

## Chapter 231

# Tests for Two Proportions 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  $2 \times 2$  *cluster-randomized cross-over design* is 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 proportions 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 Proportions 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		Cluster	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	2	0	0	1	1	1	1	1	1	1
	3	1		3	1	0	3	0	0	0	1	1	1	1	1	1
	4	0		4	0	1	4	0	0	0	0	1	1	1	1	1
	5	0		5	0	1	5	0	0	0	0	0	1	1	1	1
	6	0		6	0	1	6	0	0	0	0	0	0	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								Complete Design					
K = 6 Clusters, S = 6 Steps, T = 7 Time Periods, R = 1 Switch per Step								K = 6 Clusters, S = 3 Steps, T = 4 Time Periods, R = 2 Switches per Step					
	Cluster	Time							Cluster	Time			
		1	2	3	4	5	6			7	1	2	3
	1	0	1	1	1	1	1	1	1	0	1	1	1
	2	0	0	1	1	1	1	1	2	0	1	1	1
	3	0	0	0	1	1	1	1	3	0	0	1	1
	4	0	0	0	0	1	1	1	4	0	0	1	1
	5	0	0	0	0	0	1	1	5	0	0	0	1
	6	0	0	0	0	0	0	1	6	0	0	0	1

## Tests for Two Proportions 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
- $P_1$  The treatment group proportion, assuming the alternative hypothesis
- $P_2$  The control, standard, reference, or baseline group proportion

## 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
$\theta$	Fixed treatment effect
$\alpha_k$	Random effect for cluster $k$
$\beta_t$	Fixed effect for time $t$
$e_{ikt}$	Within-cluster error

## The Treatment Effect

$\theta$  is the treatment effect and is equal to  $D1$ , the difference between the treatment mean,  $\mu_1$ , and the control group mean,  $\mu_2$ , such that

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

In the case of testing two proportions, with  $\mu_1 = P_1$  and  $\mu_2 = P_2$  as the treatment and control group proportions, respectively, then

$$\theta = D1 = P_1 - P_2.$$

## Between- and Within-Cluster Variation

From the model above,  $\tau^2$  is the between-cluster variance and  $\sigma_w^2$  is the within-cluster variance. If  $\tau^2$  and  $\sigma_w^2$  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 Proportions 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,  $\sigma_w^2$ , to calculate the between-cluster variance,  $\tau^2$ , 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  $\mu_2 = P_2$  is the control group mean (i.e. a proportion for binary data), then the COV of outcomes for the control group is calculated as

$$COV = \frac{\tau}{\mu_2} = \frac{\tau}{P_2}$$

and can be used along with  $\mu_2 = P_2$  to calculate  $\tau^2$  as

$$\tau^2 = COV^2 \times \mu_2^2 = COV^2 \times P_2^2.$$

## Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

### Specifying the Total and Within-Cluster Variance for a Binary Response

In the case of testing two proportions where the responses,  $Y_{ikt}$ , are binary (0 or 1) and using the normal approximation, the variance can be calculated using three different methods, given as options in this procedure.

Method 1:	Null Variance	$\sigma^2 = P_2(1 - P_2)$
Method 2:	Pooled Variance	$\sigma^2 = \left(\frac{P_1 - P_2}{2}\right)\left(1 - \frac{P_1 - P_2}{2}\right)$
Method 3:	Group Variance Average	$\sigma^2 = \frac{P_1(1 - P_1) + P_2(1 - P_2)}{2}$

Method 1 is used in Hussey and Hughes (2007) and Hemming and Girling (2014). Method 2 is used in Baio et al. (2015). Method 3 is used in Hemming, Lilford, and Girling (2015).

The variance that is calculated from the two proportions,  $\sigma^2$ , may be considered to be either the total variance,  $\sigma_y^2$ , or the within-cluster variance,  $\sigma_w^2$ . If  $\sigma^2$  is considered to be the total variance,  $\sigma_y^2$ , then using the ICC, the between-cluster variance,  $\tau^2$ , can be computed as

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

and using the COV the between-cluster variance,  $\tau^2$ , can be computed as

$$\tau^2 = COV^2 \times \mu_2^2 = COV^2 \times P_2^2.$$

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

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

Otherwise, if  $\sigma^2$  is considered to be the within-cluster variance,  $\sigma_w^2$ , then

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

and using the ICC, the between-cluster variance,  $\tau^2$ , 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,  $\tau^2$ , can be computed as

$$\tau^2 = COV^2 \times \mu_2^2 = COV^2 \times P_2^2.$$

---

## Test Statistic

Hussey and Hughes (2007) suggest that a statistical hypothesis test of  $H_0: \theta = P_1 - P_2 = 0$  vs.  $H_1: \theta = P_1 - P_2 \neq 0$  can be conducted using an asymptotic Wald test. The test statistic is

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

$\hat{\theta}$  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).

## Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

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  $\tau^2$  and  $\sigma_w^2$  assumed to be known, the power for a two-sided Wald test is computed as

$$\begin{aligned} \text{Power} &= \Phi \left( \frac{\theta_A}{\sqrt{\text{Var}(\hat{\theta})}} - Z_{1-\alpha/2} \right) \\ &= \Phi \left( \frac{P_1 - P_2}{\sqrt{\{(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\}}} - Z_{1-\alpha/2} \right) \end{aligned}$$

where  $\Phi$  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.  $\theta_A$  is the value for  $\theta$  under the alternative hypothesis, i.e.  $\theta_A = P_1 - P_2$ .

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.

## Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

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						
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
	5	0	0	0	0	0	0.5	0.8

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 proportions, 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 (PI, DI, RI, ORI)*.

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 Proportions in a Stepped-Wedge Cluster-Randomized Design

---

### Test

#### Alternative Hypothesis

Specify the direction of the alternative hypothesis.

- **One-Sided (H1: P1 < P2)**  
Refers to a lower-tailed, one-sided test in which H1 is of the type  $H1: P1 - P2 = D1 < 0$ .
- **One-Sided (H1: P1 > P2)**  
Refers to an upper-tailed, one-sided test in which H1 is of the type  $H1: P1 - P2 = D1 > 0$ .
- **Two-Sided**  
Refers to a two-sided test in which H1 is of the type  $H1: P1 - P2 = D1 \neq 0$ .

For one-sided tests, the direction you select must match the values of P1 (or D1, R1, or OR1) and P2 you enter. For example, if you select  $H1: P1 < P2$ , then the value of P1 must be less than the value of P2 (or  $D1 < 0$ ,  $R1 < 1$ , or  $OR1 < 1$ ).

---

### 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 proportions 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 proportions 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),  $K = S \times R$ .

- **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.

## Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized 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.

### **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.

## Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized 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:*  $K \geq 2$

### **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.

**Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design****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**

$$R \geq 1$$

**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

- **Proportions**  
Enter P1 (Treatment Proportion) and P2 (Control Proportion)
- **Differences**  
Enter D1 (Difference = P1 - P2) and P2 (Control Proportion)
- **Ratios**  
Enter R1 (Ratio = P1/P2) and P2 (Control Proportion)
- **Odds Ratios**  
Enter OR1 (Odds Ratio = O1/O2) and P2 (Control Proportion)  
O1 = P1/(1 - P1) and O2 = P2/(1 - P2)

### P1 (Treatment Proportion)

Enter a value for the experimental or treatment group proportion under the alternative hypothesis, H1. The power calculations assume that this is the actual value of the proportion.

You can enter a single value such as *0.1* or a series of values such as *0.1 0.2 0.3* or *0.1 to 0.9 by 0.1*.

### Range

$0 < P1 < 1, P1 \neq P2$

### P2 (Control Proportion)

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

You can enter a single value such as *0.1* or a series of values such as *0.1 0.2 0.3* or *0.1 to 0.9 by 0.1*.

### Range

$0 < P2 < 1, P2 \neq P1$

### D1 (Difference = P1 - P2)

Enter a value for the difference between the two proportions under the alternative hypothesis, H1. The power calculations assume that this is the actual value of the difference, P1 - P2.

You can enter a single value such as *0.05* or a series of values such as *0.03 0.05 0.10* or *0.01 to 0.05 by 0.01*.

### Range

$-1 < D1 < 1, D1 \neq 0$

## Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

### R1 (Ratio = P1/P2)

Enter a value for the ratio of the two proportions under the alternative hypothesis, H1. The power calculations assume that this is the actual value of the ratio, P1/P2.

You can enter a single value such as 0.5 or a series of values such as 0.5 0.6 0.7 0.8 or 1.25 to 2.0 by 0.25.

#### Range

$R1 > 0, R1 \neq 1$

### OR1 (Odds Ratio = O1/O2)

Enter a value for the odds ratio for the two proportions under the alternative hypothesis, H1. The power calculations assume that this is the actual value of the odds ratio, O1/O2.

You can enter a single value such as 0.5 or a series of values such as 0.5 0.6 0.7 0.8 or 1.25 to 2.0 by 0.25.

#### Range

$OR1 > 0, OR1 \neq 1$

---

## Effect Size – Cluster Variability

### Variance Calculation Formula

The power calculations are based on the normal approximation to the binomial distribution. This option allows you to specify how the variance for the normal distribution is calculated.

The available options are:

- $\sigma^2 = P2(1 - P2)$   
The null variance is used.
- $\sigma^2 = [(P1 + P2)/2] \times [1 - (P1 + P2)/2]$   
The pooled variance is used.
- $\sigma^2 = [P1(1 - P1) + P2(1 - P2)]/2$   
The average of the two group variances is used.

### Recommendation

The 1st option,  $\sigma^2 = P2(1 - P2)$ , is used by Hussey and Hughes (2007) and Hemming and Girling (2014), so we recommend using this option.

The 2nd option,  $\sigma^2 = [(P1 + P2)/2] \times [1 - (P1 + P2)/2]$ , is used by Baio et al. (2015).

The 3rd option,  $\sigma^2 = [P1(1 - P1) + P2(1 - P2)]/2$ , is used by Hemming, Lilford, and Girling (2015).

### Use Calculated Variance as

Specify whether the variance,  $\sigma^2$ , which is calculated from P1 and P2, is to be used as the total variance ( $\sigma y^2 = \tau^2 + \sigma w^2$ ) or the within-cluster variance ( $\sigma w^2$ ).

The choices are:

- **Total Variance ( $\sigma^2 = \sigma y^2 = \tau^2 + \sigma w^2$ )**  
The calculated variance,  $\sigma^2$ , is considered to be the total variance ( $\sigma y^2$ ).  $\sigma w^2$ , which is required for the power calculations, is then computed as

$$\sigma w^2 = \sigma y^2 - \tau^2 = \sigma^2 - \tau^2$$



## Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

- **Within-Cluster Variance ( $\sigma^2 = \sigma_w^2$ )**

The calculated variance,  $\sigma^2$ , is considered to be the within-cluster variance ( $\sigma_w^2$ ). The total variance ( $\sigma_y^2$ ) is then computed as

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

The variance of cluster means ( $\tau^2$ ) can be calculated using  $\sigma_y^2$  and ICC as

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

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

The variance of cluster means ( $\tau^2$ ) can also be calculated using P2 and COV as

$$\tau^2 = \text{COV}^2 \times P2^2$$

$$\text{COV} = \tau / P2$$

### Between-Cluster Variability Input Type

Choose how to input the between-cluster variability.

The available options are

- **ICC**

Enter the Intraclass 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

$$\text{COV} = \tau / P2$$

### ICC (Intraclass Correlation Coefficient)

Enter the Intraclass 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

$$\text{COV} = \tau / P2$$

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 Hussey and Hughes (2007))

Hussey and Hughes (2007) presents an example related to a proposed program aimed at reducing the number of STDs in the state of Washington. In the study, the current prevalence of Chlamydia infection was  $P_2 = 0.05$ . The study involved  $K = 24$  counties that were to be randomized to the intervention  $R = 6$  at a time, such that the number of required time periods was  $T = 5$ , and the number of steps was  $S = 4$ . They planned to test  $m = 100$  individuals per county per time interval. What kind of power would the study have for COV values of 0.3 and 0.5 and proportion ratios ( $P_1/P_2$ ) between 0.5 and 0.8, with an alpha level of 0.05?

### 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 Proportions 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>
<b>Design Tab</b>	
Solve For .....	<b>Power</b>
Alternative Hypothesis .....	<b>Two-Sided</b>
Alpha .....	<b>0.05</b>
Design Type .....	<b>Complete</b>
Design Parameter Entry Type .....	<b>Clusters (K) &amp; Switches per Step (R)</b>
K (Number of Clusters) .....	<b>24</b>
R (Number of Clusters Switching at each Step) .....	<b>6</b>
Cluster Size Entry Type .....	<b>Subjects per Cluster per Time Period (m)</b>
m (Ave. Subjects per Cluster per Time Period) .....	<b>100</b>
Input Type .....	<b>Ratios</b>
R1 (Ratio = $P_1/P_2$ ) .....	<b>0.5 to 0.8 by 0.05</b>
P2 (Control Proportion) .....	<b>0.05</b>
Variance Calculation Formula .....	<b><math>\sigma^2 = P_2(1 - P_2)</math></b>
Use Calculated Variance as .....	<b>Within-Cluster Variance (<math>\sigma^2 = \sigma_w^2</math>)</b>
Between-Cluster Variability Input Type .....	<b>COV</b>
COV (Coefficient of Variation of Outcomes) .....	<b>0.3 0.5</b>
<b>Reports Tab</b>	
ICC, COV, and Variances Decimals .....	<b>4</b>

## Tests for Two Proportions 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: P1/P2 = 1. H1: P1/P2 = R1 ≠ 1.

Power	Design S/T/R	Clusters K	Cluster Size M/m	Sample Size N	Trt Prop P1	Ctrl Prop P2	Ratio R1	COV	Alpha
0.96458	4/5/6	24	500/100	12000	0.0250	0.0500	0.5000	0.3000	0.050
0.94839	4/5/6	24	500/100	12000	0.0250	0.0500	0.5000	0.5000	0.050
0.92361	4/5/6	24	500/100	12000	0.0275	0.0500	0.5500	0.3000	0.050
0.89805	4/5/6	24	500/100	12000	0.0275	0.0500	0.5500	0.5000	0.050
0.85387	4/5/6	24	500/100	12000	0.0300	0.0500	0.6000	0.3000	0.050
0.81900	4/5/6	24	500/100	12000	0.0300	0.0500	0.6000	0.5000	0.050
0.75065	4/5/6	24	500/100	12000	0.0325	0.0500	0.6500	0.3000	0.050
0.70974	4/5/6	24	500/100	12000	0.0325	0.0500	0.6500	0.5000	0.050
0.61788	4/5/6	24	500/100	12000	0.0350	0.0500	0.7000	0.3000	0.050
0.57680	4/5/6	24	500/100	12000	0.0350	0.0500	0.7000	0.5000	0.050
0.46947	4/5/6	24	500/100	12000	0.0375	0.0500	0.7500	0.3000	0.050
0.43445	4/5/6	24	500/100	12000	0.0375	0.0500	0.7500	0.5000	0.050
0.32539	4/5/6	24	500/100	12000	0.0400	0.0500	0.8000	0.3000	0.050
0.30041	4/5/6	24	500/100	12000	0.0400	0.0500	0.8000	0.5000	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.  $m = M/T$ .

P1 is the treatment proportion, assuming the alternative hypothesis.

P2 is the control, standard, reference, or baseline proportion.

R1 = P1/P2 is the ratio assuming the alternative hypothesis (H1).

COV is the coefficient of variation of outcomes.

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

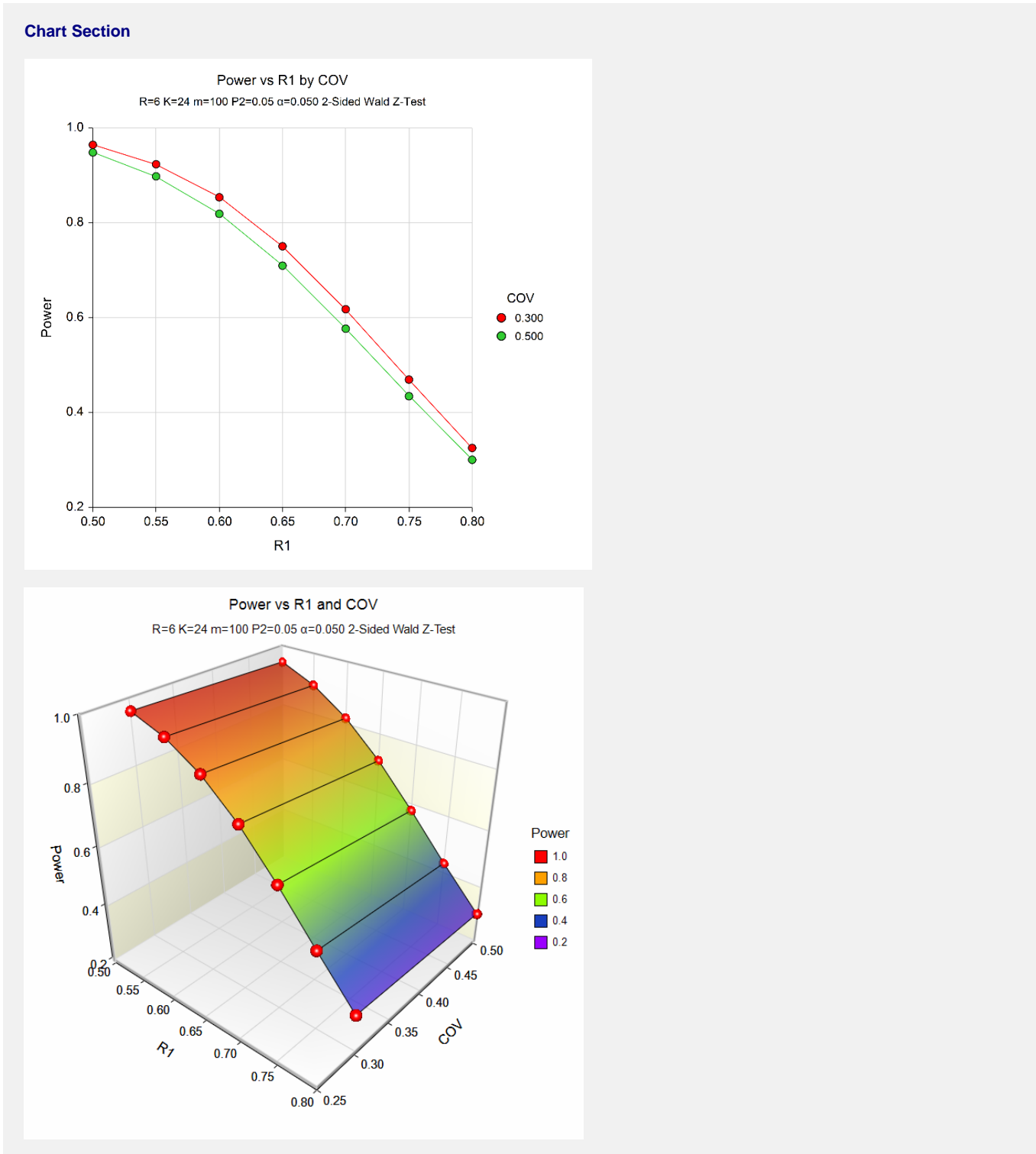
## Summary Statements

A sample of 24 clusters in a complete stepped-wedge cluster-randomized design with 5 time periods (including the baseline), 4 steps, 6 cluster(s) switching from control to treatment at each step, and an average of 500 subjects per cluster with an average of 100 subjects per cluster per time period (for a total sample size of 12000 subjects) achieves 96.458% power to detect a ratio of 0.5000. The treatment proportion is assumed to be 0.0250 under the alternative hypothesis. The control proportion is 0.0500. The test statistic used is the two-sided Wald Z-Test. The COV is 0.3000, 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 values from this table are plotted in the chart below.

Tests for Two Proportions 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 various combinations of the effect size (R1) and COV. If you look on page 187 of Hussey and Hughes (2007), you'll see that these charts look very similar to the chart in Figure 2, with the X-axis reversed. This is because the effect size in Hussey and Hughes (2007) is equal to 1 – R1, as specified here.

## Tests for Two Proportions 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 Prop P1	Ctrl Prop P2	Total Variance $\sigma_y^2$	Between-Cluster Variance $\tau^2$	Within-Cluster Variance $\sigma_w^2$	ICC	COV
0.96458	4/5/6	24	500/100	0.0250	0.0500	0.0477	0.0002	0.0475	0.0047	0.3000

## Design Pattern (Complete)

Cluster	T1	T2	T3	T4	T5
1	0	1	1	1	1
2	0	1	1	1	1
3	0	1	1	1	1
4	0	1	1	1	1
5	0	1	1	1	1
6	0	1	1	1	1
7	0	0	1	1	1
8	0	0	1	1	1
9	0	0	1	1	1
10	0	0	1	1	1
11	0	0	1	1	1
12	0	0	1	1	1
13	0	0	0	1	1
14	0	0	0	1	1
15	0	0	0	1	1
16	0	0	0	1	1
17	0	0	0	1	1
18	0	0	0	1	1
19	0	0	0	0	1
20	0	0	0	0	1
21	0	0	0	0	1
22	0	0	0	0	1
23	0	0	0	0	1
24	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) & Switches per Step (R)
K (Number of Clusters):	24
R (Number of Clusters Switching at each Step):	6
Cluster Size Entry Type:	Subjects per Cluster per Time Period (m)
m (Ave. Subjects per Cluster per Time Period):	100
Input Type:	Ratios
R1 (Ratio = P1/P2):	0.5 to 0.8 by 0.05
P2 (Control Proportion):	0.05
Variance Calculation Formula:	$\sigma^2 = P2(1 - P2)$
Use Calculated Variance as:	Within-Cluster Variance ( $\sigma^2 = \sigma_w^2$ )
Between-Cluster Variability Input Type:	COV
COV (Coefficient of Variation of Outcomes):	0.3 0.5

## Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

## 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.

## Additional Validation

Hussey and Hughes (2007) reports that for  $COV = 0.3$  and  $P1 = 0.032$ , the power is “about 80%”. To replicate this, you may make the appropriate entries as listed below, or open **Example 1b** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	Power
Alternative Hypothesis .....	Two-Sided
Alpha .....	0.05
Design Type .....	Complete
Design Parameter Entry Type .....	Clusters (K) & Switches per Step (R)
K (Number of Clusters) .....	24
R (Number of Clusters Switching at each Step) .....	6
Cluster Size Entry Type .....	Subjects per Cluster per Time Period (m)
m (Ave. Subjects per Cluster per Time Period) .....	100
Input Type .....	Proportions
P1 (Treatment Proportion): .....	0.032
P2 (Control Proportion) .....	0.05
Variance Calculation Formula .....	$\sigma^2 = P2(1 - P2)$
Use Calculated Variance as .....	Within-Cluster Variance ( $\sigma^2 = \sigma w^2$ )
Between-Cluster Variability Input Type .....	COV
COV (Coefficient of Variation of Outcomes) .....	0.3
<b>Reports Tab</b>	
ICC, COV, and Variances Decimals .....	4

## 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:  $P1/P2 = 1$ . H1:  $P1/P2 = R1 \neq 1$ .

Power	Design S/T/R	Clusters K	Cluster Size M/m	Sample Size N	Trt Prop P1	Ctrl Prop P2	Ratio R1	COV	Alpha
0.77393	4/5/6	24	500/100	12000	0.0320	0.0500	-0.0180	0.3000	0.050

PASS calculated a value of 0.77393 for the power, which matches the statement in Hussey and Hughes (2007) that the power is “about 80%”.

## Example 2 – Finding Power of a Complete Design (Validation using Hemming and Girling (2014))

Hemming and Girling (2014) presents an example where power is calculated for a complete design where  $P_1 = 0.5$ ,  $P_2 = 0.4$  (Note that in the paper, Proportion 1 is the standard proportion and Proportion 2 is the treatment proportion, so they are reversed),  $\alpha = 0.05$ ,  $\text{ICC} = 0.01$ ,  $m = 12$ ,  $R = 1$ , and  $T = 11$  (Note that they state that the number of time periods not including the baseline is 10). They use  $\sigma^2 = P_2(1 - P_2)$  as the variance calculation formula and assume it is the total variance.

They compute power for this scenario to be 0.6998, a total sample size of 1320, a  $\tau^2$  of 0.0024, a COV of 0.12, and 10 steps in the final 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 Proportions 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>
<b>Design Tab</b>	
Solve For .....	<b>Power</b>
Alternative Hypothesis .....	<b>Two-Sided</b>
Alpha .....	<b>0.05</b>
Design Type .....	<b>Complete</b>
Design Parameter Entry Type .....	<b>Time Periods (T) &amp; Switches per Step (R)</b>
T (Number of Time Periods including Baseline) .....	<b>11</b>
R (Number of Clusters Switching at each Step) .....	<b>1</b>
Cluster Size Entry Type .....	<b>Subjects per Cluster per Time Period (m)</b>
m (Ave. Subjects per Cluster per Time Period) .....	<b>12</b>
Input Type .....	<b>Proportions</b>
P1 (Treatment Proportion) .....	<b>0.5</b>
P2 (Control Proportion) .....	<b>0.4</b>
Variance Calculation Formula .....	<b><math>\sigma^2 = P_2(1 - P_2)</math></b>
Use Calculated Variance as .....	<b>Total Variance (<math>\sigma^2 = \sigma_y^2 = \tau^2 + \sigma_w^2</math>)</b>
Between-Cluster Variability Input Type .....	<b>ICC</b>
ICC (Intraclass Correlation Coefficient) .....	<b>0.01</b>
<b>Reports Tab</b>	
ICC, COV, and Variances Decimals .....	<b>4</b>



## Tests for Two Proportions 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:  $P1 - P2 = 0$ . H1:  $P1 - P2 = D1 \neq 0$ .

Power	Design S/T/R	Clusters K	Cluster Size M/m	Sample Size N	Trt Prop P1	Ctrl Prop P2	Diff D1	ICC	Alpha
0.69978	10/11/1	10	132/12	1320	0.5000	0.4000	0.1000	0.0100	0.050

## Design 1 Details

Design Type: Complete

Power	Design S/T/R	Clusters K	Cluster Size M/m	Trt Prop P1	Ctrl Prop P2	Total Variance $\sigma^2$	Between-Cluster Variance $\tau^2$	Within-Cluster Variance $\sigma_w^2$	ICC	COV
0.69978	10/11/1	10	132/12	0.5000	0.4000	0.2400	0.0024	0.2376	0.0100	0.1225

## Design Pattern (Complete)

Cluster	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11
1	0	1	1	1	1	1	1	1	1	1	1
2	0	0	1	1	1	1	1	1	1	1	1
3	0	0	0	1	1	1	1	1	1	1	1
4	0	0	0	0	1	1	1	1	1	1	1
5	0	0	0	0	0	1	1	1	1	1	1
6	0	0	0	0	0	0	1	1	1	1	1
7	0	0	0	0	0	0	0	1	1	1	1
8	0	0	0	0	0	0	0	0	1	1	1
9	0	0	0	0	0	0	0	0	0	1	1
10	0	0	0	0	0	0	0	0	0	0	1

The power value of 0.69978 calculated by **PASS** matches the result of Hemming and Girling (2014) exactly, as do the values for N,  $\tau^2$ , COV, and S. The design pattern shows exactly what the design looks like.

## Example 3 – Finding Power of an Incomplete Design with an Implementation Phase (Validation using Hemming, Lilford, and Girling (2014))

Hemming, Lilford, and Girling (2014) presents an example on page 190-191 that calculates power for an incomplete design with an implementation phase. The study design aims to evaluate whether a training scheme improves the rate at which midwives perform membrane sweeping in post-term pregnancies. Membrane sweeping has been shown to reduce the need for formal labor induction. The design involves 10 community teams (clusters) with 12 births per team per week. The rollout is staggered between 1 and 3 weeks between cluster starts. The proportions are  $P_1 = 0.5$ ,  $P_2 = 0.4$ , with  $\text{Alpha} = 0.05$  and  $\text{ICC} = 0.01$ . It is unclear what formula they use for variance, so we'll calculate using both  $\sigma^2 = P_2(1 - P_2)$  and  $\sigma^2 = [P_1(1 - P_1) + P_2(1 - P_2)]/2$  in this example. Both will be assumed to be the total variance.

The design pattern matrix shown on page 191 is as follows:

		Week																																																		
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39												
Team (Cluster)	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	•	1	1	1	1	1	1	1	1	1	1	1	1																								
	2			0	0	0	0	0	0	0	0	0	0	0	0	0	•	1	1	1	1	1	1	1	1	1	1	1	1	1																						
	3				0	0	0	0	0	0	0	0	0	0	0	0	0	•	1	1	1	1	1	1	1	1	1	1	1	1	1	1																				
	4					0	0	0	0	0	0	0	0	0	0	0	0	0	•	1	1	1	1	1	1	1	1	1	1	1	1	1	1																			
	5						0	0	0	0	0	0	0	0	0	0	0	0	0	•	1	1	1	1	1	1	1	1	1	1	1	1	1	1																		
	6							0	0	0	0	0	0	0	0	0	0	0	0	0	•	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1																
	7								0	0	0	0	0	0	0	0	0	0	0	0	0	•	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1													
	8									0	0	0	0	0	0	0	0	0	0	0	0	0	•	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1						
	9										0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	•	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
	10																	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

They state that the power for this scenario is “in the region of 78%”, with a total sample size (N) of 2880.

### 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 Proportions in a Stepped-Wedge Cluster-Randomized Design** procedure window. You may then make the appropriate entries as listed below, or open **Example 3a** by going to the **File** menu and choosing **Open Example Template**.

Option	Value
<b>Design Tab</b>	
Solve For .....	<b>Power</b>
Alternative Hypothesis .....	<b>Two-Sided</b>
Alpha.....	<b>0.05</b>
Design Type.....	<b>Incomplete (Custom)</b>
Design Parameter Entry Type .....	<b>Custom Design Pattern Matrix</b>
Design Pattern Matrix Columns.....	<b>ALL</b>
R (Number of Design Pattern Replicates) .....	<b>1</b>
Cluster Size Entry Type .....	<b>Subjects per Cluster per Time Period (m)</b>

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

Design Tab (Continued)

m (Ave. Subjects per Cluster per Time Period).....12  
 Input Type.....Proportions  
 P1 (Treatment Proportion).....0.5  
 P2 (Control Proportion).....0.4  
 Variance Calculation Formula ..... $\sigma^2 = P2(1 - P2)$   
 Use Calculated Variance as .....Total Variance ( $\sigma^2 = \tau^2 + \sigma w^2$ )  
 Between-Cluster Variability Input Type .....ICC  
 ICC (Intracluster Correlation Coefficient).....0.01

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: P1 - P2 = 0. H1: P1 - P2 = D1 ≠ 0.

Power	Design S/T/R	Clusters K	Cluster Size M/m	Sample Size N	Trt Prop P1	Ctrl Prop P2	Diff D1	ICC	Alpha
0.78826	38/39/1	10	288/12	2880	0.5000	0.4000	0.1000	0.010	0.050

The power value of 0.78826 and total sample size of N = 2880 calculated by **PASS** matches the results in Hemming, Lilford, and Girling (2014).

Now, change the variance calculation formula to  $\sigma^2 = [P1(1 - P1) + P2(1 - P2)]/2$  or open **Example 3b** by going to the **File** menu and choosing **Open Example Template**.

Option Value

Design Tab

Variance Calculation Formula ..... $\sigma^2 = [P1(1 - P1) + P2(1 - P2)]/2$

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: P1 - P2 = 0. H1: P1 - P2 = D1 ≠ 0.

Power	Design S/T/R	Clusters K	Cluster Size M/m	Sample Size N	Trt Prop P1	Ctrl Prop P2	Diff D1	ICC	Alpha
0.77997	38/39/1	10	288/12	2880	0.5000	0.4000	0.1000	0.010	0.050

The power value of 0.77997 and total sample size of N = 2880 calculated by **PASS** also matches the results in Hemming, Lilford, and Girling (2014).

## 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 probability,  $P_2$ , is 0.26 and the odds ratio,  $OR_1$ , is 0.56. 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. Baio et al. (2015) uses the variance calculation formula of  $\sigma^2 = [(P_1 + P_2)/2] \times [1 - (P_1 + P_2)/2]$  and uses it as the total variance. 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 10, 13, 12, 11, 10, and 8 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 Proportions in a Stepped-Wedge Cluster-Randomized Design** procedure window. You may then make the appropriate entries as listed below, or open **Example 4a** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Sample Size [K (Number of Clusters)]</b>
Alternative Hypothesis .....	<b>Two-Sided</b>
Power.....	<b>0.80</b>
Alpha.....	<b>0.05</b>
Design Type.....	<b>Incomplete (Custom)</b>
Design Constraint Entry Type.....	<b>Fixed Number of Time Periods (T)</b>
T (Number of Time Periods including Baseline).....	<b>6</b>
Cluster Size Entry Type .....	<b>Subjects per Cluster per Time Period (m)</b>
m (Ave. Subjects per Cluster per Time Period).....	<b>20</b>
Input Type.....	<b>Odds Ratios</b>
OR1 (Odds Ratio = O1/O2) .....	<b>0.56</b>
P2 (Control Proportion).....	<b>0.26</b>
Variance Calculation Formula .....	<b><math>\sigma^2 = [(P_1 + P_2)/2] \times [1 - (P_1 + P_2)/2]</math></b>
Use Calculated Variance as .....	<b>Total Variance (<math>\sigma^2 = \sigma_y^2 = \tau^2 + \sigma_w^2</math>)</b>
Between-Cluster Variability Input Type .....	<b>ICC</b>
ICC (Intracluster Correlation Coefficient).....	<b>0 to 0.5 by 0.1</b>

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

## Output

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

### Numeric Results

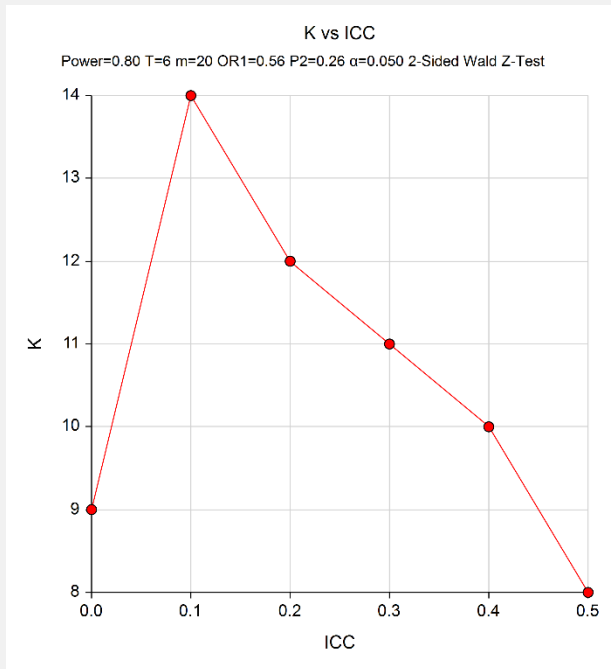
Design Type: Incomplete (Custom)

Test Statistic: Wald Z-Test

H0: O1/O2 = 1. H1: O1/O2 = OR1 ≠ 1.

Power	Design S/T	Clusters K	Cluster Size M/m	Sample Size N	Trt Prop P1	Ctrl Prop P2	O.R. OR1	ICC	Alpha
0.81965	5/6	9	120/20	1080	0.1644	0.2600	0.5600	0.000	0.050
0.82622	5/6	14	120/20	1680	0.1644	0.2600	0.5600	0.100	0.050
0.80496	5/6	12	120/20	1440	0.1644	0.2600	0.5600	0.200	0.050
0.81516	5/6	11	120/20	1320	0.1644	0.2600	0.5600	0.300	0.050
0.83368	5/6	10	120/20	1200	0.1644	0.2600	0.5600	0.400	0.050
0.81935	5/6	8	120/20	960	0.1644	0.2600	0.5600	0.500	0.050

### Chart Section



### Design 1 Details

Design Type: Incomplete (Custom)

Power	Design S/T	Clusters K	Cluster Size M/m	Trt Prop P1	Ctrl Prop P2	Total Variance $\sigma_y^2$	Between-Cluster Variance $\tau^2$	Within-Cluster Variance $\sigma_w^2$	ICC	COV
0.81965	5/6	9	120/20	0.1644	0.2600	0.167	0.000	0.167	0.000	0.000

Tests for Two Proportions 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 9, 14, 12, 11, 10, and 8, which match those given in Table 1 of Baio et al. (2015). There are two differences, however, when ICC = 0 and ICC = 0.1. When ICC = 0, Baio et al. (2015) reports a value of 10, while PASS computes a value of 8. When ICC = 0.1, Baio et al. (2015) reports a value of 13, while PASS computes a value of 14. These 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 0.81965, for example, indicates that K = 9 with ICC = 0 does achieve the desired power and is, therefore, the correct balanced-design solution.

It’s interesting to note that if you search among unbalanced design pattern matrices in PASS, the reported number of clusters for ICC = 0 and ICC = 0.1 are even lower... 8 and 13, respectively. The value of 13 for ICC = 0.1 then matches Table 1 of Baio et al. (2015), but the value for ICC = 0 is lower and even further from the Table 1 result of 10. To replicate this, you may make the appropriate changes as listed below, or open **Example 4b** by going to the **File** menu and choosing **Open Example Template**.

Option Value

**Options Tab**

Design Pattern Assignment Type for Extra Clusters ..... **Unbalanced**

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: O1/O2 = 1. H1: O1/O2 = OR1 ≠ 1.

Power	Design S/T	Clusters K	Cluster Size M/m	Sample Size N	Trt Prop P1	Ctrl Prop P2	O.R. OR1	ICC	Alpha
0.80381	5/6	8	120/20	960	0.1644	0.2600	0.5600	0.000	0.050
0.80057	5/6	13	120/20	1560	0.1644	0.2600	0.5600	0.100	0.050
0.80496	5/6	12	120/20	1440	0.1644	0.2600	0.5600	0.200	0.050
0.81516	5/6	11	120/20	1320	0.1644	0.2600	0.5600	0.300	0.050
0.83368	5/6	10	120/20	1200	0.1644	0.2600	0.5600	0.400	0.050
0.81935	5/6	8	120/20	960	0.1644	0.2600	0.5600	0.500	0.050

PASS computed K values of 8, 13, 12, 11, 10, and 8, which match those given in Table 1 of Baio et al. (2015), except for ICC = 0. The power value of 0.80381, however, indicates that K = 8 with ICC = 0 does achieve the desired power and is, therefore, the correct unbalanced-design solution.

## Example 5 – Finding Effect Size for an Incomplete Design (Validation using Hemming and Girling (2014))

Hemming and Girling (2014) presents an example (Example 2) where they calculate the effect size required to achieve 80% power in an incomplete design with an alpha level of 0.05 when  $P_2 = 0.4$  (Note that in the paper, Proportion 1 is the standard proportion and Proportion 2 is the treatment proportion, so they are reversed),  $ICC = 0.01$ , and  $m = 12$ . They use  $\sigma^2 = P_2(1 - P_2)$  as the variance calculation formula and assume it is the total variance.

The design pattern matrix shown on page 375 is as follows:

		Week																						
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
Team (Cluster)	1	0	1	1	1	1	1	1	1	1	1	1	1	1										
	2	0	0	1	1	1	1	1	1	1	1	1	1	1	1									
	3	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1								
	4	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1							
	5	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1						
	6	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1					
	7	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1			
	8	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1		
	9	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	
	10	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1

They compute the detectable difference to be 0.1096 ( $P_1 = 0.5096$  if searching above  $P_2$  and  $P_1 = 0.2904$  if searching below  $P_2$ ) and a total sample size of 2100.

### 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 Proportions 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>
<b>Design Tab</b>	
Solve For .....	<b>Effect Size (P1, D1, R1, OR1)</b>
Alternative Hypothesis .....	<b>Two-Sided</b>
Power .....	<b>0.80</b>
Alpha .....	<b>0.05</b>
Design Type .....	<b>Incomplete (Custom)</b>
Design Parameter Entry Type .....	<b>Custom Design Pattern Matrix</b>
Design Pattern Matrix Columns .....	<b>ALL</b>
R (Number of Design Pattern Replicates) .....	<b>1</b>
Cluster Size Entry Type .....	<b>Subjects per Cluster per Time Period (m)</b>
m (Ave. Subjects per Cluster per Time Period) .....	<b>12</b>

## Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

## Design Tab (Continued)

Input Type.....Proportions  
 P1 (Treatment Proportion).....Search > P2  
 P2 (Control Proportion).....0.4  
 Variance Calculation Formula ..... $\sigma^2 = P2(1 - P2)$   
 Use Calculated Variance as .....Total Variance ( $\sigma^2 = \sigma\gamma^2 = \tau^2 + \sigma w^2$ )  
 Between-Cluster Variability Input Type .....ICC  
 ICC (Intracluster Correlation Coefficient).....0.01

## 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: P1 - P2 = 0. H1: P1 - P2 = D1 $\neq$ 0.									
Power	Design S/T/R	Clusters K	Cluster Size M/m	Sample Size N	Trt Prop P1	Ctrl Prop P2	Diff D1	ICC	Alpha
0.80000	21/22/1	10	210/12	2100	0.5096	0.4000	0.1096	0.010	0.050

The detectable difference value of D1 = 0.1096, treatment proportion of 0.5096, and total sample size of N = 2100 calculated by **PASS** all match the results in Hemming and Girling (2014) exactly.

Now, change the search direction for P1 to Search < P2 or open **Example 5b** by going to the **File** menu and choosing **Open Example Template**.

### Option

### Value

#### Design Tab

P1 (Treatment Proportion).....Search < P2

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: P1 - P2 = 0. H1: P1 - P2 = D1 $\neq$ 0.									
Power	Design S/T/R	Clusters K	Cluster Size M/m	Sample Size N	Trt Prop P1	Ctrl Prop P2	Diff D1	ICC	Alpha
0.80000	21/22/1	10	210/12	2100	0.2904	0.4000	-0.1096	0.010	0.050

The detectable difference value of D1 = -0.1096, treatment proportion of 0.2904, and total sample size of N = 2100 calculated by **PASS** all match the results in Hemming and Girling (2014) exactly.



## Example 6 – Finding Effect Size for an Incomplete Design with a Transition Period (Validation using Hemming and Girling (2014))

Hemming and Girling (2014) presents an example (Example 3) where they calculate the effect size required to achieve 80% power with an alpha level of 0.05 in an incomplete design that has a transition period when  $P_2 = 0.12$  (Note that in the paper, Proportion 1 is the standard proportion and Proportion 2 is the treatment proportion, so they are reversed),  $COV = 0.3$ , and  $m = 1250$ . They use  $\sigma^2 = P_2(1 - P_2)$  as the variance calculation formula and assume it is the total variance.

The design pattern matrix shown on page 376 is as follows:

		Month							
		0-3	3-6	6-9	9-12	12-15	15-18	18-21	21-24
Hospital (Cluster)	1	0	0	•	1	1			
	2	0	0	•	1	1			
	3	0	0	•	1	1			
	4		0	0	•	1	1		
	5		0	0	•	1	1		
	6		0	0	•	1	1		
	7			0	0	•	1	1	
	8			0	0	•	1	1	
	9			0	0	•	1	1	
	10				0	0	•	1	1
	11				0	0	•	1	1
	12				0	0	•	1	1

They compute the detectable difference to be 0.0241 ( $P_1 = 0.1441$  if searching above  $P_2$  and  $P_1 = 0.0959$  if searching below  $P_2$ ) and a total sample size of 60000.

### 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 Proportions in a Stepped-Wedge Cluster-Randomized Design** procedure window. You may then make the appropriate entries as listed below, or open **Example 6a** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Effect Size (P1, D1, R1, OR1)</b>
Alternative Hypothesis .....	<b>Two-Sided</b>
Power .....	<b>0.80</b>
Alpha .....	<b>0.05</b>
Design Type .....	<b>Incomplete (Custom)</b>
Design Parameter Entry Type .....	<b>Custom Design Pattern Matrix</b>
Design Pattern Matrix Columns .....	<b>ALL</b>
R (Number of Design Pattern Replicates) .....	<b>3</b>

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

Design Tab (Continued)

Cluster Size Entry Type .....Subjects per Cluster per Time Period (m)  
 m (Ave. Subjects per Cluster per Time Period) .....1250  
 Input Type .....Proportions  
 P1 (Treatment Proportion) .....Search > P2  
 P2 (Control Proportion) .....0.12  
 Variance Calculation Formula ..... $\sigma^2 = P2(1 - P2)$   
 Use Calculated Variance as .....Total Variance ( $\sigma^2 = \tau^2 + \sigma w^2$ )  
 Between-Cluster Variability Input Type .....COV  
 COV (Coefficient of Variation of Outcomes) .....0.3

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: P1 - P2 = 0. H1: P1 - P2 = D1 ≠ 0.

Power	Design S/T/R	Clusters K	Cluster Size M/m	Sample Size N	Trt Prop P1	Ctrl Prop P2	Diff D1	COV	Alpha
0.80000	7/8/3	12	5000/1250	60000	0.1441	0.1200	0.0241	0.300	0.050

The detectable difference value of D1 = 0.0241, treatment proportion of 0.1441, and total sample size of N = 60000 calculated by **PASS** all match the results in Hemming and Girling (2014) exactly.

Now, change the search direction for P1 to Search < P2 or open **Example 6b** by going to the **File** menu and choosing **Open Example Template**.

Option Value

Design Tab

P1 (Treatment Proportion) .....Search < P2

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: P1 - P2 = 0. H1: P1 - P2 = D1 ≠ 0.

Power	Design S/T/R	Clusters K	Cluster Size M/m	Sample Size N	Trt Prop P1	Ctrl Prop P2	Diff D1	COV	Alpha
0.80000	7/8/3	12	5000/1250	60000	0.0959	0.1200	-0.0241	0.300	0.050

The detectable difference value of D1 = -0.0241, treatment proportion of 0.0959, and total sample size of N = 60000 calculated by **PASS** all match the results in Hemming and Girling (2014) exactly.

## Example 7 – Finding Power for an Incomplete Design with a Delayed Treatment Effect

Continuing with the situation described in Example 1 above, Hussey and Hughes (2007) presents an example in section 3.6 that demonstrates power calculations with a delayed treatment effect. A delayed treatment effect occurs when the full effect of the treatment does not occur in the same time period where it is introduced. Delayed treatment effects are modeled with fractional numbers in the design pattern matrix.

Motivated by the example in Hussey and Hughes (2007) but not using the same design, in this example we'll investigate the power achieved by the following design pattern matrix, with each row replicated 6 times.

		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

What is the power for the parameters  $P2 = 0.05$ ,  $R1 = P1/P2 = 0.7$ ,  $COV = 0.02$  to  $0.5$ , and  $Alpha = 0.05$ ?

### 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 Proportions in a Stepped-Wedge Cluster-Randomized Design** procedure window. You may then make the appropriate entries as listed below, or open **Example 7a** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Power</b>
Alternative Hypothesis .....	<b>Two-Sided</b>
Alpha .....	<b>0.05</b>
Design Type .....	<b>Incomplete (Custom)</b>
Design Parameter Entry Type .....	<b>Custom Design Pattern Matrix</b>
Design Pattern Matrix Columns .....	<b>ALL</b>
R (Number of Design Pattern Replicates) .....	<b>6</b>
Cluster Size Entry Type .....	<b>Subjects per Cluster per Time Period (m)</b>
m (Ave. Subjects per Cluster per Time Period) .....	<b>100</b>
Input Type .....	<b>Ratios</b>
R1 (Ratio = $P1/P2$ ) .....	<b>0.7</b>
P2 (Control Proportion) .....	<b>0.05</b>
Variance Calculation Formula .....	<b><math>\sigma^2 = P2(1 - P2)</math></b>
Use Calculated Variance as .....	<b>Within-Cluster Variance (<math>\sigma^2 = \sigma w^2</math>)</b>
Between-Cluster Variability Input Type .....	<b>COV</b>
COV (Coefficient of Variation of Outcomes) .....	<b>0.02 to 0.5 by 0.02</b>

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

## Output

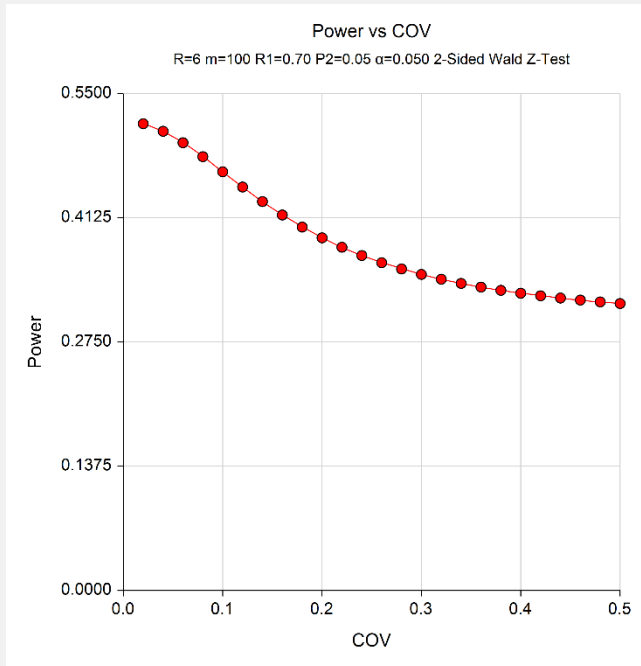
Click the Calculate button to perform the calculations and generate the following output.

### Numeric Results

Design Type: Incomplete (Custom)  
 Test Statistic: Wald Z-Test  
 H0: P1/P2 = 1. H1: P1/P2 = R1 ≠ 1.

Power	Design S/T/R	Clusters K	Cluster Size M/m	Sample Size N	Trt Prop P1	Ctrl Prop P2	Ratio R1	COV	Alpha
0.51663	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.020	0.050
0.50827	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.040	0.050
0.49562	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.060	0.050
0.48017	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.080	0.050
0.46341	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.100	0.050
0.44656	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.120	0.050
0.43046	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.140	0.050
0.41561	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.160	0.050
0.40222	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.180	0.050
0.39034	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.200	0.050
0.37990	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.220	0.050
0.37077	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.240	0.050
0.36280	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.260	0.050
0.35586	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.280	0.050
0.34980	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.300	0.050
0.34451	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.320	0.050
0.33987	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.340	0.050
0.33579	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.360	0.050
0.33220	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.380	0.050
0.32901	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.400	0.050
0.32619	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.420	0.050
0.32368	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.440	0.050
0.32143	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.460	0.050
0.31942	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.480	0.050
0.31761	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.500	0.050

### Chart Section



The results display the power achieved by the design for values of COV between 0.02 and 0.5. The power ranges from about 0.32 to 0.52.

**Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design**

To compare this against the power without a delayed treatment effect, go back and change the custom design pattern matrix by replacing all fractional numbers by 1's or open **Example 7b** by going to the **File** menu and choosing **Open Example Template**.

		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

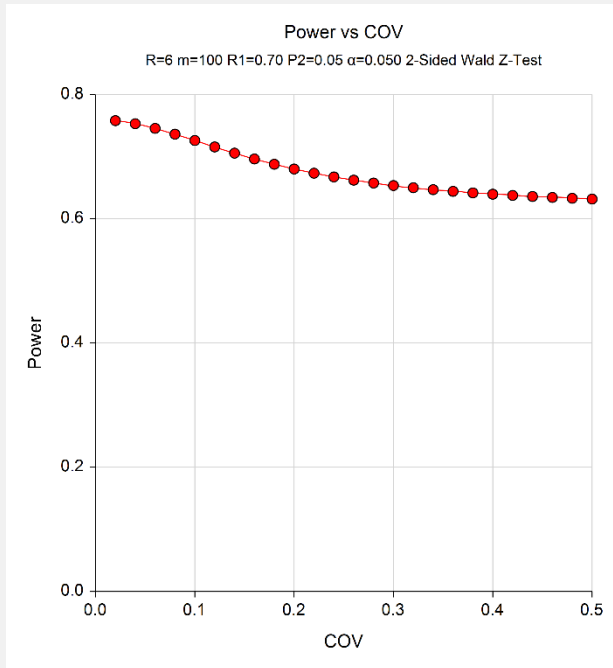
Click the Calculate button to perform the calculations and generate the following output.

**Numeric Results**  
 Design Type: Incomplete (Custom)  
 Test Statistic: Wald Z-Test  
 H0: P1/P2 = 1. H1: P1/P2 = R1 ≠ 1.

Power	Design S/T/R	Clusters K	Cluster Size M/m	Sample Size N	Trt Prop P1	Ctrl Prop P2	Ratio R1	COV	Alpha
0.75806	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.020	0.050
0.75312	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.040	0.050
0.74558	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.060	0.050
0.73629	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.080	0.050
0.72609	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.100	0.050
0.71572	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.120	0.050
0.70570	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.140	0.050
0.69634	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.160	0.050
0.68781	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.180	0.050
0.68018	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.200	0.050
0.67340	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.220	0.050
0.66743	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.240	0.050
0.66219	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.260	0.050
0.65760	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.280	0.050
0.65356	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.300	0.050
0.65002	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.320	0.050
0.64691	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.340	0.050
0.64416	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.360	0.050
0.64173	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.380	0.050
0.63957	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.400	0.050
0.63765	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.420	0.050
0.63594	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.440	0.050
0.63440	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.460	0.050
0.63303	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.480	0.050
0.63179	6/7/6	24	700/100	16800	0.0350	0.0500	0.7000	0.500	0.050

## Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

## Chart Section



The power ranges from about 0.63 to 0.76. As you can see, the power is much lower in the design with the delayed treatment effect than it is with this design.