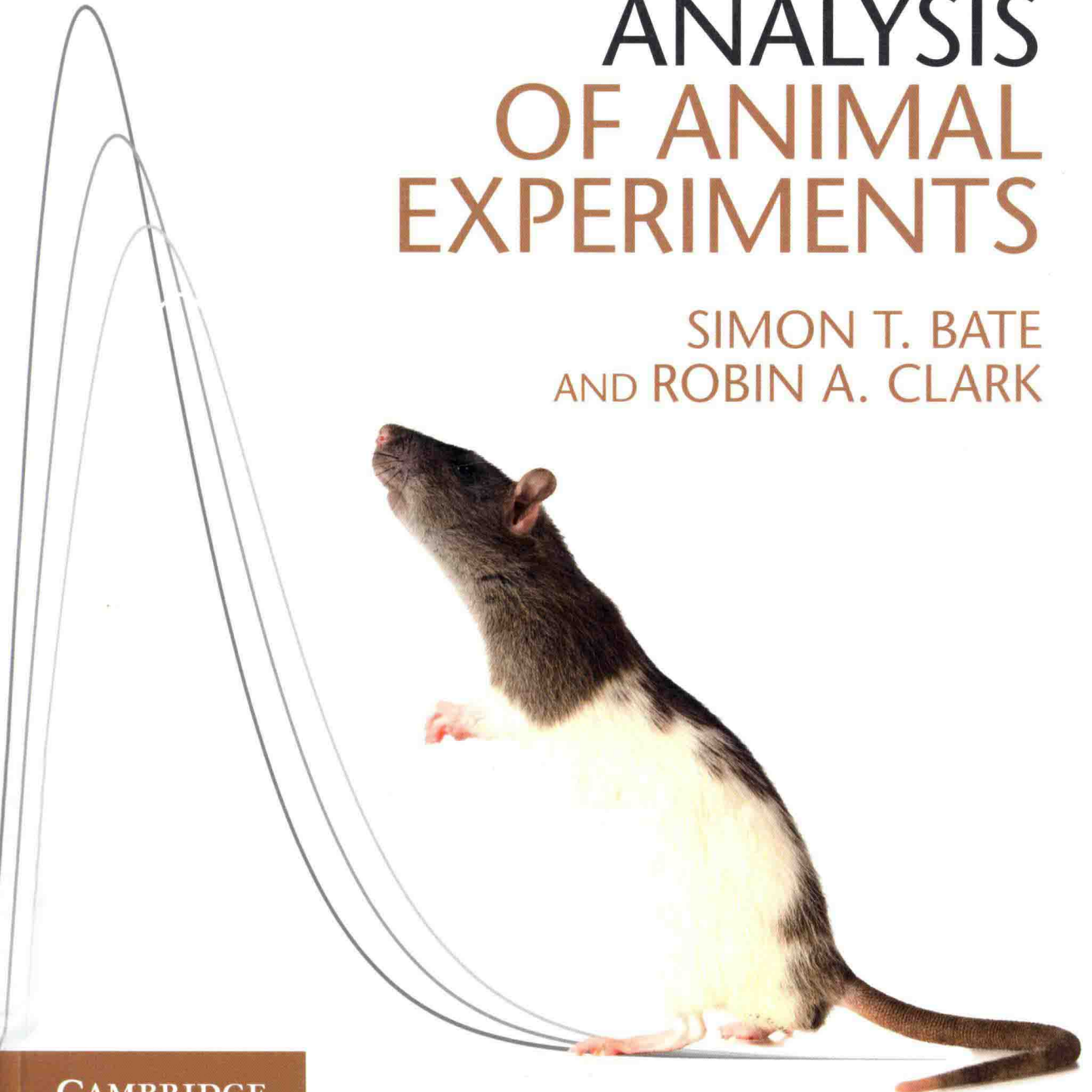


THE DESIGN AND STATISTICAL ANALYSIS OF ANIMAL EXPERIMENTS

SIMON T. BATE
AND ROBIN A. CLARK



CAMBRIDGE

The Design and Statistical Analysis of Animal Experiments

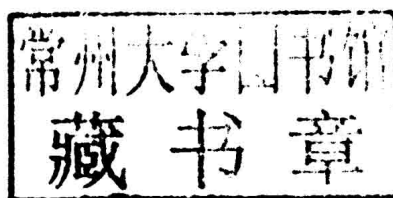
Simon T. Bate

GlaxoSmithKline, UK

and

Robin A. Clark

Huntingdon Life Sciences, UK



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Published in the United States of America by Cambridge University Press, New York

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781107030787

© S. T. Bate and R. A. Clark 2014

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2014

Printed in the United Kingdom by Clays, St Ives plc

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data

Bate, Simon T., 1975–

The design and statistical analysis of animal experiments / Simon Bate, GlaxoSmithKline, UK, Robin Clark, Huntingdon Life Sciences, UK.

pages cm.

Includes bibliographical references and index.

ISBN 978-1-107-03078-7 (hardback) – ISBN 978-1-107-69094-3 (paperback)

1. Animal experimentation–Statistical methods. 2. Experimental design. I. Clark, Robin A., software developer. II. Title. HV4930.D47 2002

590.72'4–dc23

2013033851

ISBN 978-1-107-03078-7 Hardback

ISBN 978-1-107-69094-3 Paperback

Additional resources for this publication at www.cambridge.org/9781107030787

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

The Design and Statistical Analysis of Animal Experiments

Written specifically for animal researchers, this is the first book to provide a comprehensive guide to the design and statistical analysis of animal experiments. It has long been recognised that the proper implementation of these techniques can help to optimise the number of animals used in an experiment. By using real-life examples to make them more accessible, this book explains the statistical tools that are routinely employed by practitioners.

A wide range of design types are considered in detail, including block, factorial, nested, crossover, dose-escalation and repeated measures. Alongside each design, techniques are introduced to analyse the experimental data generated. Each analysis approach is described in non-mathematical terms, helping readers without a statistical background to understand key techniques such as: *t*-tests, ANOVA, repeated measures, analysis of covariance, multiple comparison tests, non-parametric methods and survival analysis.

This is also the first text to describe technical aspects of InVivoStat, a powerful open-source software package developed by the authors to enable animal researchers to analyse their data and obtain informative results. InVivoStat can be downloaded at www.invivostat.co.uk.

Simon T. Bate is a Principal Statistician at GlaxoSmith-Kline, supporting pre-clinical research. He has spent over 12 years supporting the design and statistical analysis of animal experiments, including drug discovery research, toxicology studies and safety assessment. He presents many statistics courses around Europe, including the statistics module of the British Association for Psychopharmacology's Pre-Clinical Certificate.

Robin A. Clark is a Senior Analyst Programmer at Huntingdon Life Sciences, Alconbury. Originally qualified as a marine ecologist, he is now a software architect producing data collection and statistical applications for the pharmaceutical industry. Robin produced the user interface and designed the system architecture for InVivoStat.

To

RB, NRB, MTB, EC, EJC, ZMC

Preface

This book is aimed at practitioners who do not have a statistics degree and yet wish to apply statistics to help them arrive at valid and reliable conclusions while minimising the animal numbers required. Descriptions of the mathematical methods underpinning the topics covered in the book are purposefully kept to a minimum. If readers wish to gain a better understanding of the mathematics behind experimental design and statistical analysis then reading a more advanced textbook would help further their understanding.

The solutions to practical problems encountered when conducting animal experiments are explained using non-technical approaches. We believe that in many situations advanced statistical ideas can be employed successfully by researchers with no statistical qualification, using a combination of common sense and modern statistical analysis software packages. In our experience statistical ideas are often introduced to scientists using mathematical terminology. This can be off-putting to non-mathematicians and can leave researchers with, at best, only rudimentary statistical tools and at worst a fear of statistics.

To keep the descriptions of the statistical tools covered in this book as simple as possible, we shall occasionally give pragmatic explanations. While such explanations may not apply in all cases and in all scientific disciplines, this approach does allow us to introduce methods in a clear and concise way. By allowing ourselves the freedom to simplify the problems pragmatically, we aim to make statistical tools more accessible. The reader is invited, once they have familiarised themselves with (and hopefully found the benefit of using) the tools described in this book, to read more advanced texts on the subject.

This book is divided into seven chapters which loosely correspond to the procedure a researcher should take when planning the experimental design, running the experiment and evaluating the data generated. Following an introductory chapter and a second describing certain statistical concepts, the third chapter covers different types of designs. Designs are outlined, where possible, in simple non-technical language. This is followed by a chapter describing the randomisation of the experimental material. The fifth chapter discusses the statistical analysis of animal experiments and this is followed by a chapter describing how these methods can be applied within the statistical software package

InVivoStat. The final chapter draws some conclusions about the ideas contained within the text.

A scientist can apply all of the methodology described in this book. Certain topics covered are more advanced than others and while we aim to make all subjects accessible, the reader should be aware that the help of a professional statistician may be advisable when first implementing some of the more advanced tools. However, once the readers have familiarised themselves with the ideas contained within this book, we hope they will have a fuller appreciation of the help statistics can offer to improve the conclusions that can be made when running animal experiments.

Acknowledgments

We would like to thank all our colleagues, both past and present, for providing so much inspiration for this book. We hope this text reflects the discussions, consultations and the range of examples that we have worked on together. There are simply too many scientists to mention, but you know who you are! We hope the wide variety of interesting practical problems we have faced together is reflected throughout this book.

Special thanks also to the statisticians who have advised us over the years and shaped our views on this subject matter. Simon would especially like to thank Philip Overend and Andrew Lloyd for their input and encouragement and Janet and Ed Godolphin for their help and support over the years.

We would like to give an extra special thanks to Gillian Amphlett, Gill Fleetwood, David Willé, Clare Stanford and Nathalie Percie du Sert for giving up their time to review sections of the book, and providing ideas, advice and feedback. Their input has been most helpful and improved the text considerably.

Finally we would like to thank our friends and family for putting up with the long unsociable hours we spent writing this book.

Contents

Preface	<i>page</i> xiii
Acknowledgments	xv
1 Introduction	1
1.1 Structure of this book	3
1.1.1 Introductory sections	4
1.1.2 Approaches to consider when setting up a new animal model	4
1.1.3 Approaches to consider when generating hypotheses	5
1.1.4 Approaches to consider when testing hypotheses	5
1.2 Statistical problems faced by animal researchers	5
1.3 Pitfalls encountered when applying statistics in practice	6
1.3.1 Pitfalls with experimental design	6
1.3.2 Pitfalls with randomisation	9
1.3.3 Pitfalls with statistical analysis	10
1.3.4 Pitfalls when reporting animal experiments	13
1.4 So where does statistics fit in?	15
1.5 The ARRIVE guidelines	15
2 Statistical concepts	18
2.1 Decision-making: the signal-to-noise ratio	18
2.2 Probability distributions	19
2.2.1 The frequency distribution	20
2.2.2 The density distribution	20
2.2.3 The probability distribution	21
2.2.4 The normal distribution	21
2.2.5 The chi-squared distribution	22
2.2.6 The t -distribution	22
2.2.7 The F -distribution	23

2.3	The hypothesis testing procedure	23	3.2.8	Repeatedly measuring the animal	45
2.3.1	The null and alternative hypotheses	23	3.3	Summary of design types	46
2.3.2	The p -value	25	3.3.1	Block designs	46
2.3.3	The significance level	25	3.3.2	Factorial designs	47
2.3.4	Significant stars	26	3.3.3	Dose-response designs	47
2.3.5	Type I and Type II errors	26	3.3.4	Nested designs	47
2.4	Exploratory vs. confirmatory experiments	28	3.3.5	Split-plot designs	48
2.5	The estimation process	29	3.3.6	Repeated measures and dose-escalation designs	48
3	Experimental design	30	3.3.7	Designs applied in practice	48
3.1	Why design experiments?	30	3.4	Block designs	49
3.1.1	Practical reasons	30	3.4.1	Practical reasons to block	49
3.1.2	Statistical reasons: variability, the signal and bias	31	3.4.2	Statistical reasons to block	49
3.2	What does an experimental design involve?	32	3.4.2.1	Variance reduction	49
3.2.1	Variables to be recorded	32	3.4.2.2	Bias reduction	51
3.2.1.1	Types of response	32	3.4.3	How to block	51
3.2.1.2	Reporting responses	34	3.4.4	Complete block designs	53
3.2.1.3	Baseline responses	34	3.4.4.1	Efficiency	53
3.2.1.4	Recording conditions during the experiment	35	3.4.4.2	Randomisation	53
3.2.2	Set of treatments	35	3.4.4.3	Statistical analysis of block designs	54
3.2.3	The experimental unit and the observational unit	36	3.4.5	Incomplete block designs	54
3.2.4	Effects and factors	37	3.4.6	Balanced incomplete block design	55
3.2.4.1	Defining factor level labels	39	3.4.6.1	Efficiency	55
3.2.4.2	Defining the factors in an experimental design	39	3.4.6.2	Randomisation	55
3.2.5	Fixed and random factors	39	3.4.6.3	Statistical analysis	55
3.2.5.1	Fixed factors	40	3.4.7	More than one block: the row-column block design	56
3.2.5.2	Random factors	40	3.4.7.1	Efficiency	56
3.2.5.3	Random or fixed?	41	3.4.7.2	Randomisation	56
3.2.6	Categorical factors and continuous factors	42	3.4.7.3	Statistical analysis	56
3.2.7	Crossed factors and nested factors	42	3.4.8	Row-column block designs based on Latin squares	57
3.2.7.1	Nested factors	42	3.4.8.1	Efficiency	58
3.2.7.2	Crossed factors	43	3.4.8.2	Randomisation	58
3.2.7.3	Partially crossed factors	45	3.4.8.3	Statistical analysis	58
3.2.7.4	Designs containing nested and crossed factors	45	3.4.9	Crossover designs	59
			3.4.9.1	Complete crossover designs	59
			3.4.9.2	Incomplete crossover designs	60
			3.4.9.3	The benefits of crossover designs	61

3.4.9.4	The issues with crossover designs	62	3.6.3.3	Adding an offset to the dose	88
3.4.9.5	Treatment carry-over effects	62	3.7	Nested designs	90
3.5	Factorial design	63	3.7.1	Types of nested design	91
3.5.1	Randomisation	64	3.7.1.1	Single-order nested design	91
3.5.2	Categorical factors and interactions	64	3.7.1.2	Higher-order nested design	91
3.5.3	Small factorial designs	66	3.7.2	Sample size and power	93
3.5.4	Large factorial designs	68	3.7.2.1	Factors that influence sample size	93
3.5.4.1	Strategies when setting up a new animal model	68	3.7.2.2	Calculating sample sizes	95
3.5.4.2	Graphical representation of large factorial designs	70	3.7.2.3	When not to calculate the statistical power	97
3.5.4.3	Hidden replication	70	3.7.3	Higher-order nested designs	99
3.5.4.4	Fractional factorial designs to reduce animal use	72	3.7.3.1	Identifying nested factors	99
3.5.4.5	Two-stage procedure to reduce animal use	75	3.7.3.2	Investigating the sources of variability in higher-order nested designs	101
3.5.5	Factorial designs with continuous factors	77	3.7.3.3	Variance components: estimating the observational unit variability	102
3.5.5.1	Strategies for setting up a new animal model	78	3.7.3.4	Predicting the experimental unit variability	103
3.5.5.2	Drug combination studies	81	3.7.3.5	Investigating alternative nested designs	105
3.5.5.3	Continuous vs. categorical factors	83	3.7.3.6	Pseudo-replication	106
3.5.6	Final thoughts on factorial designs	83	3.8	Repeated measures and dose-escalation designs	110
3.6	Dose-response designs	84	3.8.1	Repeated measures designs	110
3.6.1	The four- and five-parameter logistic curves	84	3.8.1.1	The repeated factor	110
3.6.2	Experimental design considerations	85	3.8.1.2	The core experimental design	112
3.6.2.1	Increasing the number of doses	86	3.8.1.3	Nested repeated measures designs	112
3.6.2.2	Decreasing the number of animals	86	3.8.1.4	More complex repeated measures designs	114
3.6.3	Including the control group	87	3.8.2	Dose-escalation designs	116
3.6.3.1	Analysing a change from the control response	87	3.8.2.1	More complex dose-escalation designs	117
3.6.3.2	Using a dual statistical model	88			

3.9	Split-plot designs	117	5.2.1.3	The predicted mean	137
3.9.1	Animals as whole plots	117	5.2.1.4	The geometric mean	137
3.9.2	Animals as subplots	118	5.2.2	Parametric measures of spread	138
3.10	Experimental designs in practice	119	5.2.2.1	Variance	138
3.11	A good design should result in...	120	5.2.2.2	Standard deviation	138
4	Randomisation	122	5.2.2.3	Standard error of the mean	138
4.1	Practical reasons to randomise	122	5.2.2.4	Confidence intervals	139
4.1.1	Bias reduction	122	5.2.2.5	Coefficient of variation	139
4.1.1.1	Removing unforeseen trends	123	5.2.3	Non-parametric measures of location	139
4.1.1.2	Humans are systematic	123	5.2.4	Non-parametric measures of spread	140
4.1.2	Blinding	124	5.3	Graphical tools	140
4.2	Statistical reasons to randomise	124	5.3.1	Scatterplots	140
4.2.1	Estimating the variability	125	5.3.2	Box-plots	142
4.2.2	Deciding upon the statistical analysis strategy	125	5.3.3	Histograms	143
4.2.2.1	Including interactions in the statistical model	126	5.3.4	Categorised case profiles plot	144
4.2.2.2	Including blocking factors	127	5.3.5	Means with SEMs plot	145
4.2.3	Repeatedly measured responses	127	5.3.5.1	Problems with the means with SEMs plot	145
4.2.3.1	Repeated factors and randomised factors	127	5.3.5.2	Benefits of the means with SEMs plot	151
4.2.3.2	Block and dose-escalation designs	127	5.4	Parametric analysis	151
4.2.3.3	Crossover and dose-escalation designs	128	5.4.1	Parametric assumptions	152
4.2.3.4	Including interactions involving the repeated factor	129	5.4.1.1	Numeric and continuous responses	152
4.3	What to randomise	129	5.4.1.2	Normally distributed residuals	153
4.4	How to randomise	130	5.4.1.3	Homogeneity of variance	155
5	Statistical analysis	132	5.4.1.4	Independence of the responses	158
5.1	Introduction	132	5.4.1.5	Removal of outliers	159
5.1.1	InVivoStat	133	5.4.1.6	Additivity	162
5.1.2	A recommended five-stage parametric analysis procedure	133	5.4.2	The <i>t</i> -test	163
5.2	Summary statistics	135	5.4.2.1	The unpaired <i>t</i> -test	163
5.2.1	Parametric measures of location	135	5.4.2.2	When not to use an unpaired <i>t</i> -test	165
5.2.1.1	The true mean and the sample mean	135	5.4.2.3	The paired <i>t</i> -test	167
5.2.1.2	The observed mean	136	5.4.2.4	Randomisation and the paired <i>t</i> -test	168

5.4.3	Analysis of variance (ANOVA)	168	5.4.6.5	Predicted group means	203
5.4.3.1	One-way ANOVA	169	5.4.6.6	Assumptions for ANCOVA	204
5.4.3.2	Including the positive control	173	5.4.6.7	Strategy for when the independence assumption does not hold	207
5.4.3.3	Two-way ANOVA	174	5.4.6.8	ANCOVA and stratified randomisation	208
5.4.3.4	Two-way vs. one-way ANOVA	176	5.4.6.9	Change from baseline responses	208
5.4.3.5	Dealing with missing factor combinations	177	5.4.7	Regression analysis	211
5.4.4	Repeated measures analysis	179	5.4.8	Multiple comparison procedures	212
5.4.4.1	Categorised case profiles plot	181	5.4.8.1	The risk of finding false positives and false negatives	212
5.4.4.2	Analysis of summary measures	181	5.4.8.2	Choosing the family of tests	214
5.4.4.3	Repeated measures analysis	189	5.4.8.3	Unadjusted tests	215
5.4.4.4	The mixed-model approach vs. the ANOVA-based approach	191	5.4.8.4	Stepwise multiple comparison procedures that control the FDR	218
5.4.4.5	Advantages and disadvantages of the repeated measures analysis	195	5.4.8.5	Simultaneous multiple comparison procedures that control the FWE	218
5.4.5	Predicted means from the parametric analysis	196	5.4.8.6	Stepwise multiple comparison procedures based on group differences that control the FWE	222
5.4.5.1	Least square (predicted) means	196	5.4.8.7	Stepwise-based multiple comparison procedures based on p -values that control the FWE	223
5.4.5.2	Variability of the least square (predicted) means	197	5.4.8.8	The gateway ANOVA approach	224
5.4.5.3	Geometric means and confidence intervals	197	5.4.8.9	Multiple comparison procedures in statistical software packages	227
5.4.5.4	Reliability of the predicted means	198	5.4.8.10	Recommendations	228
5.4.6	Analysis of covariance (ANCOVA)	199	5.5	Other useful analyses	228
5.4.6.1	What is a covariate?	200	5.5.1	Non-parametric analyses	228
5.4.6.2	Best-fit lines and predicted lines	201	5.5.1.1	When to use a	
5.4.6.3	Categorised scatterplot	201			
5.4.6.4	Predictions from ANCOVA	202			

	non-parametric test	229		6.3.3.5 Analysis of designs with missing factor combinations	252
	5.5.1.2 Non-parametric tests	230			
5.5.2	Testing the difference between proportions	231	6.4	Repeated Measures Parametric Analysis module	252
	5.5.2.1 Analysis procedure	232		6.4.1 Analysis procedure	252
	5.5.2.2 Chi-squared test	232		6.4.2 Worked example	255
	5.5.2.3 Fisher's exact test	233		6.4.3 Technical details	255
5.5.3	Survival analysis	234	6.5	<i>P</i> -Value Adjustment module	258
	5.5.3.1 The survival function	235		6.5.1 Analysis procedure	259
	5.5.3.2 Comparing groups	236		6.5.2 Worked example	259
6	Analysis using InVivoStat	238	6.6	Non-Parametric Analysis module	260
6.1	Getting started	238		6.6.1 Analysis procedure	260
	6.1.1 Data import	238		6.6.2 Worked example	262
	6.1.1.1 Single measure format	238	6.7	Graphics module	262
	6.1.1.2 Repeated measures format	239		6.7.1 Analysis procedure	262
6.1.2	Importing a dataset into InVivoStat: Excel import	240		6.7.2 Example plots	263
6.1.3	Importing a dataset into InVivoStat: text file import	240	6.8	Power Analysis module	263
6.1.4	Data management	240		6.8.1 Analysis procedure	263
6.1.5	Running an analysis	240		6.8.2 Worked example	265
6.1.6	Warning and error messages	241	6.9	Unpaired <i>t</i> -test Analysis module	267
6.1.7	Log file	241		6.9.1 Analysis procedure	267
6.1.8	Exporting results	241		6.9.2 Worked example	271
6.2	Summary Statistics module	241	6.10	Paired <i>t</i> -test/within-subject Analysis module	272
	6.2.1 Analysis procedure	242		6.10.1 Analysis procedure	272
	6.2.2 Worked example	243		6.10.2 Worked example	276
6.3	Single Measure Parametric Analysis module	243	6.11	Dose-Response Analysis module	277
	6.3.1 Analysis procedure	243		6.11.1 Technical details on curve fitting	277
	6.3.2 Worked example	245		6.11.2 Fitting logistic curves to data	278
	6.3.3 Technical details	248		6.11.3 Analysis of quantitative assays	278
	6.3.3.1 Analysis of large factorial experiments	248		6.11.4 Analysis procedure	279
	6.3.3.2 Analysis of small factorial experiments	248		6.11.5 Worked example: a biological assay	281
	6.3.3.3 Analysis of experiments involving blocking factors	249		6.11.6 User-defined equation option	282
	6.3.3.4 Analysis of crossover trials	251	6.12	Chi-squared Test and Fisher's Exact Test module	282
				6.12.1 Analysis procedure	283
				6.12.2 Worked example	284
			6.13	R-Runner module	285
			6.14	Nested Design Analysis module	285

6.14.1	Analysis procedure	286	7	Conclusion	293
6.14.2	Worked example	289			
6.15	Survival Analysis module	289		Glossary	295
6.15.1	Analysis procedure	289		References	297
6.15.2	Worked example	291		Index	303

Introduction

Many researchers, either directly or indirectly, rely on statistical ideas when carrying out animal experiments. While some statistical tools are well known and are applied routinely, other tools are less well understood and so are less well used. The overall aim of this book is to discuss statistical methodologies that can be applied throughout the many stages of the experimental process. Researchers should be able to carry out most of the techniques described, although the advice of a professional statistician is advisable for some of the more advanced topics. Making use of these techniques will ensure that experiments are conducted in a logical and efficient way, which should result in reliable and reproducible decisions.

The particular types of study addressed in this book, as the title suggests, are studies involving animals. We attempt to cover all of the statistical tools that the animal researcher should use to run successful studies. Of course many of the problems faced by the animal researcher are common to other disciplines, and hence the ideas contained within this book can be applied to other areas. It should be noted that certain topics described in the text have been simplified to allow non-statisticians to apply the ideas without professional statistical support. Such pragmatic descriptions, while simplifying the technical details, are not universal and will not be applicable in all scientific disciplines.

There has been much interest in the use of statistics in animal research, in particular in the application of the 3Rs, replacement, reduction and refinement, as described by Russell and Burch (1959):

Every time any particle of statistical method is properly used, fewer animals are employed than would otherwise have been necessary.

Many authors have since highlighted how important the use of good experimental design is when conducting animal experiments; see Festing (1994, 2003a, 2003b) and the references contained within. Some of the more practical, as well as statistical, aspects of experimental design and statistics when applying the 3Rs are described in the book by Festing *et al.* (2002). There have also been surveys into the use of statistics in refereed journals; see McCance (1995) and more recently Kilkenny *et al.* (2009). The latter draws attention to some of the mistakes that can be made by researchers when designing and analysing animal experiments. The reliability of the reporting of animal experiments has been considered in, for example, Macleod *et al.* (2009) and Rooke *et al.* (2011). These articles highlight that papers describing experiments that do not employ suitable randomisation techniques and/or blinding may contain biased results.

The main goal of this text is to demonstrate how statistics can aid the reduction and refinement of animal studies. The efficient use of statistics, both in terms of complex experimental design and powerful statistical analysis, can reduce the number of animals required. Statistics can also help the researcher understand the processes that underpin the animal model and help identify factors that are influencing the experimental results. Such an understanding will inevitably lead to a refinement in the experimental process and a reduction in the total number of animals used.

Statistics, as a discipline, provides researchers with tools to help them arrive at valid conclusions. However, statistics, along with the application of some common sense, can also increase the understanding of the animal model through the application of graphical and mathematical techniques. For example, graphical tools play an important role in helping the researcher understand the effect of the features of the experimental design and also uncover overall patterns present in the data. The application of a formal statistical test, without first investigating the data graphically, can lead to the researcher drawing incorrect conclusions from the data. Consider the following real-life case study, which used graphical, as well as statistical, tools. If a conventional statistical analysis had been carried out, without first investigating all of the information gathered within the experiment, then the conclusions would have been misleading.

Example 1.1: Reducing blood cholesterol levels in mice

A scientist wanted to test the hypothesis that a novel compound had a beneficial effect on reducing high-density lipoprotein (HDL) cholesterol levels in a transgenic C57Bl/6J strain of mice. A blood sample was taken pre-treatment and the baseline cholesterol level for each animal measured. The mice were then randomised to either the drug treatment group or the control group and dosed with either the drug treatment or vehicle twice daily for two weeks. At the end of this period, a terminal blood sample was taken and the HDL cholesterol level measured.

As the scientist wanted to make use of the baseline information in the statistical analysis, it was decided that the percentage change from baseline would be a suitable response to investigate. This would, the scientist hoped, effectively remove the animal-to-animal differences by normalising to the baseline level. While there was evidence of a decrease in HDL cholesterol level in the group of animals administered the drug treatment (a 20% decrease from baseline in the drug treatment group compared to a 10% decrease in the control group) this was not deemed statistically significant using an unpaired *t*-test ($p = 0.191$). A means with standard errors of the mean (SEMs) plot of the data (see Section 5.3.5) is presented in Figure 1.1.

As a follow-up the scientist also analysed the terminal HDL cholesterol level. From this analysis it appeared that there was a statistically significant increase in cholesterol level in the drug-treated group compared to the control. A plot of the means with SEMs of the terminal HDL cholesterol level is presented in Figure 1.2.

Based on the results of this experiment, should we conclude the drug increases cholesterol levels? And why did the two analyses give such different conclusions? These questions can be answered

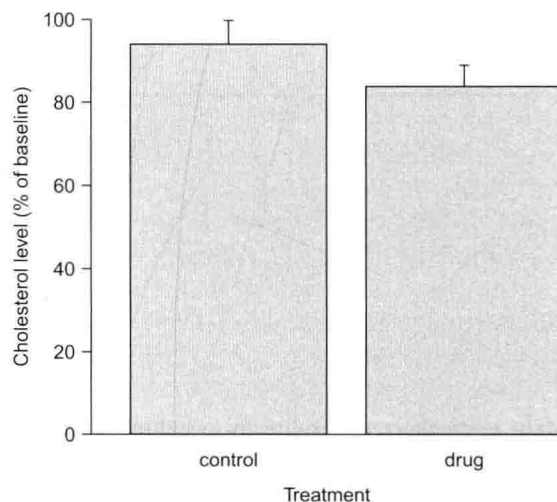


Figure 1.1. Plot of treatment means with standard errors for the percentage of baseline cholesterol response for Example 1.1.

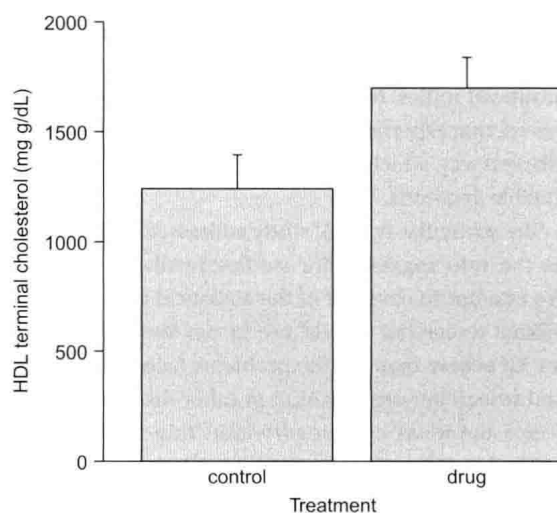


Figure 1.2. Plot of treatment means with standard errors for the terminal HDL cholesterol for Example 1.1.

by a simple scatterplot of the measured HDL cholesterol levels. If we plot terminal vs. baseline HDL cholesterol levels, an underlying problem with the experiment becomes clear. The scatterplot is presented in Figure 1.3.

From Figure 1.3 it can be seen that there are two distinct groupings along the *X*-axis. The plot reveals that, in terms of the HDL baseline cholesterol level, the animals belong to one of two