STATISTICS IN BIOPHARMACEUTICAL RESEARCH
2018, VOL. 10, NO. 2, 123–129
https://doi.org/10.1080/19466315.2017.1422799

Taylor & Francis
Taylor & Francis Group

Check for updates

Interpreting the Regulatory Perspective on Adaptive Designs

John Balser[a], Mark Chang[a,b], and Robin Bliss[a]

[a]Biostatistics Department, Veristat, Southborough, MA; [b]Biostatistics Department, Boston University School of Public Health, Boston, MA

ABSTRACT

The Food and Drug Administration (FDA) Center for Drug Evaluation and Research (CDER) and Center for Biologics Evaluation and Research (CBER) released a draft guidance (DG) on adaptive clinical trials (ACT) for drugs and biologics in February, 2010. In May, 2016, FDA Center for Devices and Radiological Heath (CDRH) and CBER issued the final guidance (FG) on adaptive medical device trials. The purpose of the FG is to provide clarity on how to plan and implement adaptive designs (AD) for clinical studies used in medical device development and to further encourage companies to use AD.

While both the device FG and drug and biologics DG provided positive review of ACT, the FG position was stronger, stating that the FDA centers "further encourage companies to consider the use of AD in their clinical trials." Both guidances emphasize the importance of preplanning to avoid Type I error inflation, strict following of the plan to minimize operational bias, and frequent and early interactions with the FDA to ensure the success of the planned ACT. Both guidances emphasize the utilities of clinical trial simulations in design of ACT and in analysis of adaptive trial data. In this article, we present our understanding the guidances.

ARTICLE HISTORY
Received December 2016
Revised November 2017

KEYWORDS
Adaptive design; Clinical trial; Clinical trial design; FDA guidance

Background

Since the Food and Drug Administration (FDA) released the critical path initiative document in 2005 (FDA 2005), a group of experts (mainly statisticians and medical doctors) from industry, academic institutions, and government positions began to promote adaptive clinical trial (ACT) designs at conferences and workshops both nationally and internationally. The FDA Center for Drug Evaluation and Research (CDER) and Center for Biologics Evaluation and Research (CBER) then released a draft guidance (DG) document on adaptive clinical trials (ACT) for drugs and biologics in February, 2010. Six years later, FDA Center for Devices and Radiological Heath (CDRH) and CBER issued the final guidance (FG) on adaptive medical device trials in May, 2016. The purpose of the FG as stated it is to provide clarity on how to plan and implement adaptive design (AD) for clinical studies used in medical device development and to further encourage companies to consider the use of AD in their clinical trials.

During the 12-year course of AD development since 2004, there are many research advancements in and implementations of ACT designs. Chow and Chang (2008) gave a review of adaptive designs. Bauer et al. (2015) and Chang and Balser (2016) provided two recent reviews on adaptive designs. Two task forces played particularly important roles in popularizing and advancing the adaptive approach are PhRMA adaptive design working group and Biotechnology Industry Organization (BIO) Adaptive Design Group. In 2016, PhRMA working group has released six white papers and the BIO working group released one white paper. In addition, multiple adaptive design books were published, including *Group Sequential Method in Clinical Trial Design* (Jennison and Turnbull 2000), *Adaptive Design Method for Clinical Trials* (Chow and Chang 2006, 2nd ed. 2011), and *Adaptive Design Theory and Implementation using SAS and R* (Chang 2007, 2nd ed. 2014) which provide a great source for understanding the concept, theory, method, and implementation of adaptive trial.

What is adaptive design? The first formal definition may have been provided by Chow, Chang, and Pong (2005): An adaptive design is a clinical trial design that allows adaptations or modifications to aspects of the trial after its initiation without undermining the validity and integrity of the trial. PhRMA adaptive design working group (Gallo et al. 2006) added the basis for adaptations, that is, using accumulative data, and defined adaptive design as *a clinical study design that uses* accumulating data *to decide how to modify aspects of the study as it continues, without undermining the validity and integrity of the trial.* The DG by FDA (2010) further stressed the importance of pre-specification of adaptations and defined an adaptive design 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. By *prospectively* the guidance intends that planning should occur before any data analysis is performed. The new guidance for adaptive design in device clinical trials provides a similar definition: "Adaptive design allows for *prospectively planned* modifications to a clinical study based on *accumulating data*, while maintaining the trial's *integrity and validity*."

To clarify the three levels of FDA policy (LaVange 2016: CDER and CBER Experiences: (1) Statutes are laws enacted by Congress, (2) regulations are binding interpretations of the law, and (3) guidance documents are nonbinding descriptions of FDA's current thinking on a topic.

CONTACT Robin Bliss robin.bliss@veristat.com Biostatistics, Veristat, 118 Turnpike Road, Suite 200, Southborough, MA 01772.
© 2018 American Statistical Association

FDA CDRH and CBER Experience on AD

Lin et al. (2016) at the FDA conducted a survey of investigational new drug (IND) and investigational device exemption (IDE) for Phase II to IV trials from 2008–2013 and reported that among 12,095 submissions during that time period, 1,225 were screened, and 140 (11.4%) were identified with AD components (Figure 1: source Lin, 2016). The AD trial characteristics by phase are summarized in Table 1. Among the 140 adaptive trials reviewed, the majority of the adaptive designs were group sequential designs (one-third) and sample size reestimation designs (20%). Eighty-two percent of AD had frequentist-based approaches, 3% were Bayesian adaptive designs, and 15% were unclear in methodology. The Bayesian adaptive trials submitted to the FDA unusually were considered low because CDRH is perceived to be a division that is more accepting of Bayesian approaches than CDER. Two-thirds of trials were Phase III and IV trials and one-third were Phase-I trials. Blinded and unblinded ADs were each about 50%. In those AD submissions, in 37% the study sponsor solicited and received FDA comments while 63% passed without the agency's comments.

Why Do We Need Adaptive Designs?

According to the FG, "When properly implemented, adaptive design can reduce resource requirements and/or increase the chance of study success." AD can also reduce the time to market, mitigate risks, and deliver the right drug to the right patient in right amount and at the right time.

The FG further elaborates: "Overall, adaptive designs may enable more timely device development decision-making and therefore, more efficient investment in resources in a clinical study. From an ethical standpoint, adaptive designs may optimize the treatment of subjects enrolled in the study and safeguard their welfare from ineffective or unsafe treatments and interventions at the earliest possible stage... Adaptive study design planning focuses on anticipated changes that may be desirable based on the data that will be accumulating during the course of the study."

The FG lays out two principles for adaptive design: (1) control the chance of erroneous conclusions (positive and negative) and (2) minimize operational bias.

Why Do We Need Prospective Planning?

The main reasons for prospectively planning the opportunities of adaptations are to maintain the integrity and validity of an adaptive trial. Our understanding of "*fully prospectively specified*" is dependent on the study design method used. It may not say in the protocol, for example, the exact increase in sample size under each possible effect size scenario, but for the interim outcomes (in this example: continue without sample size reestimation or continue with sample size reestimation), the adaptation should be uniquely defined with clear parameters, such as the calculation method for sample size reestimation and specific criteria under which the sample size would or would not be increased, so that under all possible adaptations the error and bias are controlled. As another example, a futility boundary can be arbitrary chosen prior to the time of analysis because when nonbinding rule is applied, Type I error is controlled. In general, adaptations can be made based on the cumulative data that are internal or external to the trial as long as the methodology prespecified and proved to be valid.

What Are Validity and Integrity in Adaptive Clinical Trials?

The FG clarifies: "For the purposes of this definition, integrity refers to the credibility of the results and validity refers to being able to make scientifically sound inferences." Validity, the reliability and accuracy of the results, can be understood as "internal validity" which is achieved when investigators can correctly conclude that an independent variable is, in fact, responsible for variation in the dependent variable. "Integrity," can be understood as "external validity," the generalizability of research findings to populations of subjects and across settings (Chang 2014). If (1) internal validity holds, (2) the experimental sample is representative of the intended population, and (3) experimental conditions are the same as (or similar to) outside reality, the study has both internal and external validity and we can extrapolate the experimental result to the population.

There are some threats to the validities are common to both traditional and adaptive designs. For example, if the inclusion and/or exclusion criteria of a trial change over time, the target population shifts and may introduce confounding variables that are not easy to recognize, making the associations between independent and dependent variables unclear, affecting internal validity. It may also make the target population shift over time making generalizability difficult if not impossible and thereby compromising the external validity of the trial. As another example, if multiple screenings are required to qualify subjects for a trial, the baseline response value may be distorted because, while baseline may be defined as the last observed value prior to dosing with study medication, the subject may have already undergone a number of study procedures over a number of weeks. The baseline response may not truly reflect the actual pre-study values and thereby limit internal validity. A poorly conducted trial may lead extensive missing data. Missing data will cause difficulties in statistical analysis and could lead to a considerable bias which threatens both internal and external validity.

Other threats to internal and external validity are more present in AD trials. When a response-adaptive randomization trial or "pick-the winner" design is used, patients who enrolled later in the trial will have a higher probability of randomization to the better/best treatment group(s). This could, in theory, cause heterogeneity of patients before and after the adaptation. As another common example, when the interim results from an ACT are released, it could change the sponsor's, investigator's, patients', and even the competitor's reviews of the drug and potentially bias the participation, conduct, and evaluation of subjects during the remainder of the trial.

The FG suggests a homogeneity check on the results before and after the adaptation. If a difference in the study sample or response is observed it may be an indication of study operational bias and could undermine the scientific validity and integrity of the study. This may be a problem in nonadaptive design studies as well.

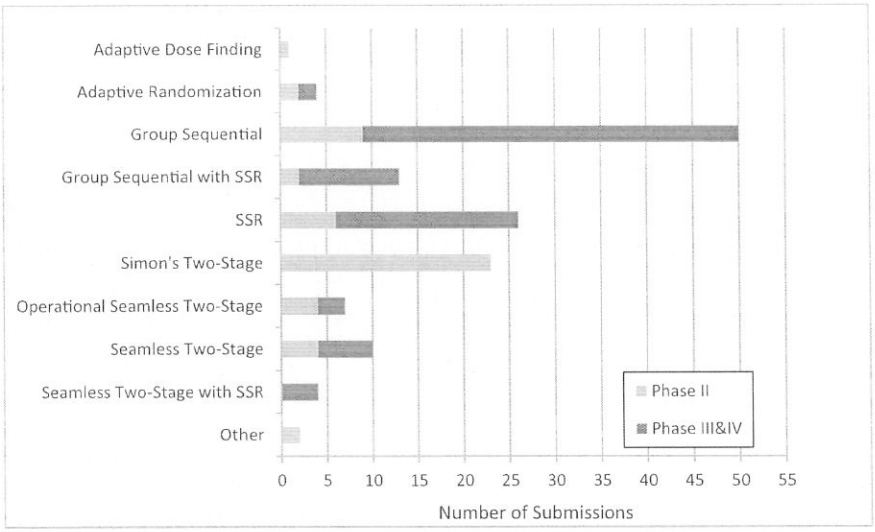

Figure 1. Type of adaptive design by trial phase.

A controversial issue surrounding internal and external validity is so-called subgroup analysis. As also occurs in many other industries, with economic globalization there are increasing numbers of international clinical trials (adaptive or non-adaptive multi-regional trials). It is possible that a drug or device may have demonstrated efficacy in some countries but not in others. While the drug may have an overall efficacious result when results from all countries are combined into one analysis, this may not reflect the country- or region-specific trends. To what extent can or should a study sponsor claim the drug is effective? Can the claim be extended to only some of the countries? To all the countries in the trial? Or can it be generalized to in all the countries in the world?

How to use Cumulative Information to make Adaptations and What Types of Adaptive Designs are Allowed

Adaptive clinical trials can be frequentist-based or Bayesian-based (FDA 2010) designs, and the cumulative information that adaptations are based on can be internal or external to the trial. From the FDA's experiences, however, there were only 8 trials using Bayesian methods (3 phase III and 5 phase I/II trials) between 2008 and 2013 and from 4 trials using Bayesian methods (2 phase II and 1 phase III/IV trials) (LaVange 2016). While both frequentist and Bayesian designs are possible to use, the vast majority of ACT to this point have been frequentist.

What types of AD are allowed? Our understanding is that any type of AD is allowed as long as it follows the spirit of the FDA's AD definition and the sponsor can prove that the design will uphold the protection of subject safety, the efficiency of the trial, and the protection of the validity and integrity of the trial. The guidance does, however, provide a list of example AD, including: group sequential design, sample size reestimation (SSR), Bayesian sample size adaptation, group sequential design with sample size reestimation, dose adjustment (DA) treatment arms, changes to randomization ratio, early stopping for both superiority and non-inferiority, adaptive enrichment design (FDA 2012), making adaptations based on total information, adaptations to device or to endpoints, and seamless design.

Table 1. Trial characteristics by phase.

	Study phases		
	II ($n = 53$)	III&IV ($n = 87$)	Overall ($N = 140$)
Trial Design			
Blinded	19 (35.8%)	58 (66.7%)	77 (55.0%)
Open label	34 (64.2%)	29 (33.3%)	63 (45.0%)
Parallel controlled	25 (47.2%)	84 (96.6%)	109 (77.9%)
Single arm	28 (52.8%)	3 (3.4%)	31 (22.1%)
Randomized	25 (47.2%)	83 (95.4%)	108 (77.1%)
Nonrandomized	28 (52.8%)	4 (4.6%)	32 (22.9%)
Method types			
Frequentist	45 (84.9%)	70 (80.5%)	115 (82.1%)
Bayesian	2 (3.8%)	2 (2.3%)	4 (2.9%)
Unclear	6 (11.3%)	15 (17.2%)	21 (15.0%)
Methods categories			
Well understood	36 (67.9%)	49 (56.3%)	85 (60.7%)
Less well understood	10 (18.9%)	22 (25.3%)	32 (22.9%)
Unclear	7 (13.2%)	16 (18.4%)	23 (16.4%)
Review decisions			
No comments	29 (54.7%)	23 (26.4%)	52 (37.1%)
Need clarification	24 (45.3%)	64 (73.6%)	88 (62.9%)

The DG divides AD into two categories from knowledge and experience perspectives: well-understood vs. less well-understood adaptive design. Most adaptive designs with blinded data review are considered well-understood. There is only one unblinded adaptive, group sequential design, that is considered well-understood. Group sequential design is singled out in part because of its long history; applications of group sequential design stretch back at least 30 years prior to the development of the DG. All other types of adaptive designs are considered less well-understood. The FG does not use terms "well-understood" and "less well-understood" to classify ADs since the terms have caused some confusions and misunderstanding. Some have inferred from the distinction of trials that the FDA would prefer to avoid less well-understood designs or may not allow them. The FG carefully avoids such inference by not referring to trials as "well-understood" and "less well-understood." With added research, the knowledge about AD is constantly growing and the available designs are increasing. As a result, the less-well known designs are becoming better understood. The FG also uses the term "seamless design" (Phase I/II trials combined or Phase II/III trials combined) that was apparently discouraged in DG, but many in the industry love to use the term since it is one of the ADs that has potentially larger gains over an traditional design. Over the past 6 years, we believe the status of knowledge and experiences of AD in FDA, industry, and academia have been improved dramatically.

The FG directly states that the FDA "further encourage companies to consider the use of adaptive design in their clinical trials." The DG, in comparison, was more encouraging of exploration into AD in earlier phase studies, such as through use of Bayesian approaches for early phase exploratory studies, including Bayesian continual reassessment method in oncology Phase I and Phase II trials. The DG also encouraged blinded AD before conducted complex unblinded AD for confirmatory trials.

Both guidances classify ADs into blinded and unblinded categories. An example of a blinded AD is an adaptation based on demographic or baseline measurements of the subjects enrolled in the study to modify the allocation rule on an individual basis to obtain better balance between the control and treatment groups. The FG further states, "While it is strongly preferred that such adaptations be preplanned at the start of the study, it may be possible to make changes during the study's conduct as well. In such instances, the FDA will expect sponsors to be able to both justify the scientific rationale why such an approach is appropriate and preferable, and demonstrate that they have not had access to any unblinded data (either by coded treatment groups or completely unblinded) and that access to the data has been scrupulously safeguarded."

The FG clarifies that "... an adaptive SSR study design is not intended to fix or salvage an already failed study, but instead can help prevent a failed study from occurring in the first place ... "It is recommended that the sponsor and the FDA reach agreement prior to study initiation on the minimal clinically important difference in treatment effect. Any decision concerning whether a particular difference is clinically important should be made at the outset and not influenced by the interim study effectiveness results." According to the FG, sample size reestimation can be based on Bayesian theory by incorporating the prior information, but the FG does not specify how to control Type I error.

The FG allows for total information designs such as to use blinded or unblinded variance to determine the sample size or the completion of study without statistical penalty. For example, a trial may stop when a fixed confidence interval width for the difference has been achieved.

Regarding seamless designs, the FG clarifies, "Prospective study planning to combine the feasibility and pivotal study phases should occur before the feasibility data are accessed in an unblinded manner; the plan needs to control the overall Type I error for the combined two stages. If the two studies are combined for operational purposes but are inferentially distinct then this would not be considered an adaptive design."

The FG also provides specific advice on AD for safety, observational, comparative, open label, and single arm studies as well as additional consideration for diagnostic devices.

From relevant sources of FDA (Lavange 2016 and Division V Director, 6,7), there is a sign to encourage the adaptive clinical develop program (Chang 2010, pp. 160–168). In FDA's terminology, the Master Protocol, which we believe is the adaptive CDP with more details.

Role of Data Monitoring Committee

Because of potential interim modifications to the adaptive trial that could risk to the validity and integrity due to operational bias, any ad hoc or interim changes to the study or the study design can be a great concern. The FG and DG stress the critical role of data monitoring committee (DMC) in preserving the validity and integrity of the AD trial by reviewing the interim study results and recommending the course of action, thereby reducing the number of personnel who are informed of the interim study results. The FG also suggests using an independent party such as clinical research organizations for computing analyses and serving on DMCs.

The FG suggest a separate statistical analysis plan for the interim analysis in an AD that is blinded to the sponsor since the detailed algorithms and interim decision rules (e.g., increase sample size) may allow the investigators to calculate the interim statistics given the decisions made after the interim. We believe that the need for such a separate SAP is actually dependent on the type of AD used. Careful review of the potential decision rules and information released to the blinded sponsor must be performed prior to the initiation of the study to ensure that any information made available will limit the amount of operational bias and very little can be presumed based on the known results. In addition, the FG makes clear that if an AD trial continues with modification after the interim analysis, it is still considered an AD and therefore the predefined alpha penalty applies.

With regarding sponsor monitoring, the FG advises sponsors to have a risk-based monitoring plan in place that focuses on specific aspects of adaptive studies that are of particular importance and may not be present in traditional (nonadaptive) trial designs. Maintaining the binding to the largest degree is expected.

The FG points out the "critical aspects [of AD] include but are not limited to: (1) assurance of a robust firewall for managing access to unblinded interim data/analysis, since DMC interactions with a sponsor have the potential to adversely impact study integrity, and (2) the shielding of investigators and study

participants as much as possible from knowledge of the adaptive changes that are implemented. The DMC charter should include a complete description of standard operating procedures relating to implementation of the adaptive design protocol. The protocol should state the role of the DMC, with particular emphasis on how the DMC will be involved in the conduct/analysis of the adaptation. A clarification on whether or not a DMC will review any interim analyses and who will conduct the adaptation of the design should be provided ... While the use of the DMC to manage the adaptations during an adaptive design clinical trial may be an acceptable option, a sponsor may instead consider assigning the responsibility for adaptation decisions to an independent statistician, a contract research organization, or some other independent clinical trial body."

Choice of Adaptive Design and the Role of Clinical Trial Simulation

Regarding when to choose an adaptive design, two questions need to be considered as suggested by the FG: (1) whether an adaptive design is feasible (e.g., if the timing of the measurement of the endpoint is appropriate for the interim analysis; if the sample size driven by safety; if the biomarker is good enough to guide an interim decision) and (2) whether an adaptive design is more appropriate than nonadaptive (conventional) design.

To determine whether or not to pursue an adaptive study design, it can help to select a number of realistic scenarios and optimize over them. The FG suggests considering "anticipated regret" by anticipating particular study outcomes that could lead to failure. All these can be done through clinical trial simulations (CTS).

Clinical trial simulation (CTS) is a process that mimics clinical trials using computer programs. In our view, CTS is particularly important in adaptive designs for several reasons: (1) the statistical theory of AD is complicated with limited analytical solutions available under many assumptions; (2) the concept of CTS is very intuitive and easy to implement; (3) CTS can be used to model very complicated situations with minimum assumptions, and Type I error can be strongly controlled; (4) using CTS, we cannot only calculate the power of an adaptive design, but we can also generate many other important operating characteristics such as expected sample-size, conditional power, and repeated confidence interval—ultimately this leads to the selection of an optimal trial design or clinical development plan; (5) CTS can be used to study the validity and robustness of an adaptive design in different hypothetical clinical settings, or with protocol deviations; (6) CTS can be used to monitor trials, project outcomes, anticipate problems, and suggest remedies before it is too late; (7) CTS can be used to visualize the dynamic trial process from patient recruitment, drug distribution, treatment administration, and pharmacokinetic processes to biomarkers and clinical responses; and finally, (8) CTS has minimal cost associated with it and can be done in a short time.

The DG and FG recognize the importance of trial simulations in determining operating characteristics of AD, justifying the selection of particular AD among alternative designs, and the understanding of the Type I error control, power, and bias in estimation of treatment effect under various scenarios (different treatment effects, models for early dropouts, etc.). The guidance states that the reporting of trial simulations should be an important component of the documentation to be submitted to FDA when a sponsor proposes the use of an AD in the development program, suggesting that the models, programs, and flow charts of AD simulation should be included.

Bayesian approach designs are encouraged in early exploratory designs. This indicates that Bayesian approaches provide a cohesive framework for describing the various choices and decisions available in an AD. The guidance states that Bayesian decision rules can be used to guide adaptations while preserving the Type I error rate in a frequentist sense; however, the guidance also states that, though trial simulations are acknowledged as useful, or even essential, for the understanding of operating characteristics of an AD, their use to establish strict control of Type I error rate in an AD is controversial and not fully understood. The implication is that while the CTS can support use of a Bayesian design and control over the Type I error rate, this may not be sufficient evidence for the regulatory review of the trial. In our opinion, from a methodological perspective, simulations can demonstrate full control the Type I error rate; however, the application of this may not be easily done. The logistical challenge is due to the infinitely many scenarios (null hypothesis configurations) that require demonstration of controlled Type I error. If the worst possible scenario cannot be identified, then we have not sufficiently demonstrated strong control over the Type I error. In such cases, practical control may be enough; meaning demonstration of control of the Type I error rate in practical situations and ignoring situations with remote probabilities.

Operation Perspectives of Adaptive Design

To minimize operation bias, the FG suggests limiting access to the data and maintaining a separate document detailing the adaptation algorithm from the traditional statistical analysis plan.

For institutional review boards, the FG suggests, "As an initial step when seeking IRB approval, sponsors should clearly describe the adaptive nature of the study and provide an informed consent document that accurately reflects the study's risks and meets other informed consent requirements. Potential planned adaptations should be described to the IRB and sponsors are encouraged to clearly articulate the circumstances under which protocol amendments will be submitted to the IRB for review." FG warns, "If prespecified adaptations were not disclosed to the IRB during the initial approval process, the sponsor risks critical IRB-related delays that can hinder study progress."

The FG provides a section on regulation considerations and states that (1) although a study sponsor may direct a New Drug Application (NDA) or Investigational Device Exemption (IDE) for Agency evaluation, the likelihood of success is increased through interactions with the relevant FDA review division and statistical staff during the study planning phase. An AD trial procedure should be strictly followed and if a deviation occurs, the FDA should be informed in a timely fashion.

Analysis of Adaptive Trial Data

The analysis of data from ACT is challenging and raises some controversial issues, such as how to compute an unbiased

estimate of treatment effect and the definitions p-value and confidence interval. As much discussion on these topics is still ongoing, the FG touches analysis only gently, mentioning estimation bias as the only topic of analysis included in the document. Here, we aim to further clarify some of the confusions that currently exist in the community.

1. Group sequential design, like sample size reestimation design, also violates the "one-person one vote" rule by allowing subjects in the first stage to "vote" multiple times in the final analysis. In other words, both study designs suffer from including subjects that contribute toward the interim analysis in the final analysis calculations.
2. Classic designs are also biased in reporting the results in the sense that only drugs that reach (or nearly reach) statistical significance will be allowed to market. Doctors and patients will see only the positive results, which involve both true positive and false positive. On average, the results reported are biased to the positive direction.
3. Whether the analysis and reporting of ACT data should be based on conditional (on the stage where the trial actually stopped) or unconditional result is a debatable issue. Should they be estimated in the same way? Or should there be some condition due to one study stopping early versus the other continuing to completion?
4. The p-value in an AD may be difficult to define and may be one of several options, none of which is as satisfied as for the classical design. For example, at an overall alpha of 0.05, a p-value of 0.006 can be very statistically significant or not significant at all depending on the choice of stopping boundary and stage of the stopped ACT trial.

Content of an Adaptive Design Submission to the FDA

Submissions to the FDA for an adaptive study design should clearly identify that the clinical study employs an adaptive design and should provide details of the proposed adaptations. Information provided should address what, when, how, and why the adaptation will be performed. The adaptation should be prospectively described at least generally in the protocol and in detail in the Statistical Analysis Plan, which should include the operating characteristics of the design. Submissions should also address key issues related to study monitoring and role of the DMC. Decision points should be delineated and documented for inclusion in the final study report to be submitted as evidence of safety and effectiveness to the FDA.

The original 15-item list suggested in the DG is not restated in the FG.

The DG suggested that the prospective specification of all aspects of the study design and key analyses are critical part of regulatory success of an AD trial. In the protocol of an adequate and well-controlled AD study, the DG generally requires more details (such as the simulation results) than are required for a classical design to allow FDA to evaluate the proposed AD. The DG specifies that the information to be included in an AD submission includes:

1. Study rationale
2. Justification of design features, including any proposed adaptations
3. Operating characteristics of the proposed design, such as Type I error rate and power
4. Plans to ensure study integrity when unblinded interim analyses are planned
5. Role of AD in overall clinical development strategy
6. Objectives and design features of the AD, all possible adaptations envisioned, assumptions, analysis methods, and quantitative justification for design choices at planning stage (typically via simulations)
7. Impact of adaptations on Frequentist operating characteristics (e.g., Type I error rate)
8. Summary of models used in planning (e.g., disease progression, dropout, dose–response)
9. Analytical derivations to demonstrate strict control of Type I error rate, if appropriate (e.g., A&WC studies)
10. Charter of personnel involved in carrying out adaptations and study monitoring
11. The example of elements that should be included in an AD SAP listed in the guidance are the following:
12. All prospectively planned adaptations
13. Statistical methods to be used to implement adaptations (e.g., how to calculate a potential increase in sample size or trial duration, rule used to select a dose)
14. Justification of Type I error control
15. Statistical approach to be used for appropriately estimating treatment effects

Conclusion

Six years have passed since FDA CDER and CBER released the draft guidance and, as the knowledge and experiences accumulate, FDA's position became more positive, as reflected in the final guidance published by CDRH and CBER in 2016. As stated, the new guidance is to further encourage companies to consider the use of adaptive design in their clinical trials. To summarize the key elements, the guidance suggests:

- Prospective trial specification is a key in the AD.
- Simulation is important to justify use of AD in the protocol.
- It is critical to have early and more intense interactions with FDA regarding the AD.
- An independent group from the study team, typically the DMC, is required for unblinded interim analyses and adaptive decision making.
- All analyses included in the final report should adhere to the statistical analysis plan.
- Well-planned, well-implemented studies are required.
- The consistency of estimated treatment effects across study stages should be explored because of concerns about shifts in patient population during the study in some of adaptive designs.

Both the final guidance and draft guidance emphasize the importance of preplanning to avoid Type I error inflation, strict adherence to the study plan to minimize the operation bias, and early and frequent interactions with the FDA to ensure the success of the planned ACT. Both guidances repeatedly emphasize the great utilities of clinical trial simulations in design of ACT and analysis of adaptive trial data. By simulating the practically possible scenarios and confirming the Type I error rates

to ensure the adaptive design (AD) will control Type I error in all practical situations. Simulations can also optimize AD by confirming the operating characteristics of different adaptive designs and estimating biases. Adaptive designed trials are not always better than traditional design and often simulations can bring this to light.

The FG clarifies that while most of the adaptations described in the guidance are more useful and appropriate for pivotal studies; adaptive designs can apply to some late feasibility studies. In general, the draft guidance from 2010 was to inform the industry that adaptive designs would be considered reasonable in many settings but was primarily for obtaining feedback from public on the designs in practice. The final guidance released in May 2016 is intended for industry to adopt.

References

Bauer, P., Bretz, F., Dragalin, V., Königa, F., and Wassmere, G. (2015), "TwentyFive Years of Confirmatory Adaptive Designs: Opportunities and Pitfalls," *Statistics in Medicine*, 35, 325–347. [123]

Food and Drug Administration (2010), *Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials*, Silver Spring, MD: FDA. [123,125]

——— (2012), *Draft Guidance—Guidance for Industry on Enrichment Strategies for Clinical Trials to Support Approval of Human Drugs and Biological Products*, Silver Spring, MD: FDA [125]

Chow C.S., and Chang, M. (2008), "Adaptive Design Methods in Clinical Trials—A Review," *Orphanet Journal of Rare Diseases*, 3, 11. [123]

Chang, M. (2007, 2014), *Adaptive Design Theory and Implementation of Adaptive Designs using SAS and R* (1st ed. 2007, 2nd ed. 2014), Boca Raton, FL: CRC Presss. [123]

——— (2010), *Monte Carlo Simulation for the Pharmaceutical industry*, Boca Raton, FL: CRC Press. [126]

——— (2014), *Principles of Scientific Methods*, Boca Raton, FL: CRC Press. [124]

Chang, M., and Balser, J. (2016). "Adaptive Design—Recent Advancement in Clinical Trials," *Journal of Bioanalysis and Biomedicine*, 1, 1–14. [123]

Chow, S.C., and Chang, M. (2006, 2010), *Adaptive Design Methods in Clinical Trials* (1st ed. 2006, 2nd ed. 2010), Boca Raton, FL: CRS Press. [123]

Chow, S. C., Chang, M., and Pong, A. (2005), "Statistical Consideration of Adaptive Methods in Clinical Development," *Journal of Biopharmaceutical Statistics*, 15, 575–591. [123]

FDA (2005), http://www.fda.gov/downloads/drugs/guidancecompliancere gulatoryinformation/guidances/ucm417627.pdf. [123]

Gallo P., et al. (2006), "Adaptive Designs in Clinical Drug Development—An Executive Summary of the PhRMA Working Group" (with discussion), *Journal of Biopharmaceutical Statistics*, 16, 275–283. [123]

Jennison, C., and Turnbull, B. W. (2000), *Group Sequential Methods with Applications to Clinical Trials*, New York: Chapman & Hall/CRC. [123]

LaVange, L. (2016), *CDER and CBER Experiences. ACDRS Special Workshop: Substantial Evidence in 21st Century Regulatory Science*, Washington, DC: UC-Washington Center. [123,125,126]

Lin, M., et al. (2016). "CBER's Experience with Adaptive Design Clinical Trials," *Therapeutic Innovation Regulatory Sciences*, 50(2), 195–203. [124]

Jun18_06_32_01