RCSI Animal Research Ethics Committee

Sample Size Justification Template Guide

Dr Finbarr Leacy, RCSI Data Science Centre
Dr Fiona Boland, RCSI Data Science Centre & HRB Centre for Primary Care Research
v1.1 (October 2017)

Contents

Notes for the applicant

1. Please complete a separate template for each and every proposed experiment and include in the study protocol document.
2. Please provide sufficiently detailed information in each table to allow for independent replication of the sample size calculation (see examples below).
3. Where possible a statistician should be involved in the sample size calculation.

Useful Resources

- 3Rs resources – Library of resources maintained by National Centre for the Replacement, Refinement & Reduction of Animals in Research (NC3Rs)
- ARRIVE – Reporting guidelines for research involving use of animals developed by NC3Rs
- Experimental Design Assistant – Free web application from the NC3Rs
- Experimental Design and Statistics in Biomedical Research – 2002 special issue of ILAR Journal focusing on design and analysis of animal experiments (open-access)
- The Design and Statistical Analysis of Animal Experiments – 2014 special issue of ILAR Journal (open-access)
- G*Power – Open-source statistical software for power analysis and sample size calculation

The following reference books are now available in the RCSI library:

- Bate ST, Clark RA. *The Design and Statistical Analysis of Animal Experiments*. Cambridge, UK: Cambridge University Press; 2014.
- Festing M. *The Design of Animal Experiments.* 2nd ed. London, UK: Sage Publications; 2016.
- Lazic SE. *Experimental Design for Laboratory Biologists: Maximising Information and Improving Reproducibility*. Cambridge, UK: Cambridge University Press; 2017.
- Ruxton GD, Colegrave N. *Experimental Design for the Life Sciences*. 4th ed. Oxford, UK: Oxford University Press; 2016.

Template

Overall Project Aim	Please state the project aim to which this experiment relates
Primary aim of experiment	Please provide a concise description of the primary aim of this experiment
Primary outcome measure (units)	Please state the primary outcome measure and its measurement unit. The primary *outcome* is the *outcome* that an investigator considers to be the most important among the many *outcomes* that are to be examined in the study.
Primary outcome measure type	Delete as appropriate: ▪ Continuous ▪ Count ▪ Categorical ▪ Ordinal ▪ Time-to-event with administrative censoring ▪ Time-to-event without administrative censoring
Experimental Design	Please provide a classification of the study design, e.g. completely randomised design, full factorial design, randomised complete block design etc.
Experimental Factors	Please list the factors/explanatory variables investigated in the study
Treatment Groups	Please provide a concise description of every treatment group investigated in this experiment. Group 1: Group 2: Group 3: Group 4: etc.
Total number of treatment groups	
Proposed primary analysis (e.g. t-test, ANOVA, log-rank test)	Please provide a concise description of the proposed primary analysis
Proposed secondary analyses, if any	Please provide a concise description of any proposed secondary analyses.
Group comparisons of primary interest	Please list all individual group comparisons for which formal statistical tests will be conducted. e.g. 1. Group 1 v Group 2 2. Group 1 v Group 3 3. Group 1 v Group 4 4. Group 2 v Group 4
Total number of group comparisons of primary interest	

Sample size justification	The number of animals should always be large enough to provide a reliable answer to questions addressed. There are many formulae and inputs used to calculate the sample size; sufficient detail should be provided here to allow for independent replication. If the sample size is determined on some other basis, then this should be made clear and justified. Please provide details of relevant inputs to sample size calculation including: a. statistical test b. one- or two-sided test c. equal or unequal group allocation d. minimum clinically important difference / biologically relevant effect size e. significance level α f. power (1-β) Other relevant inputs may include group means, standard deviations or proportions. Please describe the source of relevant input values, e.g. pilot study data, published literature (with citation) Please describe any adjustments made to account for testing of multiple group comparisons (e.g. Bonferroni correction) Please describe any adjustments made to account for potential loss of animals due to adverse events etc. Please provide details of statistical software used to perform sample size calculation, e.g. Stata, G*Power etc. *Every effort should be made to calculate sample sizes using power analysis. However, where sample size calculation is impossible (e.g. it is a pilot study and previous studies cannot be used to provide the required inputs) then please explain why the sample size to be used has been chosen. Please refer to the Resource Equation Method.*
Required sample size per group	
Total number of animals used in this experiment	

Please copy the above table for each experiment in this project

Examples

Please note that these examples are not exhaustive of the range of study designs encountered in animal research.

Example 1: Simple two-arm completely randomised design

Overall Project Aim	Project Aim 1
Primary aim of experiment	The aim of this experiment is to investigate whether a modified diet has an impact on weight.
Primary outcome measure (units)	Weight (grams - g)
Primary outcome measure type	Continuous
Experimental Design	Completely Randomised Design
Experimental Factors	Factor 1: Diet (Normal, Modified)
Treatment Groups	Group 1: Normal Diet Group 2: Modified Diet
Total number of treatment groups	2
Proposed primary analysis (e.g. t-test, ANOVA, log-rank test)	t-test for two independent group means
Proposed secondary analyses, if any	N/A
Group comparisons of primary interest	Group 1 v Group 2
Total number of group comparisons of primary interest	1
Sample size justification	The required sample size per group was calculated based on a two-sided t-test for two independent means comparing group 1 and group 2, assuming equal numbers of animals per group. The minimum clinically important difference, that is the smallest change in mass that would be deemed as important in this experiment, is 5 g. Group 1 mean (standard deviation (sd)): 20 (3) g Group 2 mean (sd): 25 (5) g Group means and standard deviations are derived from unpublished pilot data. α= 0.05 (No adjustment for multiple comparisons required) Power = 0.9 No loss of animals is anticipated in this experiment. All calculations were conducted using G*Power 3.1.9.2.
Required sample size per group	12
Total number of animals to be used in this experiment	2 groups x 12 animals per group = 24

Example 2: A three-arm completely randomized design

Overall Project Aim	Project Aim 1
Primary aim of experiment	The aim of this experiment is to investigate whether diet A and/or diet B have an impact on weight.
Primary outcome measure (units)	Weight (grams - g)
Primary outcome measure type	Continuous
Experimental Design	Completely Randomised Design
Experimental Factors	Factor 1: Diet (Normal, A, B)
Treatment Groups	Group 1: Normal Diet Group 2: Diet A Group 3: Diet B
Total number of treatment groups	3
Proposed primary analysis (e.g. t-test, ANOVA, log-rank test)	One-way ANOVA followed by t-test for two independent means
Proposed secondary analyses, if any	N/A
Group comparisons of primary interest	1. Group 1 v Group 2 2. Group 1 v Group 3
Total number of group comparisons of primary interest	2
Sample size justification	The required sample size per group was calculated based on a two-sided t-test for two independent means comparing group 1 and group 3, assuming equal numbers of animal per group. This is the group comparison of interest for which we expect to observe the smallest mean difference. The minimum clinically important difference, that is the smallest change in mass that would be deemed as important in this experiment, is 5 g. Group 1 mean (standard deviation (sd)): 20 (3) g Group 2 mean (sd): 25 (5) g Group means and standard deviations are derived from unpublished pilot data. α= 0.05/2 = 0.025 (Bonferroni correction for 2 group comparisons) Power = 0.9 No loss of animals is anticipated in this experiment. All calculations were conducted using G*Power 3.1.9.2.
Required sample size per group	14
Total number of animals to be used in this experiment	3 groups x 14 animals per group = 42

Example 3: 2 x 2 factorial design analysed by two-way ANOVA

Project Aim	Project Aim 1
Primary aim of experiment	To investigate the effects of a new drug and exercise on weight.
Primary outcome measure (units)	Weight (grams - g)
Primary outcome measure type	Continuous
Experimental Design	2 x 2 factorial design
Experimental Factors	Factor 1: Treatment (Drug A, no drug) Factor 2: Exercise (exercise, no exercise)
Treatment Groups	Group 1: No drug no exercise Group 2: Drug A, no exercise Group 3: No drug, exercise Group 4: Drug A, exercise
Total number of treatment groups	4
Proposed primary analysis (e.g. t-test, ANOVA, log-rank test)	Two-way ANOVA followed by t-test for two independent means
Proposed secondary analyses, if any	N/A
Group comparisons of primary interest	1. Group 1 v Group 2 2. Group 3 v Group 4 3. Group 1 v Group 3 4. Group 2 v Group 4
Total number of group comparisons of primary interest	4
Sample size justification	The sample size per group was calculated based on a two-sided t-test for two independent means, assuming equal numbers of animals per group. We used the mean and standard deviation of group 1 (no drug, no exercise) as the control mean and standard deviation. The minimum clinically important difference, that is the smallest difference in weight that would be deemed as important, is 2.5 grams. Group 1 mean (standard deviation (sd)): 31.2 (3.0) g Group 2 mean (sd): 28.1 (2.5) g Group 3 mean (sd): 26.4 (2.7) g Group 4 mean (sd): 23.9 (2.9) g Group means and standard deviations are derived from unpublished pilot data. $\alpha = 0.05/4 = 0.0125$ (Bonferroni correction for 4 group comparisons) Power = 0.9 The estimated sample size required per treatment group is 43. Hence, the total sample size required is 172.

	Additionally, given previous experiments, we expect a potential 2% loss of animals due to adverse events. Hence, we have increased the sample size in each group by 2 animals to account for this and thus ensure we have 43 animals per treatment group for analysis. All calculations were conducted using Stata 14.1
Required sample size per group	45
Total number of animals to be used in this experiment	4 groups x 45 animals per group = 180

Example 4: Repeated Measures analysed by two-way ANOVA

<table>
<tr><td>Project Aim</td><td>Project Aim 1</td></tr>
<tr><td>Primary aim of experiment</td><td>To detect phlebitis during the intravenous administration of a particular drug. It is believed that increased temperature in the treated ear may be an early sign of phlebitis.</td></tr>
<tr><td>Primary outcome measure (units)</td><td>Difference in temperature between the treated ear and the untreated ear (Celsius - °C)</td></tr>
<tr><td>Primary outcome measure type</td><td>Continuous</td></tr>
<tr><td>Experimental Design</td><td>Repeated Measures (over time)</td></tr>
<tr><td>Experimental Factors</td><td>Factor 1: Treatment (between-subjects factor)
Treatment 1: The drug in a solution designed to carry the drug
Treatment 2: The carrier solution only (no drug)
Treatment 3: Saline solution
Factor 2: Time since treatment administration (within-subjects factor)
0 minutes
30 minutes
60 minutes
90 minutes</td></tr>
<tr><td>Treatment Groups</td><td>Group 1: The drug in a solution designed to carry the drug
Group 2: The carrier solution only (no drug)
Group 3: Saline solution</td></tr>
<tr><td>Total number of treatment groups</td><td>3</td></tr>
<tr><td>Proposed primary analysis (e.g. t-test, ANOVA, log-rank test)</td><td>Two-way ANOVA followed by t-test for two independent means at 90 mins</td></tr>
<tr><td>Proposed secondary analyses, if any</td><td>t-tests for two independent means at 30 and 60 minutes, respectively</td></tr>
<tr><td>Group comparisons of primary interest</td><td>We are mainly interested in the treatment effect (between-subjects effect) but we are also interested in the following comparisons at 90 minutes:
Group 1 v Group 2
Group 1 v Group 3
Group 2 v Group 3</td></tr>
<tr><td>Total number of group comparisons of primary interest</td><td>3</td></tr>
<tr><td>Sample size justification</td><td>Pilot study results (unpublished) showed the following mean difference in temperature between the treated ear and the untreated ear:
<table><tr><td></td><td>0mins</td><td>30mins</td><td>60mins</td><td>90mins</td></tr><tr><td>Treatment 1</td><td>-0.25</td><td>1.30</td><td>2.01</td><td>2.50</td></tr><tr><td>Treatment 2</td><td>-0.30</td><td>-0.51</td><td>0</td><td>0.10</td></tr><tr><td>Treatment 3</td><td>-0.22</td><td>0.21</td><td>-0.54</td><td>0.23</td></tr></table>
Furthermore, from the pilot study, we assume that the variance of temperature will be 2.25 for all groups at each of the four measurements and that the correlation between the</td></tr>
</table>

	repeated measurements within subjects is 0.7. 1. We are mainly interested in the treatment effect (between-subjects effect). The sample size per treatment group was calculated based on a repeated measures ANOVA, assuming equal numbers of animals per group. α= 0.05 Power = 0.9 The required sample size per treatment was estimated to be 16 (48 in total). 2. However, we are also interested in treatment comparisons at 90 minutes. We expect there to be no difference between treatment group 2 and group 3 and the minimum clinically important difference, that is the smallest difference in temperature between the treatment groups that would be deemed as important, is 1.7 °C. The sample size per treatment group was calculated based on a two-sided t-test for two independent means and assuming equal numbers of animal per group. α= 0.05/3 = 0.017 (Bonferroni correction for 3 group comparisons) Power = 0.9 The estimated sample size per group was 23 (69 in total). As a final sample size we choose the largest of the sample sizes to ensure sufficient animal numbers (i.e. 23 per treatment group). We do not anticipate any loss of animals or data during the experiment. All calculations were conducted using Stata 14.1
Required sample size per group	23
Total number of animals to be used in this experiment	3 groups x 23 animals per group = 69

Example 5: Survival analysis with administrative censoring

Project Aim	Project Aim 1
Primary aim of experiment	To investigate whether drug A and/or drug B is associated with improved survival following tumour resection
Primary outcome measure (units)	Time to euthanasia following tumour resection (days) Administrative censoring at end of study: All animals still alive at 100 days will be euthanised.
Primary outcome measure type	Time-to-event with administrative censoring
Experimental Design	Completely randomised design
Experimental Factors	Factor 1: Drug (vehicle, A, B)
Treatment Groups	Group 1: Vehicle Group 2: Drug A Group 3: Drug B
Total number of treatment groups	3
Proposed primary analysis (e.g. t-test, ANOVA, log-rank test)	Log-rank test
Proposed secondary analyses, if any	N/A
Group comparisons of primary interest	1. Group 1 v Group 2 2. Group 1 v Group 3
Total number of group comparisons of primary interest	2
Sample size justification	The sample size per treatment group was calculated based on a two-sided log-rank test assuming equal numbers of animal per group using the Freedman method. The minimum clinically important difference, that is the smallest difference in the proportion of animals surviving to 100 days, is 0.4. Proportion of animals still alive at 100 days: Group 1: 0.4 Group 2: 0.8 Group 3: 0.9 $\alpha = 0.05/2 = 0.025$ (Bonferroni correction for 2 group comparisons) Power = 0.8 We do not anticipate any loss of animals to other, unrelated causes during the experiment. All calculations were conducted using Stata 14.1
Required sample size per group	33
Total number of animals to be used in this experiment	3 groups x 33 animals per group = 99