

Innovation in Drug Development: Adaptive Designs for Clinical Trials

Definitions, Classification, Regulatory Responses, Operational Considerations and Recommendations: Version 2.0

By: Anastassios D. Retzios, Ph.D.*

Contents

A. Abstract.....	2
B. Introduction	3
C. Definitions	4
D. Types of Adaptive Designs.....	6
1. Model-dependent/Continual Reassessment Adaptive Designs.....	8
2. Group Sequential/Sample-Size Re-Estimation Designs	9
3. Group Sequential/Response-Adaptive Designs	10
4. Randomization-Adaptive Designs	11
E. Utilization of Bayesian Statistics	11
F. Regulatory Agencies and their Response to Clinical Study Innovations	12
1. The FDA’s Critical Path Initiative and Design Issues in Clinical Studies.....	12
a. The FDA’s View of Adaptive Designs.....	13
b. The FDA Draft Guidance on Adaptive Designs	14
2. The EMA Initiative.....	16
a. The EMA’s Reflection Paper on Adaptive Designs in Confirmatory Studies	17
G. Operational Considerations.....	19
H. Recommendations.....	21
1. The Planning Stage	21
2. The Conduct Stage.....	22
3. The Imperative of Full Documentation	23
I. Conclusions	24
J. Keywords	25
K. Appendix 1: The FDA Draft Guidance in Detail	25
L. References.....	35

* Address correspondence to the author at: Bay Clinical R&D Services, 2417 Canyon Lakes Drive, San Ramon, California 94582 - Address email to: aretzios@adrclinresearch.com

A. Abstract

In the last decade there has been an increased interest in adaptive designs for clinical trials. These designs allow the modification of several provisions of the clinical study protocol following an interim analysis while preserving the validity and integrity of the study. The elements that can be modified include the sample size of the study, the type of patient enrolled, the randomization algorithm, the primary endpoint and others. These designs promise a reduction of the attrition rate in new medical entities in development and reductions in time and money required to develop new agents. The experience so far is limited and such claims remain to be verified.

A number of industry groups have attempted to define and promote the use of these designs. The Pharmaceutical Research and Manufacturers of America (PhRMA) defined adaptive designs a “*study designs that uses accumulated data to modify elements of the study without undermining the validity and integrity of the study*”, a definition that has met with wide acceptance although recently the FDA introduced a new definition that more closely matches regulatory views of these designs (see [Section C](#))

There are various classifications of these designs. The one introduced in this paper is essentially operational and it is based on the degree of design heterogeneity introduced in parts of the study after the interim analysis(es). Thus, we discern four main groups: model-dependent/continuous assessment method; group sequential/sample size re-estimation; group sequential/response adaptive and adaptive randomization designs.

The regulatory response to these designs has been surprisingly uniform in the US and Europe. The FDA and the EMA noted the promise of these designs but also sounded a note of caution regarding several practices. Neither of these agencies has produced a full official guidance at this stage. Members of the FDA and the EMA have published papers in industry journals examining aspects of these designs. The FDA has released a draft guidance for comments and the EMA has produced a reflection paper on adaptive designs used in confirmatory clinical trials. This report examines these publications and summarizes the current regulatory approach to these designs. It is expected that the regulatory position would evolve with the increased utilization of these designs.

There are a number of operational considerations in the application of these designs. The potential beneficial effect of adaptations depends highly on the information derived by data collected from each patient treated in the study (in continual assessment method designs) or at the interim analysis (for group sequential designs). Thus, it is imperative that the study management makes certain that compliance to the study is strictly enforced and that data are collected in a timely and accurate fashion. A high number of deviations or many missing data will

undermine the validity of the analysis and the quality of the decision from the small set of data at the interim. In addition, when unblinding is required for the interim analysis, the appropriate firewalls should be in place and should be copiously documented. It is imperative that no bias is introduced by either the sponsor or independent committees that render decisions on various study adaptations. This paper provides an analysis of regulatory concerns and guidelines as expressed by major agencies and certain recommendations for meeting the operational demands of these studies.

B. Introduction

In one of my previous articles on the rate of failure in Phase 3 clinical studies,¹ certain fault lines in clinical research were examined. We discussed limitations in designs and endpoints that affect the pace of development and/or the accuracy and

reliability of the collected information. After more than four decades of modern clinical research, the limitations of our “tools” are obvious, but the methods to overcome them are not.

As can be seen in Figure 1, there has been a progressive and substantial decrease in the approval rate of new medical entities (NMEs) and hefty increases in development budgets. These findings are not limited to the US. They are consistent throughout the area covered by major regulatory agencies. The reasons for this trend are multiple: The limitations of clinical development tools are certainly one them and one that has been addressed previously by the author.¹ Others are related to bottlenecks in discovery (relevant animal models, biomarkers, etc.), inadequate information sharing, risk-averse funding of new companies, industry consolidation and the concentration of resources to new medical entities (NMEs)

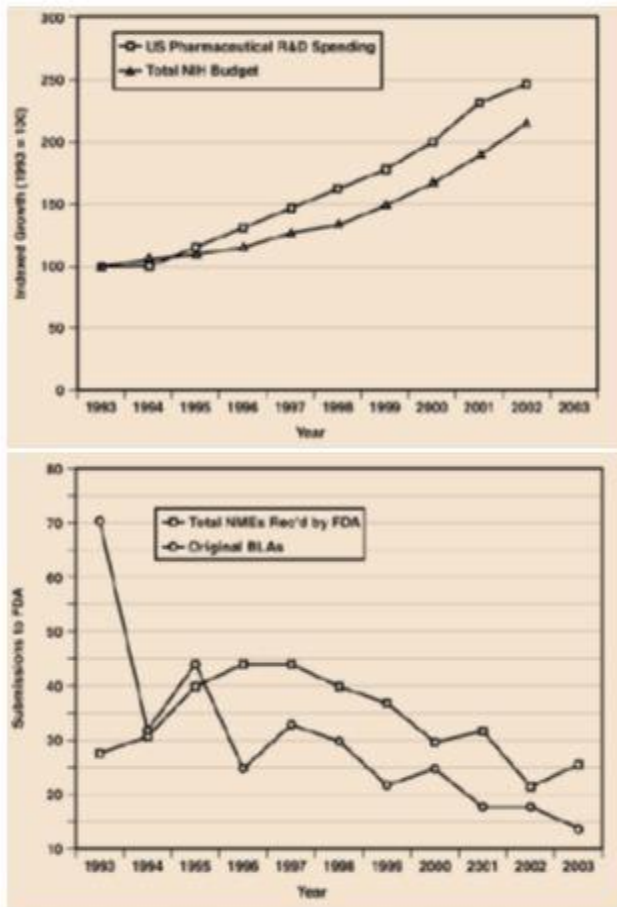


Figure 1: Increase in pharmaceutical R&D expenditures and drop in NME approvals between 1993 and 2003. The graphs and data are derived from the “Challenges and Opportunities Report” published by the FDA in March 2004.

with substantial marketing potential.

Alarmed by this trend, both the FDA and EMA have launched initiatives to jumpstart the process of innovation, as have various academic and industry forums. These initiatives address clinical study designs among other research bottlenecks. Thus, the renewed interest in clinical study designs referred to as “adaptive” or “flexible”.

The concept and practice of changes in clinical trial design based on accumulated information are by no means new. However, recent innovations in the manner and the extent of these adaptations while retaining study validity and integrity have captured the attention of many in the pharmaceutical R&D community.² They have also elicited strong interest from regulatory authorities in pursuit of efforts to increase innovation in drug development. As the utilization of these methods is still rare, it remains to be determined if adaptive designs are capable of reducing budgets, shrinking timelines and of improving attrition rates. As we will be discussing later, there is a price to be paid for adaptability. The operational efforts to obtain quality data in the interim and maintain study validity are often very demanding and probably beyond the capabilities of most R&D departments with the exception of those of large pharmaceutical companies. It is important to note here that this paper is not meant to be a detailed review of adaptive clinical trial designs. A number of recent reviews provide an exceptional level of detail.³ Although core elements of these designs will be summarized therein, my aim here is to broadly classify these designs in terms easily understood by all participants in pharmaceutical research, examine their impact on associated operational activities, highlight regulatory concerns in the US and EU, and discuss their relevance in a real world environment.

C. Definitions

Adapting clinical trials to problems encountered while patient accession and data collection is ongoing is not new. In many cases, changes to the protocol are necessitated by a variety of factors, most notably lower than expected accession rates or difficulty in collecting specific information pertinent to the endpoints of the study. Studies can also be stopped after planned interim analyses for achieving their efficacy goals or for futility. However, the term “adaptive clinical trial design” extends well beyond these practices.

In 2006, the PhRMA, a pharmaceutical industry organization, assembled a working group to provide definitions and to examine additional issues pertaining to adaptive clinical trial designs.⁴ This group defined the term adaptive design as:

“Any clinical study design that uses accumulated data to modify elements of the study without undermining the validity and integrity of the study”.

The first element of the definition is not specific enough to provide any insight. As mentioned earlier, modifications of studies (either planned or unplanned) on the basis of collected information were not unknown prior to the recent emphasis on “adaptive designs”. Interim analyses in a group sequential methodology allowed the possibility of the “early” termination of the clinical study if efficacy had been achieved and for futility while preserving the type I error (maintaining the rate of false positives at the pre-specified level). In fact, such designs also allowed a certain flexibility regarding the number of interim analyses.⁵ Obviously, the “adaptive designs” that have attracted so much attention in this decade go beyond this level of adaptation (they are discussed in detail in [Section D](#)), but no specific details are included in the PhRMA definition.

The second element of the definition, that of preserving the validity and integrity of the study, refers to the statistical and operational considerations utilized to provide the same level of statistical inference and integrity of process as classical “fixed” designs. However, as we will see, the regulatory agencies have specific concerns which will progressively define the acceptable bounds of adaptive designs and the extent of their utilization in pivotal studies. The debate between the regulatory agencies and the industry is ongoing and the regulatory approach will evolve as our experience deepens.

A missing element of the definition is the words “planned” and “prospective”. Expected adaptations in study design must be declared prospectively allowing all “stakeholders” to plan accordingly and for the regulatory agencies to provide input regarding the validity of the approach. Planning and meticulous execution of the plan is of outmost importance in maintaining the integrity of a study based on an adaptive design.

The “prospective” and “planned” element of these designs is a core element of the definition that was introduced by the FDA in its [recent draft guidance on adaptive designs](#): The definition reads as follows:

An adaptive design clinical study is defined as a study that includes a prospectively planned opportunity for modification of one or more specified aspects of the study design and hypotheses based on analysis of data (usually interim data) from subjects in the study. Analyses of the accumulating study data are performed at prospectively planned timepoints within the study, can be performed in a fully blinded manner or in an unblinded manner, and can occur with or without formal statistical hypothesis testing.

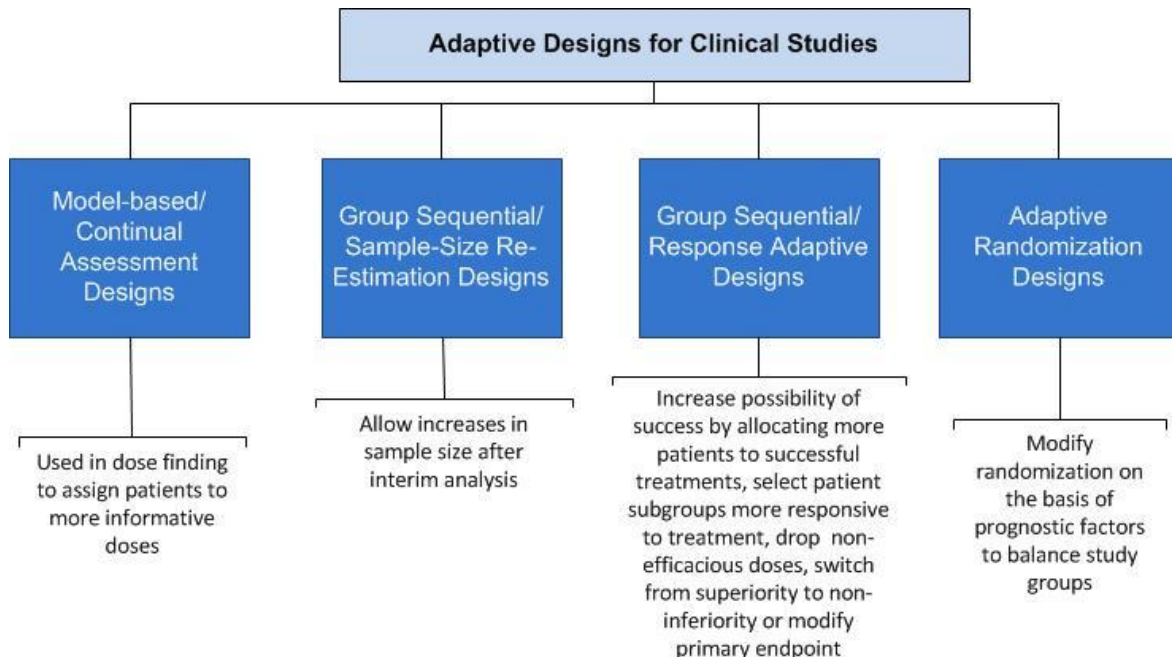
It is interesting that the FDA has side-stepped the statement “without undermining the validity and integrity of the data” in the industry-proposed definition. As we will examine later, the FDA has reservations about specific adaptations because of their potential to introduce bias and confound the interpretation of study results. The FDA draft guidance will be discussed in greater detail in [Section F.1.b](#) and in [Appendix 1](#).

D. Types of Adaptive Designs

There are adaptations that address most aspects of a clinical study. Thus, the classification of adaptive designs presents a number of challenges. A “rules” based classification is certainly possible, although somewhat cumbersome. In general, adaptive designs modify four basic elements of the study (occasionally referred to as rules) : (a) the manner in which the patients are randomized into the study (allocation rule); (b) The number of subjects that will be included (sampling rule); the rules by which a decision will be made to move to a different stage or modify elements of the study such as the primary endpoint and/or the method of analysis (selection rule); and how the study will be brought to an end (stopping rule). Thus, utilizing a matrix of rules, designs can be classified by the number and type of rules they modify. Such a classification may be detailed but it fails to provide relevant information to persons outside the biostatistics community.

Another classification scheme is based on the phase of development for which the adaptive design is best suited. Certain adaptations are utilized mostly in early development studies, others in the confirmatory and pivotal phase. Such a classification scheme is simple, but lacks specificity.

In this paper, I have classified the studies on the basis of how the results of the interim/continual analysis affect the post-interim structure and design of the study. This is essentially an “operational” classification and one that I think provides a real-world insight to these designs. The classification presented here is not drastically different from the one presented by Coffee and Kairalla,⁶ albeit simplified. The reasons for the simplification are explained below, along with the outline of this classification.



- Model-based/continuous assessment adaptive designs: A group of designs commences mostly with modeling and simulation, initiates subject treatment, continuously assesses collected data and then assigns consecutive patients to dose groups on the basis of the fit of the collected data to the model. These designs are utilized mostly in early clinical development for dose finding and ranging.
- Group sequential/sample-size re-estimation (SSR) designs: This second group consists of group sequential designs* in which the interim analysis is utilized to recalculate the sample size and possibly the number of additional interim analyses. They were introduced by Bauer and Köhne⁷ based on the premise of planning a multi-stage study as a meta-analysis of independent studies. In these designs, there are usually no changes in the groups being tested and in patient characteristics after the interim analysis. Heterogeneity between stages is less pronounced than in response adaptive designs.
- Group sequential/response adaptive (RA) designs: This third group consists of designs in which several elements of the study, such as enrollment criteria, number of arms/doses, randomization scheme, primary endpoint, stopping rules and switching between non-inferiority to superiority or vice versa, are modified in the 2nd stage after subject response in the 1st stage is assessed. The heterogeneity of the study is thus more pronounced in stage II of the clinical trial than in the SSR designs. These designs are very operationally demanding and can be used both in early and late phases of development.
- Adaptive randomization designs: In this final group belong studies that utilize an adaptive randomization scheme to balance the groups of the study for a variety of prognostic factors that may have a bearing on outcome.

One can expand the categories of adaptive designs by including hybrids and combinations of these designs. In certain classifications, seamless phase 2/3 clinical trial designs⁶ are treated as a separate category of “adaptive designs” although they are simply administrative constructs of the designs described above. Seamless Phase II/III studies may utilize a model-based/continual assessment adaptive design in phase 2 for dose finding, followed by a Phase 3 study with the doses of interest utilizing either a sample size re-estimation, an adaptive response or a classical fixed group design. Such seamless designs certainly save time and money as they decrease the logistical requirements and compress the time for starting the Phase 3 study. They do have a number of operational disadvantages including certain regulatory reticence when used in pivotal phases (Section F.) Also, sample-size re-estimation designs can include elements of response adaptation. Readers are advised to examine other classifications of these designs. Both Chow and Chang³ and Dragalin,⁸ among others, have surveyed and classified these designs in detail.

* In sequential designs, data are assessed after each patient is treated to compare the test statistic with the boundaries of the stopping rule; in group sequential designs, groups of patients are treated prior to an interim analysis to assess if the boundary has been crossed or not.

1. Model-dependent/Continual Reassessment Adaptive Designs

These designs are normally used in early development, in phase 1 and phase 2 studies. In the early stages of development, the effort consists of carefully defining the risk/benefit ratio throughout the dose spectrum. Thus, early studies attempt to define the maximum tolerable dose (MTD) and the minimum effective dose (MED). In this process, there is a moral imperative to quickly discontinue ineffective or toxic doses of a drug while obtaining reliable information. Adaptive designs can successfully address both of these requirements.

In Phase 1 safety studies, adverse reactions and toxicities are the primary endpoints. Phase 2 dose-ranging studies usually utilize surrogate efficacy endpoints or a combination of pharmacodynamic/biomarker endpoints. Although usually their correlation to the clinically beneficial endpoint is somewhat weak,[†] they can produce a statistically meaningful differential response to treatment with relatively few subjects. In the classical Phase 1 methodology, a typical dose escalation (usually 3 + 3) approach[‡] is employed until the MTD is defined. The adaptive approach is based on the creation *a priori* of a model of the toxicity curve expected; this model is continuously adjusted and refined with the results obtained from each individual patient. The dose selected for each subsequent patient depends on the presence or not of a dose-limiting toxicity (DLT) in the previous patient. If a DLT occurs, the dose is lowered; if not, it is increased. This design is based on a continual reassessment method (CRM) of subject information introduced by Quigley *et al.*⁹ A number of variants merging this method with a traditional 3 + 3 design soon appeared.^{10,11} Some of these hybrid methods do not require a prospectively defined dose toxicity curve but proceed in the manner of the classical 3 + 3 design.

Dose ranging studies in Phase 2 present similar challenges and opportunities. In the classical approach, which is still prevalent in use, the study subjects are randomized equally to a spectrum of doses[§] deduced from previous preclinical or clinical studies. The sample size is declared prospectively, calculated in the typical process that utilizes the expected differential in response between doses, the desired level of significance and power. Despite its simplicity, such a design may be inefficient: it may miss a substantial section of the dose-response curve.

[†] For the discussion of phase 2 study endpoints see: [“Why do so many phase 3 studies fail? Part 1: The Effect of Deficient Phase 2 Trials in Therapeutic Areas with High Failure Rates in Phase 3 Studies”](#)

[‡] The 3 + 3 design is based on enrolling 3 subjects per dose in an escalating manner. If none of the patients has a dose limiting toxicity (DLT) at a given dose, 3 more subjects are treated in the next dose level. If 1 subject displays a DLT, 3 more subjects are enrolled. If 2 subjects overall develop a DLT, then escalation stops and the previous dose level is declared as the Maximum Tolerable Dose (MTD)

[§] In these designs, 3 to 5 doses, placebo and/or active control is a typical configuration of the arms of the study

Even if it does not, the vagaries of sample size calculation may mean that the variance in response may be such that no statistical significance can be obtained.

Adaptive designs for dose ranging studies try to avoid these pitfalls and provide a more reliable definition of dose response although there are number of limitations and caveats in their use.¹² These designs are typically based on the continual reassessment method (CRM) outlined above for phase 1 clinical studies. These CRM-based designs can be combined with a fixed randomization design for a confirmatory study in a seamless Phase II/Phase III construct.

2. Group Sequential/Sample-Size Re-Estimation Designs

Group sequential designs were developed early on to allow a clinical study to be stopped after an interim analysis if the results showed that there has been a convincing demonstration of efficacy or a demonstrable futility of achieving a meaningful efficacy result.¹³

Group sequential designs possess a certain appeal because many of the elements that lead to the estimation of the sample size for the study are informed assumptions that may not be representative of the response in the population and/or the selected sample. It is well known that enrollment criteria substantially modify the sample from the overall disease population so that many assumptions based on previous results or clinical observations may no longer apply. More often than not, they are derived from groups or subgroups of patients with different demographics and prognostic factors from those of the prospective population of the planned study.

The major elements that go into the calculation of a sample size (beyond the desired level of significance and power) are the expected clinically beneficial effect** (δ) and the anticipated variance (standard deviation, σ). In certain cases, regulatory guidances may provide detailed information as to the size of the effect necessary to gain marketing authorization. .

In classical group sequential designs, a study is powered by the best guess for the lowest clinically beneficial effect required to obtain marketing authorization. This results in the maximum sample size to be assessed in the study. The reverse is true for adaptive group sequential/sample size re-estimation designs. In these, a more optimistic value for the clinically beneficial effect is selected, thus allowing the study to start with a relatively small number of patients and increase the sample size, if needed, based on the response assessed at the interim analysis.¹⁴

** The primary endpoint of a pivotal study has to correspond to a well-defined clinical benefit. The terms “clinically beneficial effect” and “treatment effect” is synonymous here with the primary endpoint of a pivotal study

Of course, what applies to the treatment effect applies equally as well to the variance. In this case, one may examine the variance of the primary endpoint by pooling all results and thus avoid unblinding during the interim analysis.¹⁵ In this case, the type I error rate is fully preserved and no statistical penalty needs to be paid.

Does the sample size re-estimation methodology present specific advantages? In cases in which the guess for the clinically beneficial effect is based on less than ideal information, and in which one wants to avoid expanding unduly the scope of early development to define it better, such designs may have a definitive utility. In a variety of therapeutic areas such as neurology, oncology and others, earlier phases of development utilize surrogate endpoints (e.g., brain lesions in the case of multiple sclerosis, progression-free survival in oncology). The typical endpoint for the pivotal phase must correspond to clinical benefit (e.g., time to progression to next stage in multiple sclerosis); thus prior information based on surrogate endpoints or prior clinical trials may only provide informed guesses as to the possible treatment effect and the statistical underpinnings of the study.^{††}

There are a number of concerns: the differences observed at the interim, based on small samples may be due to chance fluctuations and may mislead sample re-estimation. In addition, if the sample size in the 2nd stage is very large in comparison to the sample size of the Stage I, a remote possibility exists that a treatment may be declared beneficial if the results of stage I are highly positive but the much larger dataset of stage II is negative.¹⁶ Also, adaptation has a price: adaptive group sequential designs are not as statistically efficient under certain circumstances as the classical “fixed” approaches and may result in more patients treated prior to the conclusion of the study than it would have been the case otherwise.^{17,18} There are regulatory concerns: regulatory authorities would need to be convinced that these designs are not utilized for the sake of minimizing the Phase 2 program. In addition, designs in which not only the sample size but also the number of interim analyses are subject to change may face regulatory obstacles, as more than one interim analysis may make it difficult to convince regulators that the integrity of the study has been adequately maintained ([Section F](#)).

3. Group Sequential/Response-Adaptive Designs

These are designs in which substantial changes are introduced at the interim stage based on patients’ response to treatment with the sole purpose of

^{††} Oncology presents certain challenges for group sequential approaches (either adaptive or not) because oncology studies utilizing such designs are likely to use progression-free survival (PFS) to assess efficacy during the interim stages, not overall survival (OS). Thus, a stopping rule would have to utilize the surrogate endpoint. The problem with this approach is discussed in summary by Hung *et al.*²⁶

“amplifying” this response, concentrate resources and patients into effective treatments and generally improve the possibilities of a successful outcome. The changes may include reduction in dose groups, modification of enrollment criteria, and switching from superiority to non-inferiority. Also, randomization may become unbalanced in stage II, allowing allocation of more patients to groups with superior outcome or by the discontinuation of treatment arms with inferior outcomes (play-the-winner¹⁹ / drop-the-loser designs²⁰). In certain cases, the adaptation following the interim analysis may include a change in the primary endpoint (or components of a composite endpoint) if the clinical benefit in a certain disease is not well understood and no specific regulatory guidance applies. Thus, there is usually a substantial qualitative change between Stage I (study design prior to the interim analysis) and Stage II of the clinical trial. These studies are operationally quite demanding for both the sponsor and for the investigative sites in which the study is carried out.

Many two-stage designs are appropriately powered in stage II. Because the 1st stage may be lacking adequate power, the decision as to which group to drop/modify may be often based on precision analysis. In the final analysis of the data, methods exist that allow the combination of p-values of endpoints from various stages of the study.^{7,21}

The regulatory approaches for such designs are discussed in more detail in the discussion of the EMA’s reflection paper in [Section F](#).

4. Randomization-Adaptive Designs

Certain study designs may also modify the randomization algorithm on the basis of covariates (covariate adaptive randomization designs) if certain patient characteristics appear to be important in the response, thus balancing out prognostic factors.²² These approaches maybe valuable in studies with relatively small number of patients in which the study groups may not obtain the appropriate balance in factors deemed capable of modifying response to treatment.

E. Utilization of Bayesian Statistics

Certain of the approaches discussed above are based on the “frequentist” approach to statistics, in which hypothesis testing is based on data collected directly from the sample tested in the clinical study. Model-dependent/Continual Reassessment designs (discussed in [Section D.1](#)) and certain multistage designs depend of Bayesian statistics. Bayesian statistics allow the probability distribution of a given endpoint from previous information (earlier stages of the study, previous studies or other observations) to be combined with the probability distribution derived from the currently tested sample. Although it may be tempting or even intuitive to utilize previous information, there are substantial issues that may advise against such an

approach, (such as sample differences and methodology changes) that may introduce a bias in the interpretation of results. Thus, a lot of objections to the use of Bayesian statistics revolve around combining prior information with current observations to provide an “optimal” estimate. The obvious question here is that if that prior information is solid, why perform another experiment and if it is not, why undertake the exercise? In addition to this intrinsic problem, the computational and programming aspect of Bayesian statistical analyses presents major problems. They have been based on individual approaches that present difficulties to third parties and regulatory agencies in the evaluation of their statistical properties. Overall, there are no generally accessible programmed solutions that meet regulatory requirements, a fact that explains the regulatory reticence in the utilization of Bayesian inference in pivotal studies.

Since CRM-based studies mainly driven by Bayesian inference address the sponsors “learning” requirements in early development, regulatory inhibitions about their use in that context are few. The FDA is attempting to redress the difficulty of evaluating the statistical robustness of programming Bayesian statistics. However, the agency’s timelines are not clear. In the progress report on the Critical Path Initiative (See [Section F](#)) for 2008,²³ the FDA reports that it is working with a “large statistical software company” under a cooperative research and development agreement to produce a commercially available software package that would allow the design and analysis of Bayesian-driven clinical trials. This package should have commenced beta testing at the end of 2009.²³ The EMA is also examining the use of Bayesian methodology. In the final report from the EMA/CHMP think tank on innovative drug development, there is a commitment to investigate such methods and possibly render them usable and evaluable by the agency.²⁴

F. Regulatory Agencies and their Response to Clinical Study Innovations

1. The FDA’s Critical Path Initiative and Design Issues in Clinical Studies



The FDA initiative is called the [Critical Path Initiative](#) and it was launched in March of 2004. At its launch, the FDA issued a report,²⁵ stating that the launch of this initiative was an imperative because of the slowing pace of development of new agents. It highlighted the low productivity, rising costs of development, increased risk and a higher failure rate. It also stated that predicting success was very difficult with current tools, an issue discussed extensively in a previous article. The FDA estimated that a compound entering clinical research has only an 8% chance of reaching the market, substantially lower than the 14% chance that has been historically established before the first decade of the 21st century.²⁵ In the launching of the initiative, the FDA promised

to bring its expertise and accumulated data to bear in the development of a new toolkit for drug development.

The Critical Path Initiative essentially aims to create a number of technologies that would facilitate drug development. These “integrated technologies” consist mainly of developing better animal models of disease; extending information on useful biomarkers that can provide dependable information on disease progress and potential clinical benefit and may make earlier phases of development more predictive of the final outcome; pharmacogenomic information that can predict response to the drug; new clinical trial designs and statistical capabilities; and improved quality assessment tools. A report by the FDA at the end of 2008 highlighted its progress in a variety of these areas.²³

a. The FDA’s View of Adaptive Designs

A more detailed critique of adaptive clinical trials design was published in 2006, authored by members of the FDA’s Division of Biometrics and Office of Biostatistics.²⁶ In that paper, the authors stated several of the areas of uncertainty introduced by adaptive designs and in some cases questioned the need for them. A very strong argument was made that if the criteria for the planned adaptation are carefully defined *a priori*, a more statistically efficient “fixed” design may be available.¹⁷ and that the estimate of the treatment effect at the interim may be inadequate in providing guidance for sample size re-estimation.²⁷ The authors also noted that in study designs in which an arm of the study is dropped, reallocation of the alpha to the remaining arms may lead to a substantial inflation of the type I error. They caution that, in such a case, either the reallocation of the unused alpha should not be attempted or the planned sample size of the terminated arm should be distributed to the remaining treatment arms.²⁸ The authors also addressed the possibility of changing from superiority to non-inferiority. They concluded that no adjustment to the alpha is necessary if superiority and non-inferiority are tested with the same confidence interval for the treatment effect. However, they strongly advise setting and justifying the non-inferiority margin prospectively, and warn that a non-inferiority margin derived from interim data is not going to be interpretable. In terms of changing the primary endpoint at the interim, the FDA authors conclude that this approach, however valid the statistical test employed, probably has no advantages compared to a “fixed” design with multiple primary endpoints in which the alpha is allocated by a Bonferroni adjustment.##

In a Bonferroni adjustment, the alpha is divided by the number of primary endpoints and then each endpoint is tested against the adjusted alpha. However, other adjustment methods exist for dealing with multiple primary endpoints such as the Hochberg approach.

It was also emphasized that adaptive designs leave unclear which point estimate for the treatment effect can be reported in the product label. This appears to be a serious issue for the authors. In addition, the authors wonder what happens to the validity of multiplicity adjustments for secondary endpoints after a number of adaptations. They also present a number of case studies to highlight issues with adaptive designs and conclude with a number of logistical concerns.

Gallo and Mauer²⁹ have provided a response to this paper on behalf of the industry. They defend the overall concept of the adaptive designs because they claim that studies should not necessarily be held hostage to initial assumptions that may be proven wrong during the conduct of the study or to unfavorable chance events at an interim of a non-adaptive group sequential design. The authors also defend the estimates for the treatment effect obtained by adaptive designs, claiming that many methodologies of fixed designs, such as carrying forward the last observation or worst-case imputation, also modify such estimates included in product labels.

b. The FDA Draft Guidance on Adaptive Designs

From the onset of the Critical Path Initiative, the FDA has indicated that it is working on a number of guidances and has highlighted some issues and opportunities in presentations by its senior personnel. Members of the FDA team have presented outlines of a draft guidance in recent meetings.³⁰ The outline did not exhibit any substantive departures from the opinions voiced by Hung *et al.*²⁶ It stressed that the field is in evolution and that a final guidance, when introduced, should be understandable by a wide audience. The presentation of the early draft focused on concerns regarding the integrity of the study following the interim analysis and the difficulty of interpreting results from studies utilizing adaptive designs. Also, very much like the EMA, the FDA is concerned that adaptive designs may be used to shortchange early development, thus expanding the use of an experimental agent well before there is adequate safety information to allow such a step.

In February 2010, the FDA produced a relatively lengthy draft guidance which amplified on concerns that were expressed by the agency in presentations to various pharmaceutical development groups and the draft outline of the guidance discussed above. We have already referred to the FDA definition of a “flexible design” as worded in the draft guidance in [Section C](#) of this document. The draft guidance is certainly lengthy and many of the concerns lack specific examples and concrete suggestions. However, the document provides a good summary of areas of regulatory comfort and concern and does offer a number of suggestions to sponsors that may

facilitate the process of review and acceptance of development plans that utilize adaptive designs.

Two items of concern are repeated numerous times throughout the document: the control of the Type I error rate and the introduction of biases (operational or statistical) into elements of the study. The FDA also poses a number of interesting questions in its statistical considerations: when a number of adaptations are possible after the interim analysis that increase the possibility of success, how does one account for the multiplicity of choices in a statistically meaningful manner? The FDA does not provide an answer, but the question highlights its overall sense of discomfort with these designs.

The areas in which the FDA draft guidance provides some clear instructions are the following:

- (a) Adaptations that are well understood with “valid approaches to implementation”: These include mostly designs in which adaptations are based on blinded interim analyses and classical group sequential designs with well-defined alpha spending adjustments)
- (b) Adaptations that are not as well understood: These include almost all the adaptations discussed in [Section D](#) of this document. The draft guidance is not summarily dismissive, but it is quite clear that the onus to prove that these designs do not introduce biases and control the Type I error are solely on the sponsors.
- (c) Safety considerations: In exploratory studies, the FDA apparently favors less aggressive, hybrid designs than the pure continual reassessment methodology ([Section D.1](#)). In pivotal studies that commence without a full Phase 2 program, the FDA expects more frequent and more comprehensive safety monitoring during the conduct of the study/
- (d) Documentation that should be included in the study protocol: The proposed documentation is really to assist the agency in assessing the validity of these protocols. In this process, the FDA strongly recommends performing study simulations that should be fully documented in the protocol. In addition to documenting the statistical considerations and simulations, the “enhanced” protocol needs to contain a full description of the study teams, data management and steering committees and other personnel, their charters, responsibilities and the “firewalls” in place to prevent information dissemination
- (e) Documentation required to protect study blinding and information sharing: The draft guidance stresses the need for detailed SOPs for adaptive clinical studies, written minutes of every committee meeting, careful documentation of all considerations for design choices and

maintenance of all interim analysis databases. Although not mentioned until study reporting, the firewalls must be assessed for effectiveness and this assessment must be documented (although there are not specific guidelines for this assessment)

- (f) Interactions with the FDA: In effect, the FDA wants more frequent and more substantive interactions with sponsors utilizing adaptive designs and it would be ready to grant a number of Type C meetings to discuss potential issues, especially in “innovative products” and areas of “unmet clinical need”. Because the regulatory agencies are not involved in the interim adaptations, it is made clear that Special Protocol Assessments (SPAs) would be less “binding” in adaptive designs.
- (g) Documentation for reporting a completed study: The FDA states that more information is required than what is covered by the ICH E3 (Format and Content of Clinical Study Reports), although in reality most of the elements discussed are covered by the provisions of this guidance. However, the agency expects that all records by the committees involved in the study conduct to be submitted with the report along with results, datasets and programs for all interim analyses. It also specifies that the analysis of the study must include an exploration of heterogeneity of study population in different parts of the study as well as outcomes. It does accept that the power for these comparisons may be limited, but it expects the sponsors to amplify them but no specific methodology is suggested.

The provisions of the FDA draft guidance are provided in greater detail in [Appendix 1](#).

2. The EMA Initiative

The EMA was also alarmed with the slowing pace of innovation in drug development. Its approach is similar in broad outlines although there are differences in detail, foundational philosophy and implementation.^{§§} The EMA’s response to adaptive clinical trial designs and other issues of pharmaceutical development is an element of the Innovative Medicines Initiative (IMI) and its Strategic Research Agenda (SRA).^{***} In the process of forming a response, the EMA



^{§§} The FDA has formulated the CPI initiative and drives its implementation. The EMA pursues its agenda in partnership and collaboration with industrial groups

^{***} IMI resulted from a partnership of the European Commission’s with the European Federation of Pharmaceutical Industries and Associations (EFPIA).

assembled a think-tank group that consulted both academia and corporations on ways for speeding development; this group also examined the methods by which the agency may reorganize and modernize itself in order to streamline drug development. The findings of the EMA/CPMP drug group were published in 2007. Some of the findings echoed positions enunciated earlier by the FDA on the development of biomarkers and flexible study designs.

a. The EMA's Reflection Paper on Adaptive Designs in Confirmatory Studies

In October 2007, the EMA also released a reflection paper on adaptive (flexible) designs in confirmatory clinical trials. Since this is the first document officially released by a major regulatory agency on this issue (although it is not a guidance), it is worthy of a more detailed examination. It should be pointed out that the "Reflection Paper" is very similar in concept and occasionally utilizes the same language as the paper by Armin Koch of German drug regulatory authority which was published in 2006.³¹

The reflection paper³² mostly centers on considerations for studies with planned interim analyses and outlines practices that may be acceptable for a positive review of an application. It does not discuss statistical approaches in detail apart from general statements regarding the control of type I error, makes no mention on the acceptability of Bayesian statistical methods in these studies and reiterates a number of pre-existing guidances which it regards as still applicable to clinical trials with flexible designs.

The guidance places a lot of emphasis on the confidentiality of the results of interim analyses, a common thread in the concerns of regulatory agencies. The EMA apparently feels that the danger of compromising the study is substantial and insists that the need and the number of interim analyses should be carefully justified. The agency would need to be convinced, for example, that interim analysis for sample size recalculation is not undertaken because of the insufficiency of earlier studies. Thus, the EMA states that flexible designs should not be utilized as a method of substantially reducing the "learning" phase of clinical development. The inherent dangers of the approach can only be overcome by careful reasoning that the agency can accept. Thus, extensive consultation on this point would be crucial. The agency has further concerns about the possible bias being introduced by interim analyses. The reflection paper makes clear that an analysis of the data prior to and following the interim analysis would be necessary to show that the homogeneity of the study has been maintained. Obviously, this may be possible with adaptive group sequential designs but not with multi-stage adaptive randomization designs which, by definition, introduce heterogeneity. Because of the possible introduction of bias by the interim analysis, the reflection paper is rather negative on the introduction of more

than one interim analysis. The agency assumes that the need for more than one interim analysis indicates that conditions of the study fluctuate far more than it is acceptable for a confirmatory trial.

The EMA discusses the possible consequences of stopping a given study “early”. Many in the development community would take exception to the term “early”³³ but the agency is apparently concerned that an “early” discontinuation may compromise the safety data expected to be collected. Since the EMA feels quite strongly on this issue, the best approach would be to contemplate the first interim analysis only when at least the minimum number of patients for safety determination has been accessed as per ICH guidelines. In conclusion, the agency specifies that “interim analyses without realistic objectives should be avoided” but it fails to define what these realistic objectives may be. My guess is that these “realistic objectives” of the interim analyses would need to be decided on an individual basis in consultation with the agency and early stopping of clinical studies should be undertaken with great caution. If an “early” stop to the study has occurred, the agency is very clear that it wants to see two analyses: the analysis of the data collected at the interim analysis stage and an additional one in which the patients that were accessed and treated after the commencement of the interim analysis are included. Any discontinuities between these analyses may present problems at the review stage. It would all depend on the magnitude of difference between the interim and final analyses although no specific guidance is offered. The EMA is apparently convinced that interim analyses, on average, over-estimate the true treatment difference and this should be kept in mind when providing a rationale for them to the agency.

The EMA has a variety of reservation regarding adaptations. They stem from the belief that the need for such adaptations spring from incomplete prior development and errors on assumptions. Thus, the reflection paper states that re-assessment of sample size may also indicate that many of the assumptions about the design were simply “erroneous”. In the case of non-inferiority clinical studies, the non-inferiority margin must be re-evaluated and re-justified if the sample size has been re-estimated.

In planning the sample size, the treatment effect should be well defined prospectively and the agency would apparently frown upon justifying a treatment effect as clinically beneficial on the basis of interim results. This is in line with the EMA’s strong concerns about changing the main endpoint of the study at the interim result stage. In the reflection paper, the agency authors makes clear that the main endpoint in pivotal studies is selected on the basis of its clinical benefit and not on the basis of displaying differences between groups. Although it does not provide details, the EMA states that

rejection of null hypothesis based on results from different endpoints in multistage designs is unacceptable.

The agency appears to harbor strong reservations for designs that discontinue study arms, especially the placebo one. The reflection paper notes that study populations may vary at different stages, depending on the inclusion or exclusion of a placebo arm in these stages. It is clear, however, that discontinuing certain ineffective doses of the test drug may face less opposition than discontinuing controls or the placebo. The agency clearly indicates that it much prefers studies with unbalanced randomization to studies that discontinue treatment arms if the approach is statistically robust. Thus, in a multistage trial design contemplated to support registration in the EU, it may be necessary to continue enrolling patients in the 2nd stage in the placebo and control arms in an unbalanced randomization scheme..

As a note of caution for arm discontinuation in multistage designs, the reflection paper also clearly indicates that in studies with more than one dose it would not be sufficient to show that some dose of the drug (combining all drug doses) is effective. The selected dose should achieve this aim on its own. In addition, in a multi-stage setting, only data from the treatment that has gone through all stages would be acceptable as part of the label claim, even if arms discontinued in earlier stages showed superiority against placebo.

The EMA repeated its typical guidance in switching between superiority and non-inferiority. However, it is rather hostile to a study design that proves non-inferiority at the interim and continues treatment to show superiority. The agency prefers two independent non-inferiority studies the results of which may be combined in a metanalysis to prove superiority.

The reflection paper is negative on Phase II/Phase III combinations, if these studies are the sole element in support of a marketing authorization. The draft guidance flatly states that such studies are not going to be acceptable for filing purposes and should be used to investigate correlations between surrogate endpoints and to define the optimal dose regimen.

G. Operational Considerations

In classical fixed study designs, data are examined after the study has been completed, all data have been gathered and the blind removed. By their very definition, this is not the case with adaptive designs. As the study is planned to be modified on the basis of accumulated data, a number of operational considerations must be taken into account in order (a) to collect the data within an appropriate time frame and (b) to maintain the integrity of the study while examining that data, especially if this requires unblinding. Both of these efforts place a substantial

burden on the conduct of the clinical study and the obligations of the sponsor, the sponsor's agents and investigative sites.

In all studies in which interim analyses are scheduled, speed and accuracy of data collection is imperative; otherwise the delay in this step can be substantial. The pressure for timely and accurate data collection and auditing is even more pronounced in methods that require continual reassessment in model-dependent adaptive designs. In addition, extreme care should be taken to assure that the both the sponsor and the investigative sites remain blinded as to the group assignments and do not bias the study in its later stages. This may involve the removal of the sponsor from the committees that evaluate the data and render decisions to proceed or not with the study. These committees may include a steering committee and/or a data safety monitoring board (DSMB) with well specified responsibilities and appropriate charters. If the sponsor assigns employees to support these committees such as a biostatistician, programmer and/or data manager, and if the data accessed reside in the sponsor's databases, then the appropriate "firewalls" should be put in place and they should be copiously documented. It is obvious that for adaptive designs to be efficient, the main endpoint and other essential data should not require an excessively long period of time to be collected, otherwise treatment in stage 2 would be substantially delayed.

In response adaptive clinical study designs, additional considerations may apply. If the design addresses the possibility that the drug is successful only in subset of the population tested, these subpopulations should be carefully constructed and "nested." The appropriate sample size for each subpopulation tested at the interim analysis stage should be defined.

The operational demands of adaptive designs act as a barrier for their adoption by the organizations that may actually need them the most: the small biotech companies. In adaptive group sequential/sample-size re-estimation designs in which subject numbers may be increased substantially from the starting estimate, securing funding for the largest feasible sample size may be just too difficult. In multistage designs, the "winner" may be a smaller section of the population than originally envisaged, thus undermining pre-existing funding structures. But beyond these general considerations, the everyday operational requirements of studies with adaptive designs are too demanding for the organizations of medium to small biotech companies. The much greater emphasis in timely and accurate data collection and dissemination of information to all stakeholders, supervision of enrollment, constant examinations of the validity of the database, the maintenance and documentation of firewalls may be insurmountable obstacles for companies that are "virtual" in many of their functions and for which continuous monitoring of processes and regular training is beyond their means and beyond the expertise of their personnel. The CRO and sponsor "organizational distance" imposes limitations on information flow and agile study and data management. In addition,

devolving important decisions regarding the study to independent committees may be beyond their capacity of senior management of small biotech companies to accept. One may take the cynical view that with all the inefficiencies built into the whole R&D effort of small companies as they attain “focus”, adaptive designs are probably the least of their problems.

In larger pharmaceutical companies in which all the organizational pre-requisites exist and risk is well-apportioned to a large number of compounds, adaptive clinical studies may provide substantive money and time savings. Progressively, as these designs mature and experience in dealing with them increases, their adoption by smaller companies may be easier.

H. Recommendations

1. The Planning Stage

So, how does one proceed with a study based on an adaptive design? The answer to this question certainly varies with the capabilities of each organization. First and foremost, all elements of the study must be carefully understood and accounted for in the planning stage. Any deficiencies in the plan can spell disaster later on.

A full examination and possible revision of SOPs should be undertaken to make certain that no gaps exist and that there will be no deficiencies in compliance when dealing with management of data (blinded or unblinded) of an ongoing study. The FDA draft guidance on adaptive designs strongly recommends new SOPs for adaptive design studies; it is possible, however, to accommodate the requirements of these studies in existing SOPs. As one may want SOPs to remain somewhat general, it may be appropriate to construct a number of best practices to address information access and flow, responsibilities, firewalls and other operational requirements. In this context, both the monitoring and the data management plan should go into details and include a full risk mitigation plan that takes into account all eventualities. It is also imperative that the clinical study team should undertake the effort of “educating” major stakeholders of the corporation to the issues that may be encountered.

What should also be addressed in the planning phase and should be fully in place prior to the beginning of the study is the composition, charter, membership and of a DSMB and/or a steering committee for the study, if the study requires decision making by an independent entity. If needed and not previously employed, SOPs for the DSMB should also be compiled and all its members should receive adequate instruction in them. The information flow between the corporation and the independent committees should be well regulated on the basis of both the DSMB/steering committee and the sponsor SOPs. All

communications should be fully documented to assure regulators that study integrity was not undermined.

The study feedback loops should also be planned at this stage. Beyond the typical feedback loops that consist of site monitoring and independent audit reports, additional ones must be planned such as:

1. Metrics that assess study compliance (study deviations and violations) against specific benchmarks
2. Metrics that assess monitoring/management effectiveness (number of unresolved issues, recurrence of “corrected problems”, etc). Again, these assessments should be made against certain predetermined benchmarks. Failure to perform according to benchmarks should initiate corrective actions immediately

During the conduct of the study, these feedback loops should be continuously assessed to make certain that errors are kept to minimum, information flows as planned and no bottlenecks exist. It is very important that the study manager should be in control of the feedback loops and that the management structure is as centralized as possible. A diffuse decision making in such a program can spell disaster. Since the Stage I of all adaptive studies is crucial in further decision making, getting it right from the very beginning is imperative.

Maintenance of the blind is crucial. It is thus important to design at this stage a system of restrictions (firewalls) for the dissemination of information from interim analyses. The FDA draft guidance requires that these firewalls are assessed for effectiveness and the assessment is documented. Each R&D department must devise effective protocols for the assessment of firewalls before the study begins.

In certain cases, maintaining the blind may be challenging because of differences in test/control drug volume sizes, expected topical adverse effects, anticipated typical post-administration AEs, etc. Efforts beyond the usual should be undertaken to mask these differences during the conduct of the study. If the maintenance of the blind is expected to be present challenges, then the plans to proceed with an adaptive design should be reappraised.

2. The Conduct Stage

Since a number of decisions taken after the interim analysis would be crucial to the success of the study, the quality of the data at the interim must be high. The number of protocol deviations and violations should be kept to a minimum. Although this is a good advice in general for all clinical studies, it has a special urgency in adaptive designs. Usually, the rate of violations/deviations improves as the study goes on. Continuous corrective actions by study personnel improve the education and subsequent compliance of the investigative sites and

ineffectual or non-compliant investigators are removed. Unfortunately, in designs heavily dependent on an interim analysis, the margins are far tighter. Corrective actions should be as timely as possible and information about them should be quickly disseminated to all sites. Because quality data at the interim stage is so crucial, it is suggested that monitoring visits (and, if possible, independent audits) are conducted as soon as the first group of patients (no more than three) is enrolled and treated at each site. This enhanced monitoring should reveal any systemic or site-specific problems before they affect the majority of the data.

The study management team should be “at the top of its game.” All efforts should be expended to make certain that only the proper patients enter into the study, all tests and assessments are performed and completed within the appropriate time windows and that the protocol-mandated treatment algorithm is adhered to. Too many missing data and too many deviations at the interim may seriously bias the adaptation decisions.

In order to achieve high quality of data, a stable, well-trained and well-motivated study management and monitoring team is required; Teams with decentralized personnel/management structures and high turnover rate may be inappropriate for such an effort. Of course, the investigative site education effort should be a thorough, ongoing and unremitting process.

3. The Imperative of Full Documentation

It is imperative that all communications with the sites be documented and reviewed in as detailed a fashion as possible to make certain that any introduction of bias by the sponsor or any independent participant does not occur. Detailed standard procedures and thorough documentation of interactions with sites may avert any suspicion by regulatory bodies of inadvertent introduction of bias if some “peculiarities” are detected in the data. Rosters, attendance records and detailed minutes of all meetings of the various independent or sponsor-connected committees (i.e., DSMB, Steering Committee, Clinical Study Team, etc.) should be maintained. In fact, the FDA draft guidance on adaptive designs requires that these documents become essential elements of the clinical study report.

The following example highlights such a case of missing documentation and the problem that it creates: Koch³¹ examined a study in which three different stages were discernible (the first one was up to the interim analysis). The experimental treatment achieved virtually identical results in all three stages, while the efficacy of the standard treatment declined considerably from stage to stage (the decline was statistically significant). Koch stated that such discontinuities would raise concerns among regulators, unless “full reassurances exist that the treatments are **fully** (my emphasis) blinded to patients as well as observers.” Gallo and

Mauer,²⁹ in their reply to Koch stated that within-study “drifts” may “occur naturally” possibly because of the experience gained by investigators during the trial or because of “natural shifts” in patient population. I find this reply totally unconvincing, because if the changes in efficacy were caused by the investigators gaining experience, they would have affected the experimental rather than the standard treatment. The reverse actually happened. In addition, I am not sure what is “natural” in data drifts and patient populations and I think that the regulatory authorities would have the same uncertainty. A full documentation of what has actually transpired throughout the study can only assist sponsors when such suspicions are raised.

I. Conclusions

The allure of adaptive designs in clinical studies is based on the promise of lowering costs, speeding up development and reducing attrition rates. However, many of these promises depend on a flawless implementation of very complex procedures and methods of analysis that may not be well understood. There is skepticism and caution within large regulatory agencies regarding their large scale adoption. The field is in evolution and the regulatory agencies are slowly responding to the challenges that such designs pose. Their proponents, such as various CROs and biostatisticians, may be overselling their benefits, at least at present.

The complexity of implementation as well as regulatory caution essentially assures that progress in the field will be slow despite the excitement that has been generated in the last few years. Also, the challenges of implementation (and possibly of conception) make it obvious that these designs will be utilized mostly within the confines of large, well-funded and well-staffed pharmaceutical companies. Unfortunately, these have not been the well-springs of productivity in research as of late.

At this time, it appears that the regulatory agencies will have little problem with adaptive studies in the exploratory studies of Phase 1 or 2, although they may have safety concerns with “aggressive” model-dependent/continual reassessment methods (the FDA favors “hybrid” designs in this case). In any case, as long as the ethics guidances are met, it is really the responsibility of sponsors to “learn” at this stage of development. In fact, efforts for dose finding at these stages may provide a good test for a variety of adaptive designs and the opportunity to discover if these methods do provide more accurate information for a “go/no go” decision.

As usual, the regulatory agencies review early stage information at the end-of-phase 2 meeting (or equivalent) and its adequacy (or lack of it) shapes the regulatory feedback on the pivotal phase designs. Both the EMA and the FDA have sounded strong warnings to sponsors about proceeding with adaptive designs in the pivotal phase without an adequate safety database after the Phase 2 program; the FDA thus calls for more frequent and extensive safety monitoring in these cases. Regulatory

bodies may have a much easier time accepting adaptive designs in confirmatory studies, if surrogate or pharmacodynamic/ biomarker endpoints have been used in Phase 2. In such circumstances, there would always be a good case that the assumptions regarding the clinical beneficial endpoint may not be precise and an adaptive design (such as sample size recalculation) would be the best way of addressing this uncertainty.

For those planning the clinical development of drugs and biologics, it is important to consider that, because of regulatory caution, substantive departures from the paradigm of “two well-designed confirmatory studies” with fixed designs should be adopted (a) only when the rationale is sound and has been fully accepted by regulatory authorities and (b) when there is the organizational capability to support such an effort. The same note on caution applies to contemplating and planning confirmatory studies with adaptive designs without compiling an adequate safety experience in the early phases. For well-funded corporations with adequate personnel and organization, adaptive clinical trials in early phases may save money and move the decision process faster and better than classical designs.

J. Keywords

Pharmaceutical Development - Clinical Trial - Failure Rate - Clinical Trial Design - Adaptive Design - Group Sequential Design - Phase II - Phase III - Bayesian Statistics - Classification - Recommendations - FDA - EMA - Critical Path Initiative - Innovative Medicines Initiative - Reflection Paper - Regulatory Guidance - Good Clinical Practices - GCP - Protocol Violations

K. Appendix 1: The [FDA Draft Guidance](#) in Detail

Introduction

This is a brief introductory section about the rationale for this guidance (to assist sponsors) and a general statement about the legal framework for FDA guidances (not legally enforceable)

Background

There is little information here about the progress and challenges of adaptive designs. Most of the section provides a rationale for the compilation and structure of this guidance

Description and Motivation for Adaptive Designs

This section provides the FDA’s own definition of an adaptive design. This definition, discussed briefly in Section C, is as follows:

- An adaptive design clinical study is defined as a study that includes a prospectively planned opportunity for modification of one or more specified aspects of the study design and hypotheses based on analysis of data (usually interim data) from subjects in the study. Analyses of the accumulating study data are performed at prospectively planned time points within the study, can be performed in a fully blinded manner or in an unblinded manner, and can occur with or without formal statistical hypothesis testing.

The draft guidance also includes in this section a number of terms and explanations that for the most part, are well established in clinical research and do not require any extensive discussion. The guidance includes the term “adequate and well-controlled” (A&WC) effectiveness studies for clinical trials that would be typically referred to as pivotal or Phase 3. The term “exploratory studies” is utilized when very strict control of the Type I error is not required or when the endpoints used are not clinically relevant (even if the studies are adequate and well-controlled). The FDA dispenses with the terms “seamless designs” and “Phase II/III studies” because they do not provide useful information, very much along the lines of the argument in Section D of this document.

In addition to definitions, this section of the draft guidance goes into some extent about the “potential for counterproductive impact” by adaptive designs. These include again the regulatory concern of speeding through development without “identifying the gaps to knowledge” and the “eliminating the time to thoughtfully explore study results”. The FDA also makes the case that in “seamless” designs that fused an exploratory elements at interim analysis (i.e., an expanded dose-response study) followed by an A&WC pivotal study (i.e., testing only one or two doses vs. control), the aggregate would be regarded as a single study, not an independent replication of the primary hypothesis testing. Thus, an additional study would be required for verification and a marketing application.

General Concerns Associated with Using Adaptive Design in Drug Development

In the draft guidance, the FDA identifies its two major concerns regarding adaptations in clinical studies. These consist of (a) poor control of the Type I error rate and (b) difficulty in interpreting results even if the Type I error rate has been rigorously controlled. The agency has also substantial concerns about the potential of operational bias that may be introduced by unblinded interim analyses.

In controlling the Type I error rate, the FDA favors adaptations possible after blinded interim analyses.. However, in cases in which the analysis is unblinded, the draft guidance provides mostly warnings and not many constructive suggestions. It states that because of the multiplicity of adaptations possible at this point, the opportunities to succeed increase but this multiplicity of opportunity is “difficult to understand and account with statistical adjustments”. It would appear that the FDA would entertain proposals, involved in substantive discussions especially in

the end-of-phase-2 meetings, but sponsors may also want to limit prospectively the “multiplicity” of adjustments to improve the acceptability of their designs.

Regarding the “difficulty of interpreting results”, it appears that the FDA has two main concerns: one relates to the “point estimate”. The FDA has concerns that, despite the presence of confidence intervals around the point estimate, adaptive designs may actually affect it substantially, making it difficult to evaluate risks and benefits. The other concern relates to the complexity of revisions although concrete examples are missing. Based on the sense of FDA’s discomfort on these complex designs, a good approach consists of utilizing simpler designs for A&WC studies, at least until the FDA becomes more comfortable with more complex situations.

Within its concerns about adaptive designs, the guidance discusses “operational bias” and includes proposals for documentation of interactions between various teams and committees responsible for the study conduct. The FDA’s concern regarding “contamination” of the study by information obtained in interim analyses has been echoed in various forums and it is examined also in [Section G](#) of this document.

The problems with Type I error control, result interpretation and operational issues do not conclude the FDA concerns about adaptive designs. The draft guidance also notes that certain A&WC studies utilizing adaptive designs require a set of assumptions that may be based in inadequate prior information. Thus, the potential of failure may increase. This is definitely an issue that sponsors should carefully consider, although inadequately supported assumptions are hardly absent in fixed designs. The draft guidance also mulls the issue that “seamless” designs may remove the time for careful and thoughtful examination of the data (an issue also raised by the EMA). There are additional concerns that are discussed in some detail below.

Designs that are “Well Understood with Valid Approaches to Implementation”

In this section of the guidance the FDA lists what it terms “well-understood adaptive designs with valid approaches to implementation”: These include the following:.

- Adaptation of Study Eligibility Criteria Based on Analyses of Pretreatment (Baseline) data: Because the pretreatment data do not involved in the efficacy analysis, there are no substantial objections in the use of this adaptation
- Adaptations to maintain study power based on blinded interim analyses of aggregate data: The main reason to use such an adaptation is to calculate more accurately the event rate (or time to event) and the variance on which the assumptions for the primary endpoint are based on and adjust power accordingly. The draft guidance states that such designs may also examine subsets of patients in which variance is lower or event rate are higher and adjust

the entry criteria for the study. The draft guidance specifically encourages the use of this adaptive methodology because it is unlikely to introduce bias if correctly applied. Adaptations based on interim results of an outcome unrelated to efficacy: This is an adaptation that allows the removal of dose groups for higher than anticipated toxicity following a blinded interim analysis that does not include any efficacy parameters. However, for this adaptation not to introduce bias, the toxicity must not be connected to increased efficacy because it would then result in an uninformative risk/benefit assessment. Utilizing such an adaptation presupposes that prior to the study, the lack of connection between the toxicity to be examined and efficacy has been established.

- Adaptations using group sequential methods and unblinded analyses for early study termination because of either lack of benefit or demonstrated efficacy: We briefly discussed the classical group sequential designs in Section D.2. These have been a standard fare of clinical research for some time and the alpha-spending methodology for these has been well established. It is rather surprising that they are introduced in this draft guidance because their inclusion in adaptive designs is questionable. The ICH E9 also discusses these designs, something that this draft guidance notes. It is possible that this category may be removed from the final guidance because the draft simply repeats existing guidance and concerns regarding introduction of bias following the unblinded analysis.
- Adaptations in the data analysis plan not dependent on within study, between-group outcome differences: The draft guidance here refers to changes in the statistical analysis plan (SAP) that may be introduced following unblinded examination of the data if specific assumptions of the SAP (value distribution, missing values, etc) do not appear to conform closely to those observed in the study. Again, the ICH E9 discusses this “adaptation” as well. However, the FDA draft guidance certainly allows for a prospectively-defined analysis that may modify or change the primary endpoint of the study. It is again important to stress that any analysis that may change the primary endpoint must be blinded.

Designs Whose Properties are Less Well Understood.

In this section, the FDA lists designs that would require the enhanced processes and procedures discussed later in the document, including expanded interactions with the agency. Unfortunately, these include most of the currently discussed adaptive designs included in Section D of this document. In summary, these designs include:

- Adaptation for dose selection studies: This category includes mostly study designs discussed in [Section D.1](#) (Model-dependent/Continual Reassessment Designs). However, the draft guidance does not indicate any substantial objections in the use of this methodology in exploratory studies. It mostly refers to the possibility of taking more than two doses into an A&WC clinical studies.

The authors of the draft guidance make the point that in many pivotal studies, the dose (or doses) selection is based on inadequate information which increases the likelihood of failure. Increasing the number of doses tested in these studies with an adaptive methodology that would remove ineffective doses after interim analysis, combined with a robust method to control for the Type I error rate maybe desirable.

- Adaptive randomization based on relative treatment group responses: This category refers mostly to designs discussed in [Section D.3](#) (Group Sequential: Response-Adaptive Designs) of this document. These designs increase the possibility of success by assigning more patients to groups with more successful outcomes (play-the-winner) or removing groups with less than desired response. The draft guidance notes that such approaches have the potential of undermining the balance of the groups in terms of patient characteristics and thus introduce substantial bias in the study. The document recommends that enrollment in the placebo groups should be maintained at an appropriate level to alleviate unbalancing concerns. Also, analyzing the groups versus placebo at different stages of the study would also ease concerns regarding bias.
- Adaptation of sample size based on interim-effect size estimates: These designs have been extensively discussed in [Section D.2](#) (Group Sequential/Sample-Size Re-estimation Designs). of this document. The draft guidance states that such designs should be used to increase the sample size, not to decrease it. Decreases in sample size should be achieved with well understood classical group sequential designs. The major concerns here are based on the adjustments in the final study analysis to compensate for the increase in the Type I error rate after an unblinded interim analysis. The FDA notes that these methods depend on decreases in alpha or in differential weighing of parts of the study or combination of both. The draft guidance indicates a certain discomfort with differential weighing, especially if it is disproportionate to the number of patients in each part of the study. The FDA repeats the often-stated concern that modifications in study sample size after interim analysis of relatively small number of patients may lead to errors and the approach here should be as conservative as possible.
- Adaptation fo patient population based on treatment-effect estimates: The draft guidance refers here to designs summarized in [Section D.4](#) (Randomization Adaptive Designs). These methods identify patient characteristics (covariates) that may possibly improve outcome and enrich for these. The draft guidance states that these designs pose challenges and generally call for statistical adjustments to avoid an increase in the Type I error rate but offers no specific proposals. Section VII of the guidance provides input regarding statistical considerations in the “less well-understood” adaptive design methodsl

- Adaptation for endpoint selection based on interim estimate of treatment effect: This adaptation is a subset of designs discussed in [Section D.3](#) (Group Sequential/Response-Adaptive Designs). The FDA agrees that on several occasions the optimal endpoint for assessing the drug response may not be well understood; thus a change in primary endpoint after an unblinded interim analysis may be necessary. The point is clearly made that such designs may actually be preferable to fixed designs with multiple primary endpoints and alpha adjustments for multiplicity. The main issue here is the quality of data available at the interim analysis for each of the endpoints selected. Low quality data for certain endpoints may lead to erroneous decisions and compromise the study.
- Adaptation of multiple-study design features in a single study: This category includes again designs discussed in [Section D.3](#) in which more than one adaptation is considered after the interim analysis. Essentially, the draft guidance makes the point that more than one change may make the study very difficult to interpret. In addition, it warns that inter-dependencies among elements of the protocol may make the study more prone to failure if multiple changes are implemented.
- Adaptation in non-inferiority studies: The draft guidance sees little role for adaptations in non-inferiority studies. It makes the point that the treatment effect and the delta are usually defined from prior studies and it is thus inappropriate to change them after an interim analysis. In addition, the study population is always consistent with previous studies. Any changes to the study population would invalidate the non-inferiority margin. The only adaptation that the draft guidance regards as acceptable, is an increase in sample size to make certain that superiority can also be demonstrated, if this is desired.

Statistical Considerations for Less Well-Understood Adaptive Design Methods

This section provides broad suggestions to designers of adaptive studies. It does not provide any specific information as to what reviewers would like to see in protocols of studies utilizing adaptive designs. These are including in Section IX of the draft guidance.

- Controlling study wide Type I error: This section makes again the point that adaptations should be carefully defined prospectively and acceptability of the study depends on the control of the Type I error rate. The draft guidance warns that when there is a bias in the estimation of the treatment effect, designs that increase sample size may actually amplify the bias and result in a much higher Type I error rate than predicted.
- Statistical bias in estimates of treatment effects with study design adaptations: Since interim results may be quite variable, adaptations based on these may simply lead to erroneous choices. In addition, the method of combining the

analyses of the different parts of the study can introduce statistical biases. Substantial differences in the treatment effect between stages of the study may lead to difficulty in interpretation of the results and they should be critically examined in the final study report.

- Potential for increased Type II error: Although uncommon in such designs, the draft guidance makes the point that the possibility of an increased Type II error exists in certain designs, such as the ones in which several doses have been eliminated on the basis of relatively limited data after an interim analysis and the remainder are inadequately powered to detect a treatment effect.
- Role of clinical trial simulations in adaptive design planning and evaluation: The draft document recommends clinical simulation methodology based on Bayesian principles to assess the validity of certain adaptations
- Role of the prospective statistical analysis plan in adaptive design studies: For adaptive designs, the draft guidance includes the recommendation of preparing the statistical analysis plan (SAP) as soon as the protocol is finalized. It is proposed that the SAP should be more detailed for adaptive studies than it is the norm for fixed designs. Optimally, the SAP should be finalized prior to the first interim analysis. Later revisions will create questions regarding their impact and importance.

Safety Considerations in Adaptive Design Studies

- Safety of patients in adaptive design dose escalation studies early in drug development: The draft guidance echoes concerns with drugs with a substantial toxic profile being tested with model-dependent/ continual reassessment designs. Such designs reassign dose levels after each patient has been treated. For drugs with a possibility of serious adverse drug reactions, the FDA would prefer hybrids of these designs with more fixed methodology such as those discussed in Section D.1. In any case, the model for such designs and any simulations should be submitted to the FDA for comments prior to implementation
- Earlier design and conduct of A&WC studies with major expansion of treatment-exposed patients: This area of concern simply highlights the discomfort of the agency with “seamless” designs or other constructs (including group sequential/sample size re-estimation designs) that eliminate more extensive Phase 2 studies. In this case, the agency proposes enhanced safety monitoring but wants to amplify it with other procedures such as including a small number of patients and evaluating them thoroughly prior to full enrollment into the study. In any case, the guidance states that such designs that proceed “faster” than previously possible may have to be amplified with additional safety studies.

Content of an adaptive design study protocol

The draft guidance believes that the information included in the [ICH E3](#) guidance [“Structure and Contents of Clinical Study Reports”] regarding study protocols, is inadequate when adaptive study protocols are submitted for A&WC studies. This is somewhat perplexing, as the [ICH E3](#) does not include any specific set of protocol guidelines; references to the protocol content are only within the context of what should be included in the clinical study report (CSR). In any case, the FDA would like to see more detailed documentation explaining the adaptations, including a full analysis of the operational role of the Data Monitoring Committee (DMC) and a full description of each operational team in the study and their responsibilities.

In summary, the FDA proposes that protocols for A&WC studies should include:

- A complete description of objectives and design features
- A summary of each adaptation and its impact on the statistical analysis and hypotheses tested. To this goal, the agency encourages the use of a Bayesian framework that incorporates uncertainty in a quantitative manner (presumably, results of this framework should be documented and included in the protocol). Any models utilized should also be summarized clearly to allow their evaluation.
- Computer simulations that quantify the level of statistical uncertainty for each adaptation including the impact on the Type I error, unconditional or conditional study power, or any biases in the estimation of the treatment effect. If more than one adaptation is planned, the simulations should assess the combination of all proposed adaptations. Computer programs for simulations and graphical flowcharts of adaptive pathways including probabilities of their adoption should be included in the protocols. The FDA draft guidance goes on to provide summary descriptions of the quantitative models it wants to see included in the protocols
- Analytic calculations of the Type I error and any statistical biases in the estimation of treatment effect if obtained without simulations.

A full description of all personnel teams including compositions and charters (mainly the Data Monitoring Committee). Essentially, in this part of the protocol the FDA wants to assess the effectiveness of the “firewalls” to be utilized to protect the blinding of the study.

Interactions with the FDA when Planning and Conducting an Adaptive Design

In summary, the FDA expects to be substantially involved only in pivotal studies; It anticipates an enhanced number of interactions with the sponsor. In addition, it is fair to say that because of the FDA is not involved in the decision and implementation of adaptations, it sees limited value in Special Protocol

Assessments. In summary, the draft guidance expects that the FDA involvement would be along the following lines:

- Early and middle period of drug development: the draft guidance underlines that for exploratory studies, the FDA review is usually limited to the safety aspects of the study, but the agency, depending on workload, may be able to offer comments within the context of the Type C meeting. It is far more likely that the FDA will provide substantive input in innovating drugs and in areas of unmet clinical need.
- Late stages of drug development: For pivotal study designs that fall within those characterized as “less well understood”, the FDA suggests that sponsors do not wait for the “End-of-Phase 2” (EOP2) meeting to discuss elements of the study protocol. They are encouraged to schedule a Type C meeting or and EOP2a meeting and a subsequent EOP2 meeting. This schedule would allow time to consider if the proposed studies can be regarded as appropriate for submission. The document makes clear that since the FDA does not have access to the interim data, it cannot evaluate in real time if the adaptations chosen are in line with the agency’s thinking and the FDA’s approval of a design does not necessarily extent to the changes undertaken during its implementation.

Special protocol assessments: The FDA specifies that an SPA response for an adaptive design would likely require more than the 45-day period that the current guidelines specify. In additions, because the changes in design during implementation are the responsibility of the sponsor, it is made clear that the FDA response would include “limitations” (hedges) that do not usually accompany fixed-study designs. Thus, the SPAs for adaptive studies would have less of a value than SPAs for fixed-design studies.

Documentation and Practices to Protect Study Blinding and Information Sharing for Adaptive Designs

In the beginning of this section, the FDA discusses the documentation that should be available for preserving the operational integrity of the study. The FDA expects that the sponsors will issue specific SOPs for adaptive designs; it expects the SOPs dealing with the prevention of information dissemination/maintenance of the blind in an unblinded interim analysis to be very detailed. In addition to these SOPs, the agency is likely to require SOPs of how compliance with the adaptive study SOPs would be monitored.

The draft guidance also recommends that detailed minutes are kept for meetings each of the committees that are involved in the implementation of the study, its interim analysis and adaptation implementation. In addition, it expects that all data and analysis at the interim state should be preserved in a secure fashion.

The draft guidance directs certain “fire” towards CROs because “certain of them do not have long histories of carrying out these responsibilities”. Pointless to state here, but the same can be said for many small biotech companies with relatively new clinical development departments and a low level of expertise. Thus, the inclusion of these comments here betrays a certain “large pharmaceutical company” bias on the part of the FDA. This is really not that surprising as much of the impetus for adaptive designs has been provided by these companies.

Evaluating and Reporting a Completed Study

In this section, the draft guidance is getting more specific. It states that the ICH E3 may be inadequate for the purposes of reporting an adaptive clinical study and more detailed information should be included.

In fact, the FDA proposes the following in addition to the contents of the CSR as included in the ICH E3.

- Complete information regarding the planning of the study (although much of the items below can easily be accommodated in the Section 9 of a CSR (ICH E3)
 - Complete information regarding the study procedures including the DMC and other committee charters.
 - Supportive information that was utilized in the planning of the study including supportive information submitted to the FDA in meetings and interactions prior to the NDA.
 - Detailed rationale for using an adaptive design and the role of this study within the total development plan (although one would surmise that this information would be included in the material pertinent to the planning of the study)
 - Study simulations and analytical evaluations utilized during planning
 - Published articles regarding the adaptive methodology used
- Complete information regarding the conduct of the study
 - Compliance with the planned adaptive process and procedures for maintaining study integrity (although this provision is covered in section 10.2 of the ICH E3)
 - Descriptions of the processes and procedures actually carried out when there were any deviations from those planned (again, this information is required in section 9.8 of the CSR/ICH E3)
 - List of participants and records of all deliberations (including minutes) by committees involved in the conduct of the study.
 - Results of interim analyses

- Assessment of “firewalls” established to limit dissemination of information from unblinded interim analyses. There is no specific guidance as to content and measures of such an assessment.
- A copy of databases at each interim analysis (these, as well as the programs and results of the interim analyses should be submitted with the CSR)
- Additional information in the analysis of the study results (sections 11 and 12 of the CSR)
 - Inclusion of full information from each stage of the study
 - Examination of the consistency of treatment effects and other relevant results between study stages. The FDA accepts that statistical comparison of the results between parts of the study would have inadequate power and thus other approaches should also be utilized (although there are no concrete suggestions).
 - Comparisons of the baseline characteristics of patients as well as their outcomes from each part of the study (although the same argument for poor power of these comparisons can be made). If the evaluations indicate “shifts” in patient characteristics or outcomes, “more detailed” characterization is required (although again there are no specific guidelines for this more detailed characterization. Of course, certain adaptations are fully expected to result in “shifts” in patient baseline characteristics and outcomes.

There is here also a set of suggestions about the presentation of the study in CSRs and actual decisions taken. The FDA appears to favor a graphical representation of the study design, illustrating the various parts and processes, decision nodes and actual decisions.

L. References

- 1 [Retzios AD: “Why do so many Phase 3 studies fail: Part 1: The Effect of Deficient Phase 2 Trials in Therapeutic Areas with High Failure Rates in Phase 3 Studies” Bay Clinical R&D Services Web Site, 2009](#)
- 2 Katsnelson A: Adaptive Evolution. *New Scientist* 23: 55, 2009
- 3 Chow S_C and Chang M: Adaptive methods in clinical trials – a review. *Orphanet J Rare Dis* 3: 11 – 24, 2008
- 4 Gallo P, Chuang-Stein C, Dragalin V, *et al.*: Adaptive designs in clinical drug development – An executive summary of the PhRMA working group. *J Biopharm Stat* 16: 275 – 283, 2006

- 5 Jennison C and Turnbull BW: Group sequential methods with applications to clinical trials. Chapman & Hall/CRC, Boca Raton, 1999
- 6 Coffey CS and Kairalla JA: Adaptive Clinical Trials: Progress and Challenges. *Drugs R D* 9: 229-242, 2008
- 7 Bauer P and Köhne K: Evaluation of experiments with adaptive interim analyses. *Biometrics* 50: 1029-1041, 1994
- 8 Dragalin V: Adaptive designs: terminology and classification. *Drug Information Journal* 40, 425-435, 2006
- 9 O'Quigley JO, Pepe M, and Fisher L: Continual Reassessment Method: A Practical Design for Phase I Clinical Trials in Cancer. *Biometrics*, 46:33-48, 1990
- 10 Goodman, SN, ML Zahurak, and Piantadosi S: Some Practical Improvements in the Continual Reassessment Method for Phase I Studies. *Statistics in Medicine*, 14:1149-1161, 1995
- 11 Piantadosi S, Fisher JD and Grossman S : Practical Implementation of a modified continual reassessment method for dose-finding trials. *Cancer Chemother Pharmacol* 41:429-436, 1998
- 12 Resche-Richon M, Zohar S, and Chevret S: Adaptive designs for dose-finding in non-cancer phase II trials: Influence of unexpected outcomes. *Clin Trials* 5: 595-606, 2008
- 13 Pocock SJ: Group sequential method in the design and analysis of clinical trials. *Biometrika* 64: 191-199, 1977
- 14 Friede T and Kieser M: A comparison of methods for adaptive sample size adjustment. *Stat Med* 20: 3861-3873, 2001
- 15 Gould AL: Interim analysis for monitoring clinical trials that do not affect the type I error rate. *Stat Med* 11: 53-66, 1992
- 16 Proschan MA: Sample size re-estimation in clinical trials. *Biomet J* 51: 348-357, 2009
- 17 Jennison C and Turnbull BW: Mid-course sample size modification in clinical trials based on the observed effect of treatment. *Stat Med* 22: 971-993, 2003
- 18 Tsiatis AA and Mehta C: On the inefficiency of adaptive designs for monitoring clinical trials *Biometrika* 90: 367-378, 2003
- 19 Zelen M: Play the winner and the controlled clinical trial. *JASA* 64: 131-146, 1969
- 20 Sampson AR and Sill MW: Drop-the-Losers design: normal case. *Biomet J* 47: 257-268, 2005
- 21 Löscher C and Neuhäuser M: The statistical analysis of a clinical trial when a protocol amendment changed the inclusion criteria. *BMC Med Res Methodol* 8:16-25, 2008

- 22 Pocock SJ and Simon R: Sequential treatment assignment with balancing prognostic factors in the controlled clinical trial. *Biometrics* 31: 103-115, 1975
- 23 [FDA: Critical Path Initiative: Projects receiving critical path support – Fiscal Year 2008. April 2009](#)
- 24 [European Medicines Agency \(EMA\): Innovative Drug Development Approaches. Doc Ref EMA/127318/2007, London, 22 Mar 2007](#)
- 25 [FDA: Challenges and Opportunities Report – March 2004. March 2004](#)
- 26 Hung HMJ, O’Neil RT, Wang S_J and Lawrence J: A regulatory view on adaptive/flexible clinical trial design. *Biomet J* 48: 565-573, 2006
- 27 Bauer P and König F: The reassessment of trial perspectives from interim data – a critical view. *Stat Med* 14: 23 – 36, 2006
- 28 Hung MHJ, Wang S-J and O’Neil R: Methodological issues with adaptation of clinical trial design. *Pharm Stat* 5: 99 – 107, 2006
- 29 Gallo P and Maurer W: Challenges to Implementing Adaptive Designs: Comments on the viewpoints expressed by regulatory biostatisticians. *Biomet J* 48: 591-597, 2006
- 30 O’Neil RT: FDA’s draft guidance on adaptive designs in drug development: Current status and issues. Presentation, 21st Annual DIA Eurometing, Berlin, Germany, 2009
- 31 Koch A: Confirmatory clinical trials with an adaptive design. *Biomet J* 48: 574-585, 2006
- 32 [European Medicines Agency \(EMA\): Reflection paper on methodological issues in confirmatory clinical trials planned with an adaptive design. Doc Ref CHMP/EWP/2459/02, London, 18 Oct 2007](#)
- 33 Whitehead J: Stopping clinical trials by design. *Nature Reviews | Drug Discovery*, 3: 973-977, 2004