Sample Size Methods used in the SampSize App

1 Introduction

Sample size calculations using the SampSize App have been described in a series of three articles which were published by the journal *Pharmaceutical Statistic*s and for which the author submitted version of the paper is now available on the *White Rose Repository*. The papers describe calculations for the situation with a Normal outcome for trials looking to investigate objectives of superiority, non-inferiority and equivalence with a parallel group design.

2 Practical guide articles

2.1 Practical guide to sample size calculations: an introduction

A sample size justification is a vital step when designing any trial. However, estimating the number of participants required to give a meaningful result is not always straightforward. A number of components are required to facilitate a suitable sample size calculation. In this paper, the general steps are summarised for conducting sample size calculations with practical advice and guidance on how to utilise the SampSize app.

Flight L and Julious SA. Practical guide to sample size calculations: an introduction. *Pharmaceutical Statistics* 2016 **15(1)** 75-79

Available on the White Rose Repository at: http://eprints.whiterose.ac.uk/97115/

2.2 Practical guide to sample size calculations: superiority trials

A sample size justification is a vital part of any investigation. However, estimating the number of participants required to give meaningful results is not always straightforward. A number of components are required to facilitate a suitable sample size calculation. In this paper, the steps for conducting sample size calculations for superiority trials are summarised. Practical advice and examples are provided illustrating how to carry out the calculations by hand and using the SampSize app.

Flight L and Julious SA. Practical guide to sample size calculations: superiority trials. Pharmaceutical Statistics 2016 15(1) 80-89

Available on the White Rose Repository at: http://eprints.whiterose.ac.uk/97114/

2.3 Practical guide to sample size calculations: non-inferiority and equivalence trials

A sample size justification is a vital part of any trial design. However, estimating the number of participants required to give a meaningful result is not always straightforward. A number of components are required to facilitate a suitable sample size calculation. In this paper, the steps for conducting sample size calculations for non-inferiority and equivalence trials are summarised. Practical advice and examples are provided that illustrate how to carry out the calculations by hand and using the SampSize app.

Flight L and Julious SA. Practical guide to sample size calculations: non-inferiority and equivalence trials. Pharmaceutical Statistics 2016:15(1) 68-74

Available on the White Rose Repository at: http://eprints.whiterose.ac.uk/97113/

3 Description of the sample size methods by data type

3.1. Sample sizes for clinical trials with Normal data

A more detailed described of the methods for Normal outcomes can be found in a Tutorial in Biostatistics paper in the journal Statistics in Medicine for Normal data. The article gives an overview of sample size calculations for parallel group and cross-over studies with Normal data. Sample size derivation is given for trials where the objective is to demonstrate: superiority, equivalence, non-inferiority, bioequivalence and estimation to a given precision, for different types I and II errors. It is demonstrated how the different trial objectives influence the null and alternative hypotheses of the trials and how these hypotheses influence the calculations. Sample size tables for the different types of trials and worked examples are given.

Julious SA. Tutorial in Biostatistics: Sample Sizes for clinical trials with Normal Data. *Statistics in Medicine* 2004 **23:**1921-86

Available on the White Rose Repository at: http://eprints.whiterose.ac.uk/145474/

3.1 Sample sizes for parallel group clinical trials with binary data

For trials with Binary outcomes methods are described in the paper below. This article gives an overview of sample size calculations for a single response and a comparison of two responses in a parallel group trial where the outcome is binary. Sample size derivation is given for trials where the objective is to demonstrate: superiority, equivalence, non-inferiority and estimation to a given precision. For each type of trial the null and alternative hypotheses are described and how the impact these have on the sample size calculations. For each type of trial the calculations are highlighted through worked examples. Sample size tables for the different types of trials and worked examples are given to assist in future calculations

Julious SA and Campbell MJ. Tutorial in Biostatistics: Sample Sizes for clinical trials with binary data. *Statistics in Medicine* 2012;**31**:2904–36

3.2 Sample sizes for Parallel Group Clinical Trials with Survival Data

This article gives an overview of sample size calculations for parallel group studies where the primary outcome is a survival endpoint. It is the third in a series of articles and follows similar tutorials which described calculations for outcomes which were anticipated to be Normal and binary.

In this paper sample size derivations are given for trials where the objective is to demonstrate: superiority, equivalence, non-inferiority and estimation to a given precision. For each type of trial, the null and alternative hypotheses are described in context with the primary outcome as well as how these impact on the sample size calculations. Sample size tables for the different types of trials and worked examples are given to assist in future calculations.

Julious SA and F.M.S. Barthel. Tutorial in Biostatistics: Sample sizes for Parallel Group Clinical Trials with Survival Data. *Statistics in Medicine* (under review)

As the methods for trials with a survival outcome have not been published we will now go into a little detail

3.2.1 Superiority Trials

Suppose the event of interest is a negative: for example death or recurrence such that the primary objective of the trial is to delay the event from happening. The objective of the trial would be to slow down the time to the primary outcome and the primary analysis for such a response would be a log-rank test.

Now suppose the survival distributions for the two arms of the trial have instantaneous death rates of λ_A for treatment A and λ_B for treatment B. Now from this the hazard ratio (HR) is defined as

$HR = \lambda_A/\lambda_B.$

In terms of the hazard ratio the null and alternative hypothesis would be of the form

H_0: The survival experience for both treatment groups is the same (HR=1).

H$_1$: The survival experience for both treatment groups differs (HR$\neq$1).

If the hazard ratio does not change with time, then it can be estimated by

$$HR = \frac{\log \pi_A}{\log \pi_B},$$

where π_A and π_B are two survival rates at some fixed time point. Assuming an exponential survival an alternative formula for the Hazard Ratio is to derive it in terms of the median survival terms for each treatment

$$HR = \frac{M_B}{M_A},$$

where M$_A$ and M$_B$ are the median survival times on A and B respectively.

For a given hazard ratio (HR) the number of events, *E*, required in each patient group in SampSize is estimated from

$$E = \frac{2(Z_{1-\alpha/2} + Z_{1-\beta})^2}{(\log HR)^2}.$$

Where Z$_{1-\alpha/2}$ and Z$_{1-\beta}$ are standardised Normal values for α (the Type I error) and β (the Type II error. Here 1-β is the power of the study. The significance level is taken as two sided and so $\alpha/2$ is used.

3.2.2 Non-inferiority Trials

Assuming that a HR < 1 favours the experimental treatment the hazard ratio for the null and alternative hypothesis is now of the form

H$_0$: The survival experience for the new treatment is inferior to the control (HR$\geq$d).

H$_1$: The survival experience for new treatment is the same or favours it compared to control (HR<d).

The number of events, *E*, required in each patient group is estimated within SampSize from

$$E = \frac{2(Z_{1-\alpha} + Z_{1-\beta})^2}{(\log HR - \log(d))^2}.$$

where d is the non-inferiority limit in terms of a hazard ratio and Z1-α and Z1-β are defined as for the superiority trials. The significance level is taken as one sided and so α is used.

3.2.3 Equivalence Trials

In the case of an equivalence trial the aim is to show that the experimental treatment does not differ in efficacy (or safety) from the current treatment in either direction. An example of the use of this design is in the case of biosimilar trials. The hazard ratio for the null and alternative hypothesis is now of the form

> H_0: The survival experience for either treatment groups is inferior to the other (HR≥d or HR≤d).

> H_1: The survival experience for both treatment groups is the same (d<HR<d).

For equivalence trials a direct estimate of the sample size is not possible. However, for a given number of events the power of the study can be estimated. Hence, for given a number of events, *E*, SampSize estimates the power from the following result.

$$1 - \beta = \Phi\left(\sqrt{E}\left|\log HR - \log(d)\right|/\sqrt{2} - Z_{1-\alpha}\right) + \left(\sqrt{E}\left|\log HR - \log(d)\right|/\sqrt{2} - Z_{1-\alpha}\right) - 1.$$

where d is the equivalence limit in terms of a hazard ratio. To estimate the sample size you would need to iterate on E to obtain the required to get a E which gave the requisite power. Z1-α and Z1-β are defined as for the non-inferiority and superiority trials. As for non-inferiority trials the significance level is taken as one sided and so α is used.

Note that for the special case of HR=1 we can estimate the sample size from.

$$E = \frac{2(Z_{1-\alpha} + Z_{1-\beta/2})^2}{(\log(d))^2}.$$

3.2.4 Precision Trials

Precision trials are designs employed to consider the precision of for example a devise in diagnosing a disease of interest. The trials are less about proving there is a treatment difference than estimating plausible treatment differences with a view to undertaking a definitive trial later.

To obtain a sample size to have required precision w about the hazard ratio SampSize uses the following result

$$E = \frac{2Z_{1-\alpha/2}^2}{(\log(1-w))^2}.$$

The precision for a given n can be estimate from

$$w = 1 - exp - \left(\sqrt{\frac{2Z_{1-\alpha/2}^2}{E}} \right)$$

4 Additional References

4.1 Text book

All the methods are available in the following book

Julious, SA. Sample sizes for clinical trials. Chapman and Hall, 2009

4.2 Guidance on effect sizes

Randomised controlled trials are considered to be the best method to assess comparative clinical efficacy and effectiveness, and can be a key source of data for estimating cost effectiveness. Central to the design of a randomised controlled trial is an *a priori* sample size calculation, which ensures that the study has a high probability of achieving its pre-specified main objective. In this article, the key messages from the DELTA^2 guidance on determining the target difference and sample size calculation for randomised controlled trials. Recommendations for the subsequent reporting of the sample size calculation are also provided in the paper.

Cook JA, Julious SA, Sones W, Hampson LV, Hewitt C, Berlin JA, et al. DELTA2 guidance on choosing the target difference and undertaking and reporting the sample size calculation for a randomised controlled trial. *BMJ* 2018;**363:**k3750 http://dx.doi.org/10.1136/bmj.k3750

Available at: https://www.bmj.com/content/363/bmj.k3750

Available on the White Rose Repository at: http://eprints.whiterose.ac.uk/139500/

4.3 Sample sizes for pilot trials

Although not covered in the App a sample size justification for a pilot trial is important. The following paper describes methods for sample size estimation for pilot studies as well as useful rules of thumb

When the outcome is a continuous variable, the sample size calculation requires an accurate estimate of the standard deviation of the outcome measure. A pilot trial can be used to get an estimate of the standard deviation, which could then be used to anticipate what may be observed in the main trial. However, an important consideration is that pilot trials often estimate the

standard deviation parameter imprecisely. This paper looks at how you can choose an external pilot trial sample size in order to minimise the sample size of the overall clinical trial programme, that is, the pilot and the main trial together.

Available at: https://journals.sagepub.com/doi/full/10.1177/0962280215588241

Available on the White Rose Repository at: http://eprints.whiterose.ac.uk/87982/