

Design and Analysis of Clinical Trials in Drug Development

Tze Leung Lai, *Stanford University*

July 16, 2010

Outline

- 1 New drug development: “From bench to bedside”
- 2 Phase I clinical trials
- 3 Phase II and III clinical trials
- 4 Phase IV and postmarketing studies

New drug development: “From bench to bedside”

- Pre-clinical studies: “bench” refers to laboratory experiments to study new biochemical principles and discover novel treatments.
 - The experiments with promising results are followed by pre-clinical animal studies.
 - After understanding the effect of the treatment on animals, proceed to clinical trials.
- Clinical trials: involving human subjects. Phase I-IV reflects the sequential nature of the experiments involved

The science dealing with interactions between living systems and molecules, especially those from outside the system

- Clinical pharmacology:
 - To prevent, diagnose and treat diseases with drugs
 - Pathogenesis of disease due to chemicals in the environment
 - Drug: small molecule that alters the body's function when introduced into the body
 - Receptor: the components of the body that interacts with a drug and initiates the chain of biochemical events leading to the drug's effect

- Pharmacodynamics (PD): the study of drug effects and their mechanisms of action
 - The receptor concept has become the central focus of the investigation of PD
 - Complex relation between the drug and its observed effects
 - But in carefully controlled *in vitro* systems, the relation can be described by relatively simple mathematical models

- Pharmacokinetics (PK): the study about the concentration-time curve that is associated with the following “history” of a single administration of a drug:
 - absorption phase - transfer of the drug from its site of administration into the bloodstream
 - distribution phase - distribution of the drug to different compartments of the body
 - elimination phase - excretion of chemically unchanged drug or elimination via metabolism that converts drug into metabolites (e.g., at liver)

Dosage regimen

- Drug administration can be divided into PK and PD phase, both of which are important to the design of a dosage regimen to achieve the therapeutic objective
- Since both the desired response and toxicity are functions of the drug concentration, the therapeutic objective can be achieved only when the drug concentration lies within a “therapeutic window”, in which it is effective, but not toxic
- Drug concentrations are typically measured at the plasma
- Optimal dosage regimen: maintains the plasma concentration of a drug within the therapeutic window

- PK models
 - Mechanistic
 - Physiologic model - consider qualitative features shared by different tissues or organs, and use *a priori* knowledge of physiology, anatomy, and biochemistry
 - Compartmental model - the body is viewed in terms of kinetic compartments between which the drug distributes and elimination occurs. The kinetics is often described by a linear system of ordinary differential equations
 - Empirical - poly-exponential models of the form $\sum \alpha_i e^{-\lambda_i t}$
 - PD models - describe and quantify the steady-state relationship of drug concentration (C) at an effector site to the drug effect (E)
 - Emax model $E = e_{max} C / (C + c_{50})$

Population PK/PD models and NONMEM

Nonlinear least square to fit the poly-exponential regression model

$$y_j = \beta + \sum_{k=1}^K \alpha_k e^{-\lambda_k t_j} + \varepsilon_j$$

- ε_j are assumed to be independent with zero means
- $(\lambda_1, \dots, \lambda_k; \alpha_1, \dots, \alpha_k, \beta)$ can be estimated by weighted least squares

In many cases, data are collected from a number of subjects, some of whom may have intensive blood sampling while others only have sparse data. A primary objective is to study the PK/PD characteristics of the entire population → nonlinear mixed effect models (NONMEM)

- Sheiner & Beal (1980, PK Biopharm.); Lindstrom & Bates (1990, Biometrics); Lai & Shih (2003a,b Biometrika); Lai, Shih & Wong (2006, J.PK&PD)

Individual measurement model $y_{ij} = f_i(t_{ij}, \theta_i) + \varepsilon_{ij}$
population structure model $\theta_i = g(x_i, \beta) + b_i$

$$1 \leq j \leq n_i, 1 \leq i \leq K$$

- θ_i is the i th subject's parameters
- β is the "fixed effect" to be estimated
- $b_i \stackrel{iid}{\sim} G_\gamma$ (e.g., normal), are the random effects
- Can be solved by iterative procedure through software package **NONMEM** or the *nlme* in R

Phase I trial design

- Typical Phase I studies give drug to healthy volunteers, initiated at low doses and subsequently escalated to show safety at a level where some positive response occurs
- In cancer studies, patients are used as study subjects, and given the hoped-for benefit, aims at an acceptable level of toxic response in determining the *maximum tolerated dose* (MTD)
- 3-plus-3 design
 - 1 Treats group of 3 patients sequentially, starting with the min dose
 - 2 Escalate if no toxicity is observed in all 3 patients; o.w. an additional 3 patients are treated at the same dose level
 - 3 If 1/6 patients has toxicity, escalate; if 2/6 patients have toxicity, declare the current dose as the MTD; if more than 2/6 patients have toxicity, use the lower dose as the MTD

Phase I trial design

- “up-and-down” sequential design
- continual reassessment method (CRM) - use parametric modeling of the dose-response relationship and a Bayesian approach to estimate the MTD
- Escalation with overdose control (EWOC) - even though the response rates are low and that a large number of patients are treated at non-therapeutic dose, it is still widely used because of ethical issues

- Bartroff and Lai (2010a,b; Stat. Sci., Biometrics):
 - provide a mathematical representation of the dilemma between safe treatment of current patients in the dose-finding cancer trial and efficient experimentation to gather information about the MTD for future patients
 - a stochastic optimization problem that leads to a class of hybrid designs
 - a two-stage design whose first stage is a modified version of the 3-plus-3 design that generate data to check the parametric assumptions in the model-based hybrid design used in the second stage

Phase II and III clinical trials: General goal

- Phase II trials use the information collected and the dosage regimen determined from Phase I studies to evaluate the efficacy of the drug from particular indications in patients with the disease
- Phase III trials demonstrate effectiveness of the drug for its approval by the regulatory agency and also collect safety information from the relatively large samples of patients accrued to the trial

Phase II and III cancer clinical trials

- Phase II cancer trials: single-armed design in Simon's 2-stage design (Controlled Clin. Trials, 1989) testing tumor response, i.e., $H_0: p \leq p_0$ with significance level α and power $1 - \beta$ at an given alternative p_1 depends strongly on the prescribed p_0 and p_1 . Their uncertainty can increase the chance that a positive treatment being abandoned in Phase II or a profitless treatment proceeds to Phase III
- Phase III: time-to-event endpoint (time to bone metastasis in postate cancer, progression-free survival, overall survival)
- Vickers, Ballen, Scher (2007 Clinical Cancer Research): Uncertainty in the choice of p_0 and p_1 can increase the likelihood of failure at Phase III and of premature abandonment of potentially good treatment
- Inoue, Thall, & Berry (2002): Bayesian joint modeling of response and survival endpoints

Data monitoring: interim analysis and group sequential design

- Trial sample size is determined by the power at a given alternative, but it's often difficult to specify a realistic alternative at which sample size determination can be based
- Also, economic considerations impose constraints on the sample size.
- Solve this difficulty by sequential design that can “self-tune” its sample size to the increasing information on the unknown parameters during the course of the trial
- Fully sequential designs involve continuous examination of the data
- A compromise between fully sequential and fixed sample size design is a group sequential design involving interim analyses

Data monitoring: interim analysis and group sequential design

- Group sequential boundaries: let n_i be the total sample size up to the the i th interim analysis, $S_n = X_1 + \dots + X_n$ (sample sum)
 - Pocock (1977) uses square-root boundary $|S_{n_i}| \geq b\sigma n_i^{1/2}$
 - O'Brien and Fleming (1980) test has horizontal boundary $|S_{n_i}| \geq b\sigma M^{1/2}$
 - Wang and Tsiatis (1987) consider $\frac{|S_{n_i}|}{\sqrt{n_i}} \geq b\sigma \left(\frac{i}{k}\right)^{\delta - \frac{1}{2}}$, and choose δ to minimize the expected total sample size

Data monitoring: interim analysis and group sequential design

- The test for one-sided hypothesis stop not only when S_{n_i} exceeds an upper boundary (reject H_0) , but also when S_{n_i} falls below a lower boundary (suggesting futility)
- Since the clinical trial usually specify the calendar times of interim monitoring, group size will be unknown and unequal. Lan and DeMets (1983) proposed to use Brownian motion to solve the problem
- Lai & Shih (2004, Biometrika) developed flexible and efficient group sequential designs for exponential family, that can self-tune to the unknown parameters and can be generalized to generalized linear model. They used the alternative $\theta(M)$ implied by the maximum sample size M to construct the futility boundary.

Adaptive designs

- A.k.a “sample size re-estimation”, “trial extension”, “internal pilot studies”
- 2-stage procedure: treats the first stage as an internal pilot from which the overall sample size can be re-estimated.
 - intuitively appealing
 - doesn't adjust for the uncertainty in the first-stage parameter estimates that are used to determine the second-stage sample size.
 - perform poorly in terms of efficiency and power in comparison to group-sequential tests.

Adaptive designs

- Most literature has focused on finding ways to adjust the test statistics after mid-course sample size modification so that the Type I error probability is maintained at the prescribed level
- They also focused on the prototypical problem of testing a normal mean when the variance is known. The case of unknown variance is also considered when the “internal pilot” is used to estimate the variance

Adaptive designs

- Jennison and Turnbull (2006a) recently introduced adaptive group sequential tests that choose the j th group size and stopping boundary based on the cumulative sample size and sample sum, and that are optimal, in the sense of minimizing a weighted average of the expected sample size. They also found that standard group sequential tests with the first stage chosen optimally are as efficient as this design
- Bartroff and Lai (2008a,b) provided a unified treatment for multiparameter exponential families. It uses efficient generalized likelihood ratio (GLR) statistics, and adds a third stage to adjust for the sampling variability of the first-stage parameter estimates that determine the second-stage sample size. Extended to multi-armed and multivariate settings

Phase IV clinical studies

- The safety of the new drug is evaluated from the data obtained from all three phases of clinical trials prior to marketing approval of the drug, and continues to be evaluated through post-marketing Phase IV trials
- Recent area of active interest is post-licensure vaccine safety evaluation studies
- Post-licensure vaccine safety surveillance: Vaccine Adverse Event Reporting System, Vaccine Safety Datalink with data submitted by HMOs
- Sequential methods (Davis *et al.*, *Epidemiology*, 2005; Lieu, Kulldorff, Davis *et al.*, *Medical Care*, 2007; Shih, Lai, Heyse & Chen, 2010, *Stat. Med.*)

Conclusion and Summary

Despite the sequential nature of Phase I-III trials, the trials are often planned separately, treating each trial as an independent study whose design depends on results from previous phases.

- Advantage: the reproducibility of the results of the trial can be evaluated on the basis of the prescribed design, without worrying about the statistical variability of the results of earlier-phase trials that determine the prescribed design
- Disadvantage: the sample sizes are often inadequate because of the separate planning; inconclusive results at each phase
- Adaptation and sequential experimentation approach

Sequential Experimentation in Clinical Trials: Design and Analysis

by Jay Bartroff, Tze Leung Lai, and Mei-Chiung Shih
(Springer, 2011)

Table of Contents

- 1. Introduction

Part I *Sequential and Adaptive Designs of Clinical Trials*

- 2. Group sequential design of Phase II and III trials
- 3. Adaptive designs
- 4. Seamlessly expanding randomized Phase II into Phase III cancer trials
- 5. Sequential dose-finding designs
- 6. Confidence intervals and p-values
- 7. Multivariate outcomes and analysis of primary and secondary endpoints

Part II *Illustrative Examples and Other Topics in Sequential Monitoring and Experimentation*

- 8. The Beta-Blocker Heart Attack Trial and failure-time endpoints
- 9. Sequential approach to development and testing of biomarker-based personalized therapies
- 10. Sequential tests in vaccine safety evaluation
- 11. Dynamic treatment strategies and adaptive treatment allocation
- 12. Sequential experimentation and adaptive control in patient care and therapeutic development

- Appendix A Sequential testing theory
- Appendix B Boundary crossing probabilities for random walks and self-normalized processes
- Appendix C Optimal stopping, dynamic programming and experimental design
- Appendix D Survival analysis and time-sequential methods
- Appendix E Resampling methods, likelihood, Bayesian, and causal inference
- Appendix F Implementation and software packages
- References
- Index