

## Chapter 441

# Tests for Two Poisson Rates 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 *2x2 cluster-randomized cross-over design* as one in which each cluster receives both the treatment and control. The objective is to study difference between the two. Each cluster crosses over from the treatment group to the control group (or vice-versa). It is assumed that there is a washout period between applications during which the response returns back to its baseline value. If this does not occur, there is said to be a carry-over effect.

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

This procedure computes power and sample size for tests for the difference between two Poisson rates 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 Poisson Rates 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		Time		Time						
		1				1	2	3	4	5	6	7
	1	1	1	1	0	0	1	1	1	1	1	1
	2	1	2	1	0	0	0	1	1	1	1	1
	3	1	3	1	0	0	0	0	1	1	1	1
	4	0	4	0	1	0	0	0	0	1	1	1
	5	0	5	0	1	0	0	0	0	0	1	1
	6	0	6	0	1	0	0	0	0	0	0	1

## Stepped-Wedge Cluster-Randomized Designs

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

## Complete

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

$$K = S \times R.$$

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

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

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

## 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
$\lambda_1$	The treatment group Poisson rate, assuming the alternative hypothesis
$\lambda_2$	The control, standard, reference, or baseline group Poisson rate

## 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

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

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

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

with

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

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 Poisson rates, with  $\mu_1 = \lambda_1$  and  $\mu_2 = \lambda_2$  as the treatment and control group rates, respectively, then

$$\theta = D1 = \lambda_1 - \lambda_2.$$

## 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

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

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

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

with

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

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.

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

**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  $\mu_2 = \lambda_2$  is the control group mean (i.e., a Poisson rate for count data), then the COV of outcomes for the control group is calculated as

$$COV = \frac{\tau}{\mu_2} = \frac{\tau}{\lambda_2}$$

and can be used along with  $\mu_2 = \lambda_2$  to calculate  $\tau^2$  as

$$\tau^2 = COV^2 \times \mu_2^2 = COV^2 \times \lambda_2^2.$$

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

In the case of testing two Poisson rates where the responses,  $Y_{ikt}$ , are counts and using the normal approximation, the variance can be calculated using three different methods, given as options in this procedure.

Method 1:	Null Variance	$\sigma^2 = \lambda_2$
Method 2:	Group Variance Average	$\sigma^2 = \frac{\lambda_1 + \lambda_2}{2}$
Method 3:	Squared Group Standard Deviation Average	$\sigma^2 = \left( \frac{\sqrt{\lambda_1} + \sqrt{\lambda_2}}{2} \right)^2$

Method 3 is used in Hemming and Girling (2014) and Baio et al. (2015).

The variance that is calculated from the two rates,  $\sigma^2$ , 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

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

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

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

$$\tau^2 = COV^2 \times \mu_2^2 = COV^2 \times \lambda_2^2.$$

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

$$\tau^2 = COV^2 \times \mu_2^2 = COV^2 \times \lambda_2^2.$$

## Test Statistic

Hussey and Hughes (2007) suggest that a statistical hypothesis test of  $H_0: \theta = \lambda_1 - \lambda_2 = 0$  vs.  $H_1: \theta = \lambda_1 - \lambda_2 \neq 0$  can be conducted using an asymptotic Wald test. The test statistic is

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

$\hat{\theta}$  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

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

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

$$\begin{aligned} \text{Power} &= \Phi \left( \frac{\theta_A}{\sqrt{\text{Var}(\hat{\theta})}} - Z_{1-\alpha/2} \right) \\ &= \Phi \left( \frac{\lambda_1 - \lambda_2}{\sqrt{\{(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\}}} - Z_{1-\alpha/2} \right) \end{aligned}$$

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.,  $\theta_A = \lambda_1 - \lambda_2$ .

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.)



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

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

- **Balanced**

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

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

- **Unbalanced**

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

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

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

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.

## Example 1 – Finding Power of a Complete Design

Hemming and Girling (2014) presents an example in Section 4.4 where the goal is to evaluate an alternative to do-not-resuscitate with another approach called “the universal form of treatment.” A total of 20 wards are to be involved in the study, with each acting as a cluster. The trial will run for 12 10-week periods with 2 wards randomized with two wards randomized to the new treatment at each period. The example includes a transition period between treatment and control applications, making it an incomplete design. For the purposes of this example, however, we’ll eliminate the transition period and assume only 11 10-week time periods, making the design complete.

The effectiveness of the new treatment is to be evaluated based on the “number of harms.” The baseline rate is expected to be 21 harms per 1,000 patient days ( $\lambda_2 = 0.021$ ). The study aims to detect a 25% reduction in relative risk, such that  $RR_1 = 0.75$ . If we assume an ICC of 0.007, what power is achieved with an alpha level of 0.05 for cluster sizes for each time period ( $m$ ) between 200 and 300 (i.e., between 20 and 30 patients per cluster per time period, each with an average length of stay of 10 days)?

### Setup

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

#### Design Tab

Solve For .....	<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; Time Periods (T)</b>
K (Number of Clusters) .....	<b>20</b>
T (Number of Time Periods including Baseline).....	<b>11</b>
Cluster Size Entry Type .....	<b>Subjects per Cluster per Time Period (m)</b>
m (Ave. Subjects per Cluster per Time Period).....	<b>200 to 300 by 10</b>
Input Type.....	<b>Rate Ratios</b>
RR1 (Rate Ratio = $\lambda_1/\lambda_2$ ).....	<b>0.75</b>
$\lambda_2$ (Control Event Rate) .....	<b>0.021</b>
Variance Calculation Formula.....	<b><math>\sigma^2 = ((\sqrt{[\lambda_1]} + \sqrt{[\lambda_2]})/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 (Intraclass Correlation Coefficient) .....	<b>0.007</b>

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

## Output

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

## Numeric Reports

## Numeric Results

Solve For: **Power**  
 Design Type: Complete  
 Groups: 1 = Treatment, 2 = Control  
 Test Statistic: Wald Z-Test  
 Hypotheses:  $H_0: \lambda_1 / \lambda_2 = 1$  vs.  $H_1: \lambda_1 / \lambda_2 \neq 1$

Power	Design Parameters			Number of Clusters K	Cluster Size		Sample Size N	Rate		Ratio RR1	Intraclass Correlation Coefficient ICC	Alpha
	S	T	R		M	m		Treatment $\lambda_1$	Control $\lambda_2$			
0.66869	10	11	2	20	2200	200	44000	0.0158	0.021	0.75	0.007	0.05
0.68893	10	11	2	20	2310	210	46200	0.0158	0.021	0.75	0.007	0.05
0.70818	10	11	2	20	2420	220	48400	0.0158	0.021	0.75	0.007	0.05
0.72645	10	11	2	20	2530	230	50600	0.0158	0.021	0.75	0.007	0.05
0.74377	10	11	2	20	2640	240	52800	0.0158	0.021	0.75	0.007	0.05
0.76017	10	11	2	20	2750	250	55000	0.0158	0.021	0.75	0.007	0.05
0.77569	10	11	2	20	2860	260	57200	0.0158	0.021	0.75	0.007	0.05
0.79035	10	11	2	20	2970	270	59400	0.0158	0.021	0.75	0.007	0.05
0.80418	10	11	2	20	3080	280	61600	0.0158	0.021	0.75	0.007	0.05
0.81722	10	11	2	20	3190	290	63800	0.0158	0.021	0.75	0.007	0.05
0.82951	10	11	2	20	3300	300	66000	0.0158	0.021	0.75	0.007	0.05

Note: The variance is calculated as  $\sigma^2 = ((\sqrt{\lambda_1} + \sqrt{\lambda_2})/2)^2$  and is considered to be the total variance ( $\sigma^2 = \sigma_y^2$ ) in power computations.

Power The probability of rejecting a false null hypothesis when the alternative hypothesis is true.  
 S The number of steps in the study design.  $S = T - 1$ .  
 T The number of time periods in the study, including the baseline.  $T = S + 1$ .  
 R The number clusters switching from control to treatment at each step.  
 K The total number of clusters to be randomized.  $K = S \times R$ .  
 M The average number of subjects per cluster.  $M = m \times T$ .  
 m The average number of subjects per cluster per time period.  $M = m \times T$ .  
 N The total sample size from all clusters and time periods combined.  $m = M / T$ .  
 $\lambda_1$  The average treatment event rate per unit time, assuming the alternative hypothesis.  
 $\lambda_2$  The average control, standard, reference, or baseline event rate per unit time.  
 RR1 The rate ratio assuming the alternative hypothesis ( $H_1$ ).  $RR1 = \lambda_1 / \lambda_2$ .  
 ICC The intraclass correlation coefficient.  
 Alpha The probability of rejecting a true null hypothesis.

## Summary Statements

A complete stepped-wedge cluster-randomized design with 11 time periods (including the baseline) and 10 steps (with 2 clusters switching from control to treatment at each step) will be used to test whether the Group 1 (treatment) event rate ( $\lambda_1$ ) is different from the Group 2 (control) event rate ( $\lambda_2$ ) ( $H_0: \lambda_1 / \lambda_2 = 1$  versus  $H_1: \lambda_1 / \lambda_2 \neq 1$ ). The comparison will be made using a two-sided Wald Z-test based on the event rate difference with variance calculation formula  $\sigma^2 = ((\sqrt{\lambda_1} + \sqrt{\lambda_2})/2)^2$  (considered to be the total variance), and with a Type I error rate ( $\alpha$ ) of 0.05. The control group event rate ( $\lambda_2$ ) is assumed to be 0.021. The intraclass correlation coefficient is assumed to be 0.007. To detect an event rate ratio ( $\lambda_1 / \lambda_2$ ) of 0.75 (or  $\lambda_1$  of 0.0158) with 20 clusters with average cluster sizes of 2200 subjects per cluster and an average of 200 subjects per cluster per time period (for a total sample size of 44000 subjects), the power is 0.66869.

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

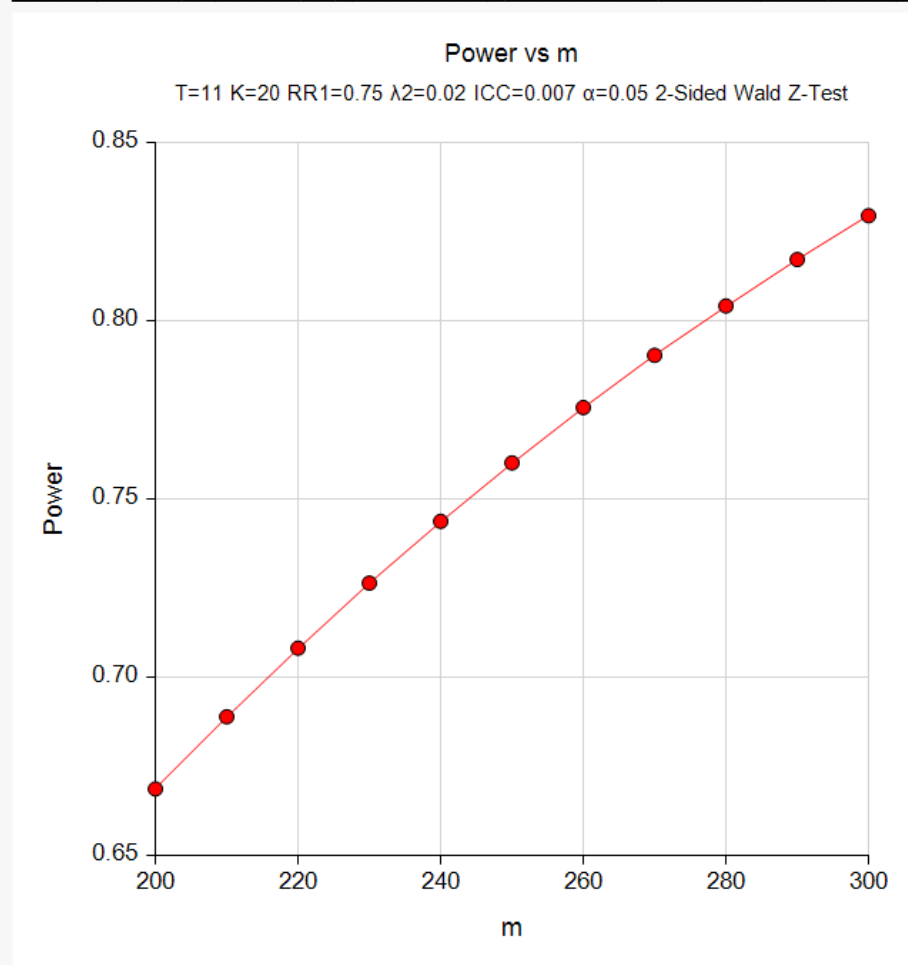
## References

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

This report shows the values of each of the parameters, one scenario per row. The values from this table are shown in the plot below.

## Plots Section

## Plots



The values from the table are displayed on the plot. This plot gives a quick look at the power that is achieved for various values of  $m$ .

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

## Design Details

## Design 1 Details

Design Type: Complete

Power	Design Parameters			Number of Clusters K	Cluster Size		Rate		Variance			Intraclass Correlation Coefficient ICC	Coefficient of Variation COV
	S	T	R		M	m	Treatment $\lambda_1$	Control $\lambda_2$	Total $\sigma^2$	Between-Cluster $\tau^2$	Within-Cluster $\sigma^2$		
0.66869	10	11	2	20	2200	300	0.0158	0.021	0.018	0	0.018	0.007	0.539

## 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	1	1	1	1	1	1	1	1	1	1
3	0	0	1	1	1	1	1	1	1	1	1
4	0	0	1	1	1	1	1	1	1	1	1
5	0	0	0	1	1	1	1	1	1	1	1
6	0	0	0	1	1	1	1	1	1	1	1
7	0	0	0	0	1	1	1	1	1	1	1
8	0	0	0	0	1	1	1	1	1	1	1
9	0	0	0	0	0	1	1	1	1	1	1
10	0	0	0	0	0	1	1	1	1	1	1
11	0	0	0	0	0	0	1	1	1	1	1
12	0	0	0	0	0	0	1	1	1	1	1
13	0	0	0	0	0	0	0	1	1	1	1
14	0	0	0	0	0	0	0	1	1	1	1
15	0	0	0	0	0	0	0	0	1	1	1
16	0	0	0	0	0	0	0	0	1	1	1
17	0	0	0	0	0	0	0	0	0	1	1
18	0	0	0	0	0	0	0	0	0	0	1
19	0	0	0	0	0	0	0	0	0	0	1
20	0	0	0	0	0	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.

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

## 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) & Time Periods (T)
K (Number of Clusters)	20
T (Number of Time Periods including Baseline)	11
Cluster Size Entry Type	Subjects per Cluster per Time Period (m)
m (Ave. Subjects per Cluster per Time Period)	200 to 300 by 10
Input Type	Rate Ratios
RR1 (Rate Ratio = $\lambda_1/\lambda_2$ )	0.75
$\lambda_2$ (Control Event Rate)	0.021
Variance Calculation Formula	$\sigma^2 = ((\sqrt{[\lambda_1]} + \sqrt{[\lambda_2]})/2)^2$
Use Calculated Variance as	Total Variance ( $\sigma^2 = \sigma_y^2 = \tau^2 + \sigma_w^2$ )
Between-Cluster Variability Input Type	ICC
ICC (Intracluster Correlation Coefficient)	0.007

**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 with a Transition Period (Validation using Hemming and Girling (2014))

Based on the scenario described in Example 1 above, Hemming and Girling (2014) calculate power for the following incomplete design pattern matrix that includes a transition period. A transition period occurs when no observation is made in a given time period.

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

With  $\lambda_1 = 0.015$ ,  $\lambda_2 = 0.021$ ,  $\alpha = 0.05$ ,  $\text{ICC} = 0.007$ , and  $m = 270$ , they compute the power to be 0.8237, a total sample size of  $N = 59,400$ , a  $\tau^2$  of 0.0001, and a COV of 0.53.



## Setup

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

### Design Tab

Solve For .....	<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>2</b>
Cluster Size Entry Type .....	<b>Subjects per Cluster per Time Period (m)</b>
m (Ave. Subjects per Cluster per Time Period).....	<b>270</b>
Input Type.....	<b>Rates</b>
$\lambda_1$ (Treatment Event Rate).....	<b>0.015</b>
$\lambda_2$ (Control Event Rate) .....	<b>0.021</b>
Variance Calculation Formula.....	<b><math>\sigma^2 = ((\sqrt{[\lambda_1]} + \sqrt{[\lambda_2]})/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 (Intraclass Correlation Coefficient) .....	<b>0.007</b>

### Reports Tab

ICC, COV, and Variances Decimals .....	<b>4</b>
----------------------------------------	----------

### Input Spreadsheet Data

Row	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
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

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

## Output

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

## Numeric Results

Solve For: [Power](#)  
 Design Type: Incomplete (Custom)  
 Groups: 1 = Treatment, 2 = Control  
 Test Statistic: Wald Z-Test  
 Hypotheses:  $H_0: \lambda_1 - \lambda_2 = 0$  vs.  $H_1: \lambda_1 - \lambda_2 \neq 0$

Power	Design Parameters			Number of Clusters K	Cluster Size		Sample Size N	Rate		Difference D1	Intraclass Correlation Coefficient ICC	Alpha
	S	T	R		M	m		Treatment $\lambda_1$	Control $\lambda_2$			
0.82367	11	12	2	20	2970	270	59400	0.015	0.021	-0.006	0.007	0.05

Note: The variance is calculated as  $\sigma^2 = ((\sqrt{\lambda_1} + \sqrt{\lambda_2})/2)^2$  and is considered to be the total variance ( $\sigma^2 = \sigma_y^2$ ) in power computations.

## Design Details

Design Type: Incomplete (Custom)

Power	Design Parameters			Number of Clusters K	Cluster Size		Rate		Variance			Intraclass Correlation Coefficient ICC	Coefficient of Variation COV
	S	T	R		M	m	Treatment $\lambda_1$	Control $\lambda_2$	Total $\sigma_y^2$	Between-Cluster $\tau^2$	Within-Cluster $\sigma_w^2$		
0.82367	11	12	2	20	2970	270	0.015	0.021	0.0179	0.0001	0.0177	0.007	0.5327

## Design Pattern (User-Entered)

Spreadsheet Columns: ALL

Cluster	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12
1	0	.	1	1	1	1	1	1	1	1	1	1
2	0	.	1	1	1	1	1	1	1	1	1	1
3	0	0	.	1	1	1	1	1	1	1	1	1
4	0	0	.	1	1	1	1	1	1	1	1	1
5	0	0	0	.	1	1	1	1	1	1	1	1
6	0	0	0	.	1	1	1	1	1	1	1	1
7	0	0	0	0	.	1	1	1	1	1	1	1
8	0	0	0	0	.	1	1	1	1	1	1	1
9	0	0	0	0	0	.	1	1	1	1	1	1
10	0	0	0	0	0	.	1	1	1	1	1	1
11	0	0	0	0	0	0	.	1	1	1	1	1
12	0	0	0	0	0	0	.	1	1	1	1	1
13	0	0	0	0	0	0	0	.	1	1	1	1
14	0	0	0	0	0	0	0	.	1	1	1	1
15	0	0	0	0	0	0	0	0	.	1	1	1
16	0	0	0	0	0	0	0	0	.	1	1	1
17	0	0	0	0	0	0	0	0	0	.	1	1
18	0	0	0	0	0	0	0	0	0	.	1	1
19	0	0	0	0	0	0	0	0	0	0	.	1
20	0	0	0	0	0	0	0	0	0	0	.	1

The power value of 0.82367 calculated by **PASS** matches the result of Hemming and Girling (2014) exactly, as do the values for N,  $\tau^2$ , and COV. The design pattern shows exactly what the design looks like.

## Example 3 – 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 event rate,  $\lambda_2$ , is 1.5 and the rate ratio, RR1, is 0.8. 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  $\sigma^2 = ((\sqrt{[\lambda_1]} + \sqrt{[\lambda_2]})/2)^2$  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 8, 11, 10, 9, and 8 for ICC values of 0, 0.1, 0.2, 0.3, and 0.4, respectively (no result is given for ICC = 0.5).

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

### Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 3** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

#### Design Tab

Solve For .....	<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>Rate Ratios</b>
RR1 (Rate Ratio = $\lambda_1/\lambda_2$ ).....	<b>0.8</b>
$\lambda_2$ (Control Event Rate) .....	<b>1.5</b>
Variance Calculation Formula.....	<b><math>\sigma^2 = ((\sqrt{[\lambda_1]} + \sqrt{[\lambda_2]})/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 (Intraclass Correlation Coefficient) .....	<b>0 to 0.5 by 0.1</b>

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

## Output

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

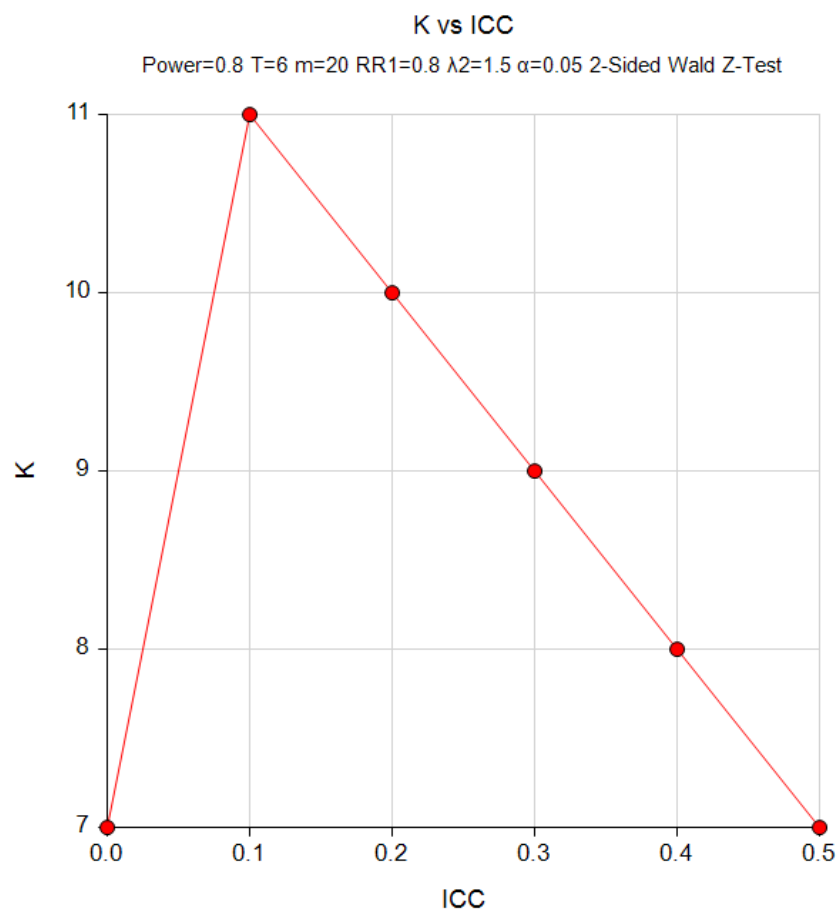
## Numeric Results

Solve For: [Sample Size \[K \(Number of Clusters\)\]](#)  
 Design Type: Incomplete (Custom)  
 Groups: 1 = Treatment, 2 = Control  
 Test Statistic: Wald Z-Test  
 Hypotheses:  $H_0: \lambda_1 / \lambda_2 = 1$  vs.  $H_1: \lambda_1 / \lambda_2 \neq 1$

Power	Design Parameters		Number of Clusters K	Cluster Size		Sample Size N	Rate		Ratio RR1	Intraclass Correlation Coefficient ICC	Alpha
	S	T		M	m		Treatment $\lambda_1$	Control $\lambda_2$			
0.82627	5	6	7	120	20	840	1.2	1.5	0.8	0.0	0.05
0.81051	5	6	11	120	20	1320	1.2	1.5	0.8	0.1	0.05
0.80654	5	6	10	120	20	1200	1.2	1.5	0.8	0.2	0.05
0.81638	5	6	9	120	20	1080	1.2	1.5	0.8	0.3	0.05
0.82780	5	6	8	120	20	960	1.2	1.5	0.8	0.4	0.05
0.84515	5	6	7	120	20	840	1.2	1.5	0.8	0.5	0.05

Note: The variance is calculated as  $\sigma^2 = ((\sqrt{\lambda_1} + \sqrt{\lambda_2})/2)^2$  and is considered to be the total variance ( $\sigma^2 = \sigma_y^2$ ) in power computations.

## Plots



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

## Design 1 Details

Design Type: Incomplete (Custom)

Power	Design Parameters		Number of Clusters K	Cluster Size		Rate		Variance			Intraclass Correlation Coefficient ICC	Coefficient of Variation COV
	S	T		M	m	Treatment $\lambda_1$	Control $\lambda_2$	Total $\sigma^2$	Between-Cluster $\tau^2$	Within-Cluster $\sigma_w^2$		
0.82627	5	6	7	120	20	1.2	1.5	1.346	0	1.346	0	0

## Design Pattern (Balanced, Optimal)

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

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

**PASS** computed K values of 7, 11, 10, 9, 8, and 7, which match those given in Table 1 of Baio et al. (2015), except for when ICC = 0. When ICC = 0, Baio et al. (2015) reports a value of 8, while **PASS** computes a value of 7. 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 0.82627, for example, indicates that K = 7 with ICC = 0 does achieve the desired power and is, therefore, the correct balanced-design solution.