nTerim

Power and Sample Size for Group Sequential Trials

Contents

Chapter 1

Introduction

nTerim is designed for the calculation of Power and Sample Size for Group Sequential Trials. Calculations are performed using the Lan-DeMets alpha spending function approach for estimating boundary values [1, 2]. Using this approach, boundary values can be estimated for O'Brien-Fleming [3], Pocock [4], Hwang-Shih-DeCani [5] and the Power family of spending functions. Calculations follow the approach of Reboussin et al [6]. Calculations can be performed for studies that involve comparisons of means, comparisons of proportions and survival studies.

1.1 Group Sequential Trials

Group sequential trials differ from fixed period trials in that the data from the trial is analyzed at one or more stages prior to the conclusion of the trial. As a result the alpha value applied at each analysis or 'look' must be adjusted to preserve the overall Type 1 error. The alpha values used at each look are calculated based upon the spending function chosen, the number of looks to be taken during the course of the trial as well as the overall Type 1 error rate. For a full introduction to group sequential methods see Jennison & Turnbull [7] and Chow et al [8].

O'Brien-Fleming	$\alpha(\tau) = 2\left(1 - \Phi\left(\frac{z_{\frac{\alpha}{2}}}{\sqrt{\tau}}\right)\right)$
Pocock	$\alpha(\tau) = \alpha \ln\left[1 + (e-1)\tau\right]$
Power	$\alpha(\tau) = \alpha\tau^{\Phi}, \Phi > 0$
Hwang-Shih-DeCani	$\alpha(\tau) = \alpha\left[\frac{(1-e^{\Phi\tau})}{(1-e^{-\Phi})}\right], \Phi \neq 0$

Table 1.1: Spending Functions

1.2 Spending Functions

There are four alpha spending functions, $\alpha(\tau)$, available to the user in nTerim as well as an option to manually input boundary values. The parameter τ represents the time elapsed in the trial. This can either be as a proportion of the overall time elapsed or a proportion of the sample size enroled. As standard all alpha spending functions have the properties that $\alpha(0) = 0$ and $\alpha(1) = \alpha$. The common element among most of the different spending functions is to use higher alpha values for the early looks. By doing this it means that the results of any analysis will only be considered significant in an early stage if it gives an extreme result.

Figure 1.1 shows a plot of the O'Brien-Fleming alpha spending function for a trial that involves four looks at the data (three interim analyses and the final analysis). The upper boundary values are 4.33263, 2.96311, 2.35902 and 2.01406, (lower boundary values are symmetrical), and these correspond to alpha values of 0.00001, 0.00304, 0.00304, 0.03070.

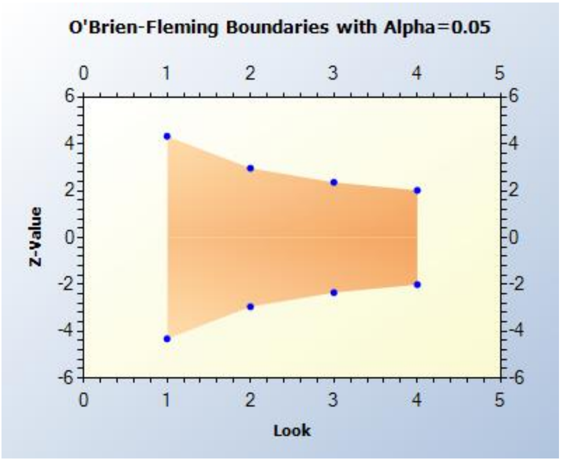


Figure 1.1: O'Brien-Fleming boundary

Other spending functions are less conservative. Table 1.2 shows a comparison of the boundary values at each look for a group sequential trial with four looks and overall alpha of 0.05.

Look	1	2	3	4
O'Brien-Fleming	4.333	2.963	2.359	2.014
Pocock	2.368	2.367	2.358	2.350
Power ($\alpha\tau^{1.5}$)	2.734	2.471	2.293	2.149
Power ($\alpha\tau^{2}$)	2.955	2.559	2.301	2.092
Hwang-Shih-DeCani ($\Phi = 0.1$)	2.485	2.401	2.322	2.255

Table 1.2: Upper boundary values for different Spending Functions

1.3 Program features

The main screen from nTerim is a spreadsheet that allows you to enter the parameters required to perform a calculation, Figure 1.2. The standard for most calculations is to solve for either sample size or power.

Whilst entering values in the table, the Help pane, Figure 1.3, located on the right of the screen will provide advice on what type of values should be entered. As part of the Help pane there is a Notes tab, Figure 1.4. The user can use this area

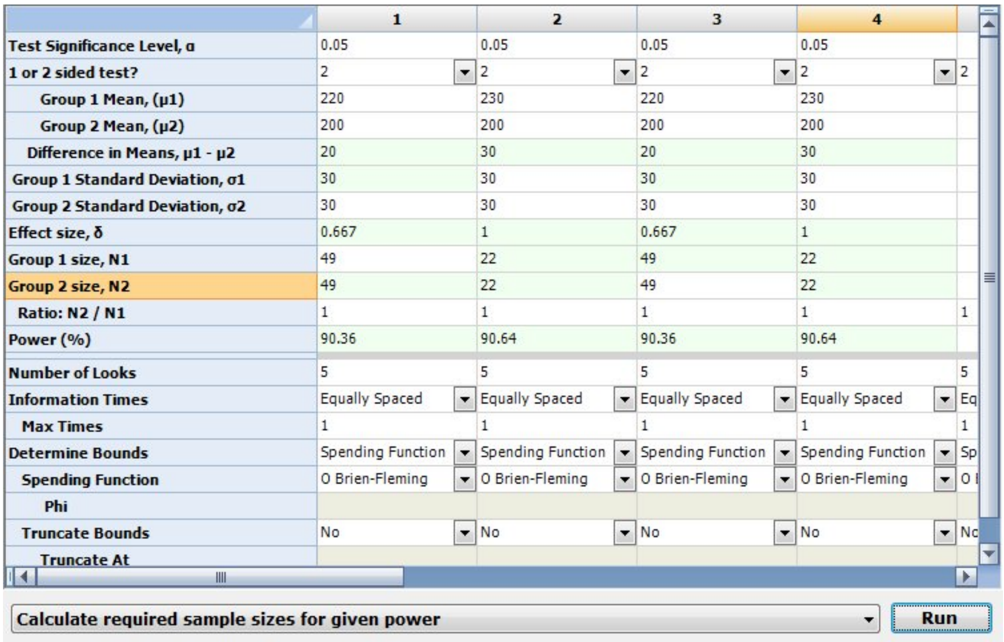


Figure 1.2: Spreadsheet interface

to record information about the calculations they are performing, perhaps details of a pilot study on which they are basing the calculation or summary notes about the study being designed.

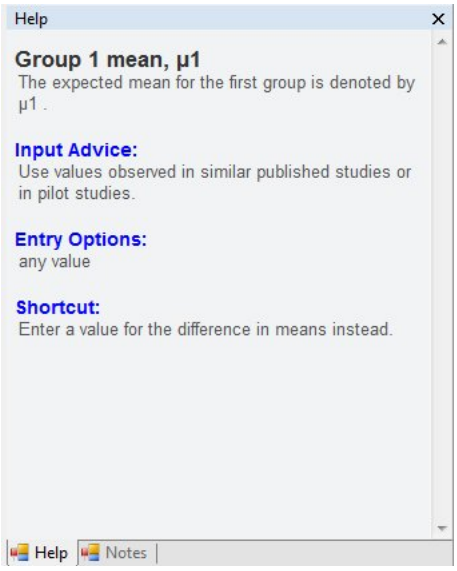


Figure 1.3: Help pane

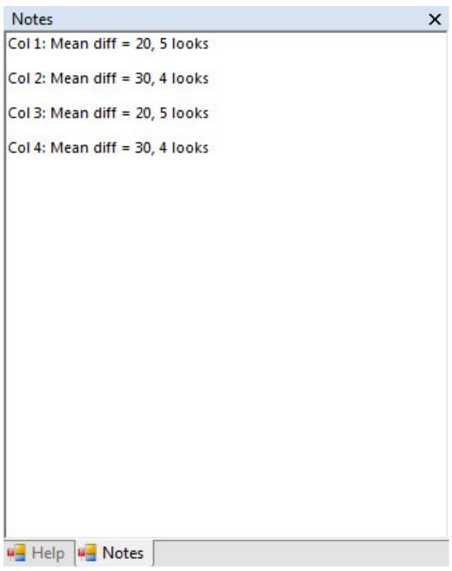


Figure 1.4: Notes pane

Once the parameters required have been entered in the table the user can select which parameter they would like to solve for from the dropdown menu below the main table. Press Run to perform the calculation. Once the calculation has completed the output value(s) will be filled in the table. Also, the table at the bottom of the screen, Figure 1.5, will be populated with details about the boundaries and alpha values calculated for each stage of the trial.

Looks

	1	2	3	4	5
Time	0.2	0.4	0.6	0.8	1
Lower Bound	-4.87688	-3.35695	-2.68026	-2.28979	-2.03100
Upper Bound	4.87688	3.35695	2.68026	2.28979	2.03100
Nominal Alpha	0.00000	0.00079	0.00736	0.02203	0.04226
Incremental Alpha	0.00000	0.00079	0.00683	0.01681	0.02558
Cumulative Alpha	0.00000	0.00079	0.00762	0.02442	0.05000
Exit Probability	0.03	10.17	35.07	29.92	15.17
Cumulative Exit Probability	0.03	10.21	45.27	75.19	90.36

Looks | Output

Figure 1.5: Boundary information

A plot of the boundaries calculated will also be displayed, Figure 1.1. By clicking on the Output tab at the bottom of the screen you can view a statement that summarizes the details of the calculation performed, Figure 1.6.

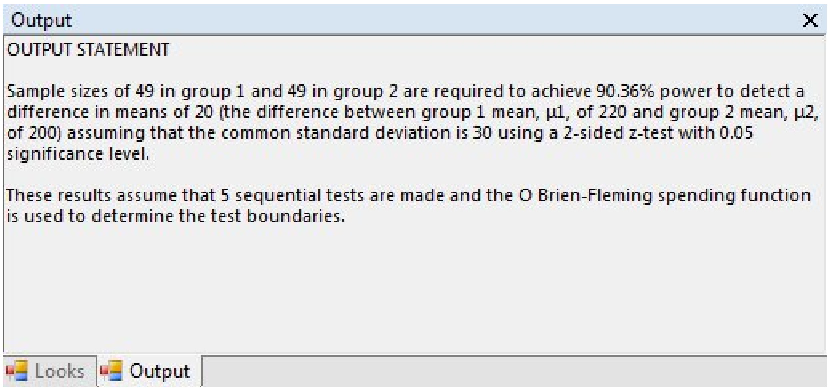


Figure 1.6: Output statement

After a calculation has been completed it is possible to create a plot of power versus sample size, Figure 1.7. Up to four columns can be produced on a single plot. To produce a plot just select any cell in the column of interest and select ‘Plot Power - Sample Size’ from the toolbar. To plot multiple columns select a cell in the first column and drag the cursor across to highlight cells in the other columns of interest and then select ‘Plot Power - Sample Size’ from the toolbar.

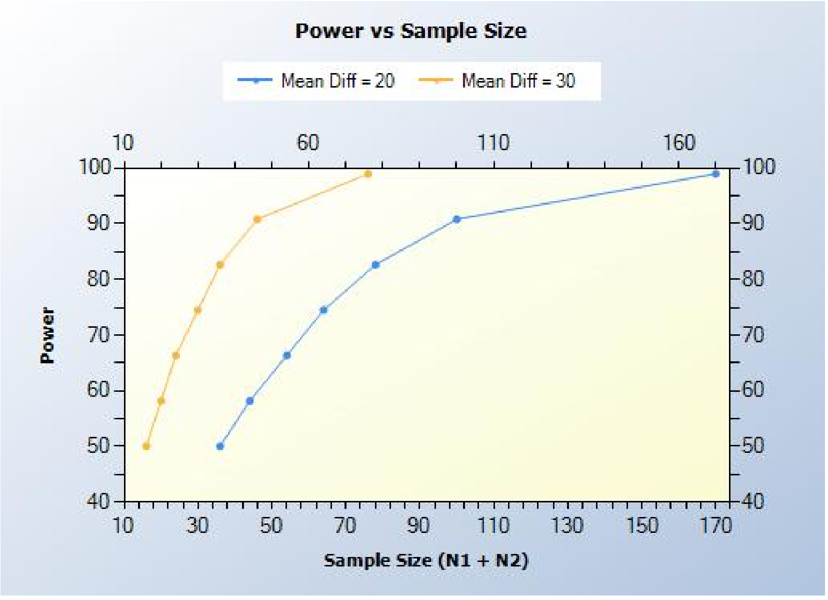


Figure 1.7: Power vs Sample Size plot

1.3.1 Exporting tables and plots

Tables can be exported to other programs once a calculation has been completed. To export the main table containing the details of the sample size calculation go to Tools > Print Main table to clipboard. You can then paste the contents of this table, include row headings, into other spreadsheet or word processing software. By the same process it is also possible to export the contents of the table containing the details of the boundary values at each interim look.

Sample Size vs Power plots and Boundary plots can be saved as either .jpg or .png files. To do this, right click on the plot and select Save image.

Chapter 2

Two Means

2.1 Example with O'Brien-Fleming spending function

This example is taken from Reboussin et al [6] using the O'Brien-Fleming spending function.

1. Open the software and select Two Means Test from the toolbar at the top, Figure 2.1 shows the Two Means table.

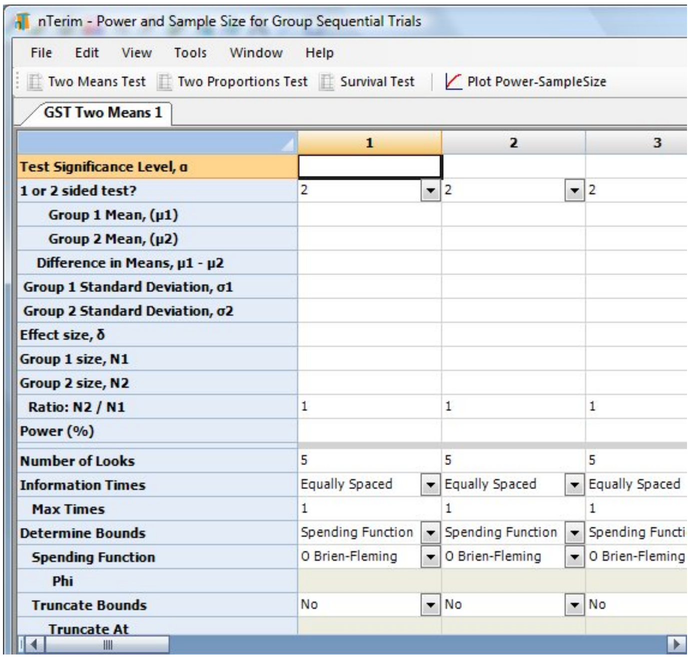


Figure 2.1: Two Means Test

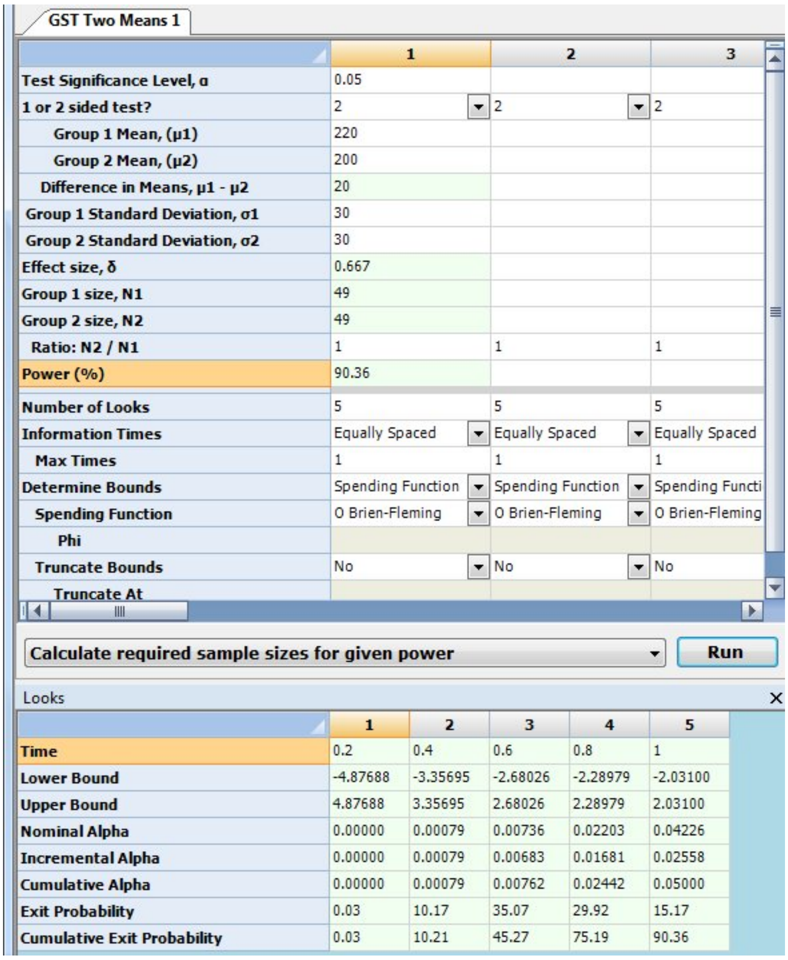
2. Enter 0.05 for alpha, 2 sided, 220 for Group 1 mean, 200 for Group 2 mean. The difference in means is calculated as 20.

3. Enter 30 for the Standard Deviation in both groups. The effect size is calculated as 0.667. We are interested in solving for sample size given 90% power so enter 90 in the Power row.

4. This study planned for 4 interim analyses. Including the final analysis this requires Number of Looks to be 5.

5. The looks will be equally spaced and the O'Brien-Fleming spending function is to be used. There will be no truncation of bounds.

6. Select: "Calculate required sample size for given power" and press Run, Figure 2.2.



GST Two Means 1

	1	2	3
Test Significance Level, α	0.05		
1 or 2 sided test?	2	2	2
Group 1 Mean, (μ1)	220		
Group 2 Mean, (μ2)	200		
Difference in Means, μ1 - μ2	20		
Group 1 Standard Deviation, σ1	30		
Group 2 Standard Deviation, σ2	30		
Effect size, δ	0.667		
Group 1 size, N1	49		
Group 2 size, N2	49		
Ratio: N2 / N1	1	1	1
Power (%)	90.36		
Number of Looks	5	5	5
Information Times	Equally Spaced	Equally Spaced	Equally Spaced
Max Times	1	1	1
Determine Bounds	Spending Function	Spending Function	Spending Functi
Spending Function	O Brien-Fleming	O Brien-Fleming	O Brien-Fleming
Phi			
Truncate Bounds	No	No	No
Truncate At			

Calculate required sample sizes for given power — Run

Looks

	1	2	3	4	5
Time	0.2	0.4	0.6	0.8	1
Lower Bound	-4.87688	-3.35695	-2.68026	-2.28979	-2.03100
Upper Bound	4.87688	3.35695	2.68026	2.28979	2.03100
Nominal Alpha	0.00000	0.00079	0.00736	0.02203	0.04226
Incremental Alpha	0.00000	0.00079	0.00683	0.01681	0.02558
Cumulative Alpha	0.00000	0.00079	0.00762	0.02442	0.05000
Exit Probability	0.03	10.17	35.07	29.92	15.17
Cumulative Exit Probability	0.03	10.21	45.27	75.19	90.36

Figure 2.2: Output with sample size and boundaries calculated

7. The boundaries calculated will be plotted also, see Figure 2.3

By clicking on the Output tab at the bottom of the screen you can see a statement giving details of the calculation:

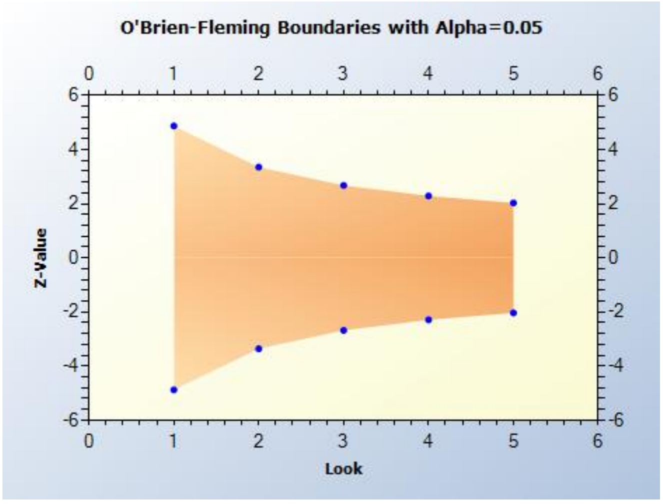


Figure 2.3: O'Brien-Fleming boundary plot

> Sample sizes of 49 in group 1 and 49 in group 2 are required to achieve 90.36% power to detect a difference in means of 20 (the difference between group 1 mean, $\mu 1$, of 220 and group 2 mean, $\mu 1$, of 200) assuming that the standard deviation in group 1 is 30 and in group 2 is 30 using a 2-sided z-test with 0.05 significance level. These results assume that 5 sequential tests are made and the O'Brien-Fleming spending function is used to determine the test boundaries. Drift: 3.29983

2.2 Example with Pocock spending function and unequal N's

This example is taken from Reboussin et al [6] using the Pocock spending function.

1. Open the software and select Two Means Test from the toolbar at the top.
2. Setup the table as in the previous example, as per Figure 2.2.
3. We will again use 5 looks but this time change the Spending Function to Pocock in the dropdown box.
4. Enter 90 in the power row and select Run and the sample size along with the boundary values will be calculated, Figure 2.4.
5. The boundaries calculated will be plotted.

Clicking on the Output tab at the bottom of the screen you can see a statement giving details of the calculation:

> Sample sizes of 57 in group 1 and 57 in group 2 are required to achieve 90.33% power to detect a difference in means of 20 (the difference between group 1 mean, $\mu 1$, of 220 and group 2 mean, $\mu 1$, of 200) assuming that the standard deviation in group 1 is 30 and in group 2 is 30 using a 2-sided

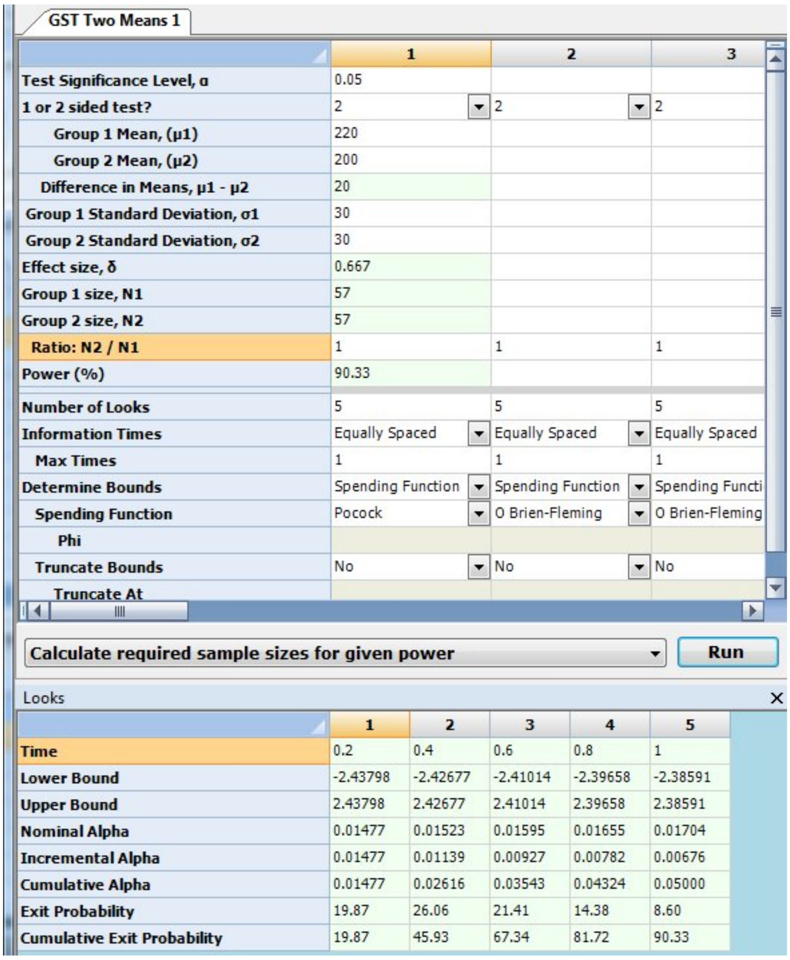


Figure 2.4: Output with sample size and boundaries calculated

z-test with 0.05 significance level. These results assume that 5 sequential tests are made and the Pocock spending function is used to determine the test boundaries. Drift: 3.55903

6. In the main table change the value in the ratio row to 2 and re-enter 90 for power.

7. Select Run and the sample size along with the boundary values will be re-calculated.

   Sample sizes of 43 in group 1 and 86 in group 2 are required to achieve 90.33% power to detect a difference in means of 20 (the difference between group 1 mean, $\mu 1$, of 220 and group 2 mean, $\mu 1$, of 200) assuming that the standard deviation in group 1 is 30 and in group 2 is 30 using a 2-sided z-test with 0.05 significance level. These results assume that 5 sequential tests are made and the Pocock spending function is used to determine the test boundaries. Drift: 3.56942

2.3 Formulae and calculations

The variables are defined as

Symbol	Description
α	Probability of Type I error
β	Probability of Type II error
$1-\beta$	Power of the test
μ_1, μ_2	group means
σ_1, σ_2	group standard deviations
N_1,N_2	group sample sizes
R	Ratio of N_2 to N_1
Θ	Drift parameter
K	Number of timepoints (looks)
α^*	Spending function (O'Brien-Fleming, Pocock etc)

2.3.1 Calculate sample sizes for a given power

Using the number of timepoints (K), timepoints ($t_1 \ldots t_K$), number of sides, type of spending function, α and power $(1-\beta)$ the drift parameter Θ can be obtained using the algorithm by Reboussin et al [6]. The test statistic is defined as

$$\Theta = \frac{\mu_1 - \mu_2}{\sqrt{\frac{\sigma_1^2}{N_1} + \frac{\sigma_2^2}{N_2}}} \tag{2.1}$$

The user supplies the means (μ_1, μ_2), and either R, N_1 or N_2. Since $R = \frac{N_2}{N_1}$ it follows that a value of $R = 1$ indicates equal sample sizes. The approach to solving this problem is dependant on what information the user supplies. Given any two of R, N_1 or N_2, the unknown is obtained by solving equation 2.1.

2.3.2 Calculate attainable power with the given sample sizes

Given α, N_1, means (μ_1, μ_2), standard deviations (σ_1, σ_2), R (or N_2), timepoints and type of spending function. The requirement is to obtain the power.

The steps are:

- Obtain Θ by solving equation 2.1 (given that N_1, R, μ_1, μ_2, σ_1 and σ_2 are known)
- Obtain power using the algorithm by Reboussin et al [6].

2.3.3 Calculate means given all other information

Given α, N_1, standard deviations (σ_1, σ_2), R (or N_2), power $(1-\beta)$, timepoints and type of spending function. The requirement is to obtain either μ_1 or μ_2, given the other.

The steps are:

- Obtain Θ by solving equation 2.1 (given that N_1, R, μ_1, μ_2, σ_1 and σ_2 are known)

- Equation 2.1 can be expressed as a quadratic in μ_1 or μ_2. The roots give the unknown μ.

By default, nTerim assumes that $\mu_1 < \mu_2$ and will select the appropriate root.

Chapter 3

Two Proportions

3.1 Example with Pocock spending function

This example is taken from Reboussin et al [6] using the Pocock spending function.

1. Open the software and select Two Proportions Test from the toolbar at the top.
2. Enter 0.05 for alpha, 1 sided test, 0.4 for Group 1 proportion, 0.6 for Group 2 proportion. The odds ratio is calculated as 2.25.
3. Select Off for the Continuity Correction. This example solved for sample size given 90% power so enter 90 in the Power row.
4. This study planned for 4 interim analyses. Including the final analysis this requires Number of Looks to be 5.
5. The looks will be equally spaced and the Pocock spending function is to be used. There will be no truncation of bounds.
6. Select run. See Figure 3.1.

Looks

	1	2	3	4	5
Time	0.2	0.4	0.6	0.8	1
Lower Bound	-8.00000	-8.00000	-8.00000	-8.00000	-8.00000
Upper Bound	2.17621	2.14371	2.11322	2.08952	2.07091
Nominal Alpha	0.01477	0.01603	0.01729	0.01833	0.01918
Incremental Alpha	0.01477	0.01139	0.00927	0.00782	0.00676
Cumulative Alpha	0.01477	0.02616	0.03543	0.04324	0.05000
Exit Probability	22.98	25.92	20.00	13.21	8.00
Cumulative Exit Probability	22.98	48.90	68.90	82.11	90.12

Looks | Output

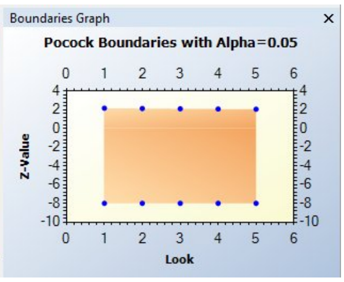


Figure 3.1: Pocock boundary values and plot

Sample sizes of at least 129 in group 1 and 129 in group 2 are required to achieve 90.12% power to detect an odds ratio of 2.25 (for proportions of 0.4 in group 1 and 0.6 in group 2) using a 1-sided z-test with 0.05 significance level. These results assume that 5 sequential tests are made and the Pocock spending function is used to determine the test boundaries. Drift: 3.21248

3.2 Example with Power Family spending function with truncated bounds

1. Open the software and select Two Proportions Test from the toolbar at the top.
2. Enter 0.05 for alpha, 2 sided, 0.41 for Group 1 proportion, 0.465 for Group 2 proportion. The odds ratio is calculated as 1.25074.
3. Select On for the Continuity Correction. We are interested in solving for power given a sample size of 1400 per group. Select 'Calculate the attainable power with the given sample sizes'. Enter 1400 in the Group 1 size, N1 row.
4. This study planned for 4 interim analyses. Including the final analysis this requires Number of Looks to be 5.
5. The looks will be equally spaced and the Power Family spending function is to be used. Enter 3 for Phi.
6. For this example we want to truncate the boundaries so as not to be over-conservative. Enter Yes for truncate bounds and then enter 3 for the value to truncate at.
7. Press Run.

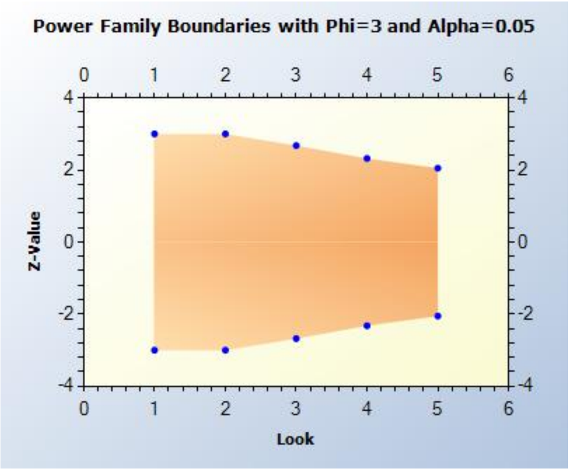


Figure 3.2: Power boundary plot

Sample sizes of at least 1400 in group 1 and 1400 in group 2 are required to achieve 81.17% power to detect an odds ratio of 1.25074 (for proportions of 0.41 in group 1 and 0.465 in group 2) using a 2-sided continuity corrected z-test with 0.05 significance level. These results assume that 5 sequential tests are made and the Power Family spending function is used to determine the test boundaries. Drift: 2.89524

3.3 Formulae and calculations

The variables are defined as

Symbol	Description
α	Probabilty of Type I error
β	Probability of Type II error
$1-\beta$	Power of the test
p_1,p_2	Group proportions
N_1,N_2	Sample Sizes
R	Ratio of N_2 to N_1
Θ	Drift parameter
K	Number of timepoints (looks)
$\overline{p}$	sample size weighted proportion
α^*	Spending function (O'Brien-Fleming, Pocock etc)

3.3.1 Calculate sample sizes for a given power

Using the number of timepoints (K), timepoints $(t_1 \ldots t_K)$, number of sides, type of spending function, α and power $(1-\beta)$ the drift parameter Θ can be obtained. The test statistic is defined as

$$\Theta = \frac{|p_1 - p_2|}{\sqrt{\frac{\overline{p}(1-\overline{p})}{N_1} + \frac{\overline{p}(1-\overline{p})}{N_2}}} \tag{3.1}$$

where $\overline{p} = \frac{N_1 p_1 + N_2 p_2}{N_1 + N_2}$. The user supplies the proportions (p_1,p_2), and either R, N_1 or N_2. Since $R = \frac{N_2}{N_1}$ it follows that a value of $R = 1$ indicates equal sample sizes and that $\overline{p} = \frac{p_1 + R * p_2}{1+R}$. The approach to solving this problem is dependant on what information the user supplies. For the case of continuity correction the formula can be written as:

$$\Theta = \frac{|p_1 - p_2| - \frac{1}{2N_1}\left(\frac{R+1}{R}\right)}{\sqrt{\frac{\overline{p}(1-\overline{p})(R+1)}{N_1 R}}} \tag{3.2}$$

as per Fleiss [9] Page 44. The validity of this formula relies on the assumption of minimum expected cell count being above a pre-specified threshold. As a rule of thumb, the normal approximation to the binomial will hold if the following conditions are met:

$p_1 N_1 > T$
$(1-p_1)\, N_1 > T$
$p_2 N_2 > T$
$(1-p_2)\, N_2 > T$

where T is a predefined threshold.

User supplies R only

The requirement is to obtain N_1 and N_2 .

Using that $N_2 = R \times N_1$ the result from equation 3.2 obtained is:

$$N_1 = \frac{\Theta^2 (R\overline{p}(1-\overline{p}) + \overline{p}(1-\overline{p}))}{R(p_1 - p_2)^2} \tag{3.3}$$

The steps involved are:

- Obtain Θ
- Solve equation 3.3 for N_1 and $N_2 = R \times N_1$

User supplies R only and selects Continuity Correction

If the user has selected to use the continuity correction then apply the formula from Fleiss et al [10]

$$N_{1cc} = \frac{N_1}{4}\left(1 + \sqrt{1 + \frac{2(R+1)}{R(N_1)|p_1 - p_2|}}\right)^2 \tag{3.4}$$

to obtain N_{1cc}. It follows that N_{2cc} is then $R \times N_{1cc}$.

If the user has NOT selected to use continuity correction then $N_1 = N_1$ and $N_2 = R \times N_1$.

User supplies N_1 only or N_2 only

When the user specifies N_1, then Equation 3.1 can be re-expressed as a quadratic in N_2 from which two roots are obtained, one less than and one greater than N_1. Similarly, if N_2 is specified the roots give the values of N_1.

3.3.2 Calculate attainable power with the given sample sizes

Given α, N_1, proportions (p_1, p_2), R (or N_2), timepoints and type of spending function. The requirement is to obtain the power.

If the user has NOT selected to use continuity correction

The steps are:

- Obtain Θ by solving equation 3.1 given that N_1, R, p_1, p_2, $\overline{p}$ are known)
- Obtain power using the algorithm by Reboussin et al [6]

If the user has selected to use continuity correction

The steps are:

- Obtain Θ by solving equation 3.2 given that N_1, R, p_1, p_2, $\overline{p}$ are known)
- Obtain power using the algorithm by Reboussin et al [6]

3.3.3 Calculate missing proportion given N_1, N_2, α, power and the other proportion.

Calculate p_1 given p_2

In order to solve for p_1 given p_2 and all other information Equation 3.1 can be re-expressed as a quadratic with respect to p_2 the roots of which give p_1. Similarly, if p_1 is specified the roots give the values of p_2.

Calculate p_1 given p_2 with Continuity Correction

In order to solve for p_1 given p_2 and all other information Equation 3.2 can be re-expressed as a quadratic with respect to p_2 the roots of which give p_1. Similarly, if p_1 is specified the roots give the values of p_2.

Chapter 4

Survival

4.1 Example with O'Brien-Fleming spending function - with Power vs Sample Size plot

1. Open the software and select Survival Test from the toolbar at the top.
2. Enter 0.05 for alpha, 2 sided, 0.3 for Group 1 proportion (this is the proportion survivng until time t) and 0.45 for Group 2 proportion. The hazard ratio is calculated as 1.508.

	1	2
Test Significance Level, α	0.05	0.05
1 or 2 sided test?	2	2
Group 1 proportion π1 at time t	0.3	0.3
Group 2 proportion π2 at time t	0.45	0.4
Hazard ratio, h=ln(π1)/ln(π2)	1.508	1.314
Survival Time Assumption	Exponential Surviva	Exponential Surviva
Total Sample Size, N	409	888
Power (%)	90.07	90.02
Number of events	256	578
Number of Looks	5	5
Information Times	Equally Spaced	Equally Spaced
Max Times	1	1
Determine Bounds	Spending Function	Spending Function
Spending Function	O Brien-Fleming	O Brien-Fleming
Phi		
Truncate Bounds	No	No
Truncate At		

Figure 4.1: Output from Survival table

3. Select Exponential Survival for the Survival time assumption.
4. We are interested in solving for sample size given 90% power so enter 90 in the Power row.

5. This study planned for 4 interim analyses. Including the final analysis this requires Number of Looks to be 5.
6. The looks will be equally spaced and the O'Brien-Fleming spending function is to be used. There will be no truncation of bounds.
7. Select run.
8. In the second column enter the same parameters as above but change the Group 2 proportion to 0.40. Select run. See Figure 4.1.
9. Click on the column title for column 1 and drag across to highlight both columns 1 and 2.
10. Select Plot Power-Sample Size from the toolbar, (it may take a moment to generate the plot as multiple calculations are performed), see Figure 4.2.

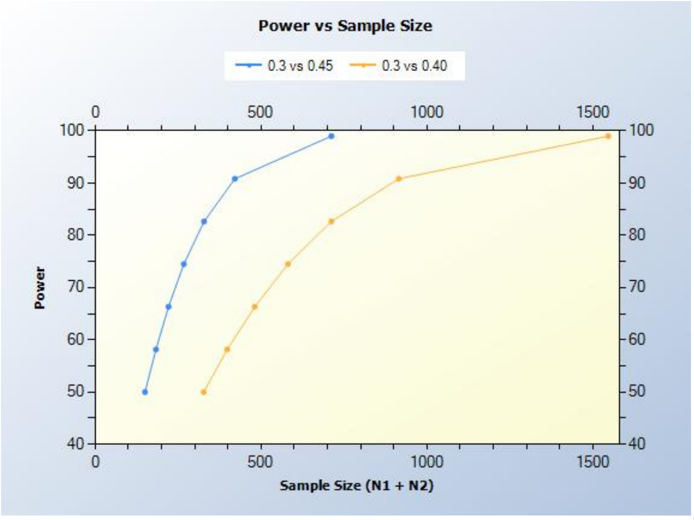


Figure 4.2: Plot of two columns from survival table

11. The legend in the plot can be customized. Just click on the legend and a dialog box will appear allowing you to enter headings appropriate to each line on the plot.

    A total sample size of at least 409 (256 events) is required to achieve 90.07% power to detect a hazard ratio of 1.508 (for survival rates of 0.3 in group 1 and 0.45 in group 2), using a 2-sided log rank test with 0.05 significance level assuming that the survival times are exponential. These results assume that 5 sequential tests are made and the O Brien-Fleming spending function is used to determine the test boundaries. Drift: 3.28269

4.2 Example with Pocock spending function - with non equally spaced looks

1. Open the software and select Survival Test from the toolbar at the top.

2. Enter 0.05 for alpha, 2 sided, 0.5 for Group 1 proportion, 0.4 for Group 2 proportion. The hazard ratio is calculated as 0.756.

3. Select Proportional Hazards for the Survival Time Assumption. We are interested in solving for power given a sample size of 1000. Select 'Calculate the attainable power with the given sample sizes' from the dropdown box under the main table. Enter 1000 in the Total Sample Size row.

4. This study planned for 4 interim analyses. Including the final analysis this requires Number of Looks to be 5.

5. The Pocock spending function is to be used, however the looks will not be evenly spaced.

6. For Information Times, select User Input. Then in the Times row in the lower table enter the values 0.1, 0.2, 0.3, 0.6 and 1, see Figure 4.3.

7. Click Run.

	1	2	3	4	5
Time	0.1	0.2	0.3	0.6	1
Lower Bound	-2.65511	-2.62320	-2.58958	-2.34880	-2.27923
Upper Bound	2.65511	2.62320	2.58958	2.34880	2.27923
Nominal Alpha	0.00793	0.00871	0.00961	0.01883	0.02265
Incremental Alpha	0.00793	0.00684	0.00602	0.01464	0.01457
Cumulative Alpha	0.00793	0.01477	0.02079	0.03543	0.05000
Exit Probability	5.20	8.79	10.68	34.58	26.06
Cumulative Exit Probability	5.20	13.99	24.68	59.26	85.32

Figure 4.3: User Input Times

A total sample size of at least 1000 (550 events) is required to achieve 85.32% power to detect a hazard ratio of 0.756 (for survival rates of 0.5 in group 1 and 0.4 in group 2), using a 2-sided log rank test with 0.05 significance level assuming that the hazards are proportional. These results assume that 5 sequential tests are made and the Pocock spending function is used to determine the test boundaries. Drift: 3.25156

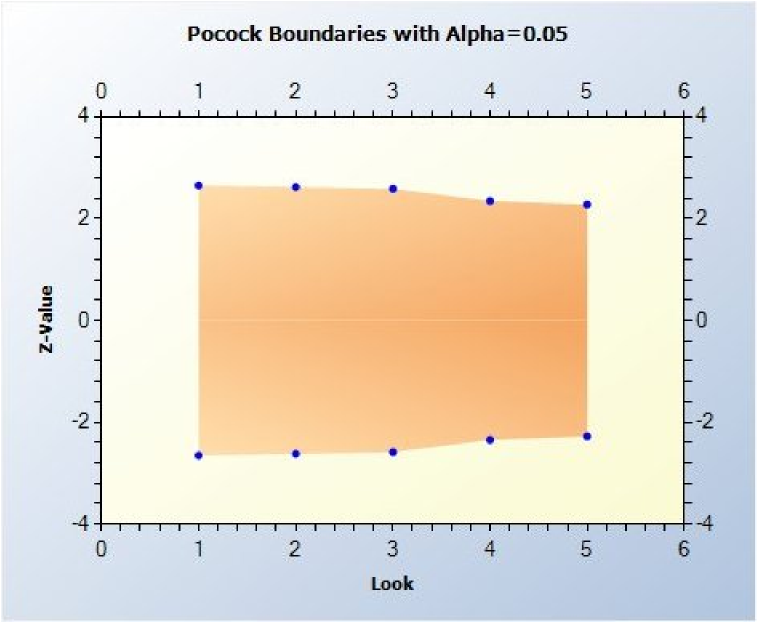


Figure 4.4: Plotted Boundary Values

4.3 Formulae and calculations

Sequential Log-Rank test of survival in two groups.

The variables are defined as

Symbol	Description
α	Probability of Type I error
β	Probability of Type II error
$1-\beta$	Power of the test
s_1,s_2	group survival proportions
d	number of events
N	sample size
Θ	Drift parameter
K	Number of timepoints (looks)
α^*	Spending function (O'Brien-Fleming, Pocock etc)

4.3.1 Calculate sample size for a given power

Using the number of timepoints (K), timepoints ($t_1 \dots t_K$), number of sides, type of spending function, α and power ($1-\beta$) the drift parameter Θ can be obtained.

$$HR = log\left(\frac{s_2}{s_1}\right) \tag{4.1}$$

For the Exponential Survival Curve

$$\Theta = \frac{log(HR)\sqrt{(d_k)}}{2} \tag{4.2}$$

This can be solved for d_k the required number of events

$$d_k = \frac{4\Theta^2}{[log(HR)]^2} \tag{4.3}$$

For the Proportional Hazards Curve

$$\Theta = \frac{1 - HR\sqrt{(d_k)}}{1 + HR} \tag{4.4}$$

This can be solved for d_k the required number of events

$$d_k = \left[\frac{(1 + HR)\Theta}{(1 - HR)}\right]^2 \tag{4.5}$$

N formula

$$N = \frac{2d_k}{2 - s_1 - s_2} \tag{4.6}$$

4.3.2 Calculate attainable power with the given sample size

Given α, N, group survival proportions (s_1,s_2), timepoints and type of spending function. The requirement is to obtain the power.

For the Exponential Survival Curve

$$\Theta = \sqrt{\frac{N(2 - s_1 - s_2)log^2(HR)}{8}} \tag{4.7}$$

For the Proportional Hazards Curve

$$\Theta = \sqrt{\frac{N(2 - s_1 - s_2)(1 - HR)^2}{2(1 + HR)^2}} \tag{4.8}$$

Chapter 5

Settings

5.1 Program settings

There are certain settings the are pre-defined in nTerim that the user can edit. To make any changes go to Tools > Settings.

Minimum expected cell count When calculating sample size or power for the comparison of two proportions the validity of the formula, 3.2 relies on the assumption of minimum expected cell count being above a pre-specified threshold. nTerim has a defined set of these limits but the user can define their own threshold limit.

Assume $\pi_1 < \pi_2$ for Proportions table During certain calculations it is necessary to specify which proportion will be greater in the parameter list.

Assume $\mu_1 < \mu_2$ for Means table During certain calculations it is necessary to specify which group mean will be greater in the parameter list.

Bibliography

[1] DeMets, D.L. and Lan, K.K.G. (1984), An overview of sequential methods and their applications in clinical trials. Communications in Statistics, Theory and Methods, 13, pages 2315-2338.

[2] DeMets, D.L. and Lan, K.K.G. (1994), Interim analysis: The alpha spending function approach. Statistics in Medicine, 13, pages 1341-1352.

[3] O'Brien, P. C. & Fleming, T. R. (1979), A multiple testing procedure for clinical trials, Biometrics 35, 549-556.

[4] Pocock, S. J. (1977), Group sequential methods in the design and analysis of clinical trials, Biometrika 64, 191-199.

[5] Hwang, I. K., Shih, W. J. & deCani, J. S. (1990), Group sequential designs using a family of type I error probability spending functions, Statistics in Medicine, 9, 1439-1445.

[6] Reboussin, D.M., DeMets, D.L., Kim, K, and Lan, K.K.G. (1992), Programs for computing group sequential boundaries using the Lan-DeMets Method. Technical Report 60, Department of Biostatistics, University of Wisconsin-Madison.

[7] Jennison, C. and Turnbull, B.W. (2000), Group sequential methods with applications to clinical trials. Chapman & Hall.

[8] Chow, S.C., Shao, J. and Wang, H. (2008), Sample size calculations in clinical research. Second Edition. Chapman & Hall.

[9] Fleiss, J.L. (1981), Statistical methods for rates and proportions. Wiley.

[10] Fleiss, J.L., Tytun, A., Ury, S.H.K. (1980), A simple approximation for calculating sample sizes for comparing independent proportions Biometrics 36 pp. 343-346