	8	1	18	30	15	540	2	1	1	2.2	0.30	0.05
0.93691	7	8	1	18	30	15	540	2	1	1	2.2	0.40	0.05
0.96669	7	8	1	18	30	15	540	2	1	1	2.2	0.50	0.05

* σ is considered to be the total standard deviation (σ = σy) in power computations.

The power values calculated by **PASS** match those from Hemming, Lilford, and Girling (2014) exactly.

Example 3 – Finding Sample Size (Number of Clusters) for an Incomplete Design (Validation using Hemming and Taljaard (2016))

Hemming and Taljaard (2016) presents an example in section 3.2.1 where they calculate the number of clusters for an incomplete stepped-wedge design when the cluster size and number of steps are both fixed. Their example computes number of clusters for a standardized effect size of 0.2, which equates to D1 = 0.2 and σ = 1 in **PASS**. What is the required number of clusters for 80% power with alpha = 0.05, *S* = 2 or 9 steps, *m* = 10 subjects per cluster per time period (*M* = 30 or 100), and ICC values of 0.01 or 0.25?

In Table 3 on page 143, they present the following calculation results for four *M*/*S*/*ICC* combinations:

M	S	ICC	K
30	2	0.01	85
30	2	0.25	85
100	9	0.01	18
100	9	0.25	18

Note that **PASS** does not solve for number of clusters using the design effect adjustment that Hemming and Taljaard (2016) uses. Instead **PASS** performs an exhaustive search of all design pattern matrix models to compute the required number of clusters.

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 3** 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	**Sample Size [K (Number of Clusters)]**
Alternative Hypothesis	**Two-Sided**
Power	**0.80**
Alpha	**0.05**
Design Type	**Incomplete (Custom)**
Design Constraint Entry Type	**Fixed Number of Steps (S)**
S (Number of Steps)	**2 9**
Cluster Size Entry Type	**Subjects per Cluster per Time Period (m)**
m (Ave. Subjects per Cluster per Time Period)	**10**
Input Type	**Differences**
D1 (Difference = μ1 - μ2)	**0.2**
μ2 (Control Mean)	**0**
σ (Standard Deviation)	**1**
Use σ (Standard Deviation) as	**Total σ ($\sigma = \sigma_y = \sqrt{[\tau^2 + \sigma_w^2]}$)**
Between-Cluster Variability Input Type	**ICC**
ICC (Intracluster Correlation Coefficient)	**0.01 0.25**

Output

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

Numeric Results

Solve For: Sample Size [K (Number of Clusters)]
Design Type: Incomplete (Custom)
Groups: 1 = Treatment, 2 = Control
Test Statistic: Wald Z-Test
Hypotheses: H0: μ1 - μ2 = 0 vs. H1: μ1 - μ2 ≠ 0

	Design Parameters		Number of Clusters	Cluster Size		Sample Size	Mean			Standard Deviation	Intracluster Correlation Coefficient	
Power	S	T	K	M	m	N	Treatment μ1	Control μ2	Difference D1	σ*	ICC	Alpha
0.80349	2	3	85	30	10	2550	0.2	0	0.2	1	0.01	0.05
0.80244	2	3	85	30	10	2550	0.2	0	0.2	1	0.25	0.05
0.80845	9	10	17	100	10	1700	0.2	0	0.2	1	0.01	0.05
0.80785	9	10	18	100	10	1800	0.2	0	0.2	1	0.25	0.05

* σ is considered to be the total standard deviation (σ = σy) in power computations.

PASS computed K values of 85, 85, 17, and 18, which match those given in Table 3 of Hemming and Taljaard (2016). There is a difference, however, when ICC = 0.01 and S = 9. This difference may be due to rounding, but it is important to note that **PASS** uses a different method to arrive at the required number of clusters. **PASS** uses the power along with a search for the optimal (balanced) design pattern matrix to find the sample size, not the design effect adjustment that Hemming and Taljaard (2016) uses. One problem with the design effect adjustment method is that it does not tell you what final model is being used and the actual power that is achieved. **PASS** always achieves the desired level of power in the result. The power value of 0.80845, for example, indicates that K = 17 with ICC = 0.01 and S = 9 does achieve the desired power and is, therefore, the correct balanced-design solution.

Example 4 – Finding Sample Size (Number of Clusters) for an Incomplete Design (Validation using Baio et al. (2015))

Baio et al. (2015) presents an example in Table 1 of various sample size calculations for the case where the baseline mean, μ2, is 0.3, the difference, D1, is -0.3785, and the total standard deviation, σ_y, is 1.55. They assume *m* = 20 individuals per cluster per time interval, *T* = 6 time intervals, and compute the required sample sizes for 80% power with ICC values from 0 to 0.5. The alpha level is 0.05. In this example we'll need to search among incomplete designs to match their results.

Using a design effect adjustment method for computing number of clusters, they report *K* values of 9, 12, 11, 10, 9, and 7 for ICC values of 0, 0.1, 0.2, 0.3, 0.4, and 0.5, respectively.

Note that **PASS** does not solve for number of clusters using the same method that Baio et al. (2015) uses. Instead **PASS** performs an exhaustive search of all design pattern matrix models to compute the required number of clusters.

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 4** 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	**Sample Size [K (Number of Clusters)]**
Alternative Hypothesis	**Two-Sided**
Power	**0.80**
Alpha	**0.05**
Design Type	**Incomplete (Custom)**
Design Constraint Entry Type	**Fixed Number of Time Periods (T)**
T (Number of Time Periods including Baseline)	**6**
Cluster Size Entry Type	**Subjects per Cluster per Time Period (m)**
m (Ave. Subjects per Cluster per Time Period)	**20**
Input Type	**Differences**
D1 (Difference = μ1 - μ2)	**-0.3785**
μ2 (Control Mean)	**0.3**
σ (Standard Deviation)	**1.55**
Use σ (Standard Deviation) as	**Total σ (σ = σy = √[τ² + σw²])**
Between-Cluster Variability Input Type	**ICC**
ICC (Intracluster Correlation Coefficient)	**0 to 0.5 by 0.1**

Reports Tab

Means and Difference Decimals	**4**

Output

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

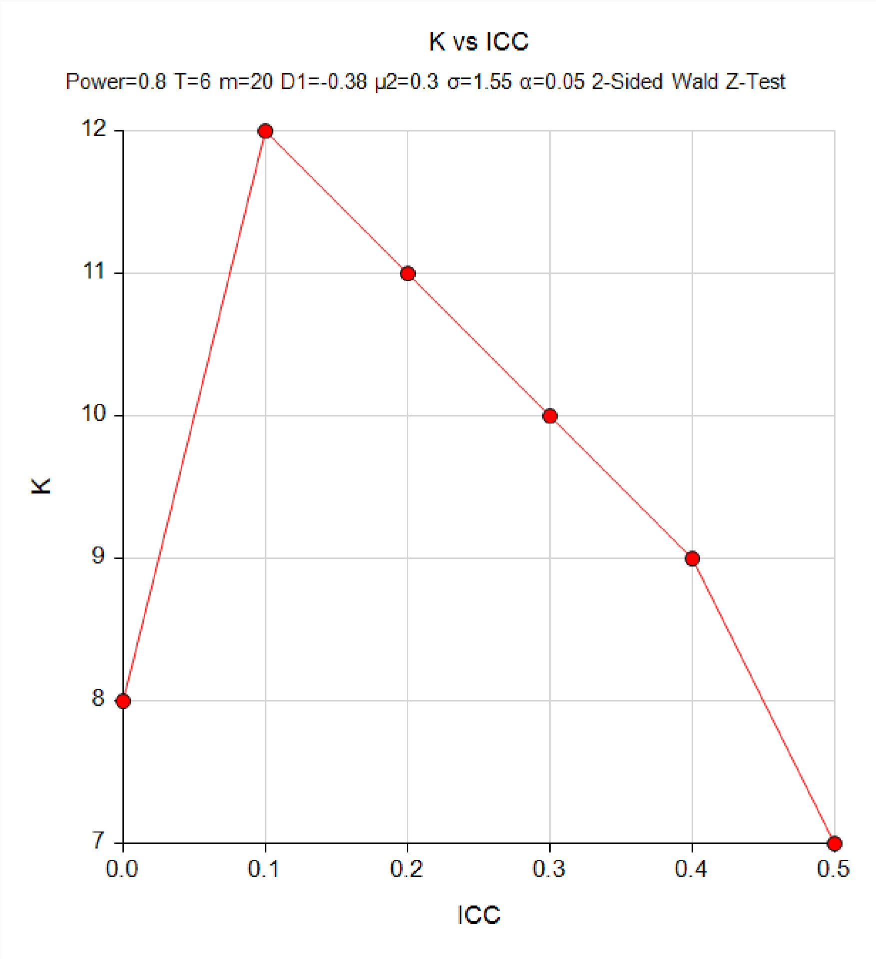
Numeric Results

Solve For: Sample Size [K (Number of Clusters)]
Design Type: Incomplete (Custom)
Groups: 1 = Treatment, 2 = Control
Test Statistic: Wald Z-Test
Hypotheses: H0: μ1 - μ2 = 0 vs. H1: μ1 - μ2 ≠ 0

Power	Design Parameters S	Design Parameters T	Number of Clusters K	Cluster Size M	Cluster Size m	Sample Size N	Mean Treatment μ1	Mean Control μ2	Difference D1	Standard Deviation σ*	Intracluster Correlation Coefficient ICC	Alpha
0.81686	5	6	8	120	20	960	-0.0785	0.3	-0.3785	1.55	0.0	0.05
0.80453	5	6	12	120	20	1440	-0.0785	0.3	-0.3785	1.55	0.1	0.05
0.80101	5	6	11	120	20	1320	-0.0785	0.3	-0.3785	1.55	0.2	0.05
0.81027	5	6	10	120	20	1200	-0.0785	0.3	-0.3785	1.55	0.3	0.05
0.82922	5	6	9	120	20	1080	-0.0785	0.3	-0.3785	1.55	0.4	0.05
0.80236	5	6	7	120	20	840	-0.0785	0.3	-0.3785	1.55	0.5	0.05

* σ is considered to be the total standard deviation (σ = σy) in power computations.

Plots



Design 1 Details

Design Type: Incomplete (Custom)

Power	Design Parameters S	Design Parameters T	Number of Clusters K	Cluster Size M	Cluster Size m	Mean Treatment μ1	Mean Control μ2	Variance Total σy^2	Variance Between-Cluster τ^2	Variance Within-Cluster σw^2	Intracluster Correlation Coefficient ICC	Coefficient of Variation COV
0.81686	5	6	8	120	20	-0.0785	0.3	2.403	0	2.403	0	0

Design Pattern (Balanced, Optimal)

Cluster	T1	T2	T3	T4	T5	T6
1	0	1	1	1	1	1
2	0	1	1	1	1	1
3	0	0	1	1	1	1
4	0	0	1	1	1	1
5	0	0	0	1	1	1
6	0	0	0	0	1	1
7	0	0	0	0	0	1
8	0	0	0	0	0	1

(A Design Details report is given for each additional scenario)

PASS computed K values of 8, 12, 11, 10, 9, and 7, which match those given in Table 1 of Baio et al. (2015). There is a difference, however, when ICC = 0. When ICC = 0, Baio et al. (2015) reports a value of 9, while **PASS** computes a value of 8. This difference may be due to rounding, but it is important to note that **PASS** uses a different method to arrive at the required number of clusters. **PASS** uses the power along with a search for the optimal (balanced) design pattern matrix to find the sample size. Baio et al. (2015) does not specify what final model is being used and the actual power that is achieved. **PASS** always achieves the desired level of power in the result. The power value of 0.81686, for example, indicates that K = 8 with ICC = 0 does achieve the desired power and is, therefore, the correct balanced-design solution.

Example 5 – Finding Cluster Size for a Complete Design (Validation using Hemming and Taljaard (2016))

Hemming and Taljaard (2016) presents an example in section 3.2.2 where they calculate the cluster size for a complete stepped-wedge design when the number of clusters and the number of steps are both fixed. Their example computes cluster size for a standardized effect size of 0.2, which equates to D1 = 0.2 and σ = 1 in **PASS**. What is the required cluster size for 80% power with alpha = 0.05, *S* = 2 or 5 steps, *K* = 30 or 60 clusters, and ICC = 0.01 or 0.25?

In Table 4 on page 143, they present the following calculation results for four *M*/*S*/*ICC* combinations:

K	S	ICC	M
30	2	0.01	96
30	2	0.25	90
60	5	0.01	30
60	5	0.25	

Note that **PASS** does not solve for cluster size using the design effect adjustment that Hemming and Taljaard (2016) uses. Instead **PASS** performs an exhaustive search of all design pattern matrix models to compute the required cluster size.

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 5** 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	**Sample Size [M and m (Cluster Size)]**
Alternative Hypothesis	**Two-Sided**
Power	**0.80**
Alpha	**0.05**
Design Type	**Complete**
Design Parameter Entry Type	**Clusters (K) & Steps (S)**
K (Number of Clusters)	**30**
S (Number of Steps)	**2**
Cluster Size Entry Type	**Subjects per Cluster (M)**
Input Type	**Differences**
D1 (Difference = μ1 - μ2)	**0.2**
μ2 (Control Mean)	**0**
σ (Standard Deviation)	**1**
Use σ (Standard Deviation) as	**Total σ (σ = σy = √[τ² + σw²])**
Between-Cluster Variability Input Type	**ICC**
ICC (Intracluster Correlation Coefficient)	**0.01 0.25**

Output

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

Numeric Results

Solve For: Sample Size [M and m (Cluster Size)]
Design Type: Complete
Groups: 1 = Treatment, 2 = Control
Test Statistic: Wald Z-Test
Hypotheses: H0: μ1 - μ2 = 0 vs. H1: μ1 - μ2 ≠ 0

	Design Parameters			Number of Clusters	Cluster Size		Sample Size	Mean		Difference	Standard Deviation	Intracluster Correlation Coefficient	
Power	S	T	R	K	M	m	N	Treatment μ1	Control μ2	D1	σ*	ICC	Alpha
0.80141	2	3	15	30	93	31	2790	0.2	0	0.2	1	0.01	0.05
0.80067	2	3	15	30	87	29	2610	0.2	0	0.2	1	0.25	0.05

* σ is considered to be the total standard deviation (σ = σy) in power computations.

PASS computed M values of 93 and 87, each 3 less than those given in Table 4 of Hemming and Taljaard (2016). This difference may be due to rounding, but it is important to note that **PASS** uses a different method to arrive at the required cluster size. **PASS** uses the power along with a search for the optimal (balanced) design pattern matrix to find the cluster size, not the design effect adjustment that Hemming and Taljaard (2016) uses. One problem with the design effect adjustment method is that it does not tell you what final model is being used and the actual power that is achieved. **PASS** always achieves the desired level of power in the result. The power value of 0.80141, for example, indicates that M = 93 with ICC = 0.01, and K = 30 and S = 2 does achieve the desired power and is, therefore, the correct balanced-design solution.

Now, go back and change K to 60 and S to 5 for the 2nd half of the table or open **Example 5b** by going to the **File** menu and choosing **Open Example Template**.

Design Tab

K (Number of Clusters) **60**
S (Number of Steps) **5**

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

Numeric Results

Solve For: Sample Size [M and m (Cluster Size)]
Design Type: Complete
Groups: 1 = Treatment, 2 = Control
Test Statistic: Wald Z-Test
Hypotheses: H0: μ1 - μ2 = 0 vs. H1: μ1 - μ2 ≠ 0

	Design Parameters			Number of Clusters	Cluster Size		Sample Size	Mean		Difference	Standard Deviation	Intracluster Correlation Coefficient	
Power	S	T	R	K	M	m	N	Treatment μ1	Control μ2	D1	σ*	ICC	Alpha
0.84118	5	6	12	60	30	5	1800	0.2	0	0.2	1	0.01	0.05
0.80507	5	6	12	60	30	5	1800	0.2	0	0.2	1	0.25	0.05

* σ is considered to be the total standard deviation (σ = σy) in power computations.

PASS computed M values of 30 for both cases. The first matches Table 4 of Hemming and Taljaard (2016). The second is left out of Table 4 for an unknown reason but is calculated by **PASS**.