Draft Guidance for Industry and Food and Drug Administration Staff

---

Adaptive Designs for Medical Device Clinical Studies

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.
Document issued on: May 18, 2015

You should submit comments and suggestions regarding this draft document within 90 days of publication in the Federal Register of the notice announcing the availability of the draft guidance. Submit written comments to the Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. Submit electronic comments to http://www.regulations.gov. Identify all comments with the docket number listed in the notice of availability that publishes in the Federal Register.

For questions regarding this document that relate to devices regulated by CDRH, contact Dr. Greg Campbell (CDRH) at 301-796-5750 or by e-mail at greg.campbell@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 or240-402-7800.

.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Devices and Radiological Health
Center for Biologics Evaluation and Research

Preface

Public Comment

Written comments and suggestions may be submitted at any time for Agency consideration to the Division of Dockets Management, Food and Drug Administration, 5630 Fishers Lane, Room 1061, (HFA-305), Rockville, MD, 20852.

Additional Copies

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.

Additional copies of this guidance document are also available from:
Center for Biologics Evaluation and Research (CBER),
Office of Communication, Outreach and Development,

10903 New Hampshire Ave., Bldg 71, Rm. 3128or by calling 1-800-835-4709 or 240-402-7800,
or email ocod@fda.hhs.gov, or from the internet at
http://www.fda.gov/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/default.htm.

Table of Contents

Draft Guidance for Industry and Food and Drug Administration Staff

Adaptive Designs for Medical Device Clinical Studies

This draft guidance, when finalized, will represent the Food and Drug Administration's (FDA's) current thinking on this topic. It does not create or confer any rights for or on any person and does not operate to bind FDA or the public. You can use an alternative approach if the approach satisfies the requirements of the applicable statutes and regulations. If you want to discuss an alternative approach, contact the FDA staff responsible for implementing this guidance. If you cannot identify the appropriate FDA staff, call the appropriate number listed on the title page of this guidance.

1. Introduction and Scope

An adaptive design for a medical device clinical study is defined as a clinical trial design that allows for prospectively planned modifications based on accumulating study data without undermining the trial's integrity and validity. Adaptive designs, when properly implemented, can reduce resource requirements 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 trials 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 trials. This guidance does not apply to clinical studies of combination products or codevelopment of a pharmaceutical product with an unapproved 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 trial design that allows for prospectively planned modifications based on accumulating study data without undermining the trial's integrity and validity.[1] In nearly all situations, in order to preserve the integrity and validity of a trial, modifications should be prospectively planned and described in the clinical trial protocol prior to initiation of the study. However, in some specific circumstances, study modifications after the trial begins can be scientifically valid if the trial design decision-makers have had no access to the outcome results by treatment.[2] The different types of adaptive trial 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 statistically 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 trial, 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 may not be approved or endorsed by FDA. Sponsors should anticipate and plan for modifications based on a variety of possible scenarios that could occur during the course of the trial.

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 to an autologous bone graft. 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 non-inferiority study compared 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 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.

C. Advantages of Adaptive Designs

An adaptive study design can have several distinct advantages when compared to an unchanged (fixed) 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 trial with a preplanned interim analysis could decide to stop early for futility.
- Adaptive designs can improve the chance of trial success by employing sample size reassessment. Based on accumulating data in the trial, planned sample size reassessment could lead to an adjustment in sample size (for example, if treatment effect is smaller than anticipated), converting an underpowered study likely to fail into a well-designed study more likely to succeed. This approach can salvage studies otherwise likely to be unsuccessful and as a result, help facilitate the timely assessment and marketing of medical devices demonstrating a reasonable assurance of safety and effectiveness.
- It can yield an improved understanding of the effect of the investigational treatment and a better understanding of benefit and risk.
- 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 Draft Guidance "Balancing Premarket and Postmarket Data Collection for Devices Subject to Premarket Approval."[3] http://www.fda.gov/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm393882.htm

- In some cases planned modifications can incur no cost in either sample size increase or false positive error inflation provided there is a strong blind to outcomes by treatment groups.
- Adaptive designs can enhance patient protection by increasing the probability that a patient is allocated 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.
- 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.

[3] As of January, 2015, the reference is a draft guidance distributed for comment purposes only and therefore not for implementation.

D. Limitations of Adaptive Designs

The following are some of the possible limitations associated with an adaptively designed 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, and be logistically difficult to carry out.
- If not done correctly, adaptive designs can introduce 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 8.B.)

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 employ statistical sample size calculations in order to draw valid conclusions. 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" http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm373750.htm.

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.J., 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

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 the results from such a flawed study may not be valid.
- Modifications based entirely on information from a source completely external to the study.

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

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

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 whether an adaptive design is feasible and advantageous compared to a fixed (non-adaptive or conventional) design.

A. When are Adaptive Designs Appropriate and When Not?

When studies enroll subjects rapidly, there may not be time to make changes to the study design. For example, if subjects are recruited quickly and reach the final follow-up at virtually the same time, it may be infeasible to adapt the sample size. In such cases sponsors may consider slowing down enrollment to allow time to learn from the accumulating data and make preplanned adaptations. Adaptive designs may not be suitable for very complex studies that have multiple primary endpoints or multiple secondary endpoints for claims. Studies with shorter endpoints but longer recruitment times may lend themselves to adaptation. Studies in which the time to the primary endpoint evaluation is long but the accrual is even longer may benefit from an adaptive design.

For a fixed (non-adaptive) design, the sample size calculation is usually based on assumed values of several parameters. A basic question is how much confidence is there 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 fixed 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 fixed design based on a more optimistic effect size would likely lead to a failed study for the sponsor. In contrast, an adaptive design planned with an interim analysis to reassess the sample size could convert what would have been an unsuccessful study into a successful one. An adaptive design can guard against these uncertainties by learning from accumulating data during the study.

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 (fixed) 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 fixed (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 fixed design. For non-adaptive designs this is usually straightforward. The topic of how to calculate these quantities for adaptive designs will be discussed later, using either analytical techniques or computer simulation. Ultimately, the decision may rest on the sponsor's confidence in the anticipated parameter values and willingness to risk a failed study such that a fixed design would be preferred over an adaptive one.

C. Anticipated Regret

It is sometimes helpful to anticipate particular study outcomes that could lead to failure so as to ask what one might have regretted in the planning. 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 have regretted 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.D.).

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.[4] These principles are crucial to assure that a clinical study produces valid scientific evidence. If the chance of erroneous positive conclusions is unacceptably large it will be very unlikely that the results will be reproducible. If the chance of erroneous negative conclusions is large, the study may fail to show the

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

device's true effectiveness. In short, studies that fail to follow these principles could generate evidence that is either inadequate or invalid. 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 provide inadequate evidence.

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 of the data 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 (or

297 false positive) error, where a Type I error would be to erroneously conclude that a device
298 was effective when in fact it was not. A Type II (or false negative) error would be failing to
299 conclude that a device was effective when in fact it was. The (statistical) power of a study is
300 the probability of correctly concluding that the device is effective and is 1 minus the
301 probability of a Type II error.

302 There are usually two approaches for evaluating the operating characteristics of
303 adaptive study designs for regulatory submissions: analytical methods and simulation
304 studies. Analytical statistical methods are often used in some frequentist adaptive study
305 designs and can provide approximate probabilities for Type I errors and for statistical power
306 for fixed and simple adaptive designs under different scenarios. Simulations can be used to
307 obtain operating characteristics for complex frequentist and Bayesian adaptive designs.
308 Analytical methods and simulation studies could be complementary to each other in
309 evaluation of the Type I error rate and power of adaptive study designs. In adherence to
310 regulatory practice, FDA strongly recommends sponsors control the Type I error rate and
311 maintain adequate power for all study designs.

312 ## *B. Minimization of Operational Bias*

313
314 One type of bias frequently encountered in studies with adaptive designs is the
315 operational bias (defined in footnote 5) which can arise in the conduct of the clinical study.
316 It is important that bias of all kinds be reduced or eliminated because the presence of bias can
317 distort the findings of a clinical study and undermine its scientific validity. For example, in a
318 two-arm study, if an interim analysis is conducted resulting in an increased sample size in a
319 preplanned manner, investigators, study subjects and/or third-party evaluators may behave
320 differently, either unconsciously or subconsciously, if the existence or siz of the increase, or
321 the reason for the increase, becomes known to them. As a consequence, bias may be
322 introduced into the clinical study. Knowledge that the size of the study has been increased
323 may help participants to estimate the magnitude of the interim treatment effect, which in
324 turn, can then affect the ongoing conduct of the study in various ways. If not blinded to the
325 patients' treatment assignment, the investigator may, unintentionally and without being

aware, change the decision about whether to enroll a subject in the study or start treating the subjects in the investigational treatment group in manner that is different from that applied to 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 analysts of the study data have access to the unblinded results of an adaptive trial during its conduct, it is vital that policies and procedures be in place 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" will be discussed in Section 9.C. 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 unblinded) are 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).

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

354 covariate adaptive randomization; it uses accumulating baseline data in an attempt to provide
355 better balance between the two groups.
356 A classic example of adaptation based on aggregate outcomes that is widely used is to
357 power a time-to-event study or a survival study not by the number of patients in the study but
358 by the total number of clinical events. The study continues until the desired number of
359 events has been observed. For such studies, the exact number of subjects cannot be planned
360 in advance. One is using the accumulating evidence from the study in the form of the
361 aggregate results, in this case the total number of events, although the number in each of the
362 comparative groups would not be revealed in either an unblinded or coded fashion to the
363 investigators. The knowledge of the total number of events could lead to changing the total
364 number of patients or to an extension of the duration of the study.
365 As another example of using aggregate results with multiple treatment groups without
366 breaking the blind, one could observe the pooled overall success rate and, assuming two
367 groups that differ by a hypothesized amount, infer that the original assumptions about the
368 control rate and the investigational rate cannot be valid and that a change in sample size is
369 merited. As yet another example, it is possible to calculate the overall variance for a
370 continuous endpoint and make a sample size adjustment based on the hypothesized
371 difference in the means.
372 In the prior two examples, the required amount of aggregate information is determined in
373 advance in order to make a prospective decision and continue the study until that information
374 is obtained.
375 If the blind is maintained so that the decision-makers have no access to the outcomes by
376 coded or unblinded treatment group in the case of a comparative study or have no access to
377 (or are firewalled off from) any outcomes if the study is unblinded in a one-arm study, then
378 such adaptive designs pose no theoretical scientific difficulty. Sponsors are encouraged to
379 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 Investigational Device Exemption, when appropriate.[5]

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 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 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 the clinical trial community. These designs usually prescribe one or more planned interim looks of unblinded data with the possibility of stopping the study at an interim look 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 fixed (non-adaptive) sample size designs while

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

controlling the overall Type I error rate of the study. 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 require prespecified 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 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 fixed 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. Sponsors often find that a Data Monitoring Committee (DMC) may be helpful 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. Another disadvantage is the possibility of operational bias after a decision to continue at an interim analysis since a trial participant could conclude that the effect size is not sufficiently large to stop the study.

B. Sample Size Adaptation

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 prespecified 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 prespecified 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 prespecified 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 prespecified manner. Such a strategy can control the chance of erroneous positive conclusions and produce scientifically valid inferences.

Adaptive designs using 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 pre-planned 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 the SSR. Some methodologies use conditional power and others predictive probability. 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 FDA reach agreement prior to study initiation on the study sample size needed to demonstrate the minimal clinically important difference (MCID) in treatment effect. The decision concerning whether a smaller effect 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; if it is done too early, it may produce an inaccurate or variable result based on relatively few patients. Analytical calculations or computer simulations performed under different scenarios can help guide the choice of optimal point(s) for the reassessment. (See Section 7.D. for more discussion on simulations.) The 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 surrogate or intermediate endpoints known to be associated with the primary endpoint. For more information on the use of surrogate and intermediate endpoints is discussed in the draft guidance "Expedited Access for Premarket Approval Medical Devices Intended for Unmet Medical Need for Life Threatening or Irreversibly Debilitating Diseases or Conditions,"[6] (http://www.fda.gov/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm393879.htm). The use of a Bayesian model that learns from the accumulating data of the surrogate or 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 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 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.

---

[6] As of January, 2015, the reference is a draft guidance distributed for comment purposes only and therefore

C. Bayesian Sample Size Adaptation

Most Bayesian designs include sample size adaptation, since several factors that determine the sample size of a Bayesian trial, such as effect size, variability of the sample, 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 borrowed increases and it increases as variability of the sample 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 and that correlation 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 then its parameters 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 a way that the sample size is adjusted as information accumulates. As noted above, this idea is referenced in the draft guidance document "Expedited Access for Premarket Approval Medical Devices Intended for Unmet Medical Need for Life Threatening or Irreversibly Debilitating Diseases or Conditions."[7]

---

not for implementation.

[7] As of January, 2015, the reference is a draft guidance distributed for comment purposes only and therefore not for implementation.

(http://www.fda.gov/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm393879.htm)

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 thus far. If the probability is sufficiently high (above a pre-specified value), the trial may stop for early success; if the probability is too low (below a pre-specified value), the trial may stop for futility; and if in between, it may warrant continuation 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 generally requires simulations for assessment of its operating characteristics; the performance of the design depends on preselected parameter values. 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.D. on simulation and for a more detailed discussion, refer to FDA "Guidance on the Use of Bayesian Statistics in Medical Device Clinical Trials."
http://www.fda.gov/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm071072.htm

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

E. Dropping a Treatment Arm

In a study in which there is more than one experimental arm, 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 in a two-arm study based on treatment outcomes is called treatment response adaptive randomization. Treatment response adaptive randomization can mitigate 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. Treatment response adaptive randomization can sometimes lead to slightly larger studies but could facilitate investigator and patient enrollment.

G. Changing the Hypothesis (Claim)

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 prespecified 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 prespecified. However, study planners may wish to incorporate a preplanned interim assessment and prespecified sample size reassessment in case mid-course results are sufficiently promising that a superiority claim may be within reach; such adaptations must be prespecified.

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 that a claim may be possible for both 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. Because there is no allowance to stop early, there is also 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 means 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 fixed 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 fixed 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 prespecified stopping rule. Intentionally overrunning the sample size can result in a variance that is smaller than agreed upon, and a misalignment of statistical significance and the MCID. As a result, statistical significance may be established that may not demonstrate clinical importance. 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 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.C. For planned changes, study planners are advised to prespecify the changes (or anticipated types of changes) and account for them in a prespecified 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. In such cases, 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 studies.

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 typically require sufficient planning and complete masking of the outcome results by treatment group, such that no one representing the sponsor (or the FDA) has access to the coded or unblinded outcome results by treatment group. A major advantage of conducting a study where the outcome by coded or unblinded treatment groups are 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 Data Monitoring Committee (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 those with access to the outcome data by treatment group, including the DMC, are not appropriate groups to propose or provide input concerning study revisions. The discussion of "firewalls" to prevent inappropriate disclosure of information is discussed further in Section 9.C. Unplanned study changes under appropriate circumstances are scientifically viable and should be discussed with 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 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 a study that requires an IDE, if the change or modification affects the validity of the data or information, the patient benefit-risk relationship, the scientific soundness of the plan, the rights, safety, or welfare of the subjects, or represents a device/manufacturing change that is a significant change in the design or basic principles of operation, then an IDE Supplement is required; otherwise a 5-day notice suffices.

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 than a pivotal study. For example, the use of pivotal study data to assess, modify, and finalize an algorithm for a diagnostic device may raise concern for biased performance due to over-fitting. In contrast, this approach may be acceptable if the finalization of the algorithm was a preplanned adaptation (for example, the choice of the threshold) with a prespecified analysis plan that adequately controls 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 (fixed) design. Simulations can provide insights into required samples sizes, operating characteristics, and interrelationships between trial design choices and patient characteristics that cannot easily be obtained in other ways.

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. Fourth, the simulations for an adaptive design are often dependent on the anticipated study subject accrual rate; therefore, the simulations should consider a variety of possible accrual patterns.

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 5%. Specifically, a group sequential approach could be used to allow for one or more interim looks and an early study termination if the serious adverse event rate was much less than 5%. Alternatively, one could develop a stopping rule that would terminate the study if there were no adverse events in a prespecified 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 least 5%.

E. Adaptive Designs for Open-Label Randomized Studies

Unlike drug trials, many scientifically valid medical device studies are not, or cannot, be masked. 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 is being used. While in some cases, patients, third-party assessors, or even the health care provider can be masked, there are many instances where this is not possible. Studies where masking 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.

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 operating characteristics checked. 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 prespecified 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 by treatment group (unblinded or coded) should be carefully restricted. A log of all incoming subjects (including those not included in the study) to each clinical site can help to reduce possible manipulation of the trial findings.

H. 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 on Design Considerations for Pivotal Clinical Investigations for Medical Devices"
http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm373750.htm and would be useful to review if considering an adaptive design for a diagnostic device. Diagnostic performance is often evaluated using estimation or 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 rapid and follow-up is not required. Thus, in some cases, the rationale for pursuing an adaptive study for a therapeutic device may not be relevant for a study of a diagnostic device.

Nevertheless, because diagnostic devices are heterogeneous in scope, there may be circumstances where an adaptive design is advantageous.

I. 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 an established clinical reference standard rather than 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. If the group making decisions about the adaptation is unblinded only to the clinical reference standard[8] results, 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 10.C.).

J. 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, some of which may require an adjustment to the confidence interval or Type I error rates.

K. Adaptation and Staged Designs

[8] The clinical reference standard is defined for this guidance 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."

For some IDE submissions, FDA may approve the number of subjects in a staged manner as described in FDA Guidance "FDA Decisions for Investigational Device Exemption (IDE) Clinical Investigations." http://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm279107.pdf

The staged approval of an IDE allows the FDA to grant IDE approval or approval with conditions for a portion of the intended study cohort. This allows timely study initiation with an opportunity for an evaluation of the safety of the early subjects in the study before exposing a large number of subjects to the investigational device. An adaptive study design could also allow for prespecified study modifications based on the accumulating effectiveness results, as long as interim effectiveness results by treatment group remain masked to those responsible for the study design modifications.

8. Two 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 the specific challenges of analysis of data from adaptively designed studies; however, a detailed discussion is beyond the scope of this guidance.

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 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. 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.

9. 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 prespecified 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 unmasked outcomes, robust prespecified and well-documented procedures must 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 decision-making related to use of the adaptation to an independent statistician, a contract research organization, or some other clinical trial body. In any case, the underlying validity and integrity of the study depends on study adaptation decision-making and implementation and must always be paramount when planning the construct of these studies.

Although the DMC may be tempted to recommend changes to the adaptive design or to the fundamental study type (e.g., from a fixed study to an adaptive one) during study conduct, once the DMC has access to coded or unmasked 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.

B. 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 prespecified adaptations were not disclosed to the IRB during the initial approval process, the sponsor risks critical IRB-related delays that can hinder study progress. 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.

C. 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 such data are restricted only to those for whom it is absolutely essential. This is required 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 removed from the protocol and placed in a separate detailed Statistical Analysis Plan for the adaptive design. 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 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 given to control the information released so that only those who have absolute necessity 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. 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 pre-planned 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.

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 must 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 is also required.

10. Regulatory Considerations

A. Interactions with FDA

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 assure that the Agency understands the sponsor's plans and device development strategy and that sponsors understand 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 "presubmission" meetings are intended to promote dialogue and interactive exchange of perspectives and allow sponsors to obtain clarity with respect to FDA expectations for a pivotal adaptive design clinical study.

The Guidance for Industry and FDA Staff entitled: "Requests for Feedback on Medical Device Submissions: The Pre-Submission Program and Meetings with Food and Drug Administration Staff " (February 18, 2014) http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM311176.pdf outlines the procedures that sponsors can follow when seeking 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 presubmission, which 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 FDA.

During the course of the conduct of an adaptive design clinical study involving a SR device, FDA should be informed of any deviations from the planned adaptive process and/or procedures for maintaining study integrity in a timely fashion. [9]FDA should also be made aware of any breeches of the study firewall that was established and described in the approved investigational protocol.

---

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

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. FDA has issued a guidance document entitled "Guidance for Industry Oversight of Clinical Investigations: A Risk-Based Approach to Monitoring" (http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM269919.pdf) in which 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 should have a pre-determined monitoring plan in place to ensure adequate monitoring if the pre-planned changes do occur. When an adaptation is planned, sponsors should consider adopting procedures such as pre-planned 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 FDA sufficient evidence of a "firewall" and documented policies and information in advance that will assure personnel are appropriately blinded/masked during the conduct of the adaptive study. 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 "Guidance for Industry, Clinical Investigators, Institutional Review Boards and Food and Drug Administration Staff on Design Considerations for Pivotal Clinical Investigations for Medical Devices."

(http://www.fda.gov/RegulatoryInformation/Guidances/ucm373750.htm), 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 FDA

Submissions to 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 10.B.) and role of the DMC (see Section 9.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 FDA.

If a firewall is part of the design, a mechanism and an implementation plan for the firewall should be provided. If a firewall is intended to provide only limited information 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 the how the adaptation was implemented, the data sets for the study, the baseline population 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.

11. 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. The anticipation of possible study changes in advance can reap great dividends for well-planned adaptive studies. Procedures to assure the proper conduct of adaptively designed studies must be put into place so the study will provide valid scientific evidence that can be relied upon by 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, and the Agency has established mechanisms to conduct such interactions in a timely and efficient manner.

12. 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.

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.

Covariate Adaptive Randomization

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

Treatment 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.

Food and Drug Administration. (2010), "Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials" (issued February 5, 2010). http://www.fda.gov/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm071072.htm