

Ton J. Cleophas
Aeilko H. Zwinderman
Toine F. Cleophas
Eugene P. Cleophas

Statistics Applied to Clinical Trials

Fourth Edition

Statistics Applied to Clinical Trials

Fourth edition

by

TON J. CLEOPHAS, MD, PhD, Professor
*Statistical Consultant, Circulation, Boston, USA,
Co-Chair Module Statistics Applied to Clinical Trials,
European Interuniversity College of Pharmaceutical Medicine, Lyon, France,
Internist-clinical pharmacologist,
Department Medicine, Albert Schweitzer Hospital, Dordrecht, The Netherlands*

AEILKO H. ZWINDERMAN, MathD, PhD, Professor
*Co-Chair Module Statistics Applied to Clinical Trials,
European Interuniversity College of Pharmaceutical Medicine, Lyon, France,
Professor of Statistics,
Department Biostatistics and Epidemiology, Academic Medical Center, Amsterdam,
The Netherlands*

TOINE F. CLEOPHAS, BSc
Department of Research, Damen Shipyards, Gorinchem, The Netherlands

and

EUGENE P. CLEOPHAS, BSc
Technical University, Delft, The Netherlands



Springer

Prof. Ton J. Cleophas
Albert Schweitzer Hospital
Dordrecht
The Netherlands

Prof. Aeilko H. Zwinderman
Academic Medical Center
Amsterdam
The Netherlands

Toine F. Cleophas
Damen Shipyards
Gorinchem
The Netherlands

Eugene P. Cleophas
Technical University
Delft
The Netherlands

ISBN 978-1-4020-9522-1

e-ISBN 978-1-4020-9523-8

Library of Congress Control Number: 2008939866

© Springer Science + Business Media B.V. 2009

No part of this work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording or otherwise, without written permission from the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work.

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

STATISTICS APPLIED TO CLINICAL TRIALS
FOURTH EDITION

PREFACE TO FIRST EDITION

The European Interuniversity Diploma of Pharmaceutical Medicine is a postacademic course of 2-3 years sponsored by the Socrates program of the European Community. The office of this interuniversity project is in Lyon and the lectures are given there. The European Community has provided a building and will remunerate lecturers. The institute which provides the teaching is called the European College of Pharmaceutical Medicine, and is affiliated with 15 universities throughout Europe, whose representatives constitute the academic committee. This committee supervises educational objectives. Start lectures February 2000.

There are about 20 modules for the first two years of training, most of which are concerned with typically pharmacological and clinical pharmacological matters including pharmacokinetics, pharmacodynamics, phase III clinical trials, reporting, communication, ethics and, any other aspects of drug development. Subsequent training consists of practice training within clinical research organisations, universities, regulatory bodies etc., and finally of a dissertation. The diploma, and degree are delivered by the Claude Bernard University in Lyon as well as the other participating universities.

The module "Statistics applied to clinical trials" will be taught in the form of a 3 to 6 day yearly course given in Lyon and starting February 2000. Lecturers have to submit a document of the course (this material will be made available to students). Three or 4 lecturers are requested to prepare detailed written material for students as well as to prepare examination of the students. The module is thus an important part of a postgraduate course for physicians and pharmacists for the purpose of obtaining the European diploma of pharmaceutical medicine. The diploma should make for leading positions in pharmaceutical industry, academic drug research, as well as regulatory bodies within the EC. This module is mainly involved in the statistics of randomized clinical trials.

The chapters 1-9, 11, 17, 18 of this book are based on the module "Medical statistics applied to clinical trials" and contain material that should be mastered by the students before their exams. The remaining chapters are capita selecta intended for excellent students and are not included in the exams.

The authors believe that this book is innovative in the statistical literature because, unlike most introductory books in medical statistics, it provides an explanatory rather than mathematical approach to statistics, and, in addition, emphasizes non-classical but increasingly frequently used methods for the statistical analyses of clinical trials, e.g., equivalence testing, sequential analyses, multiple linear regression analyses for confounding, interaction, and synergism. The authors are not aware of any other work published so far that is comparable with the current work, and, therefore, believe that it does fill a need.

August 1999

Dordrecht, Leiden, Delft

PREFACE TO SECOND EDITION

In this second edition the authors have removed textual errors from the first edition. Also seven new chapters (chapters 8, 10, 13, 15-18) have been added. The principles of regression analysis and its resemblance to analysis of variance was missing in the first edition, and have been described in chapter 8. Chapter 10 assesses curvilinear regression. Chapter 13 describes the statistical analyses of crossover data with binary response. The latest developments including statistical analyses of genetic data and quality-of-life data have been described in chapters 15 and 16. Emphasis is given in chapters 17 and 18 to the limitations of statistics to assess non-normal data, and to the similarities between commonly-used statistical tests. Finally, additional tables including the Mann-Whitney and Wilcoxon rank sum tables have been added in the Appendix.

December 2001, Dordrecht, Amsterdam, Delft

PREFACE TO THE THIRD EDITION

The previous two editions of this book, rather than having been comprehensive, concentrated on the most relevant aspects of statistical analysis. Although well-received by students, clinicians, and researchers, these editions did not answer all of their questions. This called for a third, more comprehensive, rewrite. In this third edition the 18 chapters from the previous edition have been revised, updated, and provided with a conclusions section summarizing the main points. The formulas have been re-edited using the Formula-Editor from Windows XP 2004 for enhanced clarity. Thirteen new chapters (chapters 8-10, 14, 15, 17, 21, 25-29, 31) have been added. The chapters 8-10 give methods to assess the problems of multiple testing and data testing closer to expectation than compatible with random. The chapters 14 and 15 review regression models using an exponential rather than linear relationship including logistic, Cox, and Markov models. Chapter 17 reviews important interaction effects in clinical trials and provides methods for their analysis. In chapter 21 study designs appropriate for medicines from one class are discussed. The chapters 25-29 review respectively (1) methods to evaluate the presence of randomness in the data, (2) methods to assess variabilities in the data, (3) methods to test reproducibility in the data, (4) methods to assess accuracy of diagnostic tests, and (5) methods to assess random rather than fixed treatment effects. Finally, chapter 31 reviews methods to minimize the dilemma between sponsored research and scientific independence. This updated and extended edition has been written to serve as a more complete guide and reference-text to students, physicians, and investigators, and, at the same time, preserves the common sense approach to statistical problem-solving of the previous editions.

August 2005, Dordrecht, Amsterdam, Delft

PREFACE TO FOURTH EDITION

In the past few years many important novel methods have been applied in published clinical research. This has made the book again rather incomplete after its previous edition. The current edition consists of 16 new chapters, and updates of the 31 chapters from the previous edition. Important methods like Laplace transformations, log likelihood ratio statistics, Monte Carlo methods, and trend testing have been included. Also novel methods like superiority testing, pseudo-R² statistics, optimism corrected c-statistic, I-statistics, and diagnostic meta-analyses have been addressed.

The authors have given special efforts for all chapters to have their own introduction, discussion, and references section. They can, therefore, be studied separately and without need to read the previous chapters first.

September 2008, Dordrecht, Amsterdam, Gorinchem, and Delft

FOREWORD

In clinical medicine appropriate statistics has become indispensable to evaluate treatment effects. Randomized controlled trials are currently the only trials that truly provide evidence-based medicine. Evidence based medicine has become crucial to optimal treatment of patients. We can define randomized controlled trials by using Christopher J. Bulpitt's definition "a carefully and ethically designed experiment which includes the provision of adequate and appropriate controls by a process of randomization, so that precisely framed questions can be answered". The answers given by randomized controlled trials constitute at present the way how patients should be clinically managed. In the setup of such randomized trial one of the most important issues is the statistical basis. The randomized trial will never work when the statistical grounds and analyses have not been clearly defined beforehand. All endpoints should be clearly defined in order to perform appropriate power calculations. Based on these power calculations the exact number of available patients can be calculated in order to have a sufficient quantity of individuals to have the predefined questions answered. Therefore, every clinical physician should be capable to understand the statistical basis of well performed clinical trials. It is therefore a great pleasure that Drs. T.J. Cleophas, A.H. Zwinderman, and T.F. Cleophas have published a book on statistical analysis of clinical trials. The book entitled "Statistics Applied to Clinical Trials" is clearly written and makes complex issues in statistical analysis transparent. Apart from providing the classical issues in statistical analysis, the authors also address novel issues such as interim analyses, sequential analyses, and meta-analyses. The book is composed of 18 chapters, which are nicely structured. The authors have deepened our insight in the applications of statistical analysis of clinical trials. We would like to congratulate the editors on this achievement and hope that many readers will enjoy reading this intriguing book.

E.E. van der Wall, MD, PhD, Professor of Cardiology, President Netherlands Association of Cardiology, Leiden, The Netherlands

TABLE OF CONTENTS

PREFACES	xvii
FOREWORD	xxi
CHAPTER 1 / HYPOTHESES, DATA, STRATIFICATION	
1. General considerations	1
2. Two main hypotheses in drug trials: efficacy and safety	2
3. Different types of data: continuous data	3
4. Different types of data: proportions, percentages and contingency tables	8
5. Different types of data: correlation coefficient	11
6. Stratification issues	13
7. Randomized versus historical controls	14
8. Factorial designs	15
9. Conclusions	15
10. References	16
CHAPTER 2 / THE ANALYSIS OF EFFICACY DATA	
1. Overview	17
2. The principle of testing statistical significance	18
3. The t-value = standardized mean result of study	21
4. Unpaired t-test	22
5. Null-hypothesis testing of 3 or more unpaired samples	24
6. Three methods to test statistically a paired sample	25
7. Null-hypothesis testing of 3 or more paired samples	29
8. Null-hypothesis testing with complex data	30
9. Paired data with a negative correlation	31
10. Rank testing	37
11. Rank testing for 3 or more samples	40
12. Conclusions	42
13. References	42
CHAPTER 3 / THE ANALYSIS OF SAFETY DATA	
1. Introduction, summary display	45
2. Four methods to analyze two unpaired proportions	46
3. Chi-square to analyze more than two unpaired proportions	52
4. McNemar's test for paired proportions	55
5. Survival analysis	56
6. Odds ratio method for analyzing two unpaired proportions	58
7. Odds ratios for 1 group, two treatments	61
8. Conclusions	61

CHAPTER 4 / LOG LIKELIHOOD RATIO TESTS FOR SAFETY DATA ANALYSIS

1. Introduction	63
2. Numerical problems with calculating exact likelihoods	63
3. The normal approximation and the analysis of clinical events	64
4. Log likelihood ratio tests and the quadratic approximation	66
5. More examples	68
6. Discussion	69
7. Conclusions	70
8. References	70

CHAPTER 5 / EQUIVALENCE TESTING

1. Introduction	73
2. Overview of possibilities with equivalence testing	75
3. Calculations	76
4. Equivalence testing, a new gold standard?	77
5. Validity of equivalence trials	77
6. Special point: level of correlation in paired equivalence studies	78
7. Conclusions	79

CHAPTER 6 / STATISTICAL POWER AND SAMPLE SIZE

1. What is statistical power	81
2. Emphasis on statistical power rather than null-hypothesis testing	82
3. Power computations	84
4. Examples of power computation using the t-table	85
5. Calculation of required sample size, rationale	91
6. Calculations of required sample size, methods	91
7. Testing inferiority of a new treatment (the type III error)	95
8. Conclusions	97
9. References	97

CHAPTER 7 / INTERIM ANALYSES

1. Introduction	99
2. Monitoring	99
3. Interim analysis	100
4. Group-sequential design of interim analysis	103
5. Continuous sequential statistical techniques	103
6. Conclusions	105
7. References	105

CHAPTER 8 / CONTROLLING THE RISK OF FALSE POSITIVE CLINICAL TRIALS

1. Introduction	107
2. Bonferroni test	108
3. Least significant difference test (LSD) test	109
4. Other tests for adjusting the p-values	109

5. Composite endpoint procedures	110
6. No adjustments at all, and pragmatic solutions	110
7. Conclusions	111
8. References	111

CHAPTER 9 / MULTIPLE STATISTICAL INFERENCES

1. Introduction	113
2. Multiple comparisons	113
3. Multiple variables	118
4. Conclusions	121
5. References	121

CHAPTER 10 / THE INTERPRETATION OF THE P-VALUES

1. Introduction	123
2. Renewed attention to the interpretation of the probability levels, otherwise called the p-values	123
3. Standard interpretation of p-values	124
4. Common misunderstandings of the p-values	126
5. Renewed interpretations of p-values, little difference between $p = 0.06$ and $p = 0.04$	126
6. The real meaning of very large p-values like $p > 0.95$	127
7. P-values larger than 0.95, examples (Table 2)	128
8. The real meaning of very small p-values like $p < 0.0001$	129
9. P-values smaller than 0.0001, examples (Table 3)	130
10. Discussion	131
11. Recommendations	131
12. Conclusions	133
13. References	133

CHAPTER 11 / RESEARCH DATA CLOSER TO EXPECTATION THAN COMPATIBLE WITH RANDOM SAMPLING

1. Introduction	137
2. Methods and results	138
3. Discussion	139
4. Conclusions	142
5. References	142

CHAPTER 12 / STATISTICAL TABLES FOR TESTING DATA CLOSER TO EXPECTATION THAN COMPATIBLE WITH RANDOM SAMPLING

1. Introduction	145
2. Statistical tables of unusually high p-values	147
3. How to calculate the p-values yourself	147
4. Additional examples simulating real practice, multiple comparisons	150
5. Discussion	152
6. Conclusions	153
7. References	153

CHAPTER 13 / PRINCIPLES OF LINEAR REGRESSION

1. Introduction	155
2. More on paired observations	156
3. Using statistical software for simple linear regression	159
4. Multiple linear regression	162
5. Multiple linear regression, example	164
6. Purposes of linear regression analysis	168
7. Another real data example of multiple linear regression (exploratory purpose)	169
8. It may be hard to define what is determined by what, multiple and multivariate regression	171
9. Limitations of linear regression	172
10. Conclusions	173

CHAPTER 14 / SUBGROUP ANALYSIS USING MULTIPLE LINEAR
REGRESSION: CONFOUNDING, INTERACTION, SYNERGISM

1. Introduction	175
2. Example	175
3. Model (figure 1)	176
4. (I.) Increased precision of efficacy (figure 2)	178
5. (II.) Confounding	179
6. (III.) Interaction and synergism	180
7. Estimation, and hypothesis testing	181
8. Goodness-of-fit	182
9. Selection procedures	183
10. Main conclusion	183
11. References	184

CHAPTER 15 / CURVILINEAR REGRESSION

1. Introduction	185
2. Methods, statistical model	186
3. Results	188
4. Discussion	194
5. Conclusions	196
6. References	196

CHAPTER 16 / LOGISTIC AND COX REGRESSION, MARKOW MODELS,
LAPLACE TRANSFORMATIONS

1. Introduction	199
2. Linear regression	199
3. Logistic regression	203
4. Cox regression	209
5. Markow models	212
6. Regression-analysis with Laplace transformations	213

7. Discussion	217
8. Conclusions	218
9. References	219

CHAPTER 17 / REGRESSION MODELING FOR IMPROVED PRECISION

1. Introduction	221
2. Regression modeling for improved precision of clinical trials, the underlying mechanism	221
3. Regression model for parallel-group trials with continuous efficacy data	223
4. Regression model for parallel-group trials with proportions or odds as efficacy data	224
5. Discussion	225
6. Conclusions	227
7. References	227

CHAPTER 18 / POST-HOC ANALYSES IN CLINICAL TRIALS, A CASE FOR LOGISTIC REGRESSION ANALYSIS

1. Multiple variables methods	229
2. Examples	229
3. Logistic regression equation	232
4. Conclusions	233
5. References	234

CHAPTER 19 / CONFOUNDING

1. Introduction	235
2. First method for adjustment of confounders: subclassification on one confounder	236
3. Second method for adjustment of confounders: regression modeling	237
4. Third method for adjustment of confounders: propensity scores	238
5. Discussion	241
6. Conclusions	242
7. References	243

CHAPTER 20 / INTERACTION

1. Introduction	245
2. What exactly is interaction, a hypothesized example	245
3. How to test interaction statistically, a real data example with a concomitant medication as interacting factor, incorrect method	248
4. Three analysis methods	248
5. Using a regression model for testing interaction, another real data example	252
6. Analysis of variance for testing interaction, other real data examples	254
7. Discussion	259
8. Conclusions	260
9. References	261

CHAPTER 21 / META-ANALYSIS, BASIC APPROACH

1. Introduction	263
2. Examples	264
3. Clearly defined hypotheses	266
4. Thorough search of trials	266
5. Strict inclusion criteria	266
6. Uniform data analysis	267
7. Discussion, where are we now?	275
8. Conclusions	276
9. References	276

CHAPTER 22 / META-ANALYSIS, REVIEW AND UPDATE OF METHODOLOGIES

1. Introduction	277
2. Four scientific rules	277
3. General framework of meta-analysis	278
4. Pitfalls of meta-analysis	281
5. New developments	284
6. Conclusions	285
7. References	285

CHAPTER 23 / CROSSOVER STUDIES WITH CONTINUOUS VARIABLES

1. Introduction	289
2. Mathematical model	290
3. Hypothesis testing	291
4. Statistical power of testing	293
5. Discussion	296
6. Conclusion	297
7. References	298

CHAPTER 24 / CROSSOVER STUDIES WITH BINARY RESPONSES

1. Introduction	299
2. Assessment of carryover and treatment effect	300
3. Statistical model for testing treatment and carryover effects	301
4. Results	302
5. Examples	304
6. Discussion	305
7. Conclusions	306
8. References	306

CHAPTER 25 / CROSS-OVER TRIALS SHOULD NOT BE USED TO TEST TREATMENTS WITH DIFFERENT CHEMICAL CLASS

1. Introduction	309
2. Examples from the literature in which cross-over trials are correctly used	311

3. Examples from the literature in which cross-over trials should not have been used	313
4. Estimate of the size of the problem by review of hypertension trials published	315
5. Discussion	316
6. Conclusions	317
7. References	318

CHAPTER 26 / QUALITY-OF-LIFE ASSESSMENTS IN CLINICAL TRIALS

1. Introduction	319
2. Some terminology	319
3. Defining QOL in a subjective or objective way?	321
4. The patients' opinion is an important independent-contributor to QOL	322
5. Lack of sensitivity of QOL-assessments	323
6. Odds ratio analysis of effects of patient characteristics on QOL data provides increased precision	324
7. Discussion	327
8. Conclusions	328
9. References	328

CHAPTER 27 / STATISTICAL ANALYSIS OF GENETIC DATA

1. Introduction	331
2. Some terminology	332
3. Genetics, genomics, proteonomics, data mining	334
4. Genomics	335
5. Conclusions	339
6. References	339

CHAPTER 28 / RELATIONSHIP AMONG STATISTICAL DISTRIBUTIONS

1. Introduction	341
2. Variances	341
3. The normal distribution	342
4. Null-hypothesis testing with the normal or t-distribution	344
5. Relationship between the normal-distribution and chi-square-distribution, null-hypothesis testing with chi-square distribution	346
6. Examples of data where variance is more important than mean	348
7. Chi-square can be used for multiple samples of data	349
8. Discussion	352
9. Conclusions	353
10. References	354

CHAPTER 29 / TESTING CLINICAL TRIALS FOR RANDOMNESS

1. Introduction	355
2. Individual data available	355

3. Individual data not available	362
4. Discussion	364
5. Conclusions	365
6. References	366

CHAPTER 30 / CLINICAL TRIALS DO NOT USE RANDOM SAMPLES ANYMORE

1. Introduction	367
2. Non-normal sampling distributions, giving rise to non-normal data	368
3. Testing the assumption of normality	369
4. What to do in case of non-normality	370
5. Discussion	371
6. Conclusions	373
7. References	373

CHAPTER 31 / CLINICAL DATA WHERE VARIABILITY IS MORE IMPORTANT THAN AVERAGES

1. Introduction	375
2. Examples	375
3. An index for variability in the data	376
4. How to analyze variability, one sample	377
5. How to analyze variability, two samples	379
6. How to analyze variability, three or more samples	380
7. Discussion	382
8. Conclusions	383
9. References	383

CHAPTER 32 / TESTING REPRODUCIBILITY

1. Introduction	385
2. Testing reproducibility of quantitative data (continuous data)	385
3. Testing reproducibility of qualitative data (proportions and scores)	388
4. Incorrect methods to assess reproducibility	390
5. Additional real data examples	390
6. Discussion	394
7. Conclusions	394
8. References	395

CHAPTER 33 / VALIDATING QUALITATIVE DIAGNOSTIC TESTS

1. Introduction	397
2. Overall accuracy of a qualitative diagnostic test	397
3. Perfect and imperfect qualitative diagnostic tests	399
4. Determining the most accurate threshold for positive qualitative tests	401
5. Discussion	404
6. Conclusions	404
7. References	406

CHAPTER 34 / UNCERTAINTY OF QUALITATIVE DIAGNOSTIC TESTS

1. Introduction	407
2. Example 1	407
3. Example 2	408
4. Example 3	409
5. Example 4	409
6. Discussion	410
7. Conclusion	411
8. References	411
9. Appendix 1	411
10. Appendix 2	412

CHAPTER 35 / META-ANALYSIS OF DIAGNOSTIC ACCURACY STUDIES

1. Introduction	415
2. Diagnostic odds ratios (DORs)	416
3. Bivariate model	419
4. Conclusions	420
5. References	420

CHAPTER 36 / VALIDATING QUANTITATIVE DIAGNOSTIC TESTS

1. Introduction	423
2. Linear regression testing a significant correlation between the new test and the control test	423
3. Linear regression testing the hypotheses that the α -value = 0.000 and the β -Value = 1.000	425
4. Linear regression using a squared correlation coefficient (r^2 - value) of > 0.95	426
5. Alternative methods	428
6. Discussion	429
7. Conclusions	430
8. References	430

CHAPTER 37 / SUMMARY OF VALIDATION PROCEDURES FOR DIAGNOSTIC TESTS

1. Introduction	433
2. Qualitative diagnostic tests	433
3. Quantitative diagnostic tests	437
4. Additional methods	443
5. Discussion	445
6. Conclusions	446
7. References	447