Adaptive Designs for Medical Device Clinical Studies

Guidance for Industry and Food and Drug Administration Staff

Document issued on July 27, 2016.

The draft of this document was issued on May 18, 2015.

For questions regarding this document that relate to devices regulated by CDRH, contact Dr. Gerry Gray (CDRH) at 301-796-5750 or by e-mail at Gerry.Gray@fda.hhs.gov.

For questions regarding this document that relate to devices regulated by CBER, contact the Office of Communication, Outreach and Development (CBER) at 1-800-835-4709 or 240-402-8010.





U.S. Department of Health and Human Services
Food and Drug Administration

Center for Devices and Radiological Health

Center for Biologics Evaluation and Research

GRZAN

Preface

Public Comment

You may submit electronic comments and suggestions at any time for Agency consideration to http://www.regulations.gov. Submit written comments to the Division of Dockets Management, Food and Drug Administration, 5630 Fishers Lane, Room 1061, (HFA-305), Rockville, MD 20852. Identify all comments with the docket number FDA-2015-D-1439. Comments may not be acted upon by the Agency until the document is next revised or updated.

Additional Copies

CDRH

Additional copies are available from the Internet. You may also send an e-mail request to CDRH-Guidance@fda.hhs.gov to receive an electronic copy of the guidance or send a fax request to 301-847-8149 to receive a hard copy. Please use the document number GUD1500005 to identify the guidance you are requesting.

CBER

Additional copies are available from the Center for Biologics Evaluation and Research (CBER) by written request, Office of Communication, Outreach, and Development (OCOD), 10903 New Hampshire Ave., WO71, Room 3128, Silver Spring, MD 20903, or by calling 1-800-835-4709 or 240-402-8010, by email, ocod@fda.hhs.gov, or from the Internet at http://www.fda.gov/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/default.htm.

GRZAN

Table of Contents

1. Introduction and Scope 1
2. What are Adaptive Designs? 2
   A. Definition 2
   B. Planning 3
   C. Advantages of Adaptive Designs 5
   D. Limitations of Adaptive Designs 6
   E. Adaptive Studies as a Learning Paradigm 7
   F. Study Design Changes That Are Not Adaptive 8
3. When to Choose an Adaptive Design 9
   A. Is an Adaptive Design Feasible? 9
   B. How to Decide an Adaptive Design is Advantageous 10
   C. Anticipated Regret 11
4. Principles for Adaptation in the Design of Clinical Studies 12
   A. Controlling the Chance of Erroneous Conclusions 12
   B. Minimization of Operational Bias 14
5. Adaptively-Designed Studies without the Need to Break the Blind 15
6. Adaptations Using Unblinded Data 17
   A. Group Sequential Designs 17
   B. Sample Size Reassessment 18
   C. Bayesian Sample Size Adaptation 21
   D. Group Sequential Designs with Sample Size Reassessment 22
   E. Dropping a Treatment Arm 23
   F. Changing the Randomization Ratio 23
   G. Investigating both superiority and non-inferiority 23
   H. Adaptive Enrichment 24
   I. Planning to Adapt Based on the Total Information 24
   J. Adaptation of the Device or Endpoint 25
   K. Seamless Studies 26
7. Special Considerations 26
   A. Changes to Pivotal Clinical Studies that are Not Preplanned Using Blinded Data 26
   B. Changes to Pivotal Clinical Studies that are Not Preplanned with Unblinded Data 27
   C. Simulations in Planning an Adaptive Design 28
   D. Adaptive Designs for Safety Endpoints 29
   E. Adaptive Designs for Open-Label Randomized Studies 29
   F. Adaptive Designs for Observational Comparative Studies 30
   G. Adaptive Designs for One-Arm Studies without a Control 31
8. Additional Considerations for Diagnostic Devices 31
   A. Adaptation to prevalence and the entire disease spectrum 32
   B. Blinded Sample Size Reassessment Based on Interim Estimates for the Comparator 33
9. Principles in the Analysis of Data from Adaptive Designs 33
   A. Bias Control in the Estimates 34
   B. Homogeneity of Results after a Modification 34
   C. No adaptation made 34

10. Challenges of Adaptive Studies 35
A. Data Monitoring Committees 35
B. Techniques to Minimize Operational Bias 36
C. Institutional Review Boards 38
D. Logistical Challenges 38
11. Regulatory Considerations 39
A. Interactions with FDA 39
B. Sponsor Monitoring 40
C. Best Practices to Protect Study Blinding (Masking) 41
D. Content of an Adaptive Design Submission to the FDA 41
12. Conclusion 42
13. References 42

GRZAN

Adaptive Designs for Medical Device Clinical Studies

Guidance for Industry and Food and Drug Administration Staff

This guidance represents the current thinking of the Food and Drug Administration (FDA or Agency) on this topic. It does not establish any rights for any person and is not binding on FDA or the public. You can use an alternative approach if it satisfies the requirements of the applicable statutes and regulations. To discuss an alternative approach, contact the FDA staff or Office responsible for this guidance as listed on the title page.

1. Introduction and Scope

An adaptive design for a medical device clinical study is defined as a clinical study design that allows for prospectively planned modifications based on accumulating study data without undermining the study's integrity and validity. Adaptive designs, when properly implemented, can reduce resource requirements, decrease time to study completion, and/or increase the chance of study success. This guidance provides sponsors and Food and Drug Administration (FDA) staff with guidance on how to plan and implement adaptive designs for clinical studies when used in medical device development programs.

This document addresses adaptive designs for medical device clinical studies and is applicable to premarket medical device submissions including Premarket Approval Applications (PMA), premarket notification (510(k)) submissions, de novo submissions (Evaluation of Automatic Class III Designation), Humanitarian Device Exemption (HDE) applications and Investigational Device Exemption (IDE) submissions. This guidance can be applied throughout the clinical development program of a medical device, from feasibility

studies to pivotal clinical studies. This guidance does not apply to clinical studies of combination products or co-development of a pharmaceutical product with an unapproved/uncleared diagnostic test.. However, the underlying principles may be applicable to such studies.

FDA's guidance documents, including this guidance, do not establish legally enforceable responsibilities. Instead, a guidance document describes the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word *should* in Agency guidance means that something is suggested or recommended, but not required.

2. What are Adaptive Designs?

A. *Definition*

An adaptive design for a medical device clinical study is defined as a clinical study design that allows for prospectively planned modifications based on accumulating study data without undermining the study's integrity and validity.[1] In nearly all situations, in order to preserve the integrity and validity of a study, modifications should be prospectively planned and described in the clinical study protocol prior to initiation of the study. However, in some specific circumstances, modifications to the design after the study begins can be scientifically valid if they are made without knowledge of the outcome results by treatment.[2] The different types of adaptive study design modifications (e.g., changes to the study design, study conduct, statistical hypotheses or analysis), as well as their advantages and limitations, are discussed in Section 6.

---

[1] 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.

[2] Knowledge of outcome results by coded treatment groups (e.g., outcomes known for treatments A and B), even without divulging which treatment is investigational, can undermine scientific validity.

B. *Planning*

A sound clinical study requires extensive planning, with consideration given to all elements of the study, from design to a plan for data analysis. 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. With adequate preplanning, a sponsor can use the study's accumulating data to modify various aspects of the study in a scientifically-valid manner.

However, there is a real danger that an unplanned modification to the study may weaken its scientific validity and therefore the study might not be adequate to support a future marketing application. Sponsors should anticipate and plan for modifications based on a variety of possible scenarios that could occur during the course of the study.

The following examples of adaptive modifications highlight some of the advantages of prospectively-planned adaptive study designs.

Example 1 - A sponsor conducted a randomized trial of a novel bone graft device designed to demonstrate non-inferiority of their device when compared to an autologous bone graft. That is, the study was designed to show that the novel device was no more than "delta" worse than the autologous graft, where "delta" is a pre-specified clinically determined margin. The study design assumed that fusion rates would be the same for the two treatments. An optional, prospectively planned, interim analysis to assess aggregate fusion outcomes (blinded (masked) by treatment group) was included in the study design to permit adjustment of the sample size, if necessary.

Example 2 - A randomized trial was conducted to demonstrate non-inferiority of an artificial cervical disc to the standard of care of anterior cervical discectomy and fusion. Although the study was sized for 500 patients, a planned interim look when subject number 340 reached the 24-month follow up demonstrated success. The PMA was submitted to the FDA and approved based on this smaller data set. This is referred to as "group sequential design" and, in many instances, has led to shorter and smaller trials. See Section 6.A. for more details.

Example 3 - A sponsor conducted a randomized two-arm unblinded study comparing a wound-healing device to the standard of care with a primary endpoint of time to drain removal. At study initiation, there was uncertainty about the variability in the estimated difference in mean time to drain removal (i.e., the standard error of the difference), so the sponsor chose to design the study to proceed until the estimated standard error for the difference in mean time to drain removal reached a certain agreed-upon threshold. As a result, the study needed to be conducted only until the pre-determined amount of information was acquired.. A similar approach could be taken in a study with a performance goal where the standard deviation is not known at the outset.

Example 4 – A sponsor conducts a diagnostic clinical study, in which the sponsor compares their investigational imaging system for detecting solid tumors, using an already-approved device as a comparator plus a reference standard. The goal is to establish non-inferiority to the already approved device. The primary comparisons will involve both sensitivity and specificity of the new device in comparison to the old device. The ability to anticipate sample size is limited, because it will depend on whether both new and old devices misclassify the same cases or different ones, as well as prevalence of the target condition of interest (a tumor present or absent). So an interim analysis can be used to determine if (a) a stop for futility is warranted or (b) additional subjects need to be enrolled.

Example 5 – A sponsor designs a group sequential trial for a surgical device so that at each interim analysis there are separate tests for effectiveness in diabetics and non-diabetics as well as overall. At one interim analysis, according to pre-specified rules, the device does not appear to be effective in diabetic patients. Further enrollment into the study is restricted to non-diabetic patients and the device gains approval for non-diabetic patients only. This is referred to as "adaptive enrichment" and is discussed further in Section 6.H.

Adaptive designs can use either frequentist or Bayesian statistical methodology. However, underlying principles discussed in this guidance regarding issues such as bias control, blinding, and operating characteristics of the study design are the same regardless of the statistical methods used.

C. *Advantages of Adaptive Designs*

An adaptive study design can have several distinct advantages when compared to a fixed (non-adaptive) design.

- It can be more efficient, saving time, money, and resources. This can occur in several ways. A trial with interim analyses could stop early for effectiveness in a preplanned way. A trial with two or more investigational arms could plan to drop one of them based on accumulating data. A single or multi-arm trial with a preplanned interim analysis could be stopped early for futility.
- Adaptive designs can improve the chance of trial success by employing sample size reassessment (SSR), whether for a superiority or a non-inferiority study. Based on accumulating data in the trial, planned SSR could lead to an adjustment in sample size, converting an underpowered study likely to fail into a study more likely to succeed. This approach allows for mid-course correction in trials where there is a treatment effect that is meaningful but smaller than originally anticipated. This facilitates the timely assessment and marketing of medical devices demonstrating a reasonable assurance of safety and effectiveness.
- Adaptive design may facilitate transition from premarket to postmarket follow-up. For example, a preplanned interim analysis that demonstrates favorable short-term study outcomes may result in a successful marketing application with continued follow-up relegated to the post-market stage. For further information see the FDA

Guidance "Balancing Premarket and Postmarket Data Collection for Devices Subject to Premarket Approval."[3]

- In some cases, planned modifications can occur without inflating the false positive error rate, provided there is a sufficiently strong blind to outcomes by treatment groups.
- Certain types of adaptive designs can enhance patient protection by increasing the probability that a patient is assigned to the treatment most likely to result in a better outcome for that patient.
- Adaptive designs can include a plan to modify the patient population during the study, converting what would otherwise be a failed study to one with, for example, a more targeted indication for which there are data to support both safety and effectiveness. This adaptation could help identify patients more likely to have a favorable benefit-risk profile from the use of a device. Such an adaptation effectively enriches the study population, which should be considered when interpreting the study results.
- Adaptive studies can improve decision-making at milestones during product development or increase the chance of a successful study with the potential to improve time-to-market.

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.

D. *Limitations of Adaptive Designs*

---

[3]http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM393994.pdf

The following are some of the possible limitations associated with an adaptive design study:

- Preplanned study design modifications can require more effort at the design stage, although this investment can pay great dividends during the study conduct. Adaptive study designs that are overly complicated can be difficult to plan, cost more, be logistically difficult to carry out, and provide overall results that are difficult to interpret.
- If not done correctly, adaptive designs can introduce operational or statistical bias, making it difficult to characterize the true effect of the investigational device. See Section 8.A. for additional details.
- A change to the study due to an adaptation may lead to results before the adaptation that are not sufficiently similar to those after the adaptation; this may confound the interpretation of the study results. (See Section 9.B.)
- The maximum sample size and/or study duration may be greater than a non-adaptive design.

For an in-depth discussion of the various types of planned modifications or adaptations, and their advantages and limitations, see Section 6.

E. *Adaptive Studies as a Learning Paradigm*

An adaptive design can allow for learning from the results of the study during its course and for preplanned changes to the study based on the accumulating outcome data. Such adaptation is a natural process during early feasibility studies in device development but for pivotal studies and some late feasibility studies such adaptation needs to be well-planned. Adaptive studies can be especially useful in the pivotal stage if there are uncertainties about one or two aspects of the study. In some cases, an adaptive design can obviate the need for a feasibility study (or a second feasibility study), and instead can allow the uncertainties to be scientifically addressed in an adaptive pivotal study. Generally, an adaptive study allows the planners to learn, during the study conduct, about a small number

of uncertainties and make preplanned, scientifically-valid changes based on accumulating data while maintaining study integrity. However, if there are numerous uncertainties, an adaptive design may be difficult to plan and implement. In such cases, it may actually be more efficient and increase the overall likelihood of success to conduct one (or more) additional feasibility studies to resolve some of these uncertainties before embarking on a pivotal trial.

Medical devices are often developed in a linear fashion, i.e., feasibility followed by pivotal explorations regarding clinical performance. Early feasibility studies may have a number of modifications that occur during the study, which may be unplanned. For these studies, it may not be necessary to apply formal statistical principles in order to draw useful conclusions for device development. In contrast, for some traditional (later stage) feasibility studies and for most pivotal studies, robust statistical validity is important, and unplanned modifications can undermine the study's purpose. For more general information on pivotal clinical investigations, see the FDA Guidance "Design Considerations for Pivotal Clinical Investigations for Medical Devices."[4]

While most of the adaptations described in this guidance are more useful and appropriate for pivotal studies, adaptive designs can apply to some late feasibility studies. For example, an adaptive feasibility study could increase the statistical rigor and lead to a more accurate estimate of device performance and hence enhance decision-making and the likelihood of later success at the pivotal stage. As outlined in Section 6.K., the planning of adaptations at the feasibility stage can also facilitate seamless feasibility-pivotal study transition. Sponsors may be able to productively utilize information from feasibility studies to help guide the appropriate design of pivotal studies, whether adaptive or not.

F. *Study Design Changes That Are Not Adaptive*

[4]http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm373750.htm

The following are examples of changes that are not adaptive:

- Any change or revision to a study design is post hoc and not adaptive if it is based on unplanned findings from an interim (or final) analysis in a study where the blind (mask) of outcomes by treatment groups has been broken (even if only the coded treatment group outcomes). Such modifications generally would endanger the scientific validity of the study since the false positive rate is not controlled and there is a strong possibility of operational bias[5]. The results from such a flawed study may not be valid.
- Modifications based entirely on information from a source completely external to the study are not adaptive designs according to the definition provided in this guidance.

These modifications will be discussed in detail in Section 7.B.

3. When to Choose an Adaptive Design

Several factors contribute to the decision of whether or not to choose an adaptive design. The most important considerations are, first, whether an adaptive design is feasible and second, whether an adaptive design is more appropriate than non-adaptive (conventional) design.

A. Is an Adaptive Design Feasible?

An adaptive design is usually feasible if there are a small number of endpoints on which the adaptation will take place, and if the timing of the primary outcome is such that there is time to implement the adaptation if required.

---

[5] For the purposes of this guidance, operational bias is the bias that arises if some or all participants (investigators, patients, care-givers) in the study have access to study results by treatment group and this information influences the ongoing operations of the study.

- When the time to ascertainment of the primary outcome is short relative to a longer enrollment period, a study is well situated to allow for adaptation.
- When subjects enroll rapidly there may be inadequate time to implement changes in the study design. Alternatively, sponsors could consider slowing enrollment in order to implement the study modifications.
- When enrollment and ascertainment time are both long but measurement of the primary outcome at an intermediate time point is predictive of the final time point then an adaptive design may be feasible, see section 6.C. for discussion of Bayesian adaptive designs.
- When studies are designed with multiple primary and secondary endpoints then proper statistical control of the Type I (or false positive) error under an adaptive design may be difficult to ascertain.
- When the study sample size is driven primarily by safety concerns (e.g. there are important rare adverse events associated with a medical device), sample size adaptation on effectiveness endpoints will not be feasible.

B. How to Decide an Adaptive Design is Advantageous

Given that an adaptive design is an option, there still remains the question of whether or not to choose an adaptive as opposed to non-adaptive design. The choice of an adaptive design should be considered as the sponsor plans a pivotal study. The recommendation is to select the optimal design for the particular situation, whether it is adaptive or a non-adaptive design. In order to determine whether or not to pursue an adaptive study design, it can help to select a number of realistic scenarios, some perhaps optimistic and some less so. For each scenario and a particular adaptive design, the challenge is to gauge how likely each scenario is and to calculate for that design the chance of success, the average size of the study, and the operating characteristics (probability of Type I error and the statistical power, discussed in Section 4.A.) and contrast it with the characteristics of a non-adaptive design. For non-

adaptive designs these calculations are usually straightforward. The topic of how to calculate these quantities for adaptive designs, using either analytical techniques or computer simulation, will be discussed later. Ultimately, the decision to proceed with a non-adaptive design over an adaptive one may rest on the sponsor's confidence in the anticipated parameter values and willingness to risk a failed study.

For a non-adaptive design, the sample size calculation is typically based on assumed values of several parameters. When assumptions regarding the parameters used for designing the trial are somewhat uncertain, a study may benefit from an adaptive approach to re-estimate sample sizes based on interim data. A basic issue is the amount of confidence there is in the choice of these parameter values. For example, suppose the study is planned for a somewhat optimistic treatment effect but the observed treatment effect is only 80% as large, but it is still clinically important. In a non-adaptive design powered for the optimistic effect, the chance of succeeding on the effectiveness endpoint is smaller than planned and may be unacceptably low. In this case the non-adaptive design based on a more optimistic effect size would be more likely to lead to a failed study for the sponsor. In contrast, an adaptive design with an interim analysis to reassess the sample size can correct an under-powered study, consistent with a smaller but still clinically meaningful effect size, increasing the chance of study success. An adaptive design can guard against these uncertainties by learning from accumulating data during the study.

If there is almost no knowledge of the parameters needed to design the trial then an adaptive design, although feasible, may be inefficient. In this case, a non-adaptive design will also carry high risk, and additional feasibility studies may be necessary before embarking on a pivotal trial. An alternative may be to combine feasibility and pivotal studies. See section 6.K. for discussion on seamless designs.

C. Anticipated Regret

It is sometimes helpful to anticipate particular study outcomes that could lead to failure so as to ask what planning decisions one might later regret. This concept is called "anticipated regret." For example, if a study just barely missed its objective but still had a

clinically important effect and in retrospect would have likely succeeded if the sample size had been 15% larger, that might suggest that one should have planned for an adaptive sample size design in which the sample size could be reassessed partway through the study. The ability to anticipate what one might regret and then plan to adapt can significantly increase the likelihood of study success. Adaptive designs that rely on anticipated regret can decrease the uncertainty in studies and make them much more predictable. Such planning can be thought of as insurance against possible threats to the success of the study. Using either analytical formulas or computer simulations one can calculate the costs associated with such insurance by comparing an adaptive design to a non-adaptive design. (Simulations will be discussed in Section 7.C.).

4. Principles for Adaptation in the Design of Clinical Studies

There are two underlying principles for the design of all clinical studies and of adaptive ones in particular: (1) control of the chance of erroneous conclusions (positive and negative) and (2) minimization of operational bias . These principles are crucial to assure that a clinical study produces valid scientific evidence. If the chance of erroneous positive conclusions is unacceptably large then the study will not provide sufficiently strong evidence to support a conclusion of device effectiveness. Conversely, if the chance of erroneous negative conclusions is large, the study may fail even when the device is truly effective. In short, studies that fail to follow these principles could generate evidence that is either invalid or inadequate. In the two subsections below, these principles will be further explored.

A. Controlling the Chance of Erroneous Conclusions

In order to assure scientific validity, a medical device clinical study should be designed to control the chance of erroneous conclusions. For example, in a superiority study of a new device compared to a control, an erroneous positive conclusion would be to determine that the new device is superior to the control when it is not. The inability to

minimize the chance of such erroneous conclusions threatens the scientific validity of the study and needs to be addressed. An erroneous negative conclusion would be to fail to determine that the new device is superior to the control when it is. Failure to control this type of error could lead to studies that fail to demonstrate the true effectiveness of the device.

In adaptive designs, control of the rate of false positive conclusions can be a major statistical challenge, and inflation of this error rate can arise from various sources. Most commonly, inflation of the false positive rate occurs due to "multiplicity," which arises when the study data are examined and analyzed multiple times during the study without appropriate statistical preplanning and the study is stopped at any time point where nominal statistical significance appears to have been achieved. Such multiple looks at the data may require a statistical adjustment to control the chance of erroneous positive conclusions. For adaptive designs there are other sources of multiplicity: multiple endpoints, multiple subgroups, multiple exposures (or dosages) or a combination of these features that could be dropped or added at an interim analysis. Another type of multiplicity would be an increase in sample size at an interim analysis without any statistical adjustment; this could also lead to the inability to control erroneous conclusions. With preplanning these types of error can be well controlled.

It is advantageous for both the sponsor and the FDA to understand the operating characteristics of a study design. The operating characteristics include the chances of false positive and false negative conclusions. The former is called the probability of a Type I error, where it is erroneously concluded that a device was effective when in fact it was not. A Type II (or false negative) error would be failing to conclude that a device was effective when in fact it was. The (statistical) power of a study is the probability of correctly concluding that the device is effective and is 1 minus the probability of a Type II error. Operating characteristics can also include other important properties of a design, including, for example, the expected sample size, or the probability of stopping for futility or success at an interim analysis.

There are usually two approaches for evaluating the operating characteristics of adaptive study designs for regulatory submissions: analytical methods and simulation

studies. Analytical statistical methods are often used in some frequentist adaptive study designs and can provide approximate probabilities for Type I errors and for statistical power for non-adaptive and simple adaptive designs under different scenarios. Simulations can be used to obtain operating characteristics for complex frequentist and Bayesian adaptive designs. Analytical methods and simulation studies can be complementary to each other in evaluation of the Type I error rate and power of adaptive study designs. In adherence to regulatory practice, the FDA recommends sponsors control the Type I error rate and maintain adequate power for all study designs.

B. Minimization of Operational Bias

One type of bias frequently encountered in studies with adaptive designs is the operational bias (defined in footnote 5) that can arise in the conduct of the clinical study. It is important that operational as well as statistical bias be reduced or eliminated because the presence of bias can distort the findings of a clinical study and undermine its scientific validity. For example, in a two-arm study, if an interim analysis is conducted resulting in an increased sample size in a preplanned manner, investigators at the study sites, study subjects and/or third-party evaluators may behave differently, either consciously or subconsciously, if the existence or size of the increase, or the reason for the increase, becomes known to them. As a consequence, bias may be introduced into the clinical study. Knowledge that the size of the study has been increased may help participants, site investigators, or blinded sponsor personnel to estimate the magnitude of the interim treatment effect which, in turn, can then affect the ongoing conduct of the study in various ways. Similarly, if not blinded to the patients' treatment assignment, the investigator may, unintentionally and unconsciously, change a decision about whether to enroll a subject in the study, or start treating the subjects in the investigational treatment group differently from subjects in the control group. Any of these actions can then lead to operational bias. Operational bias can be a significant threat to the scientific integrity of a clinical study and cannot be overcome by statistical adjustments to account for its presence. If designated analysts of the study data have access to the overall unblinded results of an adaptive trial during its conduct, it is vital that standard operating

procedures be followed to insulate this information from the study sponsor and investigators. Furthermore, it is important to assure regulatory authorities and other stakeholders that there are safeguards in place to ensure that those with legitimate access to unblinded data do not share information about these data with others. This concept of operational bias and firewalls[6] will be discussed in Section 10.B. of this document.

5. Adaptively-Designed Studies without the Need to Break the Blind

For a comparative study, when data blinding is unequivocally maintained, adaptations based only on the demographic characteristics of the subjects at baseline and/or on the aggregate outcome results do not pose any difficulty in terms of Type I error control or bias. On the other hand, changes based on outcomes by treatment group (whether coded or fully unblinded) may be problematic. In this section, "breaking the blind" means having access to the outcomes by treatment groups. It does not mean that one cannot know: 1) the demographic breakdown of the groups, 2) the overall combined outcomes if there are two or more groups, or 3) which subjects are assigned to which groups (as long as the outcomes by subject or by group remain masked or blinded). That is, someone with knowledge of both outcome and treatment group is not blinded.

An example of an adaptation based on demographic or baseline measurements of the subjects enrolled in the study would be to change the allocation rule on an individual basis to obtain better balance between the control and treatment groups. Note that this allows for knowledge of which individual subjects have been assigned to different treatment groups but does not allow for knowledge of any effectiveness or safety outcomes. This is called covariate adaptive randomization or covariate adjusted randomization; it uses accumulating baseline data in an attempt to provide better balance between the two groups.

---

6 For the purpose of this guidance we define a "firewall"as a process or procedure to ensure and document that sufficient restrictions on information flow to control statistical and operational bias is maintained.

A classic example of adaptation based on aggregate outcomes that is widely used is to power a time-to-event study or a survival study not by the number of patients in the study, but by the total number of clinical events. The study continues until the desired number of events has been observed. For such studies, the exact number of subjects cannot be planned in advance, although a maximum may be pre-specified. One is using the accumulating evidence from the study in the form of the aggregate results, in this case the total number of events, although the number in each of the comparative groups would not be revealed in either an unblinded or coded fashion to the investigators. The knowledge of the total number of events could lead to changing the total number of patients or to an extension of the duration of the study.

As another example of using aggregate results with multiple treatment groups without breaking the blind, one could observe the pooled overall success rate and, assuming two groups that differ by a hypothesized amount, infer that the original assumptions about the control rate and the investigational rate cannot be valid and that a change in sample size is merited. As yet another example, it is possible to calculate the overall variance for a continuous endpoint and make a sample size adjustment based on the hypothesized difference in the means.

In the prior two examples, in order to make a prospective decision, the required amount of aggregate information to be used is determined in advance, and the study is continued until that amount of information is obtained.

If the blind is maintained so that decisions regarding adaptation are made without knowledge of outcomes by coded or unblinded treatment group, in the case of a comparative study (or any outcomes in a one-arm study), then such adaptive designs pose no theoretical scientific difficulty. Sponsors are encouraged to consider adaptations that use baseline data

and aggregate outcomes for studies that do not break the blind, and it is strongly advised that such a study be conducted under an approved IDE, when appropriate.[7]

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.

6. Adaptations Using Unblinded Data

This section considers some adaptive designs that are based on accumulating unblinded results; these designs require thoughtful planning. Sponsors are encouraged to consult with the FDA prior to embarking on an adaptive design, in general, and for the types of adaptations that follow, in particular. Group sequential designs, sample size adaptation, and group sequential design with sample size reassessment are the most widely used.

A. *Group Sequential Designs*

Group sequential designs allow for interim analysis of the outcomes by treatment group and possible early stopping for success or futility. These designs have been relied upon for many years by the statistical and clinical trial community. These designs usually prescribe one or more planned interim looks at unblinded data with the possibility of stopping the study to declare either success or futility. They require prospective planning to determine the exact nature of the group sequential design, and introduce more flexibility compared to the non-adaptive sample size designs while controlling the overall Type I error rate of the study.

---

[7] An IDE is required when a sponsor intends to use a significant risk device in an investigation, when the sponsor intends to conduct an investigation that involves an exception from informed consent under 21 CFR 50.24, or when FDA notifies the sponsor that an application is required for an investigation. 21 CFR 812.20(a)(1).

Group sequential studies can be frequentist or Bayesian. If the device performs better than expected and there are sufficient safety data, this adaptive design can enable early stopping for success, saving time and resources. Such designs should include pre-specified statistical plans that account for the interim analyses and appropriate adjustments to the significance level alpha. For example, an O'Brien-Fleming plan prescribes a pre-determined fixed number of interim looks at fixed times with a prescribed fraction of the significance level alpha spent at each look. In contrast, a Lan-DeMets alpha-spending approach allows for more flexibility, since what is specified is the function for spending alpha at various time points in the trial. Once the alpha-spending function is specified at the outset, the number of looks (up to than the maximum allowed) and their timing are flexible. If there is a real possibility that the device may perform better than expected, the sponsor should consider using a group sequential design to allow for the possibility of stopping for success since in a non-adaptive design early stopping is not scientifically valid. If a sponsor believes that it is possible that a study could have results that would be so impressive at an interim look that the ethical decision would be to stop the trial, then the preferred approach would be to design an adaptive trial to allow for a scientifically-valid interim look such as in a group sequential trial. We recommend that sponsors utilize a Data Monitoring Committee (DMC) to examine the data in a secure and confidential manner and implement the group sequential design. (DMCs are discussed in Section 9.A.)

A disadvantage of any group sequential study is that a sponsor needs to accept some uncertainty because the accumulating data and study interim analyses will determine whether the study needs to enroll the entire cohort or can be stopped early for success or futility. Another disadvantage is the possibility of operational bias after a decision to continue at an interim analysis since a site or sponsor personnel could conclude that the effect size is not sufficiently large to stop the study.

B. *Sample Size Reassessment*

It is a common fallacy that simply adding more subjects or samples as an extension to a concluded study that has failed to meet its pre-specified endpoints is a scientifically-valid

way to continue a clinical investigation. Because the chance of an erroneous positive conclusion is no longer well controlled, the approach of simply extending a study at the end in a manner that is not pre-specified is neither scientifically sound nor recommended. In contrast, an adaptive design can permit sample size reassessment and appropriately control the Type I error in hypothesis testing or, correspondingly for interval estimation, the confidence coefficient. This may be accomplished through pre-specified analysis after a specified portion of the study has been completed to assess whether the planned sample size is adequate and, if not, to increase it in pre-specified manner. Such a strategy can control the chance of erroneous positive conclusions and produce scientifically-valid inferences.

Adaptive designs using such sample size reassessment (SSR) can help avoid under-powering studies, particularly in situations where substantial uncertainty exists concerning the variance or effect size. In a study design with a preplanned sample size reassessment, one or more preplanned interim looks are conducted to potentially adjust the sample size according to the comparison of the unblinded treatment group results. This is in contrast to blinded sample size reassessment that was considered in Section 5. It is crucial that the discussion concerning the clinically important effect size occurs during the study planning stage and not after outcome data are available. As a result, 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. Specifically, study planners should ask the anticipated regret question about the impact of a smaller effect size at the planning stage and incorporate a realistic, rather than overly optimistic, assessment of the investigational device's performance into their study planning.

There are a number of statistical techniques for SSR. Some methodologies use conditional power and others predictive probability (i.e., two different forms of stochastic curtailment). SSR can be done in a simple study with a single interim analysis or it can be performed more than once at pre-specified times during the study. 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.

In planning a sample size reassessment, careful consideration should be given to the reassessment time point(s). If reassessment is performed too late, it may be inefficient. Early looks may result in a suboptimal sample size determination due to inherent variability during earlier phases of study conduct. Analytical calculations or computer simulations performed under different scenarios can help guide the choice of optimal time point(s) for the reassessment. (See Section 7.C. for more discussion on simulations.) Actual control of Type I error rate will depend on the sample size adjustment methodology employed and the preplanned analysis that is used to combine the data from before and after the adaptation.

In some circumstances, if the primary endpoint takes a long time to observe (such as a two-year endpoint), the sample size adaptation may be ineffective. For such cases, sample size adaptation could instead be based on intermediate endpoints known to be associated with the primary endpoint. For more information on the use of surrogate and intermediate endpoints see the FDA guidance "Expedited Access for Premarket Approval and *De Novo* Medical Devices Intended for Unmet Medical Need for Life Threatening or Irreversibly Debilitating Diseases or Conditions."[8]

The use of a Bayesian model that learns from the accumulating data on the intermediate endpoint as well as the final endpoint is one statistical approach and is discussed in the next subsection.

In some cases, sample size reassessment is preferable to a group sequential design. Sample size reassessment is usually relatively more efficient when the increase in sample size is small. If, at the interim analysis, a large increase in sample size is required, then regardless of the statistical methodology chosen, SSR is extremely inefficient and a better strategy would have been to construct a group sequential design with some more realistic

---

[8]http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM393978.pdf

expectations about the size of the treatment effect. While the effect size is unknown at the start, if the expected range is narrow, a sample size reassessment strategy might make more sense.

C. Bayesian Sample Size Adaptation

Many Bayesian designs include sample size adaptation, since several factors that determine the trial sample size, such as effect size, variability of the sample estimate of the primary endpoint, and amount of prior information borrowed, are often not known at the design stage. Sample size decreases as the effect size and the amount of prior information increases and sample size increases as sampling variability increases.

When Bayesian hierarchical models are used to combine data from a current study with prior data, the amount of prior information borrowed is unknown before the start of the study and will depend on the similarity between the current study data and prior data, which is learned as data from the current trial accumulates. Whether there are prior data or not, a Bayesian trial design can often include a mathematical model that predicts a final clinical endpoint from earlier measurements. In that case, predictability will depend on the correlation between the earlier measurements and the final outcome, a correlation that is not known at the design stage. All these factors are learned as data accumulate and the sample size is adjusted as information is gathered.

In other cases, where a mathematical model relating results obtained in the course of the trial with the primary endpoint can be constructed and the model's parameters are estimated using accumulating data, the results can be used to predict the primary endpoint. The better the prediction, the smaller the required sample size, and a well-designed Bayesian study should be planned in such a way that the sample size is adjusted as information accumulates. As noted above, this idea is referenced in the FDA guidance document "Expedited Access

for Premarket Approval and *De Novo* Medical Devices Intended for Unmet Medical Need for Life Threatening or Irreversibly Debilitating Diseases or Conditions."[97]

Preplanned Bayesian adaptive designs could include interim analyses for sample size adaptation, for early trial success, and for futility. At the interim analyses, predictive probabilities of trial success would be calculated based on data accumulated up to that point. If the probability is sufficiently high (i.e.,above a pre-specified value), the trial may stop for early success; if the probability is too low (i.e.,below a pre-specified value), the trial may stop for futility; and if the probability falls between these values, continuation of the trial may be warranted (with (or without) termination of recruiting, if above (or below) yet another pre-specified value). Simulations are needed to determine reasonable thresholds for these actions.

A Bayesian adaptive design should consider simulations for assessment of its operating characteristics; the performance of the design depends on preselected design parameters. Simulations are used to determine the threshold values of predictive probabilities to stop for early success, futility, or for stopping recruitment of new patients. For more information on how to conduct such simulations, see Section 7.C. on simulation, and, for a more detailed discussion, refer to the FDA's "Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials."[10]

D. Group Sequential Designs with Sample Size Reassessment

A common adaptive design combines a group sequential design with interim looks, not only to stop early for success but also to re-assess the sample size and to increase it

---

[9]http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM393978

[10]http://www.fda.gov/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm071072.htm

according to a pre-specified plan. Such designs, while more complicated, offer additional advantages in certain situations.

E. Dropping a Treatment Arm

In a study in which there is more than one experimental arm, with proper planning one may plan to drop one of these experimental arms during the course of the study based on poor effectiveness performance. Dropping such an arm can increase study efficiency and focus resources on aspects of the study most likely to prove beneficial and successful.

F. Changing the Randomization Ratio

An adaptive randomization plan that allows for a change in the randomization ratio between the control and treatment arms based on treatment outcomes is called response adaptive randomization. Response adaptive randomization can mitigate some ethical concerns by reducing the probability that a patient will be exposed to products that are less effective or less safe. It can improve study efficiency (e.g. a Bayesian approach that adapts based on sufficiency of information from the control arm). Such adaptive designs can enhance patient protection by planned allocation to the treatment that, during the course of the study, is found to be either more effective or safer. Response adaptive randomization can sometimes lead to slightly larger studies but could facilitate investigator and patient enrollment.

G. Investigating both superiority and non-inferiority

It is possible to plan a study to investigate both the superiority and the non-inferiority of a new treatment to an active control. Two different strategies may be used: one is to plan the study as a superiority trial and have a fallback hypothesis of non-inferiority; the other is to plan (and size) the study originally as non-inferiority but allow for an investigation of superiority.

A superiority study designed to investigate non-inferiority in the event that the superiority hypothesis fails should be prospectively planned; in particular, the non-inferiority margin should be pre-specified and agreed upon in advance before any unblinding. Additionally, a prospective plan could incorporate sample size reassessment with the change in claim.

Generally if the original plan is for non-inferiority, investigating superiority is possible without additional preplanning since the superiority margin is already pre-specified. However, study planners may wish to incorporate a preplanned interim assessment and pre-specified sample size reassessment in case mid-course results are sufficiently promising that a superiority claim may be within reach; such adaptations should be pre-specified and appropriate analyses provided.

H. Adaptive Enrichment

Another type of adaptive design is one that plans to investigate, using unblinded data, at one or more interim looks, pre-specified patient subgroups that might have differing responses to the experimental device. Such analyses could be used in a preplanned way to modify the inclusion/exclusion criteria after an interim analysis. For example, suppose that it was anticipated that there may be a differential effect due to a demographic factor such as sex. Then at a preplanned interim look, the difference could be assessed and the trial potentially modified to include only men or women from that point onwards. Another type of adaptation would be to incorporate a sample size reassessment to ensure a sufficient sample to carry out analyses separately for men and women in the case where the interim data suggest that the two groups should not be pooled. Preplanned methods could also change the sample size based on the decision to narrow the population indication. In all cases it is important that the chance of erroneous findings (the overall probability of a Type I error) be well-controlled in a prospective manner.

I. Planning to Adapt Based on the Total Information

For this novel type of design, the stopping rule is based on the amount of information in the unblinded data and this information is usually measured in terms of the variance of the primary endpoint. That is, the study continues until a pre-specified amount of information is obtained. Because the stopping rule is based only on the amount of available information, there is no penalty associated with repeated looks. For example, the decision about when to stop could be based on the estimated standard error of the mean for the difference between the mean responses of the investigational and control groups. Typically, this would correspond to stopping when a fixed confidence interval width for the difference has been achieved. This total information approach safeguards against the misspecification of the parameters that one might have in a non-adaptive design study. The study is always correctly powered, and there is no statistical penalty for looking early. This design does not suffer from the problem of the non-adaptive study design, which sometimes is too large and other times not large enough; in fact, it can guarantee that a study is always "right-sized." This approach could be particularly helpful in some one-arm studies and some studies for diagnostic devices.

In this design it is important to meticulously abide by the pre-specified stopping rule. Intentionally overrunning the sample size can result in a variance that is smaller than agreed upon, and thus increases the chances of a statistically significant result for a clinically unimportant difference. Also, whereas this total information approach does control the Type I error rate, it would no longer to do so in the case of an overrun. (In that way, it is similar to the unplanned study extension that was discussed in Section 6.B.)

J. Adaptation of the Device or Endpoint

Preplanned device or endpoint adaptations are rare for pivotal studies. On the other hand unplanned changes to the device or the endpoint are quite common in feasibility studies, especially early feasibility ones. For unplanned changes to the device or to the endpoint, see Section 7.B. For planned changes, study planners are advised to prespecify the changes (or anticipated types of changes) and account for them in a pre-specified statistical plan with appropriate consultation with the FDA in advance.

K. Seamless Studies

Device development and evaluation plans may include a feasibility investigation that smoothly transitions to a pivotal study in a preplanned manner, if no significant changes to the device or study are made. For example the feasibility portion of a seamless study could include several possible device configurations; only the configuration that showed the most favorable results in the feasibility stage would be continued (seamlessly) into the pivotal study. In such cases, if the study is designed to be "inferentially seamless", all data may be included in the final analysis. 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.

7. Special Considerations

A. Changes to Pivotal Clinical Studies that are Not Preplanned Using Blinded Data

Under certain circumstances, a number of scientifically-valid changes to the study design can be entertained even if they are not preplanned. Such changes should utilize complete masking of the outcome results by treatment group, so that no one representing the sponsor (or the FDA) has access to the coded or unblinded outcomes by treatment group. A major advantage of conducting a study where the outcome by coded or unblinded treatment groups is fastidiously guarded is that changes to the study based entirely on outside information can be reasonably entertained. For example, if only an independent statistician and the DMC had access to the outcomes by coded or unblinded treatment groups and the sponsor could provide evidence that the results were limited to only those people, the sponsor or the Steering Committee could propose scientifically-valid modifications to the design of the study based on information entirely from outside the study. Note that it is not

appropriate for those with access to the outcome data by treatment group, including the DMC, to propose or provide input concerning study revisions. Inflation of Type I error can occur whenever an unplanned revision is suggested as a result of an interim analysis, regardless of where the suggestion arises. The discussion of firewalls to prevent inappropriate disclosure of information is discussed further in Section 10.B. Unplanned study changes under appropriate circumstances are scientifically viable and should be discussed with the FDA for approval before implementation.

B. Changes to Pivotal Clinical Studies that are Not Preplanned with Unblinded Data

If outcome results are not blinded or masked (as in an open label study), study design changes become problematic due to the fact that the scientific integrity of the study may be endangered. Sponsors are strongly encouraged not to implement such changes and to meet with the FDA if such changes are being considered.

In general, any proposed modification to the protocol or the Statistical Analysis Plan will be problematic if it will affect the validity of the data or information generated in the study or the scientific soundness of the plan.

For studies conducted under an IDE, if the validity of the data/information generated in the study or the scientific soundness of the plan are affected, such modifications will generally require prior approval though an IDE supplement. Note that there are exceptions to the requirement for prior approval for changes effected for emergency use, changes effected with notice to the FDA within 5-days, and changes submitted in an annual report, when those changes meet certain criteria as specified in 21 CFR 812.35 (a)(2)-(4). Changes to essential device functionality based on data should be limited to feasibility studies, if at all possible. There are limitations to the extent of allowable device changes for a pivotal study, as significant device modifications can undermine the scientific validity of the pivotal trial data and the legitimacy of combining pre- and post-device modification data. Sponsors are encouraged to engage the Agency regarding possible fundamental device modifications during a study, as delayed disclosure of device modifications can lead to longer review times and lower likelihood of study success. Additional complexity is introduced by "evolving" device modifications (e.g. an

evolving algorithm) that may be more appropriate for a feasibility study rather than a pivotal study. For example, the use of pivotal study data to assess, modify, and finalize an algorithm for a diagnostic device may produce biased performance results and therefore is not recommended unless a preplanned analysis is put in place to control the Type I error rate. When determining whether pooling of data from different device versions is acceptable, an analysis as to whether there is homogeneity between the outcomes (both safety and effectiveness) for the different versions of the device, as discussed more broadly in Section 8.B., is critical.

C. Simulations in Planning an Adaptive Design

Computer simulations can play a crucial role in adaptive designs and can provide the operating characteristics of the study design under different scenarios. The simulations can evaluate different scenarios with a variable number and timing of the interim analyses and can be used to weigh the advantages and disadvantages of different adaptive designs or an adaptive design compared to a non-adaptive design. Simulations can provide insights into required samples sizes, operating characteristics, and interrelationships between trial design choices and patient characteristics that cannot be obtained by analytical methods.

Computer simulations used in planning adaptive study designs have limitations. First, their utility and quality are dependent on the ability to model realistic scenarios. Second, programming mistakes by the sponsor in the simulation software code, which may be difficult to detect, can lead to poor study design choices. Third, complex study designs, such as those that involve multiple endpoints or a complicated null hypothesis boundary may be difficult to perform.

Nonetheless, simulations for adaptive design trials, although complex, are encouraged. Good programming practices and careful consideration of the trial design can help prevent errors in the simulations. When a sufficiently wide range of relevant scenarios can be covered, simulations can provide useful insight on adaptive trial design, conduct, and analysis. Simulations can provide finer coverage of the parameter space (including nuisance

parameters) and more detailed descriptions of the operating characteristics than analytical methods can usually provide.[11]

D. Adaptive Designs for Safety Endpoints

While many adaptive study designs focus on the effectiveness endpoint, it is also possible to design adaptive clinical studies for safety endpoints. For example, an adaptive design could be developed to demonstrate that a device had an overall serious adverse event rate of less than 8%. Specifically, a group sequential approach could be used to allow for one or more interim looks and an early study termination for success if the serious adverse event rate was much less than 8%. Alternatively, one could develop a stopping rule that would terminate the study if there were no adverse events in a pre-specified number of patients but would allow for continuation to a later stage with one or more events. The preplanned rule would need to demonstrate that it controlled the chance of the erroneous conclusion that the serious adverse event rate was at most 8%, for this hypothetical example.

E. Adaptive Designs for Open-Label Randomized Studies

Unlike drug trials, many scientifically-valid medical device studies are not, or cannot, be blinded. For example, the medical device may have visible parts or treatment features (e.g., electrical stimulation) that can make it obvious to the patient and the medical staff that a device (or which device) is being used. While in some cases, patients, third-party

---

[11] Some discussion on simulations in the Bayesian context, also applicable to frequentist trials, can be found in the Food and Drug Administration (2010) "Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials" (issued February 5, 2010) – Appendix, page 40.
http://www.fda.gov/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm071072.htm

assessors, or even the health care provider can be blinded, there are many instances where this is not possible. Studies where blinding does not occur are called "open-label."

Using an adaptive design for an open-label study presents additional difficulties because operational bias can be introduced when patients or trial personnel know the treatment assignment and either consciously or subconsciously change how they behave. This potential for bias is not unique to adaptive trials but rather is true of open-label studies, in general. However, extra care may be warranted with an adaptive design where there is a potential to alter the trial midcourse.

The importance of pre-specified adaptations is paramount for open-label studies that incorporate an adaptive design. At the design stage, every effort should be made to spell out in detail all possible intended changes and the corresponding adaptations, with appropriate checks on the trial's operating characteristics. For example, for a classical group sequential design, before the start of the trial, one should clearly pre-specify in the protocol the number and timing of the interim analyses, and the corresponding alpha-spending function. Although such pre-specification may not address the problem of operational biases in an open-label trial, a pre-specified protocol greatly reduces the possibility of unplanned changes being made based on interim trial findings. Unplanned modifications that were not anticipated during the planning stages can be problematic if they occur during the course of the open label study.

F. Adaptive Designs for Observational Comparative Studies

Adaptive designs may also be used in studies designed with an historical or non-randomized concurrent control. Typically, a comparison is conducted of baseline covariates in the treatment group compared to the control group. In an adaptive design, such a comparison should be pre-specified and performed in a manner such that the personnel who conduct the comparability evaluation are blinded/masked to outcomes of all arms. If the comparability evaluation indicates that the control group is not comparable to the treatment group with the investigational device, a change or modification to the control group may be possible. Even if the control group is appropriate, the sample size and power estimation

could be reevaluated and modified as long as unblinded access to the outcome data has not occurred.

G. Adaptive Designs for One-Arm Studies without a Control

Although every effort should be made to conduct a randomized concurrent controlled trial when possible, sometimes a medical device trial will compare the treatment arm to a performance goal because it is not ethical or feasible to have a placebo (sham) device or an active comparator device serve as the control arm. Although there are additional biases (including operational bias) that may be introduced by a one-arm study, a pre-specified adaptive design may still be possible. To control the operational bias, the knowledge of the outcome data should be carefully restricted. A screening log of all incoming subjects (including those not included in the study) at each clinical site can ensure that no overt selection occurred and thus help to reduce possible manipulation of the trial findings.

8. Additional Considerations for Diagnostic Devices

While issues discussed in other sections of this guidance also apply generally to diagnostic medical devices, there are some unique issues with adaptive study designs for diagnostic devices. A thorough discussion of general design considerations can be found in the FDA guidance "Design Considerations for Pivotal Clinical Investigations for Medical Devices[4] and would be useful to review if considering an adaptive design for a diagnostic device. Diagnostic performance is often evaluated using estimation and confidence interval approaches rather than hypothesis testing. The adaptive design methods described above can be translated into appropriate confidence intervals for diagnostic studies. As noted in Section I, this guidance does not apply to clinical studies of combination products or co-development of a pharmaceutical product with an unapproved diagnostic test. However, the underlying principles may be applicable to such studies.

Unlike studies of therapeutic devices where study completion may be challenged by slow enrollment or long follow-up times, many clinical performance studies of diagnostic devices are cross-sectional, in which enrollment is quite rapid. Thus, in many cases, the

rationale for pursuing an adaptive study for a therapeutic device may not be relevant for a study of a diagnostic device.

Nevertheless, diagnostic devices are heterogeneous in scope and include devices that may be invasive or pose risks to a patient (e.g., devices used during surgery). In this case an adaptive design can be advantageous since fewer patients may be exposed to an inferior or less safe device. Diagnostic clinical outcome studies that evaluate diagnostics for patients with a rare condition with slow accrual may also benefit from an adaptive design.

A. Adaptation to prevalence and the entire disease spectrum

Studies may be designed to be adaptive to the prevalence of the disease in the study. For example, disease prevalence could be monitored (using a clinical reference standard[12] , not the investigational device) until the requisite numbers of diseased and non-diseased subjects are enrolled.

In some diagnostic device studies, the frequency of certain critical subgroups may be less than expected; a prospective adaptive study can use a planned interim look to assess and adapt to assure appropriate subgroup representation. Such adaptations could entail the addition of new clinical sites to obtain a different patient mix, e.g., adding a family practice rather than a specialty clinic if more patients with early stage disease are sought. When only the clinical reference standard results are known to the group making adaptation decisions, no correction for confidence level is needed. A similar approach can be used when device performance is being estimated by hypothesis testing. As always, pre-specification and careful documentation of procedures to maintain the necessary blinding is recommended (See Section 11.C.).

---

[12] For purposes of this guidance, the clinical reference standard is defined as the best available method for establishing a subject's true status with respect to a target condition; please refer to FDA Guidance "Design Considerations for Pivotal Clinical Investigations for Medical Devices"

B. Blinded Sample Size Reassessment Based on Interim Estimates for the Comparator

Some diagnostic device studies are designed to compare a new, investigational diagnostic (or a marketed diagnostic for a new indication) to an already cleared or approved device. In some cases, an adaptive study design may increase study efficiency and the likelihood of success by prespecifying an interim analysis and potential sample size adjustment. For example, if the study or intended use population has a different prevalence from that of the population previously studied, a study adaptation may assure that there are a sufficient number of subjects with the target condition of interest. With appropriate pre-specifications and well-documented blinding, such an adaptation would not require statistical multiplicity adjustments in the calculation of confidence intervals. However, if the rationale for increasing the sample size is performance-based and not pre-specified, a multiplicity adjustment may be required to maintain scientific integrity of the study.

Other adaptive designs for studies evaluating diagnostic devices are feasible. Sponsors are encouraged to consult with the Agency in advance of any adaptive design submission.

9. Principles in the Analysis of Data from Adaptive Designs

While previous sections focused on the importance of prospective planning during the design phase of adaptive studies to control the risk of operational bias and erroneous conclusions, this section considers some of the specific challenges of analysis of data from adaptively-designed studies; however, a detailed discussion is beyond the scope of this guidance.

---

http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm373750.htm.

A. Bias Control in the Estimates

Even when the Type I error rate is well controlled, estimators of treatment effect for adaptive designs are frequently biased. For example, in a group sequential design, if the stopping boundary is crossed and the study is stopped at the interim for success, the naïve (point) estimate of the treatment effect is upwardly biased, even though the overall Type I error rate of the study is controlled. The same type of bias occurs in many confidence intervals. In some cases the amount of bias can be estimated by simulation. Efforts to adjust for this bias (unbiased estimators, bias-adjusted estimators, repeated confidence intervals) can be prospectively planned in the Statistical Analysis Plan.

B. Homogeneity of Results after a Modification

Studies that undergo modifications during their conduct, whether planned or unplanned, should be analyzed to determine whether there are detectable differences in study participants, investigational device performance, study outcomes, or other important study aspects before and after the study modifications. Although such analyses may not be able to detect subtle operational bias, they would allow detection of gross changes in patient characteristics or outcomes. Some adaptations might be expected to result in changes (e.g., when there is a change in the population of interest). In other cases, a difference before and after might be observed when no difference was expected or desired. Such a result may be an indication of study operational bias and can undermine the scientific validity and interpretation of the study. This may also be a problem in non-adaptive design studies.

C. *No adaptation made*

If no adaptation was performed during the course of the study that was designed to be adaptive, the study would still be considered adaptive by the FDA. Therefore, the study should be analyzed according to its pre-specified analysis plan and be reported as such to the FDA.

10. Challenges of Adaptive Studies

A. Data Monitoring Committees

Data Monitoring Committees (DMCs) play an important role in protecting the safety of trial participants. In some cases, the DMC may be prospectively selected as the appropriate entity to implement all pre-specified study adaptation decisions. Even in cases where another entity is charged with the logistics of the adaptation, the DMC is tasked with safeguarding the trial participants and should monitor their safety during the adaptive trial. The DMC should be appropriately constructed to assure that its members possess the necessary expertise and experience for an adaptive study design, if such adaptations are part of the study plan. In cases where adaptations are based on interim analyses of unblinded outcomes, robust pre-specified and well-documented procedures should be in place before initiation of the clinical trial or review of the data. Critical aspects 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. In any case, the underlying validity and integrity of the study depends on appropriate decision-making and implementation and should always be paramount when planning adaptive designs.

Although the DMC may be tempted to recommend changes to the adaptive design or to the fundamental study type (e.g., from a non-adaptive study to an adaptive one) during the conduct of the study, once the DMC has access to coded or unblinded outcomes, such recommendations can imperil the scientific integrity of the study. Fundamentally, the DMC is tasked to protect the subjects in the study and should always act accordingly to protect the subjects in the trial.

For further information, refer to the FDA guidance "The Establishment and Operation of Clinical Trial Data Monitoring Committees for Clinical Trial Sponsors."[13]

B. *Techniques to Minimize Operational Bias*

Operational bias is a major concern in adaptive designs. It can exist even in the group sequential setting. In general, to reduce operational bias in studies with adaptive designs, one should limit the access to outcomes by coded or unblinded treatment groups. One way to do that is to set up firewalls that guarantee that access to such data is restricted only to those for whom it is absolutely essential. This should be done if the sponsor wishes to retain the ability to suggest scientifically-valid changes to the design during the course of the study. In addition, to limit operational bias and depending on the type of adaptation, it is recommended that the precise details of the adaptation algorithm be separated from the investigator-shared protocol and placed in a detailed Statistical Analysis Plan for the adaptive design. This information may be shared with the DMC, the Institutional Review Board (IRB), and the FDA. This can help maintain the scientific integrity of the study and reduce the ability of study observers to "reverse engineer" the interim study results based on knowledge of the adaptation protocol.

Several examples illustrate the importance of avoiding operational bias. In a study with a pre-specified sample size reassessment, someone with knowledge of the sample size

---

[13]http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm127073.pdf

adjustment protocol and the sample size adjustment may be able to easily calculate the observed treatment effect at the time of adaptation. In a study with an adaptive randomization ratio, the relative performance in each treatment arm can be inferred with knowledge of the protocol and observed study modification. Even in a classical adaptive design such as a group sequential one, biases could be introduced through inference that a large treatment effect was not observed, since the study continues to the next stage instead of stopping at the interim analysis.

Although one cannot completely eliminate such information leakage, extra care should be taken to control the information so that only those who must have access to it know about the trial modification. For example, if the study sample size is increased after the interim analysis, clinical study site personnel can continue to enroll subjects and be notified that the final enrollment number has not been reached. In addition, the protocol could specify a categorized sample size change instead of a precisely calculated change to make the back calculation less informative. In an adaptive randomization design, when a centralized randomization mechanism is used, each clinical site can be notified of the treatment assignment for the next subject rather than being notified of the randomization ratio change. For a group sequential trial, not all principal investigators need to know that an interim analysis has been performed and a decision has been made to continue the trial to the next stage. A seamless analysis performed in the background ensures the study follows the protocol and minimizes the bias associated with the interim analysis. Similarly, for a trial with an adaptive selection of primary endpoints or an adaptive change of hypotheses, assuming all needed variables are collected according to the preplanned protocol, the decision of the change does not need to be communicated to each clinical site.

In the conduct of an adaptive design, an effective and well-documented firewall increases the likelihood that trial modifications will be scientifically valid, maintain integrity of the data and trial, and be acceptable for regulatory purposes.

C. Institutional Review Boards

Institutional Review Board (IRB) oversight (21 CFR part 56) is an important component of assuring that human research subjects receive adequate protections before and during study conduct. There are several steps that study sponsors can take in advance of initiating an adaptive clinical study that can minimize or avoid critical IRB-related delays during the study.

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.

An IRB's familiarity with adaptive design clinical studies may impact the efficiency with which they are able to review such studies and study modifications. For example, some IRBs may require the resubmission of the study protocol for full board review when an adaptation is made. If pre-specified adaptations were not disclosed to the IRB during the initial approval process, the sponsor risks critical IRB-related delays that can hinder study progress. Failure to disclose the adaptive nature of the study and its associated risks in the initial informed consent document may result in an IRB-mandated reconsenting of study subjects or subject notification related to the study modifications or identified risks.

Advanced planning and good communication with the IRB can mitigate these potential IRB-related issues.

D. Logistical Challenges

The conduct of an adaptive clinical study creates several logistical challenges. A robust infrastructure is needed to ensure that the adaptive design is implemented appropriately. All parties that will be involved in the management and implementation of the study should have a thorough understanding of the principles of adaptive design. Efficient and reliable data

management should be a priority. Mid-course changes to the sample size may create challenges regarding the timely availability of a sufficient number of investigational devices. A robust and comprehensive set of standard operating procedures to ensure that the outcome results remain sufficiently blinded or masked should be used.

11. Regulatory Considerations

A. Interactions with FDA

The FDA is committed to timely evaluation of clinical study protocols through its IDE program. Sponsor - FDA interactions and communication are the best and most efficient ways to ensure that the Agency understands the sponsor's plans and device development strategy and that sponsors understand the FDA's recommendations regarding maximizing study efficiency and chances for success.

Although a study sponsor may directly submit an 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. These "pre-submission" meetings are intended to promote dialogue and interactive exchange of perspectives, and allow sponsors to understand the FDA expectations for a pivotal adaptive design clinical study. The FDA guidance entitled "Requests for Feedback on Medical Device Submissions: The Pre-Submission Program and Meetings with Food and Drug Administration Staff[14]" outlines the procedures that sponsors can follow when seeking the FDA's feedback on specific questions relating to a proposed adaptive design clinical study.

Sponsors can use this pre-submission program to obtain Agency feedback on both investigational studies of significant risk (SR) devices as defined in 21 CFR 812.3 (which require FDA approval of an IDE application) as well as studies of non-significant risk (NSR)

devices (which require only IRB oversight) or device studies that will be conducted outside of the United States (OUS). For studies of SR devices conducted in the U.S., the adaptive design clinical study protocol, including the statistical analysis plan, will be recorded within the approved IDE and/or subsequent IDE supplements. In the case of certain NSR and OUS device studies, sponsors may choose to submit the final version of the study protocol as a submission that incorporates Agency feedback obtained from the pre-submission, but are not required to do so. Such documentation may assist in assuring a mutual understanding of the proposed study by the sponsor and the FDA.

During the course of the conduct of an adaptive design clinical study involving a SR device, the FDA should be informed of any deviations from the planned adaptive process and/or procedures for maintaining study integrity in a timely fashion.[15] The FDA should also be made aware of any breaches of the study firewall that was established and described in the approved investigational protocol.

B. Sponsor Monitoring

Sponsors are advised to have a risk-based monitoring plan in place which focuses on specific aspects of adaptive studies that are of particular importance and may not be present in traditional (non-adaptive) trial designs. The FDA has issued a guidance document entitled "Guidance for Industry Oversight of Clinical Investigations: A Risk-Based Approach to Monitoring"[16] in which the FDA recommends for all clinical investigations, adaptive or not, that sponsors consider adopting a risk-based monitoring approach that focuses on critical study parameters and relies on a combination of monitoring techniques (such as on-site monitoring and centralized monitoring) to oversee the study. For adaptive studies, sponsors

---

[14]http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM311176.pdf

[15] Please refer to 21 CFR 812.30, which describes when these changes must be submitted in an IDE Supplement.

[16] http://www.fda.gov/RegulatoryInformation/Guidances/ucm373750.htm

should have a pre-determined monitoring plan in place to ensure adequate monitoring if the preplanned changes do occur. When an adaptation is planned, sponsors should consider adopting procedures such as preplanned site visits scheduled to verify adequate documentation and execution of blinding procedures in order to ensure blinding was appropriately maintained. Additionally the monitoring plan should include procedures that confirm that data firewalls have not been breached and that statistical changes were made according to the study Statistical Analysis Plan.

C. Best Practices to Protect Study Blinding (Masking)

Sponsors should provide to the FDA, in advance, sufficient evidence of a firewall and documented policies and information that will ensure that personnel are appropriately blinded/masked during the conduct of the adaptive study. Unplanned changes in study design that occur after an unblinded interim analysis of study data are not considered adaptive and in many cases, may undermine the scientific validity of the study. Additional principles and details are available in "Design Considerations for Pivotal Clinical Investigations for Medical Devices,"[4] including, in particular, Section 9, "Sustaining the Quality of Clinical Studies," and the subsections on Handling Clinical Data, Study Conduct, and Study Analysis, and Anticipating Changes to the Pivotal Study.

D. 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 (see Section 11.B.) and role of the DMC (see Section 10.A.). 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.

If a firewall is part of the design, a mechanism and an implementation plan for the firewall should be provided. If the intent is to limit the information provided to the investigators, a general clinical protocol and a separate detailed Statistical Analysis Plan (SAP) could be used, with the SAP not widely distributed. Computer systems can be employed to monitor, document and limit access and can provide audit trails and firewalls.

At the conclusion of an adaptive study, the documentation that should be sent to the FDA should include a description of how the adaptation was implemented, the data sets for the study, the baseline subject characteristics for pre and post-adaptation subgroups, the pre-specified statistical analysis, and any deviations that may have occurred from the protocol's adaptive plan and how they have been addressed in additional analyses.

12. Conclusion

Adaptive clinical study designs for investigational medical devices can improve efficiency and increase the likelihood of study success when conducted in a pre-specified, thoughtful, scientifically-valid manner. Adaptive studies may also 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. A well-planned adaptive study, designed with thoughtful anticipation of situations that would warrant changes during the study period, can substantially improve the chances that the study will demonstrate device effectiveness.Procedures to ensure the proper conduct of adaptively-designed studies should be put into place so the study will provide valid scientific evidence that can be relied upon by the FDA to assess the benefits and risks of the investigational medical device. Sponsors are strongly encouraged to discuss the planning of adaptive clinical study designs with the appropriate FDA review division in advance; the Agency has established mechanisms to enable such interactions in a timely and efficient manner.

13. References

General

Campbell, G., Yue, L., Gao, Y., Ho, M., Li, H., Mukhi, V., Thompson, L. (2012) Quality Statistical Review Checklist of Investigational Device Exemption (IDE) Submissions for Therapeutic and Aesthetic Medical Devices, *Proceedings of the Joint Statistical Meetings, Biopharmaceutical Section,* Alexandria, VA, American Statistical Association, 845-861.

Group Sequential Methods

Lan, K.K.G., and DeMets, D.L. (1983), "Discrete sequential boundaries for clinical trials." *Biometrika* **70**, 659-663.

Lan, K.K.G., and DeMets, D.L. (1987), "Group sequential procedures: calendar versus information time." *Statistics in Medicine*, **8**, 1191-98.

O'Brien, P.C., and Fleming, T.R. (1979), "A multiple testing procedure for clinical trials." *Biometrics* 35, 549-556.

Pocock, S.J. (1977), "Group sequential methods in the design and analysis of clinical trials." *Biometrika,* **64**, 191-199.

Rosenberger, W.F., Stallard, N., Ivanova, A., Harper, C.N., and Ricks, M.L. (2001). "Optimal adaptive designs for binary response trials." *Biometrics* **57**, 909-913.

Sample Size Reassessment (With or Without Group Sequential)

Bauer, P., and Koenig, F. (2006), "The reassessment of trial perspectives from interim data—a critical view," *Statistics in Medicine* **25**, 23–36.

Chen, Y.H.J., DeMets, D.L., and Lan, K.K.G. (2004), "Increasing the sample size when the unblinded interim result is promising," *Statistics in Medicine* **23**, 1023–1038.

Cui, L., Hung, H.M., and Wang, S. (1999). "Modification of sample size in group sequential clinical trials." *Biometrics* **55**, 853-857.

Denne, J.S. (2001), "Sample size recalculation using conditional power," *Statistics in Medicine*, **20**, 2645–2660.

Gao, P., Ware, J.H., and Mehta, C. (2008), "Sample size re-estimation for adaptive sequential design in clinical trials." *Journal of Biopharmaceutical Statistics* **18**, 1184-1196.

Jennison, C., and Turnbull, B.W. (2006) "Adaptive and nonadaptive group sequential tests," *Biometrika* **93**(1), 1-21.

Lehmacher, W., and Wassmer, G. (1999), "Adaptive sample size calculations in group sequential trials," *Biometrics* **55**, 1286–1290.

Li, G., Shih, W.J., Xie, T., and Lu, J. (2002), "A sample size adjustment procedure for clinical trials based on conditional power," *Biostatistics* **3**, 277–287.

Mehta, C.R., and Pocock, S.J. (2011), "Adaptive increase in sample size when interim results are promising: A practical guide with examples." *Statistics in Medicine* **30**, 3267-3284.

Müller, H., and Schäfer, H. (2001), "Adaptive group sequential designs for clinical trials: Combining the advantages of adaptive and of classical group sequential approaches." *Biometrics* **57**, 886-891.

Proschan, M.A. and Hunsberger, S.A. (1995), "Designed extension of studies based on conditional power." *Biometrics* **51**, 1315-1324.

Tsiatis, A.A., and Mehta, C. (2003), "On the inefficiency of the adaptive design for monitoring clinical trials," *Biometrika* **90**, 367–378.

Information based adaptation

Mehta, C.R., and Tsiatis, A.A. (2001), "Flexible sample size considerations using information based interim monitoring," *Drug Information Journal* **35**:1095-1112.

Covariate Adaptive Randomization

Efron, B. (1971), "Forcing a sequential experiment to be balanced," *Biometrika* **58**, 403-417.

Response Randomization

Hu, F., and Rosenberger, W. (2006). *The Theory of Response-Adaptive Randomization in Clinical Trials*. New York: Wiley.

Berry, D.A., Eick, S.G. (1995), "Adaptive assignment versus balanced randomization in clinical trials: a decision analysis," *Statistics in Medicine* **14**, 231-246.

Bayesian (Adaptive) Design

Berry, S.M., Carlin, B.P., Lee, J.J., and Muller, P. (2010). *Bayesian Adaptive Methods for Clinical Trials*. Boca Raton, FL: CRC Press.

Campbell, G. (2013), "Similarities and differences of Bayesian designs and adaptive designs for medical devices: A regulatory view," *Statistics in Biopharmaceutical Research*, **5**, 356-368.

DMC

Food and Drug Administration. (2006). "The Establishment and Operation of Clinical Trial Data Monitoring Committees for Clinical Trial Sponsors" (issued March, 2006). http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm127073.pdf

Antonijevic Z, Gallo P, Chuang-Stein C, Dragalin V, Loewy J, Menon S, Miller E, Morgan C, & Sanchez M (2013). "Views on Emerging Issues Pertaining to Data Monitoring Committees for Adaptive Trials," Therapeutic Innovation & Regulatory Science, 47: 495-502.

o Sanchez-Kam M, Gallo P, Loewy J, Menon S, Antonijevic Z, Christensen J, Chuang-Stein C, & Laage T (2014). "A Practical Guide to Data Monitoring Committees in Adaptive Trials," Therapeutic Innovation & Regulatory Science, 48: 316-326.