

FOURTH EDITION

Medical Statistics *from* Scratch

An Introduction for Health Professionals



DAVID BOWERS

WILEY Blackwell

Medical Statistics from Scratch

Medical Statistics from Scratch

An Introduction for Health Professionals

Fourth Edition

David Bowers

Visiting Research Fellow

Leeds Institute of Health Sciences

School of Medicine

University of Leeds

Leeds, UK

WILEY Blackwell

This edition first published 2020
© 2020 John Wiley and Sons Ltd

Edition History

John Wiley and Sons (3e, 2014)

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by law. Advice on how to obtain permission to reuse material from this title is available at <http://www.wiley.com/go/permissions>.

The rights of David Bowers to be identified as the author of this work has been asserted in accordance with law.

Registered Office(s)

John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial Office

9600 Garsington Road, Oxford, OX4 2DQ, UK

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.

Limit of Liability/Disclaimer of Warranty

The contents of this work are intended to further general scientific research, understanding, and discussion only and are not intended and should not be relied upon as recommending or promoting scientific method, diagnosis, or treatment by physicians for any particular patient. In view of ongoing research, equipment modifications, changes in governmental regulations, and the constant flow of information relating to the use of medicines, equipment, and devices, the reader is urged to review and evaluate the information provided in the package insert or instructions for each medicine, equipment, or device for, among other things, any changes in the instructions or indication of usage and for added warnings and precautions. While the publisher and authors have used their best efforts in preparing this work, they make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives, written sales materials or promotional statements for this work. The fact that an organization, website, or product is referred to in this work as a citation and/or potential source of further information does not mean that the publisher and authors endorse the information or services the organization, website, or product may provide or recommendations it may make. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for your situation. You should consult with a specialist where appropriate. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

Library of Congress Cataloging-in-Publication Data

Names: Bowers, David, 1938– author.

Title: Medical statistics from scratch : an introduction for health professionals / David Bowers.

Description: 4th edition. | Hoboken NJ : Wiley, 2020. | Includes bibliographical references and index. |

Identifiers: LCCN 2019015126 (print) | LCCN 2019015564 (ebook) | ISBN

9781119523925 (Adobe PDF) | ISBN 9781119523949 (ePub) | ISBN 9781119523888 (pbk.)

Subjects: | MESH: Biometry | Statistics as Topic

Classification: LCC RA409 (ebook) | LCC RA409 (print) | NLM WA 950 | DDC 610.72/7–dc23

LC record available at <https://lcn.loc.gov/2019015126>

Cover Design: Wiley

Cover Image: Courtesy of David Bowers

Set in 10/12pt Minion by SPi Global, Pondicherry, India

Brief Contents

Preface to the 4th Edition	xix
Preface to the 3rd Edition	xxi
Preface to the 2nd Edition	xxiii
Preface to the 1st Edition	xxv
Introduction	xxvii
I Some Fundamental Stuff	1
1 First things first – the nature of data	3
II Descriptive Statistics	15
2 Describing data with tables	17
3 Every picture tells a story – describing data with charts	31
4 Describing data from its shape	51
5 Measures of location – Numbers R Us	62
6 Measures of spread – Numbers R Us – (again)	75
7 Incidence, prevalence, and standardisation	92
III The Confounding Problem	111
8 Confounding – like the poor, (nearly) always with us	113

IV Design and Data	125
9 Research design – Part I: Observational study designs	127
10 Research design – Part II: Getting stuck in – experimental studies	146
11 Getting the participants for your study: ways of sampling	156
V Chance Would Be a Fine Thing	165
12 The idea of probability	167
13 Risk and odds	175
VI The Informed Guess – An Introduction to Confidence Intervals	191
14 Estimating the value of a <i>single</i> population parameter – the idea of confidence intervals	193
15 Using confidence intervals to compare two population parameters	206
16 Confidence intervals for the <i>ratio</i> of two population parameters	224
VII Putting it to the Test	235
17 Testing hypotheses about the <i>difference</i> between two population parameters	237
18 The Chi-squared (χ^2) test – what, why, and how?	261
19 Testing hypotheses about the <i>ratio</i> of two population parameters	276
VIII Becoming Acquainted	283
20 Measuring the association between two variables	285
21 Measuring agreement	298
IX Getting into a Relationship	307
22 Straight line models: linear regression	309
23 Curvy models: logistic regression	334
24 Counting models: Poisson regression	349

X	Four More Chapters	363
25	Measuring survival	365
26	Systematic review and meta-analysis	380
27	Diagnostic testing	393
28	Missing data	400
	Appendix: Table of random numbers	414
	References	415
	Solutions to exercises	424
	Index	457

Contents

Preface to the 4th Edition	xix
Preface to the 3rd Edition	xxi
Preface to the 2nd Edition	xxiii
Preface to the 1st Edition	xxv
Introduction	xxvii
I Some Fundamental Stuff	1
1 First things first – the nature of data	3
Variables and data	3
Where are we going ...?	5
The good, the bad, and the ugly – types of variables	5
Categorical data	6
Nominal categorical data	6
Ordinal categorical data	7
Metric data	8
Discrete metric data	8
Continuous metric data	9
How can I tell what type of variable I am dealing with?	10
The baseline table	11
II Descriptive Statistics	15
2 Describing data with tables	17
Descriptive statistics. What can we do with raw data?	18
Frequency tables – nominal data	18
The frequency distribution	19
Relative frequency	20
Frequency tables – ordinal data	20
Frequency tables – metric data	22
Frequency tables with discrete metric data	22
Cumulative frequency	24
Frequency tables with continuous metric data – grouping the raw data	25
Open-ended groups	27

Cross-tabulation – contingency tables	28
Ranking data	30
3 Every picture tells a story – describing data with charts	31
Picture it!	32
Charting nominal and ordinal data	32
The pie chart	32
The simple bar chart	34
The clustered bar chart	35
The stacked bar chart	37
Charting discrete metric data	39
Charting continuous metric data	39
The histogram	39
The box (and whisker) plot	42
Charting cumulative data	44
The cumulative frequency curve with discrete metric data	44
The cumulative frequency curve with continuous metric data	44
Charting time-based data – the time series chart	47
The scatterplot	48
The bubbleplot	49
4 Describing data from its shape	51
The shape of things to come	51
Skewness and kurtosis as measures of shape	52
Kurtosis	55
Symmetric or mound-shaped distributions	56
Normalness – the Normal distribution	56
Bimodal distributions	58
Determining skew from a box plot	59
5 Measures of location – Numbers R us	62
Numbers, percentages, and proportions	62
Preamble	63
Numbers, percentages, and proportions	64
Handling percentages – for those of us who might need a reminder	65
Summary measures of location	67
The mode	68
The median	69
The mean	70
Percentiles	71
Calculating a percentile value	72
What is the most appropriate measure of location?	73
6 Measures of spread – Numbers R us – (again)	75
Preamble	76
The range	76
The interquartile range (IQR)	76
Estimating the median and interquartile range from the cumulative frequency curve	77
The boxplot (also known as the box and whisker plot)	79
Standard deviation	82
Standard deviation and the Normal distribution	84

Testing for Normality	86
Using SPSS	86
Using Minitab	87
Transforming data	88
7 Incidence, prevalence, and standardisation	92
Preamble	93
The incidence rate and the incidence rate ratio (IRR)	93
The incidence rate ratio	94
Prevalence	94
A couple of difficulties with measuring incidence and prevalence	97
Some other useful rates	97
Crude mortality rate	97
Case fatality rate	98
Crude maternal mortality rate	99
Crude birth rate	99
Attack rate	99
Age-specific mortality rate	99
Standardisation – the age-standardised mortality rate	101
The direct method	102
The standard population and the comparative mortality ratio (CMR)	103
The indirect method	106
The standardised mortality rate	107
III The Confounding Problem	111
8 Confounding – like the poor, (nearly) always with us	113
Preamble	114
What is confounding?	114
Confounding by indication	117
Residual confounding	119
Detecting confounding	119
Dealing with confounding – if confounding is such a problem, what can we do about it?	120
Using restriction	120
Using matching	121
Frequency matching	121
One-to-one matching	121
Using stratification	122
Using adjustment	122
Using randomisation	122
IV Design and Data	125
9 Research design – Part I: Observational study designs	127
Preamble	128
Hey ho! Hey ho! It's off to work we go	129
Types of study	129
Observational studies	130
Case reports	130

Case series studies	131
Cross-sectional studies	131
Descriptive cross-sectional studies	132
Confounding in descriptive cross-sectional studies	132
Analytic cross-sectional studies	133
Confounding in analytic cross-sectional studies	134
From here to eternity – cohort studies	135
Confounding in the cohort study design	139
Back to the future – case–control studies	139
Confounding in the case–control study design	141
Another example of a case–control study	142
Comparing cohort and case–control designs	143
Ecological studies	144
The ecological fallacy	145
10 Research design – Part II: getting stuck in – experimental studies	146
Clinical trials	147
Randomisation and the randomised controlled trial (RCT)	148
Block randomisation	149
Stratification	149
Blinding	149
The crossover RCT	150
Selection of participants for an RCT	153
Intention to treat analysis (ITT)	154
11 Getting the participants for your study: ways of sampling	156
From populations to samples – statistical inference	157
Collecting the data – types of sample	158
The simple random sample and its offspring	159
The systematic random sample	159
The stratified random sample	160
The cluster sample	160
Consecutive and convenience samples	161
How many participants should we have? Sample size	162
Inclusion and exclusion criteria	162
Getting the data	163
V Chance Would Be a Fine Thing	165
12 The idea of probability	167
Preamble	167
Calculating probability – proportional frequency	168
Two useful rules for simple probability	169
Rule 1. The multiplication rule for independent events	169
Rule 2. The addition rule for mutually exclusive events	170
Conditional and Bayesian statistics	171
Probability distributions	171
Discrete versus continuous probability distributions	172

The binomial probability distribution	172
The Poisson probability distribution	173
The Normal probability distribution	174
13 Risk and odds	175
Absolute risk and the absolute risk reduction (ARR)	176
The risk ratio	178
The reduction in the risk ratio (or relative risk reduction (RRR))	178
A general formula for the risk ratio	179
Reference value	179
Number needed to treat (NNT)	180
What happens if the initial risk is small?	181
Confounding with the risk ratio	182
Odds	183
Why you can't calculate risk in a case-control study	185
The link between probability and odds	186
The odds ratio	186
Confounding with the odds ratio	189
Approximating the risk ratio from the odds ratio	189
VI The Informed Guess – An Introduction to Confidence Intervals	191
14 Estimating the value of a <i>single</i> population parameter – the idea of confidence intervals	193
Confidence interval estimation for a population mean	194
The standard error of the mean	195
How we use the standard error of the mean to calculate a confidence interval for a population mean	197
Confidence interval for a population proportion	200
Estimating a confidence interval for the median of a single population	203
15 Using confidence intervals to compare two population parameters	206
What's the difference?	207
Comparing two <i>independent</i> population means	207
An example using birthweights	208
Assessing the evidence using the confidence interval	211
Comparing two <i>paired</i> population means	215
Within-subject and between-subject variations	215
Comparing two <i>independent</i> population proportions	217
Comparing two <i>independent</i> population medians – the Mann-Whitney rank sums method	219
Comparing two <i>matched</i> population medians – the Wilcoxon signed-ranks method	220
16 Confidence intervals for the <i>ratio</i> of two population parameters	224
Getting a confidence interval for the <i>ratio</i> of two independent population means	225
Confidence interval for a population risk ratio	226
Confidence intervals for a population odds ratio	229
Confidence intervals for hazard ratios	232

VII	Putting it to the Test	235
17	Testing hypotheses about the <i>difference</i> between two population parameters	237
	Answering the question	238
	The hypothesis	238
	The null hypothesis	239
	The hypothesis testing process	240
	The <i>p</i> -value and the decision rule	241
	A brief summary of a few of the commonest tests	242
	Using the <i>p</i> -value to compare the means of two independent populations	244
	Interpreting computer hypothesis test results for the difference in two independent population means – the two-sample <i>t</i> test	245
	Output from Minitab – two-sample <i>t</i> test of difference in mean birthweights of babies born to white mothers and to non-white mothers	245
	Output from SPSS: two-sample <i>t</i> test of difference in mean birthweights of babies born to white mothers and to non-white mothers	246
	Comparing the means of two paired populations – the matched-pairs <i>t</i> test	248
	Using <i>p</i> -values to compare the medians of two independent populations: the Mann–Whitney rank-sums test	248
	How the Mann–Whitney test works	249
	Correction for multiple comparisons	250
	The Bonferroni correction for multiple testing	250
	Interpreting computer output for the Mann–Whitney test	252
	With Minitab	252
	With SPSS	252
	Two matched medians – the Wilcoxon signed-ranks test	254
	Confidence intervals versus hypothesis testing	254
	What could possibly go wrong?	255
	Types of error	256
	The power of a test	257
	Maximising power – calculating sample size	258
	Rule of thumb 1. Comparing the means of two independent populations (metric data)	258
	Rule of thumb 2. Comparing the proportions of two independent populations (binary data)	259
18	The Chi-squared (χ^2) test – what, why, and how?	261
	Of all the tests in all the world – you had to walk into my hypothesis testing procedure	262
	Using chi-squared to test for related-ness or for the equality of proportions	262
	Calculating the chi-squared statistic	265
	Using the chi-squared statistic	267
	Yate’s correction (continuity correction)	268
	Fisher’s exact test	268
	The chi-squared test with Minitab	269
	The chi-squared test with SPSS	270
	The chi-squared test for trend	272
	SPSS output for chi-squared trend test	274
19	Testing hypotheses about the <i>ratio</i> of two population parameters	276
	Preamble	276
	The chi-squared test with the risk ratio	277
	The chi-squared test with odds ratios	279
	The chi-squared test with hazard ratios	281

VIII	Becoming Acquainted	283
20	Measuring the association between two variables	285
	Preamble – plotting data	286
	Association	287
	The scatterplot	287
	The correlation coefficient	290
	Pearson’s correlation coefficient	290
	Is the correlation coefficient statistically significant in the population?	292
	Spearman’s rank correlation coefficient	294
21	Measuring agreement	298
	To agree or not agree: that is the question	298
	Cohen’s kappa (κ)	300
	Some shortcomings of kappa	303
	Weighted kappa	303
	Measuring the agreement between two metric continuous variables, the Bland–Altman plot	303
IX	Getting into a Relationship	307
22	Straight line models: linear regression	309
	Health warning!	310
	Relationship and association	310
	A causal relationship – explaining variation	312
	Refresher – finding the equation of a straight line from a graph	313
	The linear regression model	314
	First, is the relationship linear?	315
	Estimating the regression parameters – the method of ordinary least squares (OLS)	316
	Basic assumptions of the ordinary least squares procedure	317
	Back to the example – is the relationship statistically significant?	318
	Using SPSS to regress birthweight on mother’s weight	318
	Using Minitab	319
	Interpreting the regression coefficients	320
	Goodness-of-fit, R^2	320
	Multiple linear regression	322
	Adjusted goodness-of-fit: \bar{R}^2	324
	Including nominal covariates in the regression model: design variables and coding	326
	Building your model. Which variables to include?	327
	Automated variable selection methods	328
	Manual variable selection methods	329
	Adjustment and confounding	330
	Diagnostics – checking the basic assumptions of the multiple linear regression model	332
	Analysis of variance	333
23	Curvy models: logistic regression	334
	A second health warning!	335
	The binary outcome variable	335
	Finding an appropriate model when the outcome variable is binary	335
	The logistic regression model	337

Estimating the parameter values	338
Interpreting the regression coefficients	338
Have we got a significant result? statistical inference in the logistic regression model	340
The Odds Ratio	341
The multiple logistic regression model	343
Building the model	344
Goodness-of-fit	346
24 Counting models: Poisson regression	349
Preamble	350
Poisson regression	350
The Poisson regression equation	351
Estimating β_1 and β_2 with the estimators b_0 and b_1	352
Interpreting the estimated coefficients of a Poisson regression, b_0 and b_1	352
Model building – variable selection	355
Goodness-of-fit	357
Zero-inflated Poisson regression	358
Negative binomial regression	359
Zero-inflated negative binomial regression	361
X Four More Chapters	363
25 Measuring survival	365
Preamble	366
Censored data	366
A simple example of survival in a single group	366
Calculating survival probabilities and the proportion surviving: the Kaplan–Meier table	368
The Kaplan–Meier curve	369
Determining median survival time	369
Comparing survival with two groups	370
The log-rank test	371
An example of the log-rank test in practice	372
The hazard ratio	372
The proportional hazards (Cox’s) regression model – introduction	373
The proportional hazards (Cox’s) regression model – the detail	376
Checking the assumptions of the proportional hazards model	377
An example of proportional hazards regression	377
26 Systematic review and meta-analysis	380
Introduction	381
Systematic review	381
The forest plot	383
Publication and other biases	384
The funnel plot	386
Significance tests for bias – Begg’s and Egger’s tests	387
Combining the studies: meta-analysis	389
The problem of heterogeneity – the Q and I ² tests	389

27 Diagnostic testing	393
Preamble	393
The measures – sensitivity and specificity	394
The positive prediction and negative prediction values (PPV and NPV)	395
The sensitivity–specificity trade-off	396
Using the ROC curve to find the optimal sensitivity versus specificity trade-off	397
28 Missing data	400
The missing data problem	400
Types of missing data	403
Missing completely at random (MCAR)	403
Missing at Random (MAR)	403
Missing not at random (MNAR)	404
Consequences of missing data	405
Dealing with missing data	405
Do nothing – the wing and prayer approach	406
List-wise deletion	406
Pair-wise deletion	407
Imputation methods – simple imputation	408
Replacement by the Mean	408
Last observation carried forward	409
Regression-based imputation	410
Multiple imputation	411
Full Information Maximum Likelihood (FIML) and other methods	412
Appendix: Table of random numbers	414
References	415
Solutions to Exercises	424
Index	457

Preface to the 4th Edition

I noticed when looking at the prefaces to each of the first three editions, that there seems to be a gap of five years between each one. And this fourth edition is no exception – it's five years since I wrote the preface to the third edition, in 2013. Spooky or what! But anyway, this new edition has given me an opportunity to completely refresh the book. I have added three shiny new chapters covering important areas in medical statistics:

- On incidence, prevalence and standardisation.
- On Poisson regression.
- And on missing data.

I have also made numerous additions in most of the other chapters. These include, for example:

- In the chapter on measures of spread, I have added material on testing for Normality using the Shapiro–Wilk and Kolmogorov–Smirnov tests, as well as the graphical QQ plot.
- In the chapter on describing data with charts, I have added more material on the scatterplot, and introduced the bubble plot.
- In the second of the chapters on research design I have added brief notes on *per protocol* analysis.
- In the confounding chapter I have introduced the idea of confounding by indication, residual confounding, and propensity score matching.
- In the chapter on comparing population parameters using confidence intervals, I have revisited the interpretation of confidence intervals and the size of corresponding samples to try and make it easier to understand. I wanted to emphasise that we should take a flexible approach to the interpretation of confidence intervals and p-values.
- In the hypothesis chapter I have introduced the notion of correcting for multiple comparisons.

- In the logistic regression chapter I have added new material on goodness-of-fit to include the Cox and Snell, and Nagelkerke's *pseudo* goodness-of-fit measures.
- In the chapter on systematic review I have added a brief discussion of two statistical tests of funnel plot asymmetry: Begg's test and Egger's test. I have also added a little something on the trim and fill method for dealing with funnel plot asymmetry.

Apart from the above additions, I have added many new examples and exercises, and have rewritten much of the existing text to (hopefully) make it easier and more accessible.

Finally, I want to thank all of those who helped me in the writing of this book, and Barbara Noble in particular for bravely (and very competently) copy-editing and proofreading the material. Any remaining errors are of course mine.

As always I am happy to receive comments and criticism of my book, and I welcome any suggestions for improvement.

“Reader, I married him.” Whoops! Wrong book.

David Bowers, 2019

Preface to the 3rd Edition

The 1st edition of this book was published in 2002 and the 2nd edition in 2008. I was surprised when I discovered it was quite such a long time ago. Where did the time go! Anyway, over the course of the last five years, I have received many favourable comments from readers of my book, which of course is immensely gratifying. I must be doing something right then.

This edition contains a completely new chapter (on diagnostic tests), there is a quite a lot of new material and most of the chapters have received an extensive re-write. I have also updated virtually all of the examples drawn from the journals and added many new exercises. I hope that this gives the book a fresh feel – as well as a new lease of life.

The book should appeal, as before, to everybody in health care (students and professionals alike) including nurses, doctors, health visitors, physiotherapists, midwives, radiographers, dieticians, speech therapists, health educators and promoters, chiropractors and all those other allied and auxiliary professionals. It might possibly also be of interest to veterinary surgeons, one of whom reviewed my proposal fairly enthusiastically.

My thanks to Jon Peacock and all the others at Wiley who have shepherded me along in the past and no doubt will do so in the future. I must also thank Barbara Noble, who patiently acted as my first-line copyeditor. She read through my manuscript, discovered quite a few errors of various sorts and made many valuable suggestions to improve readability. Any remaining mistakes are of course mine.

I also want to acknowledge my great debt to Susanne, who always encourages me, enthusiastically, in everything I attempt.

Finally, I would like to mention another book which might be of interest to any readers who are thinking of embarking on research for the first time – *Getting Started in Health Research*, Bowers et al., Wiley, 2012. This book covers both quantitative and qualitative research. It will guide you through the research process, from the very first idea to the interpretation of your results and your conclusions.

David Bowers, 2013

Preface to the 2nd Edition

This book is a “not-too-mathematical” introduction to medical statistics. It should appeal to anyone training or working in the health care arena – whatever his or her particular discipline is – who wants either a simple introduction to the subject or a gentle reminder of stuff that they might have forgotten. I have aimed the book at:

- students doing either a first degree or a diploma in clinical and health care courses
- students doing post-graduate clinical and health care studies
- health care professionals doing professional and membership examinations
- health care professionals who want to brush up on some medical statistics generally or who want a simple reminder of a particular topic
- anybody else who wants to know a bit of what medical statistics is about.

The most significant change in this edition is the addition of two new chapters, one on measuring survival and the other on systematic review and meta-analysis. The ability to understand the principles of survival analysis is important, not least because of its popularity in clinical research and consequently in the clinical literature. Similarly, the increasing importance of evidence-based clinical practice means that systematic review and meta-analysis also demand a place. In addition, I have taken the opportunity to correct and freshen the text in a few places, as well as adding a small number of new examples. My thanks to Lucy Sayer, my editor at John Wiley & Sons, for her enthusiastic support, to Liz Renwick and Robert Hambrook and all the other people in Wiley for their invaluable help and my special thanks to my copyeditor Barbara Noble for her truly excellent work and enthusiasm (of course, any remaining errors are mine).

I am happy to get any comments from you. You can e-mail me at: d.bowers@leeds.ac.uk.

Preface to the 1st Edition

This book is intended to be an introduction to medical statistics but one which is not too mathematical – in fact, it has the absolute minimum of maths. The exceptions however are Chapters 17 and 18, which have maths on linear and logistic regressions. It is really impossible to provide material on these procedures without some maths, and I hesitated about including them at all. However, they are such useful and widely used techniques, particularly logistic regression and its production of odds ratios, which I felt they must go in. Of course, you do not *have* to read them. It should appeal to anyone training or working in the health care arena – whatever his or her particular discipline is – who wants a simple, not-too-technical introduction to the subject. I have aimed the book at:

- students doing either a first degree or a diploma in health care-related courses
- students doing post-graduate health care studies
- health care professionals doing professional and membership examinations
- health care professionals who want to brush up on some medical statistics generally or who want a simple reminder of a particular topic
- anybody else who wants to know a bit of what medical statistics is about.

I intended originally to make this book as an amalgam of two previous books of mine, *Statistics from Scratch for Health Care Professionals* and *Statistics Further from Scratch*. However, although it covers a lot of the same material as in those two books, this is in reality a completely new book, with a lot of extra stuff, particularly on linear and logistic regressions. I am happy to get any comments and criticisms from you. You can e-mail me at: slothist@hotmail.com.

Introduction

My purpose in writing this book was to provide a guide to all those health care students and professionals out there who either want to get started in medical statistics or who would like (or need) to refresh their understanding of one or more of the topics which are common in medical statistics. I have tried to keep the mathematics to a minimum, although this is a bit more difficult with the somewhat challenging material on modelling in later chapters.

I have used lots of appropriate examples drawn from clinical journals to illustrate the ideas, as well as many relevant exercises which the readers may wish to work through to consolidate their understanding of the material covered (the solutions are at the end of this book).

I have included some outputs from SPSS and Minitab which I hope will help the readers interpret the results from these statistical programs.

Finally, for any tutors who are using this book to introduce their students to medical statistics, I am always very pleased to receive any comments or criticisms they may have which will help me improve the book for any possible future editions. My e-mail address is: d.bowers@leeds.ac.uk.

Enjoy!

David Bowers

