

## Chapter 129

# Equivalence Tests for the Ratio of Two Means (Normal Data)

## Introduction

This procedure calculates power and sample size of statistical tests for *equivalence* tests from parallel-group design with two groups when the data are assumed to follow the normal distribution (so the log transformation is not used). This routine deals with the case in which the statistical hypotheses are expressed in terms of mean ratios rather than mean differences.

The details of this analysis are given in Hauschke *et al.* (1999) and Kieser and Hauschke (1999).

Note that when the data follow a log-normal distribution rather than the normal distribution, you should use another PASS procedure entitled *Equivalence Tests for the Ratio of Two Means (Log-Normal Data)* to obtain more accurate results.

## Equivalence Testing Using Ratios

PASS follows the *two one-sided tests* approach described by Schuirmann (1987) and Phillips (1990). It will be convenient to adopt the following specialized notation for the discussion of these tests.

<u>Parameter</u>	<u>PASS Input/Output</u>	<u>Interpretation</u>
$\mu_T$	Not entered directly	<i>Treatment mean.</i> This is the treatment mean.
$\mu_R$	Not entered directly	<i>Reference mean.</i> This is the mean of a reference population.
$R_L, R_U$	RL, RU	<i>Equivalence Limits.</i> These limits define an interval of the ratio of the means in which their difference is so small that it may be ignored.
$\phi$	R1	<i>Actual ratio.</i> This is the value of $\phi = \mu_T/\mu_R$ at which the power is calculated.

Note that the actual values of  $\mu_T$  and  $\mu_R$  are not needed. Only the ratio of these values is needed for power and sample size calculations.

With  $R_L < 1$  and  $R_U > 1$ , the null hypothesis of non-equivalence is

$$H_0: \phi \leq R_L \text{ or } \phi \geq R_U.$$

The alternative hypothesis of equivalence is

$$H_1: R_L < \phi < R_U.$$

## Coefficient of Variation

The coefficient of variation (COV) is the ratio of the standard deviation to the mean of the control group. This parameter is used to represent the variation in the data. That is,  $COV = \frac{\sigma}{\mu_C}$ .

## Power Calculation

An exact method for the calculation of power and sample size for testing equivalence based on the mean ratio when the data are assumed to be normally distributed (untransformed) was presented by Hauschke et al. (1999). This computation requires the evaluation of the noncentral bivariate t-distribution. Kieser and Hauschke (1999) present a close approximation for the case when  $RL = 1 / RU$  (the equivalence limits are the inverse of each other). This approximation is used in PASS. Kieser and Hauschke (1999) show that the approximation is very close to the exact method.

## Model and Test Details

Suppose a comparison is to be made between two groups: a treatment (T) and a control (C). The response of interest is assumed to follow the normal distribution with (possibly different) means  $\mu_T$  and  $\mu_C$  and common variance  $\sigma^2$ . To carry out the comparison, a random sample of  $N_1 = N_2$  subjects is to be obtained from each group. The parameters of the study will be presented in terms of the mean ratio  $\mu_T / \mu_C$ .

The equivalence hypotheses are

$$H_0: R1 \leq RL \text{ or } R1 \geq RU \text{ versus } H_1: RL < R1 < RU$$

where  $R1 = \frac{\mu_T}{\mu_C}$ .

The null hypothesis  $H_0$  is rejected in favor of the alternative if a two-sided  $100(1 - 2\alpha)\%$  Fieller confidence interval is included completely between  $RL$  and  $RU$ .

## Power Approximation

As mentioned above, PASS uses the approximate formulas given by Kieser and Hauschke (1999). These formulas are stated in terms of the group sample size  $N1$ , but they can be rearranged to give formulas for power as well.

If  $R1 = 1$ , use

$$N1 \geq (1 + RL^2)(t_{\alpha, 2N1-2} + t_{\beta, 2N-2})^2 \left( \frac{COV}{1 - RL} \right)^2$$

If  $RL < R1 < 1$ , use

$$N1 \geq (1 + RL^2)(t_{\alpha, 2N1-2} + t_{\beta, 2N-2})^2 \left( \frac{COV}{R1 - RL} \right)^2$$

If  $1 < R1 < RU$ , use

$$N1 \geq (1 + RU^2)(t_{\alpha, 2N1-2} + t_{\beta, 2N-2})^2 \left( \frac{COV}{R1 - RU} \right)^2$$

Note that here,  $t_{\alpha, 2N1-2}$  denotes the  $100(1 - \alpha)\%$  percentile of the central t distribution with  $2N - 2$  degrees of freedom. Also,  $\beta$  is the probability of a type II error.

---

## Procedure Options

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

---

### Design Tab

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

---

#### Solve For

##### Solve For

This option specifies the parameter to be solved for from the other parameters. Under most situations, you will select either *Power* or *Sample Size*.

---

#### 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 nonequivalent means when in fact the means are equivalent.

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 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. In this procedure, a type-I error occurs when you reject the null hypothesis of non-equivalent means when in fact the means are nonequivalent.

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*.

You can enter a range of values such as *0.05, 0.10, 0.15* or *0.05 to 0.15 by 0.01*.

---

#### Sample Size

##### Sample Size Per Group

The Sample Size Per Group is the number of items or individuals sampled from each of the Group 1 and Group 2 populations. Since the sample sizes are the same in each group, this value is the value for  $N1$ , and also the value for  $N2$ .

The Sample Size Per Group must be integers  $\geq 2$ .

You can enter a single value or a series of values.

## Equivalence Tests for the Ratio of Two Means (Normal Data)

### Effect Size

#### Equivalence Limit Input Type

Specify whether to specify the upper limit or the lower limit. The other limit will be calculated as the inverse of the entered limit.

Most experts recommend that the two limits be the inverses of each other so that the limits are invariant to the order of the means in the ratio:  $\mu_1 / \mu_2$  or  $\mu_2 / \mu_1$ .

Possible choices are

- **Enter Upper Equivalence Limit RU**

You enter the upper equivalence limit RU. PASS will calculate the corresponding lower limit, RL, using the formula  $RL = 1 / RU$ .

- **Enter Equivalence Lower Limit RL**

You enter the lower equivalence limit RL. PASS will calculate the corresponding upper limit, RU, using the formula  $RU = 1 / RL$ .

#### RU (Upper Equivalence Limit)

Enter the upper equivalence limit for the ratio of the two means. The other limit, RL, is calculated using  $RL = 1 / RU$ .

When the mean ratio is between this value and RL, the two means are said to be “equivalent”.

You can enter a single value or a list of values.

This value must be  $> 1$ . A popular choice is 1.25.

#### RL (Lower Equivalence Limit)

Enter the lower equivalence limit for the ratio of the two means. The other limit, RU, is calculated using  $RU = 1 / RL$ .

When the mean ratio is between this value and RU, the two means are said to be “equivalent”.

You can enter a single value or a list of values.

This value must be  $< 1$ . A popular choice is 0.80.

#### R1 (Actual Ratio)

This is the value of the ratio of the two means at which the power is to be calculated.

Often, the ratio will be set to one. However, some authors recommend using a ratio slightly different than one, such as 0.95 (when higher values are "good") or 1.05 (when higher values are "bad"), since this will require a larger sample size.

#### COV (Coefficient of Variation)

The coefficient of variation is the ratio of the standard deviation and the mean of the reference (control) group. It is used to specify the variability (standard deviation) of both groups since these are assumed to be equal.

Note that this COV is defined on the original (not logarithmic) scale. This value must be determined from past experience or from a pilot study.

You can enter a single value or a list of values.

The permissible range is  $COV > 0$ .

Equivalence Tests for the Ratio of Two Means (Normal Data)

## Example 1 – Finding Power

A company has developed a generic drug for treating rheumatism and wants to show that it is equivalent to the standard drug. A parallel-group design will be used to test the equivalence of the two drugs.

Following standard procedure, researchers set the lower limit of equivalence at 0.80 and the upper limit to 1.25. Past experience leads the researchers to set the COV to 1.50. The significance level is 0.05. The power will be computed assuming that the true ratio is 0.9, 0.95, 1.0, 1.05, or 1.1. Sample sizes between 500 and 2500 will be investigated.

### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the 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**.

Option	Value
<b>Design Tab</b>	
Solve For .....	<b>Power</b>
Alpha.....	<b>0.05</b>
Sample Size Per Group .....	<b>500 1000 1500 2000 2500</b>
Equivalence Limit Input Type .....	<b>Enter Upper Equivalence Limit RU (RL = 1 / RU)</b>
RU (Upper Equivalence Limit) .....	<b>1.25</b>
R1 (Actual Ratio) .....	<b>0.9 0.95 1 1.05 1.1</b>
COV (Coefficient of Variation) .....	<b>1.5</b>

### Annotated Output

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

**Numeric Results**  
 R = Treatment Mean / Reference Mean  
 Hypotheses: H0:  $R \leq RL$  or  $R \geq RU$  vs. H1:  $RL < R < RU$

				Lower Equiv Limit	Upper Equiv Limit				
Power	N1	N2	N	RL	RU	R1	COV	Alpha	
0.31538	500	500	1000	0.800	1.250	0.900	1.500	0.050	
0.50054	1000	1000	2000	0.800	1.250	0.900	1.500	0.050	
0.64480	1500	1500	3000	0.800	1.250	0.900	1.500	0.050	
0.75275	2000	2000	4000	0.800	1.250	0.900	1.500	0.050	
0.83096	2500	2500	5000	0.800	1.250	0.900	1.500	0.050	
0.54030	500	500	1000	0.800	1.250	0.950	1.500	0.050	
0.79511	1000	1000	2000	0.800	1.250	0.950	1.500	0.050	
0.91607	1500	1500	3000	0.800	1.250	0.950	1.500	0.050	
0.96761	2000	2000	4000	0.800	1.250	0.950	1.500	0.050	

(Report continues)

**References**  
 Kieser, M. and Hauschke, D. 1999. 'Approximate Sample Sizes for Testing Hypotheses about the Ratio and Difference of Two Means.' Journal of Biopharmaceutical Studies, Volume 9, No. 4, pages 641-650.  
 Hauschke, D., Kieser, M., Diletti, E., Burke, M. 1999. 'Sample Size Determination for Proving Equivalence Based on the Ratio of Two Means for Normally Distributed Data.' Statistics in Medicine, Volume 18, pages 93-105.  
 Hauschke, D., Steinijans, V., Pigeot, I. 2007. Bioequivalence Studies in Drug Development. John Wiley and Sons. New York.  
 Blackwelder, W.C. 1998. 'Equivalence Trials.' In Encyclopedia of Biostatistics, John Wiley and Sons. New York. Volume 2, 1367-1372.

## Equivalence Tests for the Ratio of Two Means (Normal Data)

### Report Definitions

Power is the probability of rejecting a false null hypothesis.

N1 and N2 are the number of items sampled from each population.

N is the total sample size, N1 + N2.

RL, RU are the lower and upper equivalence limits, respectively, and are the maximum allowable ratios that still result in equivalence.

R1 is the actual ratio of the means at which power is calculated.

COV is the coefficient of variation on the original scale.

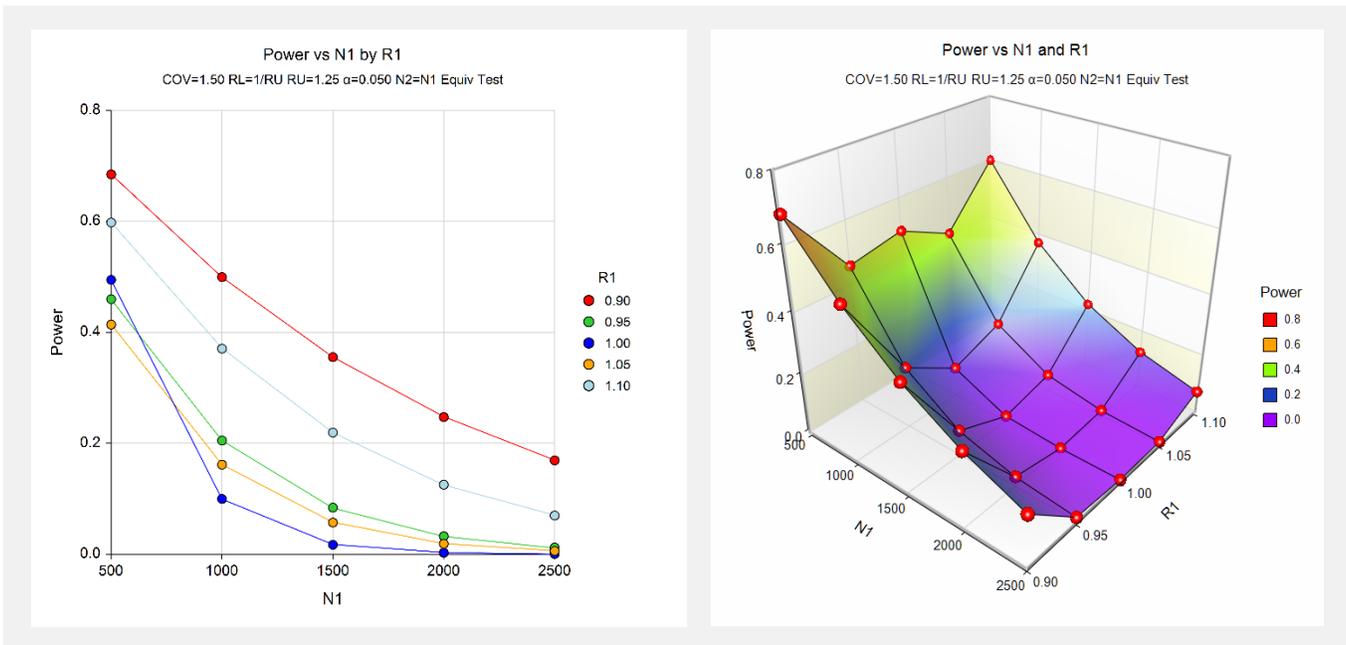
Alpha is the probability of rejecting a true null hypothesis.

### Summary Statements

An equivalence test of the ratio of two group means using two, one-sided tests on untransformed data from a parallel-group design with sample sizes of 250000 in the reference group and 500 in the treatment group achieves 32% power at a 5.0% significance level when the true ratio of the means is 0.900, the coefficient of variation on the original, unlogged scale is 1.500, and the equivalence limits of the mean ratio are 0.800 and 1.250. This test assumes that the original data in each group follows a normal distribution with (possibly) differing means and common variance.

This report shows the power for the indicated scenarios.

### Plot Section



These plots show the power versus the sample size for values of R1.

## Example 2 – Validation using Kieser and Hauschke (1999)

Kieser and Hauschke (1999) page 648 present a table of sample sizes for various scenarios. We will use one example to validate this procedure. Let the actual ratio be 1.0, the coefficient of variation be 0.35, the equivalence limits be 0.80 and 1.25, the power be 80%, and the significance level is 0.05. The calculated group sample size is 44.

### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the 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>Sample Size</b>
Power.....	<b>0.8</b>
Alpha.....	<b>0.05</b>
Equivalence Limit Input Type .....	<b>Enter Upper Equivalence Limit RU (RL = 1 / RU)</b>
RU (Upper Equivalence Limit) .....	<b>1.25</b>
R1 (Actual Ratio) .....	<b>1</b>
COV (Coefficient of Variation) .....	<b>0.35</b>

### Output

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

#### Numeric Results

<b>Numeric Results</b>									
R = Treatment Mean / Reference Mean									
Hypotheses: H0: $R \leq RL$ or $R \geq RU$ vs. H1: $RL < R < RU$									
Target Power	Actual Power	N1	N2	N	Lower Equiv Limit RL	Upper Equiv Limit RU	R1	COV	Alpha
0.80	0.80349	44	44	88	0.800	1.250	1.000	0.350	0.050

PASS also calculates the per group sample size to be 44, which validates the procedure.