Statistics for Biology and Health

*Series Editors:*
M. Gail, K. Krickeberg, J. Samet, A. Tsiatis, W. Wong
Statistics for Biology and Health

Bacchieri/Della Cioppa: Fundamentals of Clinical Research
Buzkowsky/Molenberghs/Buyse: The Evaluation of Surrogate Endpoints.
Evetitt/Rabe-Hesketh: Analyzing Medical Data Using S-PLUS.
Hougaard: Analysis of Multivariate Survival Data.
Klein/Moeschberger: Survival Analysis: Techniques for Censored and Truncated Data, 2nd ed.
Lange: Mathematical and Statistical Methods for Genetic Analysis, 2nd ed.
Manton/Singer/Suzman: Forecasting the Health of Elderly Populations.
Moyé: Multiple Analyses in Clinical Trials: Fundamentals for Investigators.
Nielsen: Statistical Methods in Molecular Evolution.
Parmigiani/Garrett/Irizarry/Zeger: The Analysis of Gene Expression Data: Methods and Software.
Simon/Korn/McShane/Rademacher/Wright/Zhao: Design and Analysis of DNA Microarray Investigations.
Sorensen/Gianola: Likelihood, Bayesian, and MCMC Methods in Quantitative Genetics.
Sun: The Statistical Analysis of Interval-censored Failure Time Data.
Therneau/Grambsch: Modeling Survival Data: Extending the Cox Model.
Ting: Dose Finding in Drug Development.
Wu/Ma/Casella: Statistical Genetics of Quantitative Traits: Linkage, Map and QTL.
Zhang/Singer: Recursive Partitioning in the Health Sciences.
Zuur/Ieno/Smith: Analyzing Ecological Data
Fundamentals of Clinical Research
Bridging Medicine, Statistics and Operations
To our parents, in particular in memory of Antonio Bacchieri
In memory of Tiziana Mistrali
Statistical methodology is an essential component of clinical (and biological) research. Therefore it is not surprising that many textbooks aiming at “explaining” statistical methods to researchers have been published and continue to appear in print. The complexity of the basic issue, that of communication between statisticians and researchers, is illustrated and discussed very well by the authors of this book in their Introduction. Thus it would be of no benefit to dwell further on this theme and on the difficulty of producing truly effective material.

I prefer to start from a personal episode. Way back in 1959, a freshman of the faculty of Statistical Sciences, I found by chance on a stand, the book “Metodi statistici ad uso dei ricercatori”, which was the Italian translation of the famous book by RA Fisher. The textbook for the first course of statistics used that year at of our University had left me perplexed. A student like myself, with a good high school scientific education, was fascinated by the wealth of real life examples, but was unable to reconstruct the thread of logical-mathematical reasoning, especially the inductive one. The unexpected access to Fisher’s legendary book raised my hopes that I could finally get to the heart of the matter. However, after several attempts at reading it, I reached the conclusion that experimental statistics required first and foremost a spirit of obedience: one was to use formulae which were incomprehensible, but justified by some higher authority. Especially mysterious was the frequently appearing concept of “degrees of freedom”, a number which seemed to be reachable only through vague analogical reasoning. Fortunately, my conclusion was a temporary one, because, while still a student, I had the fortune of being exposed to much more constructive critical approaches. But what could be the reaction of the researcher, in principle the target of the book? Most likely that of accepting statistics as a
price to pay to make his/her research publishable, definitely not as an important conceptual tool to understand problems and define useful strategies for the gaining of knowledge. RA Fisher’s forward reads: “The author was impressed with the practical importance of many recent mathematical advances, which to others seemed to be merely academic refinements. He felt sure, too, that workers with research experience would appreciate a book which, without entering into the mathematical theory of statistical methods, should embody the latest results of that theory, presenting them in the form of practical procedures appropriate to those types of data with which research workers are actually concerned. The practical application of general theorems is a different art from their establishment by mathematical proof. It requires fully as deep an understanding of their meaning, and is, moreover, useful to many to whom the other is unnecessary”. In other words: the recent developments in mathematics, which are truly revolutionary as they allow experimental results to be looked at in a totally new way, are the theoretical results obtained by Fisher himself. The methods derived from them should be used by researchers without worrying too much as to their meaning. I do want to emphasize that I am fully convinced of RA Fisher’s scientific greatness, but I do find it essential that the above mentioned “results” (which basically are the use of the sample distribution and the theory of pure significance) be accompanied by a crisp explanation of the general context, as the one presented in the following book “Statistical Methods and Scientific Inference” (1956) in which the meaning and the use of the likelihood function are masterfully explained and the limits of the theory of significance are at least partially described. A curious detail is that the second book is reasonably clear and does not hide anything relevant to the reader, whilst leaving several problems open, including some interpretative ones. However it is not meant for researchers…

The pedagogic idea underlying the book “Statistical Methods for Research Workers” is certainly extreme, whereas recent didactic literature usually seeks a compromise. It cannot be denied, however, that sometimes the suggestion to operate without wasting time to give thought to the useless theory does reappear, in the most diverse contexts. This theme could also be discussed with reference to the possible interpretations of recent reforms of the Italian university system, but then we would definitely wander from the subject.

I can now get to the point. The basic idea behind this book is happily at the opposite pole. The Authors, deliberately, intend to explain everything. In particular, they try to prevent bio-medical researchers from accepting and applying statistical procedures without understanding their meaning and therefore without applying a critical control, made possible by the subject matter knowledge which typically the statistician does not have. Obviously, the theoretical discussions have to be limited, but everything essential from a logical point of view can indeed be explained and, with some effort, understood, even using only high school level mathematics. In this respect, and also in that of trying to avoid any jargon of obscure logical roots, the book that I have the pleasure of presenting is a proposal of great interest and usefulness. Also, it must be stressed that this
book goes well beyond the typical presentations of statistical methods (although this is the aspect on which I am more inclined to comment). The practice of applied research is analyzed and discussed in its complexity, including of course ethical aspects, as well as financial, organizational ones, and so on.

My involvement in clinical research is related to my interest in applications of statistical methods of so called “Bayesian” approach, which are still not fully accepted as standards. So, such methods pose a double challenge: for a comparison between methodologies to have a true meaning, the logic behind both must be clear to the reader, otherwise all one achieves is to propose to replace one blind obedience with another blind obedience. The careful reader of this book will not feel pushed toward the Bayesian approach. Instead, the reader is given the tools to reason and, after an appropriate comparative discussion, is in a position to give his/her “informed consent” to one or the other “school” (or to recognize the merits of both). In fact, the important thing is to understand the general sense of the issue, without giving the illusion (never favoured by the Authors) that statistical methodology is established once and for all. Only with such an open attitude, in my view, can statistical education play a constructive role and not become a boring ritual.

Therefore, I do hope that this book will have the fortune it deserves and that it will stimulate many bright minds to reflect further on important aspects of the methodology of research.

Ludovico Piccinato
La Sapienza University, Rome
Acknowledgements

Writing this book has taken us several years of early mornings (GDC), late nights (AB) and weekends (GDC and AB). We are indebted to all those to whom we denied our time.

A special thank you from both of us goes to:

• Alfredo Ardia for guiding us through the first steps in this profession, teaching us creativity in reasoning and staying our friend for all these years.

• Adelchi Azzalini, Paolo Bruzzi and Ludovico Piccinato for pushing us in the adventure of writing this book, supporting us, and finding the time and energy to comment our draft. Many of their suggestions set us on the right path. Some we ignored. The heavy responsibility of such choices we accept.

• Lidia Merli and Silva Tommasini, for their help with the graphs.

• Our friends, especially those of the “Associazione Laboratorio Nuovo Alpha-beto” for supporting us in our effort at intellectual honesty, reminding us how important this is in every “public act”, even the very small one of writing a book on the methodology of clinical research.

A special thank you from AB goes to:

• Marco Corsi, who has always supported my academic activities and has taught me a great deal on how to build constructive peer-relationships between statisticians and physicians

Finally, a special thank you from GDC goes to:

• Stephen Senn, who has shown me that rigorous thinking and discipline can coexist with fun and a sense of humor.

• James Shannon, who, over the years, with little talk and much example, has taught me many things essential to the practice of clinical research: focus on the essential, calm under pressure, logical thinking and much more.
XII Acknowledgements

English is a language of multiple layers of complexity. The task of translating a book like this was daunting. We are deeply grateful to Elaine Hopper for giving us many hours of her time before and after work to review the translation. Any remaining errors are, of course, ours.

September 2006
The authors
Contents

Foreword ................................................................. VII
Acknowledgements ....................................................... XI
Contents ................................................................. XIII
Introduction ............................................................... XIX

1. Variability of Biological Phenomena and Measurement Errors 1
   1.1. Phenotypic Variability ........................................ 1
   1.2. Temporal Variability ......................................... 2
   1.3. Measurement-Related Variability ............................ 3
       1.3.1. The Measurement ....................................... 3
       1.3.2. Measurement Errors ..................................... 7
   1.4. Variability of Diagnostic Tests ............................. 10
   Summary ............................................................ 13

2. Distinctive Aspects of a Biomedical Study. Observational and Experimental Studies ................................. 14
   2.1. Distinctive Features of Biomedical Studies ............... 14
   2.2. The Study Protocol .......................................... 18
   2.3. Observational Studies ....................................... 21
   2.4. Experimental Studies ....................................... 23
   Summary ............................................................ 27
3. Observational Studies .............................. 28
   3.1. Basic Designs of Observational Studies .......... 29
       3.1.1. Prospective or Cohort Studies .......... 30
       3.1.2. Retrospective Studies or Control Cases .... 36
       3.1.3. Sample Size ............................ 40
   3.2. Bias and Confounding .......................... 40
       3.2.1. Control of Bias in Epidemiology .......... 44
       3.2.2. Control of the Phenomenon of Confounding . 49
   3.3. Advantages and Disadvantages of the Different Types
       of Observational Studies ....................... 53
   Summary ........................................... 56

4. Defining the Treatment Effect ....................... 58
   4.1. From the Single Measurement to the Signal .......... 58
   4.2. Identification and Quantification of the End-Points
       (Individual Subject Level) ...................... 64
       4.2.1. Methodological Characteristics of the End-Point . 64
       4.2.2 Discriminating between Primary and Secondary
            End-Points and between Efficacy and Safety/Tolerability
            End-Points .................................. 66
   4.3. Identification and Quantification of the Signals (Group Level) . 69
   4.4. Statistical Considerations ........................ 70
   4.5. Practical, Regulatory, Marketing and Pharmaco-Economic
        Considerations ................................ 73
   4.6. Selection and Characterization of the Primary End-Point
       and Signal: an Example .......................... 75
       4.6.1. Stage One: Define the Main Therapeutic Level .... 75
       4.6.2. Stage Two: Define the Primary End-Point
            (Individual Patient Level) .................... 77
       4.6.3. Stages Three and Four: Define the Group Indicator,
            the Signal, and the Threshold of Clinical Relevance
            (Treatment Group and Study Levels) .......... 79
   4.7. More Than One Question in the Same Study: the Problem
       of Multiple Statistical Tests ..................... 80
   4.8. Validation of Measurement Scales ................... 84
   4.9. Special Types of End-Points ........................ 85
   Summary ........................................... 88

5. Probability, Inference and Decision Making ............. 90
   5.1. Probability .................................... 91
       5.1.1 Definitions .................................. 91
       5.1.2. Probability Distribution and Probability Density Function 92
       5.1.3. Normal or Gaussian Distribution ............ 99
5.2. Basic Concepts of Inference ............................................. 100
  5.2.1. Hypothesis Testing and Statistical Formulation of the Medical Question ............................................. 103
  5.2.2. Statistical Estimation as the Tool for Evaluation of Clinical Relevance ............................................. 105
5.3. Statistical Inference in the Frequentist and the Bayesian Approaches ............................................. 106
5.4. Two Digressions: Measures of Variability and Likelihood Function ............................................. 110
  5.4.1. Measures of Variability ............................................. 110
  5.4.2. Likelihood Function ............................................. 114
5.5. Frequentist (Classical) Analysis of a Clinical Trial ............................................. 119
  5.5.1. Hypothesis Testing: the Frequentist Solution ............................................. 120
  5.5.2. Estimation of the Effect: the Frequentist Solution ............................................. 135
5.6. Bayesian Analysis of a Clinical Trial ............................................. 138
  5.6.1. Hypothesis Testing: the Bayesian Solution ............................................. 138
  5.6.2. Estimation of the Effect: the Bayesian Solution ............................................. 144
5.7. Some Additional Considerations on the Frequentist and Bayesian Approaches ............................................. 146
5.8. Parametric and Non-Parametric Inference ............................................. 149
5.9. Statistical Decision Making in the Medical Field ............................................. 150
  5.10. Evidence-Based Medicine ............................................. 152
Summary ............................................. 154

6. The Choice of the Sample ............................................. 157
  6.1. Which Subjects Should Form the Sample? ............................................. 157
    6.1.1. Characteristics of the Patients to be Enrolled in the Study ............................................. 157
    6.1.2. Mechanism of Subject Selection ............................................. 163
  6.2. How Many Subjects Should Form the Sample? ............................................. 164
    6.2.1. Statistical Considerations ............................................. 164
    6.2.2. Medical and Practical Aspects ............................................. 169
Summary ............................................. 171

7. The Choice of Treatments ............................................. 172
  7.1. Study Treatments ............................................. 172
    7.1.1. How Many Treatments ............................................. 175
    7.1.2. What Treatments ............................................. 176
    7.1.3. Blinding of the Study Treatments ............................................. 178
    7.1.4. Packaging and Logistics ............................................. 178
  7.2. Concomitant Treatments ............................................. 180
Summary ............................................. 182
8. Experimental Design: Fallacy of “Before-After” Comparisons in Uncontrolled Studies .................................................. 183
  8.1. Experimental Design: Introductory Concepts .................. 183
  8.2. Before-After Comparison in a Single Group of Subjects ...... 185
  8.3. Temporal Variations of the Disease ............................. 186
  8.4. Temporal Variations of Staff, Equipment and Environment .. 188
  8.5 Statistical Regression Toward the Mean ....................... 189
     8.5.1. The Basic Principle ........................................ 189
     8.5.2. Areas of Biomedical Experiments Affected by Regression Toward the Mean ........................................ 191
     8.5.3. How to Minimize the Effect of Regression Toward the Mean ....................................................... 193
  8.6. Learning Effect .................................................. 195
  8.7. Psychological Effect ............................................. 196
  8.8. The Before-After Design Without Control Group in Oncology . 197
Summary ........................................................................ 198

9. Experimental Design: the Randomized Blinded Study as an Instrument to Reduce Bias ............................................. 200
  9.1. Introduction ...................................................... 200
  9.2. Randomization as Antidote Against Selection Bias ............ 203
     9.2.1. Definition and Conceptual Framework ................... 203
     9.2.2. Types of Randomization ..................................... 206
     9.2.3. Other Methods for Assigning Patients to Treatments ... 215
  9.3. Blinding of Treatments as Antidote Against Assessment Bias .. 216
  9.4. A Priori Definition of the Statistical Methods and Populations as Antidote Against the Analysis Bias .................. 221
     9.4.1. Methods of Statistical Analysis ........................... 221
     9.4.2. Analysis Populations ......................................... 222
  9.5. Comparison Between an Observational and an Experimental Study .......................................................... 224
Summary ........................................................................ 227

10. Experimental Designs ............................................... 228
  10.1. Introduction ...................................................... 228
  10.2. Parallel Group Design .......................................... 233
     10.2.1. Characteristics .............................................. 233
     10.2.2. Advantages and Disadvantages ........................... 233
     10.2.3. Conditions of Applicability ................................ 234
  10.3. Variants of the Parallel Group Design .......................... 234
     10.3.1. Completely Randomized Parallel Group Design ........ 234
     10.3.2. Stratified Parallel Group Design ........................... 235
     10.3.3. Parallel Group Randomized Block Design .............. 239
     10.3.4. Balanced Incomplete Block Design ...................... 242
10.4. Other Designs with Comparison Between Subjects:
   Dose-Escalation and Dose-Titration .......................... 244
   10.4.1. Dose-Escalation Design ............................. 244
   10.4.2. Dose-Titration Design .............................. 246

10.5. Complete Cross-Over Design ................................ 247
   10.5.1. Characteristics ........................................ 247
   10.5.2. Advantages and Disadvantages ......................... 252
   10.5.3. Conditions of Applicability ............................ 255

10.6. Variants of the Cross-Over Design .......................... 256
   10.6.1. Variants Based on the Type of Randomization ........ 256
   10.6.2. Incomplete Cross-Over Designs ......................... 257

10.7. Other Designs with Within-Subject Comparisons:
   Simultaneous Treatments and Single Patient Designs ........ 261
   10.7.1. Simultaneous Treatments Design ...................... 261
   10.7.2. Cross-Over Design on a Single Patient (or “N of 1” Design) ......................... 261

10.8. Factorial Designs ............................................ 263
   10.8.1. Characteristics ........................................ 263
   10.8.2. Advantages and Disadvantages ......................... 268
   10.8.3. Conditions of Applicability ............................ 270

10.9. Split-Plot Design .............................................. 271
   10.9.1. Characteristics ........................................ 271
   10.9.2. Conditions of Applicability ............................ 273

10.10. Non-Controlled Designs in Phase II Oncology Studies .... 273

Summary ......................................................................... 275

11. Study Variants Applicable to More than One Type of Design:
   Equivalence Studies, Interim Analyses, Adaptive Plans and Repeated Measurements ................................. 277

11.1. Equivalence and Non-Inferiority Studies .................... 277
   11.1.1. Characteristics ............................................ 277
   11.1.2. The Statistical Analysis of an Equivalence Study .... 280
   11.1.3. Planning and Implementation Problems ................ 281
   11.1.4. Analysis and Interpretation Problems .................. 285

11.2. Studies with Interim Analyses and Sequential Designs .... 287
   11.2.1. Definitions and Classification ............................ 288
   11.2.2. Conditions of Applicability ............................. 290
   11.2.3. Choice of the End-Points ................................. 292
   11.2.4. Data Management Issues ................................. 293
   11.2.5. Statistical Issues and Decision Making Criteria .... 294
   11.2.6. Conflict of Interest and Confidentiality Issues .... 298

11.3. Adaptive (Flexible) Designs .................................... 299

11.4. Studies with Repeated Measurements ......................... 301

Summary ......................................................................... 302
# 12. The Drug Development Process and the Phases of Clinical Research

## 12.1. Overview of the Preclinical Development Process

## 12.2. The Phases of Clinical Development

### 12.2.1. Introduction

### 12.2.2. Phase I

### 12.2.3. Phase II

### 12.2.4. Phase III

### 12.2.5. Registration Dossier

### 12.2.6. Phase IV

### 12.2.7. Project Management

## 12.3. The Phases of Clinical Development for Oncology Compounds

### 12.3.1. Phase I

### 12.3.2. Phase II

### 12.3.3. Phase III

## 12.4. Accelerating Clinical Development

## Summary

### Appendix: Areas under the Curve of the Standard Normal Distribution

### References

### Analytical Index
In recent years many introductory textbooks on clinical trial methodology have been published, some of which are excellent, in addition to a very extensive specialist literature. Nevertheless, we decided to embark on the adventure of writing together a new book on methods and issues in clinical research. The objectives we set for ourselves, which we hope will justify our effort, can be summarized in three points.

1. Integrate medical and statistical components of clinical research. This is the primary objective of the book. The authors are a statistician (AB) and a physician (GDC) with years of experience in multidisciplinary project teams. In a clinical study (and any biomedical experiment) collaboration between representatives of the “biological” disciplines (physician, biologist, pharmacologist, etc.) and representatives of the “mathematical” disciplines (statistician, data management specialist, etc.) must be continuous and include absolutely all aspects of the planning and implementation of the study and of the analysis, interpretation and publication of results. The more troublesome this collaboration, the more at risk are the ultimate goals of our work as researchers, i.e. to ask relevant questions and provide scientifically sound answers. Unfortunately, however, there is often a complex communication problem between disciplines of biological and mathematical orientation, which can express itself at different levels.

A first, and in our opinion crucial, level of miscommunication is one of language. It is striking to note how frequently many of us make use of jargon, behind which we usually hide relatively simple and accessible concepts. This coded language is often an insurmountable wall for colleagues from different back-
grounds. What depths of mathematical reasoning lie behind the term “statistical model”? In fact, the term is often generically used to mean “approach”. What complex medical concept is hidden in the expression “differential hematological count”? It quite simply means type and number of cells in the blood.

A second level concerns the necessity to “divulge” (etymologically, to “make accessible to the people”). In order to communicate between different disciplines, it is necessary to simplify and, to some extent, trivialize concepts which are complex and rich with nuance (and fascinating to us for these very reasons). Specialists often have an inborn aversion toward divulging specialist knowledge. We subconsciously perceive it as a form of humiliation, as giving up the depth of knowledge and insight we acquired from years of study and experience. A genuine attitude toward disclosure and simplification is a welcome but rare quality, which requires profound knowledge combined with didactic intuition and empathy with the audience. We strongly believe that such an attitude is absolutely indispensable in order to accomplish interdisciplinary collaboration. We are convinced that the main reason for the limited success of some techniques and methods lie in the impossibility (rarely) or incapability (frequently) of making them accessible to a public of non-specialists. The Bayesian statistical methods are a case in point.

A third level that complicates communication between different disciplines is somewhat philosophical in nature and concerns the very way individuals from different backgrounds think. What is essential for one person may be trivial for another. The statistician often considers the “mechanistic obsession” of the biomedical researcher ridiculous: any result gets immediately fitted (or forced) into a plausible biological explanation, and so does any result reaching opposite conclusions. On the other hand, the statistician’s attention to the assumptions behind a certain method is often considered pedantic by the physician/biologist. Understandably, but mistakenly, we tend to think that “our” discipline is slightly closer to the “truth” than any other.

We hope that, through this book, we have made a small step in the right direction. Each and every medical term and concept had to pass the test of a non-physician. Likewise, each formula, statistical term and argument had to be understood and accepted by a non-statistician with only basic mathematical knowledge. Most importantly, the logical flow of each chapter had to make sense to both authors. This was achieved through endless debates and multiple rewrites of almost all chapters (some of which have been completely undone and redone four or five times). The logic behind the methods used in clinical research is the center of this book, not computational procedures or mathematical demonstrations, nor specialist medical or biological issues. Such logic must be equally accessible, through a common language, to the physician/biologist and to the statistician/data manager.

2. Do justice to the operational and practical requirements of clinical research. The decisions on the sample size, the choice of the dose(s) to be used in large phase III studies, the type and number of outcome variables, the
degree of blinding of treatments, are just a few examples where operational and practical requirements are of such importance as to prevail at times over the methodological ones.

How often do non-statistical considerations, such as the prevalence of the disease, the geographical distribution of researchers or financial constraints influence the sample size of a study? Isn’t it true that in these situations we tend to give a “statistical” justification to a sample size selected on practical grounds? To this end, we initiate a sterile retrospective process and screen the literature for any paper which could help us to justify our magic number; we inflate or deflate the magnitude of the “clinically significant difference” with the same aim, etcetera. Wouldn’t it be better to acknowledge the practical limitations and to estimate the power of the study for a range of differences, given the achievable number of patients?

The choice of the highest dose to be used in dose-response studies is frequently dictated not by pharmacological or toxicological considerations, but by limitations in pharmaceutical formulation technology or by the outcome of market research.

When choosing the number of measurements and end-points of a clinical study, a statistically pure, extremely restrictive approach to the problem of multiple comparisons fails to take into account the practical consideration that large clinical trials take years to complete, cost a fortune and very often represent unique opportunities to obtain essential information.

It is a common occurrence that scientifically impeccable protocols demand the impossible of patients and research staff: very frequent visit schedules, measurements taking many hours, repeated invasive procedures, “double-dummy” blinding solutions with dangerously complex dosing regimens. How many researchers seriously stop to consider whether all this truly contributes to improving the quality of data?

A competent researcher must not belittle “non-scientific” issues to the level of bothersome hurdles that get in the way of the perfect experiment, to be delegated to others (the “operational” staff). The real skill of a bio-medical researcher resides in designing a methodologically valid study that gives due consideration to real life and its many limitations. In this book we endeavored to elevate the operational aspects of clinical research to the level of importance they deserve.

3. **Give space to the ethical implications of methodological issues in clinical research.** Considerable progress has been made in the field of clinical research ethics in recent years. The dignity and rights of patients undergoing experimental procedures are sanctioned and regulated by international as well as national guidelines and regulations (e.g. the Declaration of Helsinki [106]). Ethics Committees (also known as Institutional Review Boards) are now operating almost everywhere. Fraud in research is a recognized problem that is policed both by the institutions where research is conducted (universities, hospitals, pharmaceutical industry) and by the institutions where research is evalu-
ated and regulated (FDA, other National Health Authorities, EMEA). However, a more insidious dimension of the ethical problem exists, which is very close to every researcher: the devious use of techniques and methods. Although this component of the ethical problem is well known to Health Authorities, researchers often ignore it. Whereas the fabrication of data is abhorred by the great majority of researchers, the fabrication of results from real data through incorrect use of methodological and statistical techniques is often considered no big deal and is unfortunately common practice. “After the fact” ("post hoc") selection of objectives based on results, poor use and interpretation of significance testing, subgroup analyses that prevail over the primary ones, statistical analyses that do not take into account the distribution of data nor attempt to verify the basic assumptions, preferential publication of “positive” data over “negative” ones. These are just some of the many ways in which results may be fabricated from real data. Is there such a big difference between fabrication of data and fabrication of results through incorrect methods? We are convinced that the answer is no. Inappropriate use of techniques and methods is immoral, as immoral as the fabrication of data and mistreatment of patients. The damage is just slower in manifesting itself, more difficult to identify and easier to justify with convoluted argumentation. Throughout this book we try to do justice to the ethical implications of poor research methodology. Clearly, the ethics of clinical research is a much larger subject, which goes far beyond the boundaries of our book.

On several occasions, we make reference to the thorny problem of conflicts of interest. Conflicts of interest are everywhere and we must acknowledge this fact. Researchers in academia must publish to obtain grants and to progress in their careers. Scientific journals benefit from new, surprising and unexpected results more than from results confirming previous research or “negative” results. The pharmaceutical industry is centered on financial profit and each company tries to “demonstrate” the advantages of its drugs and the disadvantages of the drugs of others. The large universities and teaching hospitals are most likely to support projects of great public awareness, not necessarily on scientific grounds (obviously to the disadvantage of other projects), in order to attract students and patients. How can research bear scientifically valid fruit in this jungle of conflicting interests? We believe there is no single answer or solution. A large component of the answer lies in the conscience and conscientiousness of researchers. However, much of the answer also lies in the mutual control, critical examination of publications, and repetition of experiments by different groups in diverse professional, cultural and social contexts, and in the use of methodologically correct techniques in the context of a truly interdisciplinary collaboration between researchers of different backgrounds and complementary skills.

At the end of this introduction, we wish to quote one of the reviewers of the English edition of our book, who, in our view, has captured very well its overall strengths and limits. “The authors have a great deal of practical experience and this experience gives the book an authority that is sadly lacking in some other
However, a weakness is that, however much experience the authors may have in writing and publishing articles on the results of clinical research, they have not really contributed to the methodological literature themselves. [...] This means that what one has, in my opinion, is a series of well-meditated reflections grounded in experience on methodological arguments and positions that have been expounded or developed by others”. The assessment is spot on and we gratefully accept it.

General Outline of the Book

The scope of clinical research is to evaluate the effect of a treatment on the evolution of a disease in the human species.

The treatment can be pharmacological, surgical, psychological/behavioral or organizational/logistic. The disease, intended as an impairment of a state of well-being or a condition capable of provoking such impairment over time, can be universally accepted as such (e.g. a cancer or a bone fracture) or perceived as such only by limited groups of individuals in a given cultural context (e.g. hair loss or weight gain). The course of the disease that one wishes to change can be the one with no intervention or, more frequently, the one observed with the available treatment.

The evaluation of the effect of a treatment on the course of a disease is a lengthy process, which progresses in increasingly complex stages.

As we will see in chapter 12, the clinical development of a pharmacological agent is conventionally broken down into four phases. Phase I (typically lasting six months to 1 year), generally carried out on healthy volunteers, has as its main goals the evaluation of the tolerability of increasing doses of the compound and the definition of its pharmacokinetic profile. The main goals of phase II (duration: ~1-2 years), carried out on selected groups of patients, are the proof of the pharmacodynamic activity and the selection of the dose(s) to test in the following phase. Phase III (duration: ~2-5 years), carried out on hundreds or even thousands of patients who (as a group) are to be as representative as possible of the general patient population, comprises the so-called pivotal studies, which are designed to demonstrate the therapeutic efficacy, tolerability, safety and at times also the socio-economic value of the compound. Finally, phase IV encompasses all the studies conducted after regulatory approval and marketing of the compound, within the approved indication(s). For non-pharmacological treatments the sequence of phases is generally less standardized.

The assessment of the effect of a treatment is a conceptually and methodologically complex process. It is useful to begin this book by asking ourselves why this may be the case. Why isn’t it enough to administer the treatment to one subject and then document the outcome?

The basis of this complexity resides in the ever-present variability of all biological phenomena, at times partially predictable, but often totally unpre-
dictable. It represents the “background noise” that must be overcome to recognize and measure the “signal” resulting from the treatment. In fact, clinical research can also be defined as a series of techniques and procedures aimed at separating the signal from the background noise, in order to decide if a change observed after a treatment belongs to the latter or to the former. The sources of biological and measurement-related variability are covered in chapter 1.

In order to separate the signal from the background noise, the clinical researcher can decide to conduct an experiment. What is an experiment? What are its characterizing elements that distinguish it from other forms of scientific investigation? Chapter 2 is dedicated to the characteristics that define a clinical experiment and to the fundamental distinction between observational and experimental studies (the latter typically referred to as clinical trials or clinical studies). Chapter 3 provides a brief introduction to observational studies. This chapter stands somewhat alone compared to the rest of the book. Nevertheless, we decided to include it for completeness and in order to better illustrate to the reader the differences between clinical and observational studies.

A key aspect of any experiment is that it must be carefully planned. The planning process must be complete before the experiment begins. Changes after the start of the experiment are indeed sometimes necessary, but must be the exception, not the rule, because they are complex to implement, have statistical consequences requiring careful consideration, and may have an impact on the credibility of results, as it is difficult to prove that a change has not been made to favor the results hoped for by researchers. The study plan must be documented in detail in the so-called study protocol, a document that requires many months to complete. First and foremost, a clear objective must be defined. Thereafter, a useful approach is to organize the planning of the experiment (and the writing of the study protocol) into six “blocks”.

- Definition and quantification of the treatment effect(s): what measurements should be carried out and how many; how to summarize them within each study subject (end-point); how to summarize the end-point within each treatment group (group indicator); how to express the overall effect in comparative terms between groups (signal); what is the minimum magnitude of the signal that must be shown to declare success of one treatment over another.
- Definition of the group of subjects on which the experiment is to be conducted: what features should they have and how many should they be.
- Definition of the experimental and concomitant treatments.
- Definition of the experimental design.
- Definition of the procedures for assessing results (statistical analysis and decision-making rules).
- Definition of logistic, administrative and legal issues.

Chapters 4, 5, 6 and 7 are dedicated to the definition and quantification of the treatment effects being studied, to the logical foundations of the statistical analysis, to the study sample and to the study treatments, respectively. Chapters 8, 9, 10 and 11 give an overview of the most common experimental designs, with emphasis on common methodological errors.
A detailed coverage of the logistic, administrative and legal aspects of clinical research is outside the scope of this book. However, throughout the book we keep reminding the reader of these aspects because, as already mentioned, we firmly believe they have a crucial role in determining the success of a study. The history of clinical research is paved with relics of studies started with great pomp, riding great ideas and great hopes, which drowned miserably because of inadequate logistical preparation. In our experience, the excessive complexity of a clinical trial is the single most frequent cause of failure: the study is perfect on paper, but impossible to implement by patients and staff alike. The distance between the principal investigators and the reality of clinical research in its day-to-day practice is often the main cause of such disasters. We warmly encourage everyone involved in clinical research to get involved in the logistics of a study, learning from colleagues responsible for its practical conduct (clinical research associates, data managers, etc.) and to take part, in person, in the practical implementation of a trial before attempting to design a study protocol.

The book ends with chapter 12, devoted to a brief description of the drug development process and to the phases of clinical development.

The Authors’ views expressed in this book do not necessarily reflect those of their employers.

Linguistic and Editorial Conventions

Each term included in the analytical index is highlighted in bold in the chapter where it is most extensively covered (in the index all terms appear as singular and, when necessary, have in parenthesis explanatory words which do not appear in the text). Sentences in bold and italics font are used to separate the longer sections into sub-sections. In formulae, the sign “x” is used to indicate multiplication, but omitted when it is obvious that a multiplication is being carried out.