Springer Series in Pharmaceutical Statistics

Christy Chuang-Stein
Simon Kirby

Quantitative Decisions in Drug Development

Second Edition

Springer

Springer Series in Pharmaceutical Statistics

Recent advances in statistics have led to new concepts and solutions in different areas of pharmaceutical research and development. "Springer Series in Pharmaceutical Statistics" focuses on developments in pharmaceutical statistics and their practical applications in the industry. The main target groups are researchers in the pharmaceutical industry, regulatory agencies and members of the academic community working in the field of pharmaceutical statistics. In order to encourage exchanges between experts in the pharma industry working on the same problems from different perspectives, an additional goal of the series is to provide reference material for non-statisticians. The volumes will include the results of recent research conducted by the authors and adhere to high scholarly standards. Volumes focusing on software implementation (e.g. in SAS or R) are especially welcome. The book series covers different aspects of pharmaceutical research, such as drug discovery, development and production.

More information about this series at http://www.springer.com/series/15122

Christy Chuang-Stein · Simon Kirby

Quantitative Decisions in Drug Development

Second Edition

Springer

Christy Chuang-Stein
Kalamazoo, MI, USA

Simon Kirby
Ramsgate, UK

ISSN 2366-8695 ISSN 2366-8709 (electronic)
Springer Series in Pharmaceutical Statistics
ISBN 978-3-030-79730-0 ISBN 978-3-030-79731-7 (eBook)
https://doi.org/10.1007/978-3-030-79731-7

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To Parents, Bruce, Terresa and Isaac

Christy Chuang-Stein

To Dad, Mum, Katie, Theo, Helen, Rob and Andy

Simon Kirby

Preface

By all counts, 2020 was an unusual year. COVID-19 lockdowns and pandemic-related travel restrictions had placed huge limits on our mobility. Many clinical trials were adversely impacted, and some were placed on hold due to uncertainties surrounding patient recruitment and care delivery. On the other hand, the unprecedented speed with which COVID-19 vaccines were developed and approved for emergency use was truly breathtaking. For the first time in modern product development history, the identification and the commercialization of a game-changing innovative medicine occurred in less than one year. Who would have thought this to be possible before 2020?

Confined to our respective locations and deprived of frequent family outings, we have decided to use the time wisely and revise the first edition of our book. The timing was good since a lot of new advances have been made since the publication of our first edition. The revision also allowed us to correct a few errors that went undetected when the first edition was published and a mistake spotted by an observant reader.

By and large, the second edition follows the structure of the first edition. We have decided to devote an entire chapter (Chap. 13) to the topic of adaptive designs, resulting in 14 chapters in the revision.

Chapter 1 offers a high-level overview of clinical testing and regulatory review of a pharmaceutical product. Chapter 2 reviews the Frequentist approach to the testing of hypotheses and in particular the two-action decision problem. In the context of drug development, the two actions correspond to progressing or not progressing a drug for further development. We discuss multiplicity and selective inference as well as how their indiscriminate uses have contributed to the replication crisis and the P-value controversy. Chapter 3 discusses the metrics commonly used to characterize the performance of a diagnostic test. Chapter 4 draws an analogy between successive trials conducted during the clinical testing of an investigational product and a series of diagnostic tests. Under this analogy, the condition to diagnose by a clinical trial is the existence of a clinically meaningful effect for the investigational product. We have found this analogy particularly useful to explain to our clinical colleagues why replication is such an important concept in drug development and to show why replication is not as easy as many people might hope.

The predictive power of a diagnostic test depends on the existing information concerning the prevalence of the condition to be diagnosed in a relevant population. Similarly, the predictive power of a clinical trial depends on available prior knowledge concerning the investigational product. Articulating such prior knowledge is the topic of Chap. 5. In the past 10 years, we have witnessed an increasing use of historical control data in the design and analysis of clinical trials, both when concurrent control data are available and when they are completely absent. We expanded Chap. 5 to include this topic and include an example on the challenge and use of historical control data when the trial providing evidential basis for an orphan drug approval included no concurrent control. In Chap. 6, we describe metrics that are useful to evaluate designs and associated decision rules for efficacy assessment at the various stages of pre-marketing development. The focus on efficacy is due to the generally well-defined endpoints to decide upon the beneficial effect of a new drug. Chapter 7 covers the proof-of-concept stage, while Chaps. 8 and 9 cover the dose–response and confirmatory stage, respectively. Throughout Chaps. 7–9, we have added new examples where the primary endpoint is binary.

Chapter 10 focuses on assessing the design of a trial for comparative effectiveness assessment. By comparative effectiveness, we mean the comparison of different active treatments to determine which treatment works best. This focus reflects the increasing importance of these comparisons in the market place due to the need to justify the price and to qualify for reimbursement.

The metrics used in Chaps. 7–10 do not include any cost consideration explicitly. But, cost is an integral part of drug development strategy optimization. Incorporating cost into design consideration is the topic of Chap. 11 with two example approaches. The first one optimizes a benefit–cost efficiency score that measures the cost-effectiveness of a proof-of-concept trial design. The second approach combines costs and potential commercial returns to assess drug development options. The chapter includes a detailed discussion on the calculation of the expected net present value which could be of interest to readers without much exposure to product valuation.

In Chap. 12, we examine the bias that can be produced by use of Phase 2 results that have been selected because of a favorable outcome. We have hinted at this source of bias in earlier chapters and have dedicated Chap. 12 to this issue. We have expanded this chapter to include an approach based on the maximum likelihood estimation of a truncated Normal distributions and new approaches we published in the literature in recent times. We compared different adjustment methods with respect to bias and other measures such as the probability of launching a Phase 3 trial and the average statistical power of the launched trial. We offer some recommendations based on the comparative results.

Chapter 13 is new, starting with classes of adaptive designs discussed in the finalized guidance on adaptive design issued by the Food and Drug Administration (FDA) in the United States (USA). We have included four examples of adaptive designs for, a group sequential trial with an interim analysis for futility, an adaptive dose–response study, a Phase 3 trial with a pre-planned sample size re-estimation plus population enrichment and a seamless Phase 2/3 trial.

In the final chapter of the book, we include selected topics that affect design and decision choices at all stages of drug development. Examples include sequences of trials with a correlated treatment effect, benefit–risk and economic assessment. These are all active research areas. Even though we offer some references, it is not our intent to cover these areas in detail in this book.

While the majority of the book is dedicated to trial planning and setting up decision rules, we have also included the analyses of completed trials to share insight and lessons learned. Examples with this objective include two dose–response studies of tofacitinib for moderate and severe rheumatoid arthritis in Chap. 8 and three adaptive designs in Chap. 13.

We have included numerous guidances published by the US FDA, the European Medicines Agency (EMA) and the International Council for Harmonisation (ICH). Instead of providing URL links to these guidances which could become obsolete over time, we are offering paths to locate these guidances. For guidances issued by the US FDA, readers can use the "Search All FDA Guidance" option at the path of www.fda.gov -> Drugs -> Guidance, Compliance and Regulatory Information. Guidances issued by the ICH could be reached by selecting the "Work Products" tab followed by "All Guidelines" option at the ICH home page www.ich.org. As for the EMA guidelines, readers can use the "Search for scientific guidelines" link in the middle of the page reached by paging through www.ema.europa.eu -> Human Regulatory -> Research and Development -> Scientific Guidelines.

As we stated in the Preface for the first edition, developing a new drug is a high-risk and high-reward enterprise. The high risk is reflected by the generally low success rate of turning a new molecular entity into an approved drug. The success rate has fluctuated over time and has also varied across therapeutic areas. While the success rate has improved in recent years for cancer drugs due to the advent of targeted therapies, the rate has been disappointingly low for certain disorders such as Alzheimer's disease.

In addition to the high risk, the cost of developing a new drug has increased at a pace faster than inflation. Tuft's Center for the Study of Drug Development has published a series of reports examining the average pre-tax industry cost to bring a new medicine to market. The most recent report, published in 2016, estimated an average cost around $2.56 billion USD in 2013 money. By comparison, in 2003, the cost was about $1.04 billion in 2013 dollars, based on the same method of calculation. While some researchers have questioned these figures, these reports nevertheless show a substantial increase in the cost of drug development over a few decades.

The low success rate and the high cost have motivated many pharmaceutical companies to look for better methods to make portfolio decisions including whether to invest in a particular new molecular entity and how to make Go/No-Go decisions. Since the majority of development programs are likely to fail, it is important to be able to terminate a program with a low probability of success as early as possible.

At Pfizer, where both of us worked for many years, the journey to quantitative decisions began during the first decade of the twenty-first century. The implementation began with proof-of-concept studies. Teams designing these early studies were required to present, to a technical review committee, the operating characteristics

of their trials/decision rules with respect to the target product profile. The move to assess the probability of success in the late-stage trials was firmly in place by the year 2010 with the establishment of a Probability of Technical and Regulatory Success (PTRS) Council.

Many statisticians and scientists played a critical role in the above journey. The input from commercial colleagues helped solidify the need to quantitatively incorporate the target product profile when designing a trial and setting up the subsequent decision rule. We have learned a great deal from the early pioneer advocates at Pfizer. Their work inspired us to write this book. We are particularly indebted to Mike Brown, Alan Clucas, Vlad Dragalin, Wayne Ewy, Bradley Marchant, Ken Kowalski, Mike K. Smith, Jonathan French, Cyrus Hoseyni, Richard Lalonde, Scott Marshall, Peter Milligan, Mohan Beltangady, Phil Woodward, Joanna Burke, Neal Thomas and Liam Ratcliffe for their scientific and organizational leadership.

The book is written for readers with a broad range of responsibilities in drug development. While the book contains a lot of technical details for quantitative scientists, it also contains plenty of concepts presented in a unified framework which, we believe, can help less quantitative readers make more quantitative decisions.

We hope you will enjoy reading the book as much as we did writing and revising it. Try as we may to ensure that the contents of the book are correct, there is always the possibility that some mistakes have gone undetected. If you spot one of these mistakes, we would be grateful if you could let us know by mailing us at christyazo@gmail.com (Christy Chuang-Stein) or s.kirby1.kirby@btinternet.com (Simon Kirby). Alternatively, we can be reached at www.linkedIn.com.

To close, we hope that we will all come out of the pandemic with new insights on trial conduct and what public–private partnerships could accomplish in expediting the development of life-saving medicines.

Kalamazoo, USA — Christy Chuang-Stein
Ramsgate, UK — Simon Kirby
March 2021

About This Book

Quantitative Decisions in Drug Development, 2nd edition, focuses on important decision points and evidence needed for making decisions at these points during the development of a new drug. It takes a holistic approach toward drug development by incorporating explicitly the knowledge learned from the earlier part of the development and available historical information into decisions at later stages. In addition, the book shares lessons learned from several select examples published in the literature since the publication of the first edition.

In particular, the book

- Shows the parallel between clinical trials and diagnostic tests and how this analogy is used to emphasize the importance of replication in drug development.
- Describes how to incorporate prior knowledge into study design and decision making at different stages of drug development.
- Explains metrics useful to address the objectives of the different stages of drug development and how to compare design options based on these metrics.
- Demonstrates why overestimation is a common problem in drug development and how adjustment should be considered to correct the overestimation.

Contents

About the Authors

Christy Chuang-Stein received a bachelor degree in mathematics from the National Taiwan University and a Ph.D. in statistics from the University of Minnesota. She retired from Pfizer as Vice President and Head of the Statistical Research and Consulting Center in July 2015, after 30 years in the pharmaceutical industry and 5 years in academia (University of Rochester). Currently, she is the owner and Principal Consultant of Chuang-Stein Consulting, LLC, and consults broadly in the areas of pharmaceutical development and evaluation.

She is a fellow of the American Statistical Association (ASA) and received the ASA's Founders' Award in 2012. She was a recipient of the Distinguished Achievement Award of the International Chinese Statistical Association in 2013 and the Distinguished Service Award from the National Institute of Statistical Sciences in 2020. She is also a repeat recipient of the Drug Information Association's Donald Francke Award for Excellence in Journal Publishing and the Thomas Teal Award for Excellence in Statistics Publishing. She is a founding editor of the journal *Pharmaceutical Statistics.*

Simon Kirby received a B.Sc. in Economics and Economic Policy from Loughborough University, an M.Sc. in Statistics from the University of Kent, a Ph.D. in Statistics from the University of Edinburgh and a BA in Mathematics from the Open University. He retired from Pfizer in 2018 after almost 20 years working as Principal Statistician, Clinical Statistics Head, Therapeutic Area Statistics Head and Consultant in the Statistical Research and Consulting Center. He is the owner of SKSTATS Limited for which he does occasional statistical consultancy.

Simon is Fellow and Chartered Statistician of the Royal Statistical Society. He previously worked as Lecturer, Senior Lecturer then Principal Lecturer in Statistics at Liverpool John Moores University and as Statistician at the UK's Institute of Food Research, Rothamsted Experimental Station and Revlon Healthcare.

Chapter 1
Clinical Testing of a New Drug

Nearly 60 percent of Americans—the highest ever—are taking prescription drugs.
—Washington Post, Nov 3 2015

1.1 Introduction

A research study reports an increase in the overall use of prescription drugs among adults (those ≥20 years old) between 2011 and 2012 from that between 1999 and 2000 in the United States (USA) (Kantor et al., 2015). In 1999–2000, an estimated 51% of US adults reported using any prescription drug. The estimated figure for 2011–2012 is 59%. During the same period, the prevalence of polypharmacy (use of ≥5 prescription drugs) increased from 8.2% to 15%. Many factors contribute to this increase, factors such as better disease prevention and management, lifestyle change, an aging population and an increase in the percentage of people who are either overweight or obese. The number of new prescription drugs developed and approved for public use every year has also greatly contributed to this increase.

Developing a new drug is a high-risk and high-reward enterprise. The high risk is reflected by the low success rate of turning a new molecular entity (NME) into an approved drug. The success rate has fluctuated over time and varied across therapeutic areas. For example, the US Food and Drug Administration published the Critical Path Initiative document in 2004 (FDA, 2004), in which FDA quoted a "current" success rate around 8% and a historical success rate of 14%.

Understandably, the success rate varies substantially across therapeutic areas (DiMasi et al., 2010, 2013). For example, the success rate of drugs for treating common bacterial infections is generally higher than that for drugs treating disorders of the central nervous system. This is in part due to the heavy use of the minimum inhibitory concentration (MIC) to help determine the appropriate dose and schedule for an NME for bacterial infections. For a microorganism studied in vitro, the MIC for an antibacterial agent is the lowest concentration of the agent which prevents

C. Chuang-Stein and S. Kirby, *Quantitative Decisions in Drug Development*, Springer Series in Pharmaceutical Statistics,
https://doi.org/10.1007/978-3-030-79731-7_1

detectable growth of the organism in agar or broth media under standardized conditions (Clinical and Laboratory Standards Institute, 2003). In addition to the MIC, animal models can be used to predict human response to an NME for many infections (Craig, 2003; Leggett et al., 1989). So, if an NME could deliver the desired MIC coverage without causing unacceptable side effects and if the animal model shows promising results, the NME will likely become a viable treatment option.

The success rate discussed above pertains to the clinical testing of an NME in humans. However, after an NME is synthesized, it will first be screened for biological and pharmacological activity. Preclinical testing in animals follows the biological and pharmacological screening. Preclinical testing is necessary before an NME can be tested in humans. Besides the need to understand the pharmacokinetic (PK) profile of the NME in animals, preclinical evaluation assesses the NME for its general toxicity, cardiac liability, carcinogenicity and reproductive toxicity. Some of the assessment could be done in vitro, but most is done in vivo using different animal species. The International Council for Harmonisation (ICH) has published a series of guidance documents on the technical requirements for preclinical safety evaluation of pharmaceuticals for human use. If preclinical testing suggests a reasonable PK and toxicity profile at doses likely to be used by target patients, then the NME will enter into the clinical testing stage.

Researchers have offered substantially different estimates for the success rates for the discovery and preclinical testing stages. For example, Bains (2004) estimated an approximately 30% cumulative success rate for discovery/preclinical testing combined while Hill (2008) stated a <1% success rate. Despite the difference, it is clear that the failure rate during the preclinical stage of drug development is not negligible.

In addition to the high risk, the cost of developing a new drug has increased at a faster pace than inflation. A study released by the Tuft's Center for the Study of Drug Development suggests that the average pre-tax industry cost to bring a new medicine to market was around $2.56 billion USD in 2013 money (DiMasi et al., 2016). The study included 106 investigational new drugs from ten mid- to large-sized pharmaceutical companies, and the drugs were first tested in humans during 1995–2007. Cost included clinical development up to 2013. By comparison, in 2003, the cost was about $1.04 billion in 2013 dollars. While some researchers have questioned the validity of these figures, the latest study used the same approach as that used in the previous one (DiMasi et al., 2003) in estimating the development cost. The latest study shows a substantial increase in the drug development cost over a 10-year period.

The low success rate and the high cost have motivated many pharmaceutical companies to look for better methods to make portfolio decisions. Such decisions include whether to invest in a particular NME and how to make Go/No-Go decisions concerning a particular development program. Since most development programs are likely to fail, it is important to be able to terminate a program that has a low probability to succeed as early as possible. Making efficient decisions requires designing efficient trials to acquire the needed evidence. Developing innovative designs that can enable good quantitative decisions at the earliest time has been the focus of much research in recent years.

Many books have been written about clinical trial designs to support drug development. Therefore, we will focus on methods for making quantitative decisions in this book. Because of the inseparable relationship between designs and decisions, we will also spend a good portion of this book on clinical trial designs.

In this chapter, we will offer a high-level review of clinical testing of a pharmaceutical product. We will first discuss in Sect. 1.2 the four distinct phases of clinical testing under a traditional development plan. We will discuss deviations from the traditional development plan and new regulatory approval pathways in Sect. 1.3. Section 1.4 offers some examples of recent advances in clinical trial designs. In Sect. 1.5, we briefly discuss real-world data and evidence before reflecting on the changing times in Sect. 1.6. We will conclude the chapter with a short summary in Sect. 1.7.

1.2 Clinical Development

Clinical testing of an NME to support its marketing authorization is often characterized by four phases as shown in Fig. 1.1. With some exceptions described in Sect. 1.3, three of the four phases occur before the NME is approved for marketing (pre-marketing) and the remaining one is afterward (postmarketing). The four phases are conveniently labeled as Phase 1, Phase 2, Phase 3 and Phase 4. A good description of the four phases can be found in an FDA guidance document (FDA, 1997) (Fig. 1.1).

1.2.1 Phase 1

Phase 1 trials are where an NME is first tested in human subjects. These trials are designed to investigate what the human body does to an NME in terms of absorption,

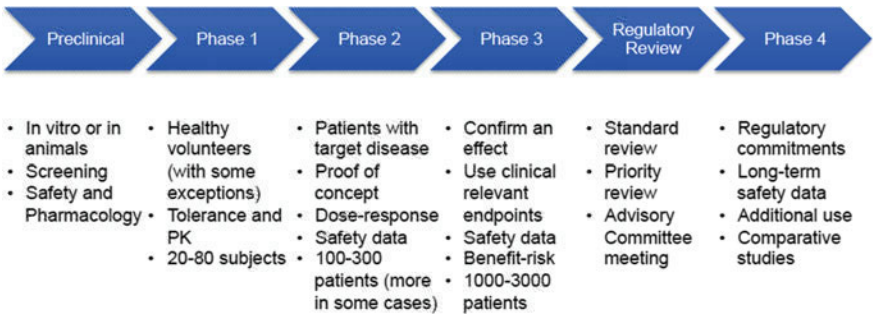

Fig. 1.1 Four phases of clinical testing

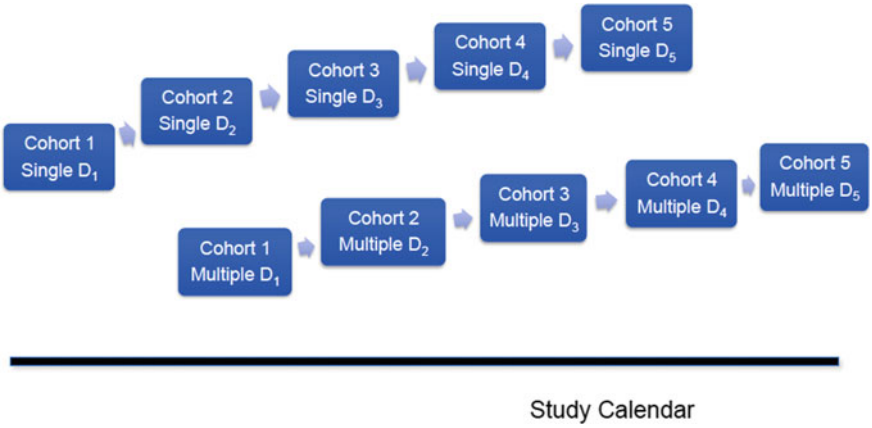

Fig. 1.2 Interwoven single- and multiple-ascending dose studies

distribution, metabolism and excretion (ADME). These are the pharmacokinetic (PK) properties of the NME. The investigation is typically conducted in healthy human volunteers, except for cytotoxic drugs. For cytotoxic drugs, Phase 1 is generally conducted in patients with very few therapeutic options due to the anticipated toxicities. When a drug is designed to target a receptor, Phase 1 trials can also include an investigation of what the NME does to the receptor.

Phase 1 trials in healthy subjects generally consist of single and multiple ascending dose cohorts. Trials studying the effect of a single dose on subjects typically precede trials studying the effect of multiple doses. Some development plans stack single-dose and multiple dose studies in such a way that there is a lag between exposing subjects to a single dose and exposing separate patients multiple times at the same dose. This strategy is shown in Fig. 1.2.

Besides collecting blood samples for PK analysis, Phase 1 trials investigate the common adverse reactions to an NME and what would be the NME's dose-limiting toxicities. We use the word "common" because the small number of subjects at this stage does not offer much opportunity to observe rare drug reactions. A typical ascending dose trial (single dose or multiple dose) randomizes subjects to a fixed dose or a control within a cohort. Observations from a cohort will be assessed to decide if another cohort should be recruited to investigate the next higher dose in a pre-specified dose range. The allowed dose range for Phase 1 testing is determined by the doses studied and adverse reactions observed in animal models.

If the NME's overall safety profile observed in Phase 1 is judged to be acceptable relative to its potential (and yet to be observed) benefit, the development will progress to the second stage (Phase 2). The number of volunteers included in Phase 1 single and multiple ascending dose studies typically ranges between 20 and 80, but could be higher if Phase 1 includes an assessment of the NME's mechanism of action or an early investigation of the NME's efficacy. The latter is a frequent feature of Phase 1 cancer trials. In these trials, a cohort of patients is often recruited at the maximum

tolerated dose (MTD) to assess the NME's efficacy once the MTD is established. A good reference on designs for Phase 1 cancer trials is the book by Cheung (2011).

Other trials with a strong PK focus conducted early in the development process include bioavailability studies, drug–drug interaction studies, food effect studies and PK studies in special populations such as subjects with impaired hepatic or renal functions. Understanding an NME's PK properties in individuals with hepatic or renal function is particularly important when an NME is excreted from the body through the liver or kidneys. Understanding how the body reacts to the NME under many different, yet important, conditions is important to the planning of subsequent trials.

1.2.2 Phase 2

Phase 2 investigates what a drug does to a patient with a target disorder (i.e., the pharmacodynamics of the drug). Clinical trials at this stage are also designed to determine dose(s) whose benefit–risk profile warrants further investigation later in a confirmatory setting. Multiple doses within the dose range identified from Phase 1 are studied at this stage.

Phase 2 is typically the time when a manufacturer first learns of the beneficial effect of an NME. This stage has the highest attrition rate among the three pre-marketing phases. Therefore, if an NME is not likely to become a treatment option, it will be best to recognize this fact as soon as possible and stop further testing of the NME for the disorder already investigated. This objective plus fewer regulatory requirements at this stage offer opportunities for out-of-the-box thinking.

Testing in Phase 2 can be further divided into two stages. The first stage aims to establish the proof of concept (POC) of the NME, using a high dose (e.g., the maximum tolerated dose identified in Phase 1) to investigate the NME's efficacy. Occasionally, a sponsor may use a biomarker to verify the conjectured mechanism of the NME in a proof of mechanism (POM) study. If the study cannot establish a positive POM or POC, the development of the NME in its current formulation for the indication under investigation will stop. Because an NME is often created with the objective to treat multiple disorders, discontinuing the development for one disorder does not necessarily mean terminating the development altogether. We have seen this in the oncology area where an NME may be targeted for multiple cancer types (e.g., breast, lung and renal).

Following a positive POC, an NME will be further tested in a dose-ranging study. A dose-ranging study typically includes a control and multiple doses of the NME. A placebo is often used as the control at this stage. The new NME and the placebo could be used alone as a monotherapy or added to a patient's background therapy.

This two-step process is often referred to as Phase 2a and Phase 2b (Sheiner, 1997). To minimize the work necessary to initiate sites and obtain approvals from multiple institutional review boards, some sponsors have opted to combine the POC and the dose–response studies into one study with an unblinded interim analysis at

the end of the POC stage. The sponsors will review results from the POC stage and may choose to use only data from the second stage to estimate the dose–response relationship. This strategy has the potential to increase operational efficiency by reducing the waiting period between Phase 2a and Phase 2b.

Depending on the target disorders, Phase 2 testing for a single disorder may consist of 100–300 patients. Despite strong advocacy by researchers like Sheiner (1997) to use a modeling approach to analyzing dose–response data, some sponsors continue to rely on pairwise comparisons to design and analyze dose–response studies. There have been renewed emphases from experts that the selection of dose(s) should be regarded as an estimation problem and handled by a modeling approach (EMA Dose Response Workshop, 2014). Recent research (Pinheiro et al., 2010; Thomas et al., 2014) has shown that 300 patients in a dose-ranging study may not be enough to adequately identify the optimal dose based on a preset criterion.

Ideally, Phase 2 studies should use the same endpoints to assess the benefit associated with a dose as those to be used later in Phase 3. Unfortunately, this is not always possible because the endpoint needed for Phase 3 such as survival and serious morbidity may take a longtime to obtain. In such a case, Phase 2 trials will often rely on a short-term endpoint that hopefully can predict the long-term clinical endpoint. An example is the use of progression-free survival as the endpoint in Phase 2 and overall survival in Phase 3 cancer trials.

Occasionally, a sponsor may have to conduct more than one study if the doses chosen in the initial dose–response study are not adequate to estimate the dose–response relationship. This could occur if the doses selected initially are too high (e.g., near the plateau of the dose–response curve) or not low enough. To reduce the chance of having to repeat a dose–response study, Pinheiro (2014) recommends including 4–7 doses in a wide dose range (e.g., the ratio of the maximum dose to the minimum dose ≥ 10) in the dose-finding study.

At times, different dose–response studies may need to be conducted for different diseases because a refractory disease may require a higher dose than a milder form of the same disease that has not been previously treated. Similarly, higher doses may be necessary to treat diseases considered to be harder to treat than diseases more responsive to treatment.

1.2.3 Phase 3

If the NME meets the efficacy requirement and passes the initial benefit–risk assessment, it will be further tested to confirm its efficacy. This is the final stage of clinical testing before an application is filed with regulatory agencies for marketing authorization. By this time, a commercial formulation of the NME should be available so the final testing could be conducted with the intended formulation. In the rare cases when the commercial formulation differs from the formulation used in Phase 3, a PK study will be required to show that the new formulation is bioequivalent to the

previous formulation in important PK properties. For convenience, we will refer to the NME as a drug candidate (or simply a drug) from this phase on.

The US FDA generally requires two well-controlled trials to confirm a drug's effect for a target disease. This means two independent Phase 3 trials, or in some cases, a Phase 3 trial plus a well-conducted high-quality Phase 2 dose-ranging study. The primary reason for requiring two "confirmatory" trials is to ensure that a beneficial result could be replicated.

There are situations, however, when one large well-controlled Phase 3 trial is considered adequate to support marketing approval. This occurs when the first study yields highly persuasive and robust results on a clinical endpoint (e.g. mortality and serious morbidity), and it is deemed unethical by the medical community to repeat a similar study. Here, robust results mean low P-values (described in Chap. 2) for the primary (clinical) and key secondary endpoints, consistent results across multiple subgroups and few issues associated with the conduct of the studies. Interested readers should consult with the FDA guidance (FDA, 1998) on providing clinical evidence of effectiveness for human drug and biological products.

Compared with previous phases, Phase 3 enrolls a greater number of patients who are more heterogeneous in their demographic and baseline disease status. Currently, nearly all Phase 3 studies are conducted in multiple countries and in multiple geographic regions. It is at this stage that the majority of pre-marketing safety data are collected. Since a major objective of Phase 3 trials is to confirm a drug's effect, analyses focus on testing pre-specified hypotheses with adequate control for the chance of making a false positive decision. Operations at this stage require carefully protecting a trial's integrity so that trial results could be trusted. The number of patients included at this stage typically ranges between 1000 and 5000. More patients will be needed if the drug is developed for multiple disorders simultaneously. An example for multiple indications is the development of antibiotics for multiple infections.

Drugs designed to reduce the risk of a clinical endpoint may require thousands, if not tens of thousands of patients. On the other hand, drugs for orphan diseases will enroll many fewer patients. An orphan disease in the USA is defined as a condition that affects fewer than 200,000 people nationwide. Orphan diseases include well-known diseases such as cystic fibrosis and Lou Gehrig's disease (also called amyotrophic lateral sclerosis, or ALS) and less well-known rare diseases such as Duchenne muscular dystrophy (DMD). DMD affects 1 in 3600 boys.

After a drug's effect is confirmed and benefit–risk assessment supports its use in the target population, the manufacturer will file a marketing application with regulatory agencies, typically in multiple countries. Nearly, all applications are for the adult population initially. If the drug is likely to be used in the pediatric population, a manufacturer often has an ongoing pediatric development program or has a plan to initiate pediatric trials at the time of the initial marketing application. The initial marketing application may be for a single indication or for multiple indications. Once the application is approved, the drug can be made available to the public.

As explained earlier, Phase 3 is the time when the majority of safety data are collected. Safety data are crucial for sound benefit–risk assessment. The International

Council for Harmonisation describes the extent of population exposure to assess clinical safety for drugs intended for long-term treatment of non-life-threatening conditions (ICH E1). For these conditions, ICH E1 expects 1500 individuals be exposed to the drug during the clinical development program. Among the 1500 individuals, 100 patients should have been exposed to the drug for at least one year. The exposure should be at the dose levels to be marketed. So, for a new drug with a large treatment effect, the need for a reasonable safety database will likely drive the sample size decisions for confirmatory trials.

1.2.4 Phase 4

The manufacturer of a marketed drug may choose to conduct additional studies to further (1) investigate the drug in the indicated population(s) or in pediatric patients with the indicated disorder(s); (2) compare the drug head-to-head with an approved drug for the same disorder(s); (3) investigate the effect of the drug at a lower/higher dose or with different administration schedules (e.g., once a day instead of twice a day); (4) study the drug in combination with other drugs or (5) test the drug for other indications. Sometimes, a manufacturer conducts Phase 4 studies as a postmarketing commitment for regulatory approval. For example, the manufacturer may be asked to conduct additional safety studies in vulnerable populations such as elderly, pediatric, obese or pregnant patients.

Another way to characterize the four phases of drug development is by the type of studies conducted during these four phases (see ICH E8, 1997). The types of studies conducted can be described as human pharmacology studies (Phase 1), therapeutic exploratory studies (Phase 2), therapeutic confirmatory studies (Phase 3) and therapeutic use studies (Phase 4).

In addition to the aforementioned documents, the ICH has published many other documents relevant to the clinical assessment of NMEs. The ICH is always working to expand the topics on which to offer internationally harmonized guidance and to amend existing documents as science and knowledge evolve. A major recent amendment was the addition of the addendum ICH E9(R1) (2019) to ICH E9 which focuses on statistical principles for clinical trials. The addendum, on estimands and sensitivity analysis, presents a structured framework to link trial objectives to a suitable trial design and tools for estimation and hypothesis testing. The central issue is how to handle missing data in the analysis of trial data to answer the primary objectives sought by the trials.

1.3 Regulatory Review

Section 1.2 describes a traditional clinical development process. It usually takes many years for an NME to go through the first three phases. Once a marketing

application is submitted, the manufacturer waits for the outcome of the regulatory review. Regulators often send queries to the manufacturer during this period for clarification or additional analyses.

In the USA, the FDA often arranges advisory committee meetings to publicly discuss submissions of NMEs or submissions that include unusual or controversial findings. Advisory committees will offer their recommendations to the agency. While these recommendations are not binding, the FDA often chooses to follow them. Before the turn of this century, the waiting period for a regulatory decision in the USA could be substantial. The review time has been significantly reduced since the beginning of the twenty-first century.

In Sects. 1.3.1 through Sect. 1.3.5, we will discuss deviations from the traditional review process that can help bring a drug with a clinically meaningful effect on serious conditions to the market faster in the USA.

In Sect. 1.3.6, we will review briefly procedures for drug approvals in the European Union.

1.3.1 Accelerated Approval

In 1992, the FDA instituted the *Accelerated Approval* regulations, allowing drugs for serious conditions that filled an unmet medical need to be approved based on a surrogate endpoint. A surrogate endpoint in this context is a measure of effect that may correlate with a real clinical endpoint but does not necessarily have a guaranteed relationship with the clinical endpoint.

Under the accelerated approval regulations, adequate and well-controlled studies that demonstrate a drug's effect on a surrogate or intermediate clinical endpoint could provide the necessary evidence for the initial marketing approval. This is the path for most cancer drug approvals in the past two decades. Even though the ultimate goal of a cancer treatment is to prolong survival, the initial approval of a cancer drug has been tumor shrinkage. The effect on tumor shrinkage is typically studied in Phase 2 trials. Some of these Phase 2 studies include only patients receiving the NME (i.e., a single arm) and rely on historical data to determine if the NME has a beneficial effect on tumor shrinkage.

With an accelerated approval, the manufacturer of a new NME for cancer still needs to conduct studies to confirm the ability of the drug to prolong survival. For this reason, accelerated approval is sometimes called *conditional* approval since there is a condition associated with the approval. A common industry practice is to start the clinical endpoint study once the effect of the NME on tumor shrinkage is confirmed. Safety data from the ongoing clinical endpoint study can be used to help augment the safety database to assist the initial regulatory review. The use of interim safety data in this fashion requires special care to protect the integrity of the clinical endpoint study.

Once a confirmatory trial verifies the clinical benefit, the FDA will generally remove the requirement. If the confirmatory trials fail to demonstrate a clinical

benefit, the accelerated approval may be withdrawn. A manufacturer often has a chance to conduct multiple studies to confirm the clinical benefit before the agency takes the step to withdraw the approved indication. Even if the approval is allowed to remain for the indication, the product label will be modified to clarify that trials failed to verify clinical benefit.

1.3.2 Breakthrough Therapy

In July 2012, the US Congress signed the FDA Safety and Innovation Act. The act allows FDA to designate a drug as breakthrough therapy if (1) the drug, used alone or in combination with other drugs, is intended to treat a serious or life-threatening disease or condition; (2) preliminary clinical evidence indicates that the drug may demonstrate substantial improvement over existing therapies on at least one clinically significant endpoint. A manufacturer can submit a request to the FDA to designate a drug as breakthrough therapy. The agency has 60 days to grant or deny the request. The submission should be done prior to the meeting with the agency to review Phase 2 results.

The breakthrough therapy designation allows the manufacturer to receive intensive guidance from the agency on the drug development program. It also signals the agency's commitment to the drug program at the senior management level including an expedited review of the drug's marketing application.

Having a drug designated as a breakthrough therapy is highly desirable. In addition to a quicker agency's response to requests for feedback and a faster review timeline, a breakthrough designation increases the prestige of a drug. A requirement for a breakthrough therapy is preliminary clinical evidence of substantial improvement over existing therapies on at least one clinically meaningful endpoint. The preliminary clinical evidence could come from an early trial in a small number of subjects. Pereira et al. (2012) reported findings from an empirical investigation on how often very large treatment effects were replicated in subsequent trials of the same comparison, disease and outcome. They concluded that most large treatment effects observed in small studies became much smaller when additional trials were performed. This is a point that we will return to in later chapters of this book.

1.3.3 Priority Review

In the USA, the Prescription Drug User Act (PDUFA) came into effect in 1992. Under the Act, manufacturers of prescription drugs pay a fee when submitting an application to market the drugs. In return, the FDA agreed to improve the drug review time with specific goals. The FDA also created a two-tiered review system timeline—*standard review* and *priority review*. The act is renewable every 5 years. The 2002 amendments to PDUFA (2nd renewal) set a goal that a standard review of a new drug

application be accomplished within ten months and a priority review be completed within 6 months.

A priority review designation is granted to drugs that, if approved, would contribute significantly to the treatment, diagnosis or prevention of serious conditions.

In the USA, a priority review voucher is awarded to any company that has obtained approval for a treatment for a neglected tropical disease and, in some cases, treatments with a rare pediatric disease designation (FDA, 2019a). The voucher, allowed under a provision of the Food and Drug Administration Amendments Act (H.R., 2007), is intended as an incentive to encourage companies investing in new drugs and vaccines for neglected tropical diseases or rare pediatric diseases. The voucher is transferrable.

The awarding of a priority review voucher has created an interesting phenomenon in the USA, that is, the selling of the voucher by its holder to the highest bidder in the open market. In some cases, the price paid for a voucher is hundreds of millions of US dollars. The purchaser can use the voucher toward any drug under regulatory review, hoping to get the drug to the market 6 months earlier or ahead of a rival drug that is being reviewed for the same indication contemporaneously.

1.3.4 Fast Track

Another designation that a manufacturer could seek of the FDA for their drug is *Fast Track*. A manufacturer could initiate the request at any time during the development process. The FDA will review the request and make a decision within 60 days based on whether the drug fills an unmet medical need in a serious condition.

A drug receiving the fast-track designation can expect to enjoy more frequent and timely interactions with the FDA. The manufacturer of a fast-track drug can submit sections of the new drug application for the agency review as they are being completed (rolling submission). A fast-tracked drug is eligible for accelerated approval and priority review, if other required criteria are also met (see Sect. 1.3.1 for accelerated approval and Sect. 1.3.3 for priority review). Because of more frequent communications and faster resolutions of issues, a fast-track designation often leads to earlier drug approval and access.

1.3.5 Orphan Drug

The US Congress passed the Orphan Drug Act in 1983 to provide incentives for developing treatments for orphan diseases (Kesselheim, 2010). The incentives include (1) federal funding of grants and contracts to perform clinical trials of orphan products; (2) a tax credit of 50 percent of clinical testing costs; (3) an exclusive right to market the orphan drug for 7 years from the date of marketing approval; (4) priority review by the FDA; (5) waiver of the drug application fees.

Within the class of orphan drugs, the amount of data submitted to support regulatory approval varies greatly. For example, on October 23, 2015, the FDA approved Strensiq® (asfotase alfa) as the first approved treatment for perinatal, infantile and juvenile-onset hypophosphatasia. Asfotase alfa, administered via injection three or six times a week, works by replacing the enzyme responsible for forming an essential mineral in normal bone. The latter has been shown to improve patient overall clinical outcomes.

The initial approval was based on the results from 99 patients who received asfotase alfa treatment for up to 6.5 years in four prospective, non-randomized studies. Study results showed that patients with the target condition and treated with asfotase alfa had improved overall survival compared with control patients selected from a natural history study group.

1.3.6 Drug Approval in the European Union (EU)

The first EU legislation on human medicine, triggered by the Thalidomide catastrophe and adopted in 1965, was Council Directive 65/65 on the approximation of the law relating to medicinal products. This was followed by two Council Directives in 1975. The first was on approximation of the Laws of Member States relating to analytical, pharmacotoxicological and clinical standards and protocols with respect to the testing of proprietary medicinal products. The second was on the approximation of provisions laid down by law, regulation and administrative action relating to medicinal products. The latter directive established a Committee on Proprietary Medicinal Products as an advisory committee and introduced the procedure now known as the mutual recognition procedure (Rägo & Santoso, 2008). A further directive introduced the procedure known today as the centralized procedure. In 1995, the European Medicines Agency was founded to harmonize the work of existing national medicine regulatory bodies and to protect public and animal health by assessing medicines to rigorous standards and providing partners and stakeholders with independent, science-based information on medicines (EMA: History of EMA, 2015).

There are currently two main routes for authorizing medicines in the EU (EMA: Authorization of medicines). The first is the centralized authorization procedure, whereby a manufacturer submits a single marketing authorization application to EMA. For new products, two rapporteurs are appointed from the Member States. The rapporteurs write scientific evaluation reports which are circulated to all other Committee for Medicinal products for Human Use (CHMP) members for comment. The CHMP reaches an opinion on the benefit–risk assessment by consensus or majority. If a decision is reached by a majority, then all CHMP members must accept the opinion. The second route is to make use of individual country national authorization procedures. If a manufacturer wishes to request marketing authorization in several EU Member States for a medicine that is outside the scope of the centralized procedure, it may use the mutual recognition procedure or the decentralized