

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 the all 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.)

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

Three Types of Cluster-Randomized Designs

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

Parallel			2x2 Cross-Over			Stepped-Wedge										
	Cluster	Time		Cluster	Time		Time									
		1			1		2	1	2	3	4	5	6	7		
	1	1		1	1	0	1	0	1	1	1	1	1	1	1	1
	2	1		2	1	0	0	0	1	1	1	1	1	1	1	1
	3	1		3	1	0	0	0	0	1	1	1	1	1	1	1
	4	0		4	0	1	0	0	0	0	1	1	1	1	1	1
	5	0		5	0	1	0	0	0	0	0	1	1	1	1	1
	6	0		6	0	1	0	0	0	0	0	0	1	1	1	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								
	Cluster	Time						
		1	2	3	4	5	6	7
	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					
	Cluster	Time			
		1	2	3	4
	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

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

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						
	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	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								
	Time							
	1	2	3	4	5	6	7	
Cluster	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							
	Time						
	1	2	3	4	5	6	
Cluster	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 ($i = 1, \dots, m$) in cluster k ($k = 1, \dots, K$) at time t ($t = 1, \dots, T$). The linear mixed model can then be written as

$$\begin{aligned} Y_{ikt} &= X_{kt}\theta + \alpha_k + \beta_t + e_{ikt} \\ \alpha_k &\sim N(0, \tau^2) \\ e_{ikt} &\sim N(0, \sigma_w^2) \end{aligned}$$

with

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

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

$$\begin{aligned} \bar{Y}_{kt} &= X_{kt}\theta + \alpha_k + \beta_t + e_{kt} \\ \alpha_k &\sim N(0, \tau^2) \\ e_{kt} &\sim N\left(0, \frac{\sigma_w^2}{m}\right) \end{aligned}$$

with

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

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

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

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{V}_k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{V}_K \end{bmatrix},$$

where each $T \times T$ block matrix, \mathbf{V}_k , represents a single cluster and has the form

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

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

ICC (Intraclass Correlation Coefficient)

The correlation structure for the stepped-wedge cluster-randomized design is often characterized by the intraclass 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 be 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

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

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

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 $H_0: \theta = \mu_1 - \mu_2 = 0$ vs. $H_1: \theta = \mu_1 - \mu_2 \neq 0$ can be conducted using an asymptotic Wald test. The test statistic is

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

$\hat{\theta}$ is the estimated treatment effect from a weighted least-squares analysis. $\text{Var}(\hat{\theta})$ 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 $\text{Var}(\hat{\theta})$ can be written in closed form as

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

where

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

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

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

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

$$\begin{aligned} \text{Power} &= \Phi \left(\frac{\theta_A}{\sqrt{\text{Var}(\hat{\theta})}} - Z_{1-\alpha/2} \right) \\ &= \Phi \left(\frac{\mu_1 - \mu_2}{\sqrt{\{(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\}}} - Z_{1-\alpha/2} \right) \end{aligned}$$

where Φ is the cumulative standard Normal distribution, $\{(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\}$ 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 $(R \times S) < K$ and $((R + 1) \times S) > K$. (Note: If $(R \times S) = K$, 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

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

- **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 $\text{Var}(\hat{\theta})$ cannot be used. In this case, $\text{Var}(\hat{\theta})$ must be calculated using matrix operations. The delay has the overall effect of reducing the power.

Procedure Options

This section describes the options that are specific to this procedure. These are located on the Design and Options tabs. For more information about the options of other tabs, see the Procedure Window chapter.

Design Tab

The Design tab contains the parameters associated with this test such as the means, sample sizes, alpha, and power.

Solve For

Solve For

This option specifies the parameter to be solved for using the other parameters. The parameters that may be selected are *Power*, *Sample Size [K (Number of Clusters)]*, *Sample Size [M and m (Cluster Size)]*, and *Effect Size (μ_1 , DI)*.

Select *Sample Size [K (Number of Clusters)]* when you want to calculate the number of clusters needed to achieve a given power and alpha level for a given effect size.

Select *Power* when you want to calculate the power of an experiment.

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

Test

Alternative Hypothesis

Specify the direction of the alternative hypothesis.

- **One-Sided (H1: $\mu_1 < \mu_2$)**
Refers to a lower-tailed, one-sided test in which H1 is of the type H1: $\mu_1 - \mu_2 = D1 < 0$.
- **One-Sided (H1: $\mu_1 > \mu_2$)**
Refers to an upper-tailed, one-sided test in which H1 is of the type H1: $\mu_1 - \mu_2 = D1 > 0$.
- **Two-Sided**
Refers to a two-sided test in which H1 is of the type H1: $\mu_1 - \mu_2 = D1 \neq 0$.

For one-sided tests, the direction you select must match the values of μ_1 (or D1) and μ_2 you enter. For example, if you select H1: $\mu_1 < \mu_2$, then the value of μ_1 must be less than the value of μ_2 (or $D1 < 0$).

Power and Alpha

Power

This option specifies one or more values for power. Power is the probability of rejecting a false null hypothesis, and is equal to one minus Beta. Beta is the probability of a type-II error, which occurs when a false null hypothesis is not rejected. In this procedure, a type-II error occurs when you fail to reject the null hypothesis of equal means when in fact they are different.

Values must be between zero and one. Historically, the value of 0.80 (Beta = 0.20) was used for power. Now, 0.90 (Beta = 0.10) is also commonly used. A single value may be entered here or a range of values such as *0.8 to 0.95 by 0.05* may be entered.

Alpha

This option specifies one or more values for the probability of a type-I error. A type-I error occurs when a true null hypothesis is rejected. For this procedure, a type-I error occurs when you reject the null hypothesis of equal means when in fact they are equal.

Values must be between zero and one. Historically, the value of 0.05 has been used for alpha. This means that about one test in twenty will falsely reject the null hypothesis. You should pick a value for alpha that represents the risk of a type-I error you are willing to take in your experimental situation.

You may enter a range of values such as *0.01 0.05 0.10* or *0.01 to 0.10 by 0.01*.

Stepped-Wedge Cluster Design and Sample Size

Design Type

Select the type of stepped-wedge cluster design that you want to analyze. The choices refer to the way that the design pattern matrix is organized.

The options are

- **Complete**
In complete stepped-wedge designs, clusters sequentially switch from control to treatment in a balanced fashion over a fixed number of time points (T). Once a cluster switches from control to treatment, the cluster continues to receive the treatment 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), $K = S \times R$.

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

- **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 K and S (or 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.

Design Parameter Entry Type (Solve For ≠ Sample Size [K (Number of Clusters)] and Design Type = Complete)

Select the parameters to use to generate the complete stepped-wedge design to analyze.

The options are

- **Clusters (K) & Steps (S)**

Enter the number of clusters (K) and the number of steps (S) in the design. K must be greater than or equal to and a multiple of S (e.g. if S = 4, then K can be 4, 8, 12, etc.)

- **Clusters (K) & Time Periods (T)**

Enter the number of clusters (K) and the number of time periods (T) in the design. K must be greater than or equal to and a multiple of T - 1, the number of steps (S) (e.g. if T = 5, then K can be 4, 8, 12, etc.)

- **Clusters (K) & Switches per Step (R)**

Enter the number of clusters (K) and the number of switches per step (R) in the design. K must be a multiple of R (e.g. if R = 2, then K can be 2, 4, 6, etc.)

- **Steps (S) & Switches per Step (R)**

Enter the number of steps (S) and the number of switches per step (R) in the design.

- **Time Periods (T) & Switches per Step (R)**

Enter the number of time periods (T) and the number of switches per step (R) in the design.

Design Constraint Entry Type (Solve For = Sample Size [K (Number of Clusters)] and Design Type = Complete)

This option allows you to specify the parameter to fix when solving for the Sample Size [K (Number of Clusters)]. Only sample sizes for complete designs are considered.

The options are

- **Fixed Number of Steps (S)**

Enter the fixed number of steps (S) in the design.

- **Fixed Number of Time Periods (T)**

Enter the fixed number of time periods (T) in the design.

- **Fixed Number of Switches per Step (R)**

Enter the fixed number of switches per step (R) in the design.

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

Design Parameter Entry Type (*Solve For ≠ Sample Size [K (Number of Clusters)] and Design Type = Incomplete (Custom)*)

Select the parameters used to generate the incomplete or custom stepped-wedge design to analyze.

PASS searches among all possible complete and incomplete design configurations that satisfy these design constraints 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 options are

- **Clusters (K) & Steps (S)**
Enter the number of clusters (K) and the number of steps (S) in the design.
- **Clusters (K) & Time Periods (T)**
Enter the number of clusters (K) and the number of time periods (T) in the design.
- **Custom Design Pattern Matrix**
Enter a custom design pattern matrix using the input spreadsheet. This option allows complete flexibility in the design.

Design Constraint Entry Type (*Solve For = Sample Size [K (Number of Clusters)] and Design Type = Incomplete (Custom)*)

This option allows you to specify the parameter to fix when solving for the Sample Size [K (Number of Clusters)].

While solving for sample size, PASS searches among all possible complete and incomplete design configurations that satisfy this design constraint 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 options are

- **Fixed Number of Steps (S)**
Enter the fixed number of steps (S) in the design.
- **Fixed Number of Time Periods (T)**
Enter the fixed number of time periods (T) in the design.
- **Fixed Base Design Pattern Matrix**
Enter a custom base design pattern matrix using the input spreadsheet. This option allows complete flexibility in the design. When this option is chosen, PASS searches for the appropriate number of replicates (R) of the base design pattern to achieve the desired power.

K (Number of Clusters)

Enter the number of clusters in the stepped-wedge design.

You can enter a single value such as 10 or a series of values such as 10 11 12 or 10 to 12 by 1.

Range

The allowable values depends on the type of design:

Complete Designs: $K \geq 2$, $K \geq S$, $K \geq (T - 1)$, K must be a multiple of the number of steps ($S = T - 1$), K must be a multiple of the number of switches per step (R)

Incomplete (Custom) Designs: $K \geq 2$

Tests for Two Means in a Stepped-Wedge Cluster-Randomized Design**S (Number of Steps)**

Enter the number of steps in the stepped-wedge design.

You can enter a single value such as *10* or a series of values such as *10 11 12* or *10 to 12 by 1*.

Range

The allowable values depends on the type of design:

Complete Designs: $S \geq 1$, $S \leq K$ (if entering K), K must be a multiple of S (if entering K)

Incomplete (Custom) Designs: $S \geq 1$

T (Number of Time Periods including Baseline)

Enter the number of time periods in the stepped-wedge design, including the baseline time period where all clusters are given the control (if applicable).

You can enter a single value such as *10* or a series of values such as *10 11 12* or *10 to 12 by 1*.

Range

The allowable values depends on the type of design:

Complete Designs: $T \geq 2$, $T - 1 \leq K$ (if entering K), K must be a multiple of $T - 1$ (if entering K)

Incomplete (Custom) Designs: $T \geq 2$

R (Number of Clusters Switching at each Step) (Complete Designs Only)

Enter the number of clusters switching at each step in the stepped-wedge design.

You can enter a single value such as *1* or a series of values such as *1 2 3* or *1 to 3 by 1*.

Range

$R \geq 1$, K must be a multiple of R (if entering K)

Design Pattern Matrix Columns (Incomplete (Custom) Designs Only)

Specify the columns on the input spreadsheet that contain the design pattern matrix you want to analyze.

ALL

If you enter *ALL* here, **PASS** will use all columns (including empty columns) up to the last non-empty column.

R (Number of Design Pattern Replicates) (Incomplete (Custom) Designs Only)

Enter the number of time to repeat each row of the custom design pattern matrix.

This option is given for convenience in specifying large design pattern matrices. You can simply specify the base design pattern and then use this option to augment it as needed.

You can enter a single value such as *1* or a series of values such as *1 2 3* or *1 to 3 by 1*.

Range

$R \geq 1$

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

Cluster Size Entry Type

This option allows you to choose how the average cluster size will be specified.

The available entry options are

- **Subjects per Cluster (M)**

Enter the average total number of subject in each cluster (M). M is used to calculate the average number of subjects per cluster per time period (m) by dividing M by the number of time periods at which observations are recorded for each cluster. For complete designs, $m = M/T$.

- **Subjects per Cluster per Time Period (m)**

Enter the average number of subjects per cluster per time period (m). m is used to calculate the total number of subjects per cluster (M) by multiplying m by the number of time periods at which observations are recorded for each cluster. For complete designs, $M = m \times T$.

M (Average Number of Subjects per Cluster)

Enter the average total number of subject in each cluster (M). M is used to calculate the average number of subjects per cluster per time period (m) by dividing M by the number of time periods at which observations are recorded for each cluster. For complete designs, $m = M/T$.

You can enter a single value such as *10* or a series of values such as *10 11 12* or *10 to 12 by 1*.

Range

$$M \geq 2$$

m (Average Subjects per Cluster per Time Period)

Enter the average number of subjects per cluster per time period (m). m is used to calculate the total number of subjects per cluster (M) by multiplying m by the number of time periods at which observations are recorded for each cluster. For complete designs, $M = m \times T$.

You can enter a single value such as *10* or a series of values such as *10 11 12* or *10 to 12 by 1*.

Range

$$m \geq 2$$

Effect Size

Input Type

Indicate what type of values to enter to specify the effect size. Regardless of the entry type chosen, the test statistic used in the power and sample size calculations is the same. This option is simply given for convenience in specifying the effect size.

The choices are

- **Means**

Enter μ_1 (Treatment Mean) and μ_2 (Control Mean)

- **Differences**

Enter $D1$ (Difference = $\mu_1 - \mu_2$) and μ_2 (Control Mean)

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

μ_1 (Treatment Mean)

Enter a value for the experimental or treatment group mean under the alternative hypothesis, H_1 . The power calculations assume that this is the actual value of the mean.

You can enter a single value such as *10* or a series of values such as *10 20 30* or *5 to 50 by 5*.

Range

μ_1 can be any value (positive, negative, or zero), $\mu_1 \neq \mu_2$

μ_2 (Control Mean)

Enter a value for the standard or control group mean. The power calculations assume that this is the actual value of the mean.

You can enter a single value such as *10* or a series of values such as *10 20 30* or *5 to 50 by 5*.

Range

μ_2 can be any value (positive, negative, or zero), $\mu_2 \neq \mu_1$

D1 (Difference = $\mu_1 - \mu_2$)

Enter a value for the difference between the two means under the alternative hypothesis, H_1 . The power calculations assume that this is the actual value of the difference, $\mu_1 - \mu_2$.

You can enter a single value such as *10* or a series of values such as *10 20 30* or *5 to 50 by 5*.

Range

D1 can be any non-zero value (positive or negative), $D1 \neq 0$

Effect Size – Cluster Variability

σ (Standard Deviation)

Enter the subject-to-subject standard deviation. This standard deviation applies for both the treatment and control groups.

Use σ (Standard Deviation) as

Specify whether the standard deviation, σ , is to be used as the total standard deviation ($\sigma_y = \sqrt{[\tau^2 + \sigma_w^2]}$) or the within-cluster standard deviation (σ_w).

The choices are:

- **Total σ ($\sigma = \sigma_y = \sqrt{[\tau^2 + \sigma_w^2]}$)**

The standard deviation, σ , is considered to be the total standard deviation (σ_y). σ_w^2 , which is required for the power calculations, is then computed as

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

- **Within-Cluster σ ($\sigma = \sigma_w$)**

The standard deviation, σ , is considered to be the within-cluster standard deviation (σ_w). The total variance (σ_y^2) is then computed as

$$\sigma_y^2 = \tau^2 + \sigma_w^2 = \tau^2 + \sigma^2$$

The variance of cluster means (τ^2) can be calculated using σ_y^2 and ICC as

$$\tau^2 = \text{ICC} \times \sigma_y^2 = \text{ICC} \times \sigma_w^2 / (1 - \text{ICC})$$

$$\text{ICC} = \tau^2 / \sigma_y^2 = \tau^2 / (\tau^2 + \sigma_w^2)$$

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

The variance of cluster means (τ^2) can also be calculated using μ^2 and COV as

$$\tau^2 = \text{COV}^2 \times \mu^2$$

$$\text{COV} = \tau/\mu^2$$

Between-Cluster Variability Input Type

Choose how to input the between-cluster variability.

The available options are

- **ICC**

Enter the Intracluster Correlation Coefficient. The ICC is calculated as

$$\text{ICC} = \tau^2/\sigma_y^2 = \tau^2/(\tau^2 + \sigma_w^2)$$

- **COV**

Enter the Coefficient of Variation of Outcomes for the control group. (This is not the COV of cluster sizes!)
The COV is calculated as

$$\text{COV} = \tau/\mu^2$$

ICC (Intracluster Correlation Coefficient)

Enter the Intracluster Correlation Coefficient. 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 response that can be accounted for by the between-cluster variation.

The ICC is calculated as

$$\text{ICC} = \tau^2/\sigma_y^2 = \tau^2/(\tau^2 + \sigma_w^2)$$

You can enter a single value such as *0.01* or a series of values such as *0.1 0.2 0.3* or *0.1 to 0.3 by 0.1*.

Range

$-1 < \text{ICC} < 1$

COV (Coefficient of Variation of Outcomes)

Enter the Coefficient of Variation of Outcomes for the control group. (This is not the COV of cluster sizes!)

The COV is calculated as

$$\text{COV} = \tau/\mu^2$$

You can enter a single value such as *0.3* or a series of values such as *0.1 0.2 0.3* or *0.1 to 0.3 by 0.1*.

Range

$\text{COV} \geq 0$

Options Tab

The Options tab contains options relevant to searching among incomplete (custom) design pattern matrices.

Incomplete (Custom) Design Pattern Matrix Search Options

Design Pattern Assignment Type for Extra Clusters

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. This option controls how design patterns are assigned to individual clusters.

Design Pattern Assignment

PASS utilizes the following logic when creating a design pattern matrix using K and S or T :

1. Sequentially assign R complete cluster sets to a design pattern sub-matrix, where R is the largest integer that satisfies $(R \times S) < K$ and $((R + 1) \times S) > K$. (Note: If $(R \times S) = K$, then there are no “Extra” clusters and this option does not apply.)
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.
 - **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.
 - **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.
3. Compute the power for all possible assignment combinations.
4. Return the design pattern matrix with the highest power.

Maximum Combinations for Incomplete Design Pattern Search

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 searches among all possible design pattern combinations to find the particular design that maximizes the power. This can take some time for larger design pattern matrices. This option specifies the largest number of possible combinations for which the search will be allowed. If Design Pattern Assignment Type for Extra Clusters is set to “Unbalanced”, and the number of possible combinations is larger than this value, then a search among “Balanced” designs is conducted instead. If a “Balanced” design search results in a number of possible combinations that is larger than this value, then “Sequential” assignment is used and no search is conducted.

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 $\sigma = 1$ 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:

<i>ICC</i>	<i>m</i>	<i>Power</i>
0.01	17	55%
0.1	17	49%
0.01	50	91%
0.1	50	90%

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests for Two Means in a Stepped-Wedge Cluster-Randomized Design** procedure window. You may then make the appropriate entries as listed below, or open **Example 1** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
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 = $\mu_1 - \mu_2$).....	0.2
μ_2 (Control Mean).....	0
σ (Standard Deviation)	1
Use σ (Standard Deviation) as	Total SD ($\sigma = \sigma_y = \sqrt{\tau^2 + \sigma_w^2}$)
Between-Cluster Variability Input Type	ICC
ICC (Intracluster Correlation Coefficient).....	0.01 0.1

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

Output

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

Numeric Results

Numeric Results

Design Type: Complete

Test Statistic: Wald Z-Test

H0: $\mu_1 - \mu_2 = 0$. H1: $\mu_1 - \mu_2 = D1 \neq 0$.

Power	Design S/T/R	Clusters K	Cluster Size M/m	Sample Size N	Trt Mean μ_1	Ctrl Mean μ_2	Diff D1	Std Dev σ	ICC	Alpha
0.54844	5/6/2	10	102/17	1020	0.20	0.00	0.20	1.00	0.010	0.050
0.48864	5/6/2	10	102/17	1020	0.20	0.00	0.20	1.00	0.100	0.050
0.91489	5/6/2	10	300/50	3000	0.20	0.00	0.20	1.00	0.010	0.050
0.90211	5/6/2	10	300/50	3000	0.20	0.00	0.20	1.00	0.100	0.050

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.

Report Definitions

Power is the probability of rejecting the null hypothesis when the alternative is true. It should be close to one.

S is the number of steps in the study design. $S = T - 1$.

T is the number of time periods in the study, including the baseline. $T = S + 1$

R is the number clusters switching from control to treatment at each step.

K is the total number of clusters to be randomized. $K = S \times R$.

M is the average number of subjects per cluster. $M = m \times T$.

m is the average number of subjects per cluster per time period. $M = m \times T$.

N is total sample size from all clusters and time periods combined. $m = M/T$.

μ_1 is the treatment mean, assuming the alternative hypothesis.

μ_2 is the control, standard, reference, or baseline mean.

$D1 = \mu_1 - \mu_2$ is the difference assuming the alternative hypothesis (H1).

σ is the subject-to-subject standard deviation.

ICC is the intracluster correlation coefficient.

Alpha is the probability of rejecting the null hypothesis when it is true.

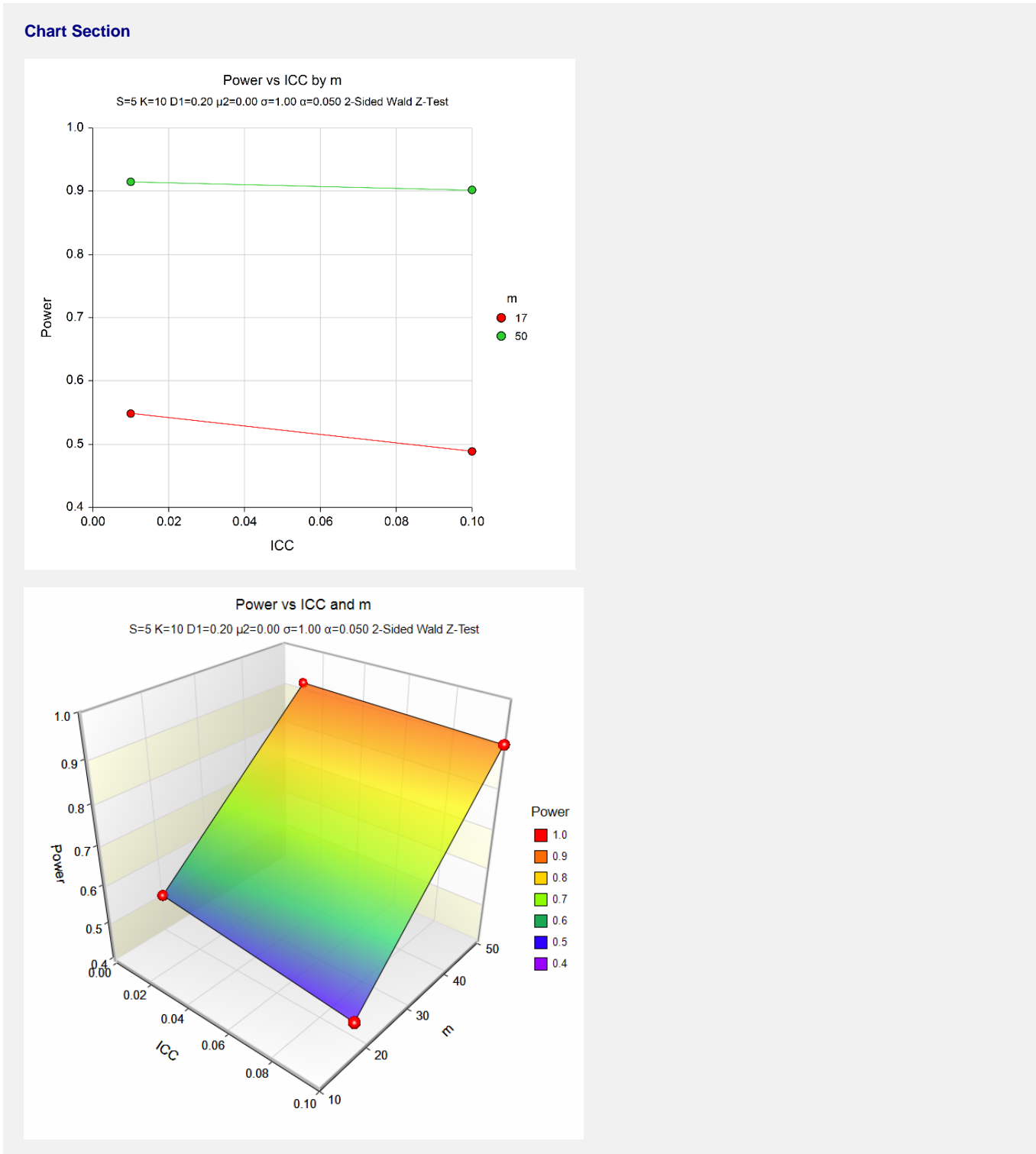
Summary Statements

A sample of 10 clusters in a complete stepped-wedge cluster-randomized design with 6 time periods (including the baseline), 5 steps, 2 cluster(s) switching from control to treatment at each step, and an average of 102 subjects per cluster with an average of 17 subjects per cluster per time period (for a total sample size of 1020 subjects) achieves 54.844% power to detect a difference between means of 0.20. The treatment mean is assumed to be 0.20 under the alternative hypothesis. The control mean is 0.00. The standard deviation is 1.00. The test statistic used is the two-sided Wald Z-Test. The ICC is 0.010, and the significance level of the test is 0.050.

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 plotted in the chart below.

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

Chart Section



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

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

Design Details

Design 1 Details

Design Type: Complete

Power	Design S/T/R	Clusters K	Cluster Size M/m	Trt Mean μ_1	Ctrl Mean μ_2	Total Variance σ_y^2	Between-Cluster Variance τ^2	Within-Cluster Variance σ_w^2	ICC	COV
0.54844	5/6/2	10	102/17	0.20	0.00	1.000	0.010	0.990	0.010	

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 = $\mu_1 - \mu_2$):	0.2
μ_2 (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								
		0	1	2	3	4	5	6	7	8
Cluster	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:

<i>ICC</i>	<i>Power</i>
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

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

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests for Two Means in a Stepped-Wedge Cluster-Randomized Design** procedure window. You may then make the appropriate entries as listed below, or open **Example 2** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
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 = $\mu_1 - \mu_2$).....	1
μ_2 (Control Mean).....	1
σ (Standard Deviation)	2.2
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).....	0.05 0.1 0.15 0.2 0.3 0.4 0.5

Output

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

Numeric Results

Design Type: Incomplete (Custom)

Test Statistic: Wald Z-Test

H0: $\mu_1 - \mu_2 = 0$. H1: $\mu_1 - \mu_2 = D1 \neq 0$.

Power	Design S/T/R	Clusters K	Cluster Size M/m	Sample Size N	Trt Mean μ_1	Ctrl Mean μ_2	Diff D1	Std Dev σ	ICC	Alpha
0.89096	7/8/1	18	30/15	540	2.00	1.00	1.00	2.20	0.050	0.050
0.87035	7/8/1	18	30/15	540	2.00	1.00	1.00	2.20	0.100	0.050
0.86936	7/8/1	18	30/15	540	2.00	1.00	1.00	2.20	0.150	0.050
0.87723	7/8/1	18	30/15	540	2.00	1.00	1.00	2.20	0.200	0.050
0.90459	7/8/1	18	30/15	540	2.00	1.00	1.00	2.20	0.300	0.050
0.93691	7/8/1	18	30/15	540	2.00	1.00	1.00	2.20	0.400	0.050
0.96669	7/8/1	18	30/15	540	2.00	1.00	1.00	2.20	0.500	0.050

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 $\sigma = 1$ 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:

<i>M</i>	<i>S</i>	<i>ICC</i>	<i>K</i>
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

This section presents the values of each of the parameters needed to run this example. First, from the **PASS** Home window, load the **Tests for Two Means in a Stepped-Wedge Cluster-Randomized Design** procedure window. You may then make the appropriate entries as listed below, or open **Example 3** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
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 = $\mu_1 - \mu_2$)	0.2
μ_2 (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 (Intraclass Correlation Coefficient)	0.01 0.25

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

Output

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

Numeric Results

Design Type: Incomplete (Custom)

Test Statistic: Wald Z-Test

H0: $\mu_1 - \mu_2 = 0$. H1: $\mu_1 - \mu_2 = D1 \neq 0$.

Power	Design S/T	Clusters K	Cluster Size M/m	Sample Size N	Trt Mean μ_1	Ctrl Mean μ_2	Diff D1	Std Dev σ	ICC	Alpha
0.80349	2/3	85	30/10	2550	0.20	0.00	0.20	1.00	0.010	0.050
0.80244	2/3	85	30/10	2550	0.20	0.00	0.20	1.00	0.250	0.050
0.80845	9/10	17	100/10	1700	0.20	0.00	0.20	1.00	0.010	0.050
0.80785	9/10	18	100/10	1800	0.20	0.00	0.20	1.00	0.250	0.050

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, μ_2 , is 0.3, the difference, D_1 , 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

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests for Two Means in a Stepped-Wedge Cluster-Randomized Design** procedure window. You may then make the appropriate entries as listed below, or open **Example 4** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
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
D_1 (Difference = $\mu_1 - \mu_2$).....	-0.3785
μ_2 (Control Mean).....	0.3
σ (Standard Deviation)	1.55
Use σ (Standard Deviation) as	Total σ ($\sigma = \sigma_y = \sqrt{\tau^2 + \sigma_w^2}$)
Between-Cluster Variability Input Type	ICC
ICC (Intraclass Correlation Coefficient).....	0 to 0.5 by 0.1
Reports Tab	
Means and Difference Decimals.....	4

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

Output

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

Numeric Results

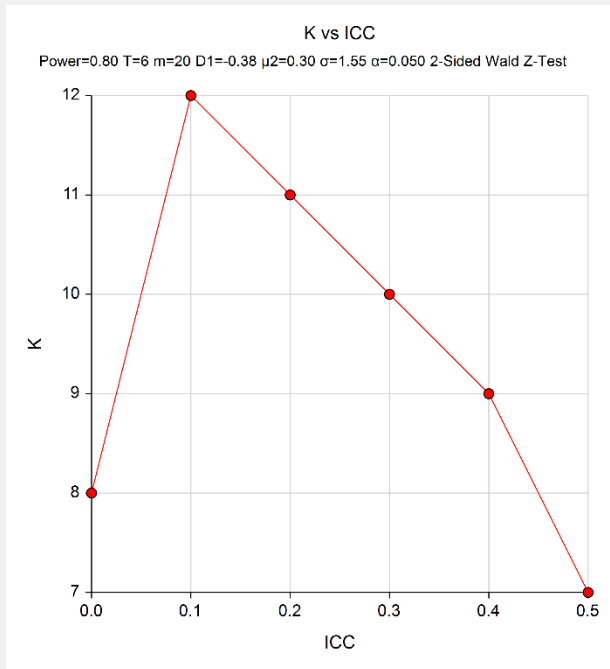
Design Type: Incomplete (Custom)

Test Statistic: Wald Z-Test

H0: $\mu_1 - \mu_2 = 0$. H1: $\mu_1 - \mu_2 = D1 \neq 0$.

Power	Design S/T	Clusters K	Cluster Size M/m	Sample Size N	Trt Mean μ_1	Ctrl Mean μ_2	Diff D1	Std Dev σ	ICC	Alpha
0.81686	5/6	8	120/20	960	-0.0785	0.3000	-0.3785	1.55	0.000	0.050
0.80453	5/6	12	120/20	1440	-0.0785	0.3000	-0.3785	1.55	0.100	0.050
0.80101	5/6	11	120/20	1320	-0.0785	0.3000	-0.3785	1.55	0.200	0.050
0.81027	5/6	10	120/20	1200	-0.0785	0.3000	-0.3785	1.55	0.300	0.050
0.82922	5/6	9	120/20	1080	-0.0785	0.3000	-0.3785	1.55	0.400	0.050
0.80236	5/6	7	120/20	840	-0.0785	0.3000	-0.3785	1.55	0.500	0.050

Chart Section



Design 1 Details

Design Type: Incomplete (Custom)

Power	Design S/T	Clusters K	Cluster Size M/m	Trt Mean μ_1	Ctrl Mean μ_2	Total Variance σ_y^2	Between-Cluster Variance τ^2	Within-Cluster Variance σ_w^2	ICC	COV
0.81686	5/6	8	120/20	-0.0785	0.3000	2.403	0.000	2.403	0.000	0.000

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

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	1	1
8	0	0	0	0	0	1
9	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 $\sigma = 1$ 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:

<i>K</i>	<i>S</i>	<i>ICC</i>	<i>M</i>
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

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests for Two Means in a Stepped-Wedge Cluster-Randomized Design** procedure window. You may then make the appropriate entries as listed below, or open **Example 5a** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
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 = $\mu_1 - \mu_2$)	0.2
μ_2 (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 (Intraclass Correlation Coefficient)	0.01 0.25

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

Output

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

Numeric Results											
Design Type: Complete											
Test Statistic: Wald Z-Test											
H0: $\mu_1 - \mu_2 = 0$. H1: $\mu_1 - \mu_2 = D1 \neq 0$.											
Power	Design S/T/R	Clusters K	Cluster Size M/m	Sample Size N	Trt Mean μ_1	Ctrl Mean μ_2	Diff D1	Std Dev σ	ICC	Alpha	
0.80141	2/3/15	30	93/31	2790	0.20	0.00	0.20	1.00	0.010	0.050	
0.80067	2/3/15	30	87/29	2610	0.20	0.00	0.20	1.00	0.250	0.050	

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**.

Option	Value
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											
Design Type: Complete											
Test Statistic: Wald Z-Test											
H0: $\mu_1 - \mu_2 = 0$. H1: $\mu_1 - \mu_2 = D1 \neq 0$.											
Power	Design S/T/R	Clusters K	Cluster Size M/m	Sample Size N	Trt Mean μ_1	Ctrl Mean μ_2	Diff D1	Std Dev σ	ICC	Alpha	
0.84118	5/6/12	60	30/5	1800	0.20	0.00	0.20	1.00	0.010	0.050	
0.80507	5/6/12	60	30/5	1800	0.20	0.00	0.20	1.00	0.250	0.050	

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.