INTERNATIONAL COUNCIL FOR HARMONISATION OF TECHNICAL REQUIREMENTS FOR PHARMACEUTICALS FOR HUMAN USE

ICH HARMONISED GUIDELINE

ADAPTIVE DESIGNS FOR CLINICAL TRIALS
E20

Draft version

Endorsed on 25 June 2025

Currently under public consultation

At Step 2 of the ICH Process, a consensus draft text or guideline, agreed by the appropriate ICH Expert Working Group, is transmitted by the ICH Assembly to the regulatory authorities of the ICH regions for internal and external consultation, according to national or regional procedures.

Explanatory Note for ICH E20: The draft guideline makes statements on adaptive design approaches for clinical trials. The draft guideline acknowledges the high potential for adaptive designs to accelerate the process of drug development and to allocate resources more efficiently without lowering scientific and regulatory standards. Some of the approaches may affect the nature and timing of interactions between industry and regulators at confirmatory trial planning and assessment. The final guideline will indicate key adaptive design principles and approaches for which discussion of adaptive design features, and the rationale for their use, are particularly critical at the planning stage. To inform guideline finalization, specific feedback is sought on adaptive design principles and approaches and their impact on industry-regulatory interactions. Until a final guideline is agreed under Step 5 of the ICH process, the draft guideline should not be understood as confirming full regulatory acceptance from ICH parties of its contents, nor superseding current regional guidance, which remains valid. Public consultation comments on the draft guideline are sought.

E20
Document History

Code	History	Date
E20	Endorsement by the Members of the ICH Assembly under *Step 2* and release for public consultation.	25 June 2025

ICH HARMONISED GUIDELINE

ADAPTIVE DESIGNS FOR CLINICAL TRIALS

E20

ICH Consensus Guideline

TABLE OF CONTENTS

1. INTRODUCTION AND SCOPE

This document provides guidance on confirmatory clinical trials with an adaptive design intended to evaluate a treatment for a given medical condition within the context of its overall development program. For the purpose of this guideline, an adaptive design is defined as a clinical trial design that allows for prospectively planned modifications to one or more aspects of the trial based on interim analysis of accumulating data from participants in the trial. The term prospectively planned means that the potential trial adaptations are pre-specified in the clinical trial protocol prior to initiation of the trial. The scope of this guideline does not include trials with unplanned modifications to the design, such as a protocol amendment proposed by an independent data monitoring committee (IDMC) based on unexpected interim results. It also does not include design changes based entirely on emerging information from a source external to the trial. Routine monitoring of operational aspects such as the enrollment rate, data quality, or extent of participant withdrawal is also out of scope.

The focus of this guideline is on principles for the planning, conduct, analysis, and interpretation of trials with an adaptive design intended to confirm the efficacy and support the benefit-risk assessment of a treatment. The emphasis is on principles that are critical to ensuring the trials produce reliable and interpretable information and that require specific considerations with use of an adaptive design. This guideline does not discuss the use of specific statistical methods. Although the guideline primarily focuses on confirmatory clinical trials, the principles outlined are relevant to all phases of clinical development.

2. ADVANTAGES AND CHALLENGES OF ADAPTIVE DESIGNS

At the planning stage of confirmatory trials, uncertainty may remain regarding design aspects such as the appropriate sample size, even after careful planning and conduct of earlier phases of drug development. Yet, with a non-adaptive design, these aspects have to be determined before the trial starts and cannot be changed during trial execution. Adaptive designs provide flexibility and the ability to safeguard against inaccurate assumptions by taking advantage of the accumulating information from trial participants and allowing pre-specified modifications to design aspects during the trial.

This added flexibility can lead to a variety of advantages. First, adaptive designs can provide ethical advantages. For example, a group sequential design with the potential for early trial stopping if there is convincing evidence the treatment is efficacious and has a positive benefit-

risk profile can reduce the number of participants exposed to an inferior control. Second, adaptive designs can improve the efficiency of a trial, for example, by increasing its power for a given expected sample size. Third, adaptive designs can help improve understanding of treatment effects and decision-making. For example, a confirmatory two-stage adaptive design with selection between two doses at an interim analysis may reduce uncertainty about the dose with the better benefit-risk profile while also allowing for confirmation of the efficacy of the selected dose.

However, adaptive designs also present challenges, as they may add complexities and uncertainty related to the key principles discussed in Section 3. For example, use of an adaptive design may add logistical difficulties in maintaining confidentiality of interim results and introduce risks to trial integrity which, if not properly addressed, may lead to unreliable results and complications with their interpretation at trial end. In addition, appropriate planning for and assessment of a trial with an adaptive design can be more complex and may require more time than for a trial without an adaptive design. In particular, use of conventional analysis methods that would apply in non-adaptive designs usually lead to an increased Type I error probability and biased treatment effect estimate. For example, in a design with an interim analysis to modify the target sample size based on the estimated treatment effect, the Type I error probability can be more than doubled when using analysis methods that do not account for the adaptation. As another example, the potential for early stopping for efficacy may lead to biased treatment effect estimates because the trial will be stopped preferentially when extreme data have been observed. Therefore, special analysis methods for hypothesis testing and estimation that account for the adaptive design usually need to be used. In addition, some trials with adaptive designs may provide less information about safety, potentially leading to more uncertainty during benefit-risk assessments. Also, adaptive designs may not be beneficial in all clinical trial settings. For example, adaptive designs may not be favored if there is fast enrollment of participants relative to the assessment time of the endpoint on which the adaptation is based, or if data cannot be made available quickly enough to facilitate reliable adaptation decisions at an interim analysis.

The decision to use or not use a specific adaptive design in a clinical trial will depend on many factors, including the ones described above. There can be a tension between the confirmatory nature of a late-stage clinical trial and the proposal to adapt aspects of the trial while it is ongoing. In planning an adaptive design, it is therefore essential to carefully justify the need to

adapt the trial and assess potential implications of the type, number, and complexity of the adaptations involved. The justification should include both clinical and statistical considerations. It should weigh the advantages of the design against the extent to which the adaptations being considered add uncertainty about the trial's ability to produce reliable and interpretable results. For example, the addition of a carefully planned interim analysis to potentially stop a trial early for efficacy or futility using appropriate pre-specified stopping rules and ensuring sufficient information for safety and benefit-risk assessment, along with use of an IDMC to maintain trial integrity, may add minimal uncertainty. On the other hand, a complex design involving adaptations to multiple trial features may add considerable uncertainty related to maintaining trial integrity. This could include uncertainty about the adequacy of information flow and data access specifications, or the potential impact of the adaptation itself on trial conduct and the trial's ability to provide interpretable treatment effects. This can lead to challenges in assessing results and in regulatory decision-making about the efficacy and benefit-risk profile of a proposed dose of a treatment for a specific patient population. A proposed adaptive design requires a clear and compelling justification. This justification should discuss how the proposed design addresses inherent needs of the clinical setting and should provide an evaluation of advantages and limitations as compared to alternative designs (including non-adaptive designs), including a comparison of important trial operating characteristics (e.g., power, expected sample size, reliability of adaptation decisions) between candidate designs.

3. KEY PRINCIPLES

For the purpose of this guideline, a principle refers to a characteristic of a trial design that is critical to ensure the reliability and interpretability of the results. This section describes principles that require specific considerations with an adaptive design. The focus is on proposals for confirmatory trials with an adaptive design. All of these principles should be followed regardless of the type of adaptation and statistical approach (e.g., frequentist or Bayesian methods).

3.1 Adequacy Within the Development Program

It is important that clinical trials are properly designed, conducted, and analyzed to address the clinical research question(s) of interest within the context of an overall development program. A stepwise program with careful analysis and evaluation of completed exploratory trials helps inform the goals and design choices for subsequent confirmatory trials and ultimately generate

data necessary for regulatory decision-making. A complete development program should seek to, among other aspects: characterize the dose-response relationship with respect to favorable and unfavorable effects; identify an appropriate patient population for treatment; select clinically meaningful and sensitive endpoints; and reliably confirm efficacy and support the assessment of safety and benefit-risk in the intended patient population.

The number and complexity of adaptations at the confirmatory stage should generally be limited. Increasing either of them, as a replacement for a sequence of multiple trials, can impair the ability to answer important clinical questions and limit the opportunity to carefully reflect on prior results to design a development program most effectively. Before planning a confirmatory trial with multiple adaptations, sponsors should discuss whether additional exploratory trials are necessary to investigate the question(s) addressed by the proposed adaptation(s).

For example, consider a confirmatory two-stage adaptive clinical trial design with selection between two doses at an interim analysis, and confirmation of efficacy of the selected dose. In a setting where a dose-ranging trial has been conducted with remaining uncertainty about the most appropriate of two candidate doses, such a design may help ensure identification of the dose with the better benefit-risk profile in the intended patient population. However, if a proper dose-ranging trial was not conducted in earlier stages of the development program, the selection of two doses for the confirmatory trial(s) may not be well supported, adding risk that the program may fail to identify an appropriate dose. An adaptive design should generally not serve as a replacement for a proper dose-ranging trial. It is generally expected that the sponsor has completed the necessary trials to evaluate a wider range and number of doses before proceeding to the confirmatory trial(s) intended to confirm efficacy and assess benefit-risk.

3.2 Adequacy of Trial Planning

Adequate planning is important for all clinical trials to ensure the design is pre-specified, conduct and analysis are appropriate, and results are reliable and interpretable. If a confirmatory clinical trial is planned with an adaptive design, the number and complexity of adaptations should generally be limited and there should be a justification for adapting aspects of the trial at this stage of drug development. Prior to initiation of a trial with an adaptive design, further aspects should be specified and justified in addition to the typical components of trial planning. These include the number and timing of interim analyses, type of adaptation, statistical methods for producing interim results, anticipated rule governing the adaptation

decision, statistical methods for the primary analysis aligned to each targeted estimand, and approaches to maintain trial integrity. For adaptive designs with a planned selection of an estimand at an interim analysis, such as treatment- or population-selection designs (Sections 4.3 and 4.4), all candidate estimands should be fully pre-specified and clinically relevant.

Some types of adaptive designs may require more planning than others. For example, a design with unblinded sample size adaptation warrants additional approaches to maintain trial integrity than one with blinded sample size adaptation. If simulations are critical to understand operating characteristics of an adaptive design, the simulation study should be carefully planned, conducted, and reported (Section 5.2). All relevant details pertinent to the planning of an adaptive trial design should be appropriately documented (Section 6).

Adequate planning facilitates the evaluation of the appropriateness of the statistical approach for many types of adaptations. For example, Type I error probability control requires the pre-specification of criteria for early efficacy stopping or rules for combining evidence across stages. As another example, specifying a blinded sample size adaptation in the protocol, together with the adaptation rule, increases confidence that an adaptively selected sample size was not influenced by unblinded data. Adequate planning also facilitates the evaluation of trial operating characteristics and enables informed discussions with the IDMC (if involved in the adaptations). Sponsors should discuss the type of adaptations and anticipated adaptation rules in detail with the IDMC to confirm its understanding and support. This ensures the IDMC is prepared to review interim results and make adaptation recommendations during the trial while also protecting individual trial participants' safety.

There should always be a clear description of the anticipated rule on which the adaptation will be based. The extent to which the anticipated rule governing the adaptation decision needs to be adhered to at an interim analysis, however, can vary depending on the type of adaptation and the statistical inferential methods being used. It is generally recommended to use analysis methods that provide valid inference while allowing flexibility to deviate from the anticipated adaptation rule based on the overall benefit-risk assessment at an interim analysis. For example, consider a confirmatory two-stage adaptive clinical trial with selection between two doses at an interim analysis, with the objective to confirm the efficacy and support the benefit-risk assessment of the selected dose. At the trial planning stage, an efficacy-based rule for the interim dose selection may be planned given that no meaningful safety issues are expected. There is a chance, however, that interim data will suggest similar efficacy between the two

doses, with an unexpected safety concern for the higher dose. When using statistical methods that allow for the flexibility to incorporate such benefit-risk considerations at the interim analysis, the pre-specified plan should acknowledge the possibility of deviations from the rule and outline factors that may lead to such deviations. If the planned statistical methods instead require strict adherence to the rule governing the interim decision to ensure valid inference (e.g., Type I error probability control), the importance of adhering to the rule should be documented in the trial protocol.

3.3 Limiting the Chances of Erroneous Conclusions

It is important to limit the chances of erroneous conclusions about the efficacy, safety, and benefit-risk profile of a proposed treatment. An essential element of regulatory decision-making is controlling the chances of false positive efficacy conclusions (i.e., conclusions that truly inefficacious treatments are efficacious). The common approach is to limit the probability of false positive efficacy conclusions within a trial by using frequentist methods that control the Type I error probability for a hypothesis test of the primary estimand at a pre-specified threshold (ICH E9).

For most adaptive designs, it is necessary to use specific methods to control the Type I error probability. For example, if a design includes an interim analysis with the potential for early stopping for efficacy, appropriate pre-specified stopping rules are needed. When an adaptive trial design includes multiple testing approaches to control the Type I error probability across multiple primary and/or secondary endpoints, those approaches should additionally address the potential for an increased Type I error probability due to the proposed adaptation.

Although the predominant approaches to the design and analysis of clinical trials have been based on frequentist statistical methods, other approaches may be appropriate when the reasons for their use are clear and when the resulting conclusions are sufficiently robust (ICH E9). Section 5.3 describes important considerations for limiting the chances of false positive efficacy conclusions in adaptive designs using Bayesian methods.

It is also important to understand how a proposed adaptive design may impact the potential for other types of erroneous conclusions. This includes the need for the trial to provide sufficient information on safety, important secondary efficacy endpoints, and relevant patient subgroups to inform a reliable benefit-risk assessment. For example, when planning a trial with the potential to stop early for an efficacy conclusion, it is important to justify that the sample size

and duration of follow-up at an interim analysis can adequately support a reliable benefit-risk assessment. This also includes evaluation of the impact of adaptive designs on conclusions made at interim analyses, and the risk that the adaptive design may be inadequate to fulfill the trial objectives. For example, sponsors should evaluate the ability of an adaptive dose-selection design to select the better out of two doses at an interim analysis based on efficacy and benefit-risk considerations. Finally, adaptations can impact the chance of a false negative efficacy conclusion (i.e., lack of evidence of an effect for a truly efficacious treatment) such that it is important to evaluate whether the trial achieves adequate power.

3.4 Reliability of Estimation

Controlling the chances of false positive efficacy conclusions is expected in a confirmatory clinical trial (Section 3.3). In addition, reliable estimation of treatment effects for the primary efficacy endpoint and other key efficacy and safety outcomes is important to facilitate the benefit-risk assessment and inform regulatory decision-making. The primary analysis of a trial with an adaptive design should therefore provide an estimate of the treatment effect that is reliable and aligned with the estimand of interest. Sponsors should evaluate bias and variability of treatment effect estimates, including measures such as the mean squared error. In the trade-off between bias and variance, the expectation is generally for limited to no bias in the primary estimate of the treatment effect. The primary analysis should also support calculation of accurate measures of uncertainty such as confidence intervals with targeted coverage probabilities.

If a trial with an adaptive design uses approaches for estimation in the primary analysis that do not account for the adaptive nature of the design, unreliable treatment effect estimates and incorrect estimates of uncertainty (e.g., incorrect confidence interval coverage) may arise. For example, selecting the treatment with the largest estimated effect from among several treatments at an interim analysis will, on average, lead to an overestimation of that treatment's effect. This holds true even if selection is based on an endpoint expected to be predictive of efficacy rather than the primary endpoint itself. Similarly, treatment effect estimates for secondary endpoints may be biased in the presence of adaptations. Adaptive design proposals should therefore evaluate bias and variability of treatment effect estimates and provide support of their reliability. In some cases, bias and variability can be calculated analytically. In other cases, the evaluation has to rely on simulations. For some designs, specific estimation methods have been derived with improved reliability, and these should be used. As one example,

methods are available in group sequential designs for adjusting estimates to reduce or remove bias associated with the potential for early stopping and to improve performance on measures such as the mean squared error.

In addition to ensuring reliable estimation of the treatment effect in the primary analysis, it is also important to support that estimates at interim analyses can facilitate reliable adaptation decisions. For example, conducting an interim analysis in an adaptive dose-selection design at an early time point may result in highly variable estimates and the selection of an inferior dose. Sponsors should therefore evaluate the overall operating characteristics of the design (e.g., probability of selecting the better dose) to inform careful selection of the timing of an interim analysis and the adaptation rules.

3.5 Maintenance of Trial Integrity

It is important that the integrity of a trial is maintained such that it achieves its objectives in a reliable, ethical, and timely manner. The impact of trial adaptations on the statistical validity of trial results is discussed in Sections 3.3 and 3.4. Maintenance of trial integrity also relies on appropriate execution of the trial and careful assessment of the potential impact of envisaged adaptations on trial conduct, which is the focus of this section.

Knowledge by the sponsor, investigators, or trial participants about individual treatment assignments, accumulating data, or certain trial changes can impact trial integrity by affecting expectations and behaviors in ways that are difficult to predict and impossible to adjust for. Such knowledge can introduce subtle changes in trial conduct, such as changes in the pace and characteristics of participants enrolled, specific details of the administration of the study treatment or other medications, or endpoint assessments, that may impact the interpretation of trial results. For example, knowledge by investigators and trial participants of a small or unfavorable estimated treatment effect based on accumulating data during an ongoing trial could be misinterpreted as reliable evidence of no effect, causing decreased enrollment, adherence, and retention of trial participants, ultimately leading to unreliable results and difficulties with their interpretation at trial end. The recommended approach is to blind participants, investigators, and the sponsor to individual treatment assignments and to accumulating summary-level data in which treatment groups are identified (either with the actual treatments or with labels such as A and B), therefore limiting the risk for occurrence of conscious and unconscious changes in trial conduct arising from such knowledge.

A fundamental aspect of many types of adaptive designs is the need for some level of access to unblinded interim results. Personnel having access to accumulating unblinded data should generally be independent in the sense that they do not have conflicts of interest or any role in trial activities and are external to the sponsor. To achieve this, an IDMC should be in place to review unblinded interim data when such access is needed as part of the adaptive design. In confirmatory trials, an IDMC will often already be planned to assure the safety of trial participants and to protect the scientific integrity of the trial. In this case, the IDMC can have an additional role of reviewing interim data for the purpose of implementing the planned adaptations. If an IDMC is not already planned, one can be set up with objectives and member expertise targeted toward implementing the adaptive design. Standard operating procedures and confidentiality agreements should be put in place to limit access to unblinded interim results beyond the IDMC. Additional discussion about the IDMC and other data monitoring considerations is available in Section 5.1.

Even the knowledge of an adaptation itself can lead to unwanted changes in behavior on the part of investigators or trial participants or can potentially reveal information about unblinded interim results. For example, if an unblinded sample size adaptation is implemented, where the revised sample size is a function of an interim treatment effect estimate, someone who understands the adaptation rule and knows the revised sample size can infer the interim effect estimate. Therefore, measures should be implemented to minimize the information that can be inferred, while maintaining ethical standards (e.g., adequate informed consent forms) and ensuring operational feasibility (e.g., adequate drug supply); see further discussion of operational considerations in Section 5.6. One particular approach to limit the knowledge that can be inferred during the trial is to use adaptation rules where a sufficiently large range of interim estimates leads to the same change (e.g., with a sample size adaptation rule that includes only a small number of potential adaptively selected sample sizes). Details of the adaptation rule could be reserved for a specific document rather than the protocol, such as a confidential appendix to the IDMC charter, that is only accessible to designated sponsor personnel separated from the team managing and conducting any aspects of a clinical trial. Additionally, sponsor personnel, investigators, and trial participants could be shielded from knowledge of specific adaptive changes. For example, trial sites could be informed after a sample size adaptation that the targeted enrollment has not been reached, or notified of site- or region-specific targets, rather than notified of the overall sample size target.

Sponsors should discuss with regulators at the planning stage the potential implications of the adaptations on trial conduct, including the type of participants enrolled, and on the interpretation of the results at trial end. This should include a discussion of the sufficiency of the size of the trial stages for assessing the impact of adaptations. Sponsors should implement approaches for maintaining trial integrity. Processes should be documented to increase adherence to these approaches and to provide transparency to relevant stakeholders (e.g., regulatory authorities and participating investigators). Appropriate training and careful planning are needed to prevent compromises to the extent possible. Because even the most rigorous processes may not fully guarantee trial integrity, the interpretation of results at trial end should involve consideration of any heterogeneity between results from different stages of the trial, the nature of the adaptive design (e.g., the number and type of adaptations and the size of the stages of the trial), the processes in place and who had access to different kinds of data and information during the trial, and any notable changes in trial conduct before and after an interim analysis (e.g., changes in the types of participants enrolled). Unexpected heterogeneity findings should be discussed by the sponsor and may impact the interpretation of the trial results.

The principles for maintaining trial integrity discussed above are particularly critical in open-label trials in which each participant's individual treatment assignment is known to the participant and/or investigator. Notably, even though individual participant assignments are known in such trials, it is feasible and strongly recommended to ensure that participants, investigators, and the sponsor do not have access to accumulating summary-level data by treatment group.

4. TYPES OF ADAPTATIONS

This section discusses common types of adaptations, with a focus on specific considerations relevant to the principles in Section 3. This section also illustrates some of the advantages and challenges of adaptive designs outlined in Section 2. The discussion focuses on designs using frequentist approaches for statistical analysis. For special considerations related to adaptive designs using Bayesian methods, see Section 5.3.

4.1 Early Trial Stopping

During the conduct of a clinical trial, accruing data can provide information that makes it no longer appropriate to continue the trial. To address this, sponsors can consider a trial design

that includes prospectively planned sequential analyses of accumulating unblinded data with anticipated rules for stopping when there is compelling evidence of efficacy (stopping for efficacy) or when the trial is unlikely to demonstrate efficacy (stopping for futility). A clinical trial design that allows such sequential analyses for early efficacy stopping based on accumulating observations of groups of participants at pre-specified points throughout the trial is called a group sequential design.

When planning a trial design that allows for early efficacy stopping, appropriate stopping boundaries should be planned for the sequential analyses such that the Type I error probability is controlled. The timing of interim analyses and specific stopping rules should be justified based on factors such as the required persuasiveness of early results to stop the trial, the probability of early stopping, and the expected and maximum sample sizes or numbers of events that may be accrued. Approaches may be considered that allow deviation from the anticipated timing of interim analyses. For example, this could help accommodate the scheduling of IDMC meetings at specific calendar times, such that the actual sample size at an interim analysis may differ slightly from the pre-specified target. In addition, methods for calculating the primary treatment effect estimate and associated confidence interval that adjust for the interim analyses should be planned to limit bias and improve performance on measures such as the mean squared error (Section 3.4).

A trial that is stopped early for efficacy will provide less information (e.g., because of a smaller sample size and/or shorter duration of follow-up) for the evaluation of safety, important secondary efficacy endpoints, and relevant patient subgroups, which are important for the overall benefit-risk assessment. Therefore, the timing of interim analyses should be selected such that the sample size is large enough and the duration of follow-up is long enough to ensure sufficient information is available for decision-making. There usually is a limit on how early interim analyses should occur or whether they should occur at all because a minimum sample size and/or duration of follow-up is expected for a sufficient evaluation of safety. This is often a relevant criterion, for example, in preventive vaccine trials and to meet regulatory standards for the extent of population exposure for treatments intended for long-term treatment of non-life-threatening conditions (ICH E1). Furthermore, interim analyses with the potential for early stopping are more often considered in circumstances where there are compelling ethical reasons (e.g., the primary endpoint is survival), and efficacy stopping rules typically require

highly persuasive results in terms of both the magnitude of the estimated treatment effect and the strength of evidence of an effect.

In the case that a stopping rule at an interim analysis is met and a decision is made to stop the trial for efficacy, additional data beyond those included in the interim analysis may continue to accumulate on participants in the trial prior to the final database lock. This can occur as a result of a time lag between data collection and interim analysis during which data adjudication and cleaning are carried out. Sponsors should ensure this additional information is appropriately documented and should report results from the interim analysis and from the analysis based on all available data, which are both important for regulatory decision-making. For example, a change in the estimated treatment effect between these two analyses that may affect the benefit-risk assessment would warrant investigation of potential explanations and may make interpretation of the trial results challenging.

When a trial design incorporates the potential for futility stopping, while anticipated futility rules should be pre-specified and justified, it is generally recommended to use nonbinding futility rules. This means that the futility stopping criteria serve as guidelines that can be deviated from based on the interim results without increasing the Type I error probability. This flexibility is important because decision-making about whether to stop for futility or continue is usually not an algorithmic process and may need to incorporate additional information beyond the primary efficacy endpoint, such as safety or other efficacy data. In contrast, there have been proposals to use binding futility rules and adjust the efficacy decision criteria for the planned futility criteria. These approaches have the disadvantage of requiring that sponsors adhere to the pre-specified futility stopping criteria, as otherwise the Type I error probability is not controlled and the interpretation of trial results can be compromised.

4.2 Sample Size Adaptation

Even after a carefully planned and conducted early-phase development program, a considerable degree of uncertainty might exist in the parameter assumptions that affect the sample size calculations for a clinical trial. One source of uncertainty are assumptions about the nuisance parameters that are not of primary interest but may affect the sample size of a trial. Examples of nuisance parameters include the standard deviation of a continuous outcome and the probability of response of the control arm for a binary outcome, which can be highly variable across trials in certain disease settings. In such cases where a sound rationale exists, sponsors may consider incorporating the potential for modification of the initial sample size based on

interim estimates of nuisance parameter values to ensure the trial is adequately powered. Another source of uncertainty at the planning stage are assumptions about the anticipated treatment effect size. In cases where there is justification based on residual uncertainty (e.g., after appropriate exploratory trials; see Section 3.1), sponsors may consider a sample size adaptation based on an interim treatment effect estimate. The goal would be to ensure sufficient power under a range of plausible and clinically meaningful treatment effect sizes.

Appropriate planning of any design incorporating sample size adaptation should include pre-specification and justification of the minimum and maximum potential sample sizes, the anticipated sample size adaptation rule, and the statistical analysis method. It is important that the minimum sample size still provides sufficient information for benefit-risk assessments (e.g., for evaluating safety, secondary endpoints, and subgroup analyses), similar to considerations for early stopping (Section 4.1).

Adaptations to the sample size based on nuisance parameter estimates should be carried out using blinded data as this approach does not incorporate information about treatment assignment, thus minimizing risks for trial integrity. The anticipated sample size adaptation rule should be pre-specified to increase confidence that an adaptively selected sample size was not influenced by unblinded data. Such pre-specification also facilitates evaluation of trial operating characteristics (e.g., power and expected sample size). Sponsors should propose and justify a testing approach that controls the Type I error probability. In some cases, conventional analysis methods that would apply in non-adaptive designs can be used for the primary analysis if there is justification (e.g., in a reasonably sized two-arm superiority trial with a continuous endpoint). In other cases (e.g., a two-arm non-inferiority trial with a continuous endpoint), the use of these conventional methods may lead to an increase in the Type I error probability and different approaches are needed.

Trials with sample size adaptations based on interim effect estimates should use an IDMC and adequate processes to maintain trial integrity, given that the adaptations are based on unblinded data. This should include steps to minimize the information that can be inferred from the interim sample size selection (Section 3.5). Given that such designs typically allow for an increase in sample size compared to the initially planned sample size, statistical significance can be achieved with weaker observed effects than initially planned. When planning such a design, it is therefore important to judge the magnitudes of effects that would be clinically

meaningful, justify the added participant exposure, and ensure that the potential sample sizes under the adaptive design are sensible from a clinical perspective.

It is generally recommended to use sample size adaptation methods that do not require adherence to the anticipated adaptation rule, such as hypothesis testing based on pre-specified weights for combining the information across trial stages. Still, the anticipated adaptation rule should be pre-specified to facilitate the evaluation of trial operating characteristics (e.g., expected sample size and power) and ensure that the IDMC understands and is in agreement with the anticipated adaptation rule.

For most designs involving adaptations to the sample size based on interim treatment effect estimates, conventional testing methods for non-adaptive designs are not appropriate and specific statistical methodology needs to be used to ensure Type I error probability control. In addition, conventional point estimates of the effect size may be biased, and conventional confidence intervals may have incorrect coverage probabilities. Therefore, it is recommended to evaluate the reliability of these estimates at the trial planning stage. This evaluation may inform the acceptability of the proposed adaptive design or the interpretation of trial results. In some cases, methods are available that adjust estimates to reduce or remove bias associated with the adaptation and these are preferred.

4.3 Population Selection

In certain settings, there may be remaining uncertainty about the patient population who should be treated with a new treatment. For example, a treatment may be expected to benefit a certain targeted subset of the overall population, while the benefit in the non-targeted (complementary) subset may be unclear. This targeted subpopulation could be defined, for example, by demographic characteristics or by a genetic or pathophysiologic marker that is assumed to be related to the treatment's mechanism of action. If the treatment were truly efficacious in the targeted subpopulation but not efficacious or minimally efficacious in the complementary subpopulation, conducting a trial in the overall population might have insufficient power to establish a treatment effect and might unnecessarily expose participants to a treatment from which they will not receive benefit. On the other hand, if the treatment were truly efficacious in the overall population, a trial in only the targeted subpopulation would not provide data on the effects of the treatment in the complementary subpopulation and would result in restricting the indication for the treatment to only a subset of the overall population that would benefit.

Such uncertainty would usually be investigated in an exploratory trial. However, in some cases there also may be consideration to conducting a confirmatory trial in the overall population, with an analysis plan that includes evaluation of efficacy in a targeted subpopulation (e.g., with a multiple testing approach to control the Type I error probability across analyses in the overall population and in the subpopulation). Alternatively, it may be more efficient to consider a design for a confirmatory trial with the option for adaptations to the patient population based on unblinded interim results. A trial might enroll participants from the overall population up through an interim analysis, at which time a decision would be made whether to continue enrollment in the overall population or to restrict future enrollment to a targeted subpopulation. If enrollment continues in the overall population, a decision would then need to be made whether to evaluate in the analysis at trial end the treatment effect in only the overall population, or in both the overall population and the targeted subpopulation. If enrollment is restricted to the targeted subpopulation, the analysis at the end of the trial would focus on the treatment effect in that subpopulation. In such settings, data accumulated both before and after the interim analysis should be appropriately combined to draw inference on the treatment effect in the selected population(s).

Adequate planning of such designs should include pre-specification of the candidate population(s) that may be selected at the interim analysis to be the target of future enrollment, the decisions to be made at the interim analysis regarding the population(s) for statistical inference and how they will be analyzed at the end of the trial, and the anticipated adaptation rules. There should also be a plan for managing participants from a population for which further enrollment and evaluation is stopped based on an interim analysis. In designs that select population(s) for enrollment and analysis based on interim treatment effect estimates, specific statistical methodology is typically needed to control the Type I error probability. Methods are generally recommended that allow flexibility in deviating from the anticipated adaptation rule, as considering the totality of information available at the interim analysis helps ensure appropriate population selection. Sponsors should also ensure that interim estimates can facilitate reliable population selection, including planning the interim analysis at an appropriate time point. Furthermore, given that such a design tends to select population(s) with more favorable interim results, conventional treatment effect estimates at trial end may be biased. The reliability of the treatment effect estimates in the different populations should be evaluated, and adjusted estimates that reduce or remove bias should be considered.

It is important that a trial with population adaptation has a sound scientific rationale. For example, a trial in the overall population that includes an interim analysis to potentially focus future enrollment and analysis on a particular subpopulation should be motivated by results from previous trials and/or biologic evidence that the benefit-risk profile may be meaningfully more favorable in the targeted subpopulation. With such a trial, it is also important to ensure that the design facilitates reliable decision-making in the scenario in which enrollment in the overall population continues after the interim analysis. This includes ensuring that the trial will provide adequate information on the benefit-risk profile in the complementary subpopulation. It also includes specifying criteria, including criteria for the estimated treatment effect in the complementary subpopulation, that would justify a conclusion of benefit in the overall population. If the baseline characteristic that may be used to define subpopulations is not binary in nature, justification should be provided at the planning stage for any threshold(s) used to define the subpopulations.

4.4 Treatment Selection

Some trials are conducted with the intent to evaluate more than one treatment. The multiple treatments might be different drugs or different doses of a single drug. For example, there might be uncertainty remaining at the end of the exploratory development phase about the benefit-risk profile of two likely efficacious doses of a certain drug. A confirmatory trial might then compare these two doses against control with the objective to confirm their efficacy and to select the most appropriate dose(s) at trial end. In such a setting, it may be conceivable to design a trial with the option for dose selection based on an interim analysis of accumulating unblinded data. Participants would initially be randomized to either of the two doses or control. At the interim analysis, one or both doses would be selected for continued randomization in the second stage. The analysis at the end of the trial would then aim to confirm efficacy and assess benefit-risk of the selected dose(s) based on data across both trial stages.

Adequate planning of a trial with adaptive treatment selection should involve specification of the treatments that will be evaluated, the decisions to be made at the interim analysis, and the anticipated rules for the selection process, including any implications for the randomization scheme and overall sample size. There should also be a plan for managing participants who are receiving a treatment for which further evaluation is stopped based on an interim analysis. In a design that potentially selects one (or more) treatments based on interim effect estimates, specific statistical methodology is needed to control the Type I error probability. It is generally

recommended to use methods that allow for flexibility in deviating from the anticipated adaptation rule. Such flexibility enables consideration of the full scope of information available at the interim analysis, helping to support more informed and appropriate treatment selection decisions. Sponsors should also ensure that interim estimates can facilitate reliable treatment selection, including planning the interim analysis at an appropriate time point. Given that such a design tends to select treatment(s) with more favorable interim results, conventional treatment effect estimates at trial end may be biased. The reliability of estimates should be evaluated, and adjusted estimates that reduce or remove bias should be considered.

4.5 Adaptation to Participant Allocation

In a randomized trial, participants are typically allocated to treatment arms according to fixed randomization probabilities. Alternatively, there are different approaches that can be considered to incorporate adaptations to the allocation scheme, where the assignment of participants to treatment arms depends on the data of earlier trial participants. These include covariate-adaptive approaches where assignment depends on accumulating baseline covariate data and response-adaptive approaches where assignment depends on accumulating outcome data. This section focuses on response-adaptive randomization (RAR) approaches where incoming participants are randomized to treatments according to probabilities that depend on previous unblinded outcome data. The key idea is to assign new participants with greater probability to treatment arms that have had, to that point, more positive outcomes than to other treatment arms.

RAR is sometimes valued for advantages to trial participants such as exposure of fewer participants to an inferior treatment or reduction in the expected number of participant treatment failures in a trial with a binary response endpoint. However, RAR procedures also bring challenges in ensuring valid statistical inference. Perhaps most concerning, RAR designs are susceptible to bias and inflation of the Type I error probability in the presence of overall time trends. For example, a RAR design would more likely show a false positive treatment effect if earlier-enrolled participants are both more likely to be assigned to control and to have a poor prognosis (e.g., because of changes in background care or participant characteristics over time) than later-enrolled participants. In addition, the use of efficacy-based algorithmic modifications to the randomization scheme could lead to an insufficient sample size to support decision-making on a treatment that may have lesser efficacy but a better benefit-risk profile.

Any proposal to use RAR should address these potential issues. The specific RAR procedure should be pre-specified and justified. There should be careful specification of analysis methods that provide Type I error probability control and reliable estimates. The proposal should address the potential for confounding due to time trends. The degree of such confounding may depend on factors such as the expected duration of the trial and the likelihood of changes in background care or prognostic factors over time (e.g., such changes may be likely in a rapidly evolving infectious disease setting). One approach that controls the Type I error probability is to allow randomization ratio adaptation at only a single or small number of interim analyses, while utilizing adaptive hypothesis testing based on pre-specified weights for combining the information across trial stages. Time trends may also be addressed by using specific methodology (e.g. re-randomization tests), but an RAR design using such tests might be less powerful than a design with a fixed randomization scheme.

An approach that implements the changes to the randomization scheme over time without sponsor involvement should be planned to reduce the risk to trial integrity. Given that knowledge of the RAR procedure and the adaptively selected randomization ratio could reveal information about the interim treatment effect estimate, steps should be taken to minimize what can be inferred from the adaptations (Section 3.5). Finally, there can be additional challenges such as ensuring the timely availability of high-quality interim data on an ongoing basis and integrating the algorithm into the randomization system.

There are also non-randomized, deterministic adaptations to participant allocation such as in a two-arm trial where a response results in assigning the next participant to the same treatment, while a non-response leads to assigning the next participant to the alternative treatment. Such deterministic procedures are discouraged (ICH E9) due to the high risk of bias and the potential for predicting the next treatment allocation.

5. SPECIAL TOPICS AND CONSIDERATIONS

This section expands on some special topics for adaptive designs, including data monitoring, simulations, use of Bayesian methods, time-to-event endpoints, exploratory trials, and operational execution.

5.1 Further Considerations on Data Monitoring

This section discusses further considerations related to data monitoring in confirmatory trials with adaptive designs that include interim analyses based on accumulating unblinded data. An

IDMC for a trial with an adaptive design should contain, as a group, all expertise needed for making adaptation recommendations in addition to meeting its usual responsibilities (i.e., protecting individual participants' safety while maintaining trial integrity). It should include at least one statistician knowledgeable and experienced in interim monitoring and in statistical methodologies relevant to the proposed adaptive design and analysis. The IDMC should generally have access to unblinded efficacy and safety data. Operational aspects should be outlined in a designated charter to document details such as content and frequency of reports to be prepared, meeting schedule and logistics, procedures to maintain confidentiality, statistical aspects of the monitoring plan, and processes for making recommendations. It is important that sponsors align upfront with the IDMC on the trial objectives and design, expectations for the IDMC (including those that go beyond the usual responsibilities), type and implications of adaptations, and anticipated adaptation rules.

An independent statistical group that conducts analyses of accumulating unblinded data and produces interim reports for the IDMC should be in place. It should not include members of the monitoring committee and should not support other trial activities. Trial integrity will be best protected when this statistical group having access to unblinded data is external to and independent from the sponsor. The statisticians and programmers that comprise this group should have the appropriate expertise to carry out the analyses needed to implement the adaptive design and to support the IDMC. They should have access to all trial data needed to carry out their responsibilities. It is strongly recommended that the independent statistical group and IDMC have sole access to unblinded interim data and results. Appropriate processes for maintaining confidentiality (e.g., standard operating procedures and confidentiality agreements) should be in place.

Upon reviewing the unblinded interim results, the IDMC should provide adaptation recommendations to designated sponsor personnel separated from the trial team. In the specific case that the IDMC has made a recommendation to stop a trial early, sufficient information may then be communicated to the sponsor (e.g., key efficacy and safety results) to allow sponsor decision-making about whether to stop the trial. In general, however, the adaptations should be planned such that the sponsor can implement the IDMC recommendations regarding trial adaptations without having access to any unblinded interim results. For example, this would be the case when the IDMC recommends continuing the trial in a group sequential design or when it selects a specific sample size in a sample size adaptation design. This requires

extensive planning and discussion between the sponsor and the IDMC at the planning stage to ensure a common understanding of the monitoring processes and anticipated adaptation rule.

Risks to trial integrity are most easily minimized by completely restricting sponsor access to unblinded interim results. However, sponsors may propose some degree of access to unblinded data in certain circumstances. This should be made explicit at the planning stage. Any proposal for sponsor access needs to be supported by a compelling rationale. In this case, there also should be planned steps to protect trial integrity such as minimizing the number of individuals with access, ensuring individuals with access are independent from those involved in trial conduct, and implementing processes to maintain confidentiality. All information regarding who accessed what data should be recorded in detail so that regulators assessing trial results before and after the adaptation can be reassured at the end of the trial that trial integrity was not compromised.

5.2 Planning, Conducting, and Reporting Simulation Studies

Simulation studies often play an important role in the planning of a trial with an adaptive design. A simulation is the repeated execution of a large number of hypothetical clinical trials to understand operating characteristics of a trial design under a series of specific configurations of assumptions (scenarios). Simulations can be used to investigate operating characteristics of a proposed adaptive design in different scenarios, such as under different treatment effect and nuisance parameter assumptions, in the presence of varying dropout or enrollment rates, or with a specific sample size when analytical properties of an analysis approach rely on large sample sizes. For example, the probability of a false positive conclusion can be estimated by calculating the proportion of simulated clinical trials that would lead to a false positive conclusion that a treatment is effective when data have been simulated under the assumption of no beneficial treatment effect. Simulations can facilitate comparisons of adaptive and non-adaptive designs, comparisons of different adaptive design options, and comparisons of different drug development programs (i.e., a comparison of a sequence of trials). Simulations can also inform internal sponsor decision-making on trial logistics such as site selection and drug supply. This section focuses on principles for the appropriate planning, conduct, and reporting of simulations when they are critical for understanding the operating characteristics of a trial with an adaptive design.

It is important to clearly define and focus on the key objectives the simulation study is designed to address. These should be specific, relevant, and directly related to the decisions that will be

made as a result of the simulation study. To address the objectives, a range of clinical trial designs and analysis options should be carefully selected. These should include a benchmark design and analysis approach, i.e., a design with well-understood operating characteristics such as a non-adaptive or group sequential design. This range of designs may also include, for example, different choices for the number and timing of interim analyses, stagewise sample sizes, types of adaptations, stopping and adaptation rules, and statistical methods for testing and estimation. The choice of design options may be an iterative process as operating characteristics are explored and should be sufficiently broad to allow a comprehensive assessment of the selected adaptive design. The evaluation of the advantages and limitations of all design options included in the simulation study is critical to understand the tradeoffs in the selection of the proposed design.

It is also important to define and assess key operating characteristics that align with the questions the simulation study is designed to address. These operating characteristics should generally include the Type I error probability, expected sample size, expected trial duration, power, coverage of confidence intervals, and bias and mean squared error of treatment effect estimates. Other operating characteristics such as the probability of stopping for futility or efficacy at an interim analysis may also be of interest, depending on the trial design and setting. Considerations around operating characteristics for adaptive designs using Bayesian methods are discussed in Section 5.3. Sometimes, operating characteristics beyond a single trial may be of interest, such as the probability of selecting an appropriate dose and subsequently confirming its efficacy. While it is relevant to summarize the average of the results across the simulated trials (repetitions), it may also be important to evaluate the variability, minimum and maximum, or other aspects of the distribution of results (e.g., the sample size distribution in a trial with the potential for early stopping or sample size adaptation).

The scenarios included in the simulation study should cover the plausible range of assumptions to ensure a robust assessment of the performance of the proposed adaptive design. This includes assumptions about the treatment effects and nuisance parameters, such as the standard deviation for a continuous outcome, and operational assumptions for which a sponsor may have greater control (e.g., enrollment or dropout rates). The adequacy of the assumptions should be justified based on clinical and statistical considerations, with documentation of the supporting knowledge. This information can come from a variety of sources, including data from previous trials, publications, results from extrapolations, and expert input. All relevant sources of

information available to the sponsor should be used, and attempts should be made to quantify uncertainty and identify potential biases. Using a grid of assumptions (e.g., discrete set of assumptions across a specific range) should be supported by justification based on existing clinical knowledge that the range evaluated in the grid covers all plausible scenarios. It is also important to justify (e.g., based on monotonicity arguments) that the grid is fine enough (i.e., that a sufficient number of different assumptions are included within the range) to provide a reliable estimate of the operating characteristics of interest. Sources of information based on robust evidence and understandable from a clinical perspective will make the simulation study results more interpretable and convincing.

It is essential that the simulated scenarios comprehensively cover the plausible range of nuisance parameter configurations. For example, in using simulations to investigate the Type I error probability, it is impossible to simulate under every nuisance parameter configuration consistent with no beneficial treatment effect, even in the simplest trial designs. Thus, there is additional uncertainty for designs in which simulations are critical to understand the Type I error probability. Given the additional uncertainty, additional justification is expected to support such designs.

Implementation details of the simulation study should be described and justified. This includes clear specification of the data-generating process. In many cases, a simple statistical model, such as a normal distribution with mean and variance obtained from previous trials, may be appropriate. In other cases, a more complex model fit based on earlier trial results (e.g., longitudinal patient profiles) may be considered. This also includes determining the number of repetitions needed to get sufficient precision in the estimation of important operating characteristics. More precision may be needed for certain operating characteristics or scenarios. For example, it may be important to use 100,000 or more repetitions per scenario to ensure sufficient precision for estimating the Type I error probability, whereas fewer repetitions may suffice for other operating characteristics such as power. Algorithms should be documented and random numbers should be generated in a reproducible way, such as using a documented seed.

Finally, it is important to document the design, results, and conclusions of the simulation study. A comprehensive and structured report of the simulations should be included in regulatory submissions prior to conducting the trial (Section 6.1). There should be explicit links between

clinical and statistical assumptions and results of the simulations. The report should align with the considerations outlined in this section and include the following:

1. Key questions the simulation study is designed to address.

2. The clinical trial design and analysis options evaluated in the simulation study.

3. The choice of operating characteristics assessed in the simulation study.

4. Existing knowledge, and any supporting data or references, to inform the simulation scenarios.

5. The set of parameter configurations used for the simulation scenarios, along with a clinical justification based on existing knowledge that the set adequately covers the plausible range of values for the different parameters.

6. Implementation details, including the data-generating process and the number of repetitions for each scenario, along with justifications for these choices.

7. Software package used for simulations and, if custom software was used, the simulation code. When code is provided, it should have adequate comments with detailed instructions on how to execute the code (e.g., an example call and the starting seed).

8. A summary providing overall results, interpretations, and conclusions. This should include a detailed discussion of the proposed adaptive design and its estimated operating characteristics under the various scenarios. Summarizing results in interactive graphs, where possible, can help make the results more accessible.

9. A description of relevant examples of single simulated clinical trials with different adaptations and conclusions. For example, in a design with sample size adaptation, this might include trials with different sample size modifications at the interim analysis and with positive or negative primary analysis results to facilitate a better understanding of potential interim decisions and their impact on the trial results.

10. A description of any aspects that limit the interpretation of the simulation results (e.g., uncertainty in assumptions or extrapolations).

11. A clinical discussion about if and to what extent the simulation results address the key questions.

The careful documentation of simulation studies is also critical because the validity of the simulations and associated conclusions will be part of the regulatory review of results at the end of the trial.

5.3 Adaptive Designs Using Bayesian Methods[1]

ICH E9 notes that the use of Bayesian methods in clinical trials may be considered when the reasons for their use are clear and when the resulting conclusions are sufficiently robust. Bayesian methods refer to a wide range of statistical approaches that combine a prior probability distribution with current trial data to obtain a posterior probability distribution for a quantity of interest (e.g., the treatment effect or estimand). Bayesian methods are potentially applicable to a variety of adaptive designs. The principles outlined in Section 3 should be followed regardless of the specific statistical approach. There are different types of application of Bayesian methods to clinical trials with an adaptive design, each with different considerations.

Bayesian methods can be used to inform adaptations in a trial where decision criteria for the primary analysis are chosen to ensure that the Type I error probability is controlled. For example, a trial might include interim analyses with pre-specified non-binding futility stopping rules based on a scale such as the posterior probability that the treatment is inefficacious or the predictive probability of rejecting the null hypothesis at trial end, where the primary efficacy analysis is performed with a frequentist hypothesis test at a pre-specified significance level. For such designs, expectations for operating characteristics are the same as for adaptive designs that do not involve Bayesian methods. Sponsors should justify that the prior distribution, decision criteria, and adaptive design elements (e.g., number and timing of interim analyses and adaptation rules) can achieve targeted operating characteristics (e.g., power, expected sample size, reliability of adaptation decisions) while maintaining Type I error probability control.

[1] This section on Bayesian methods for adaptive designs is not fully harmonized. The broad use of Bayesian methods may not be justified in all situations for regulatory decision-making. As noted in ICH E9 and in this draft guideline, the use of Bayesian methods in clinical trials may be considered when the reasons for their use are clear and when the resulting conclusions are sufficiently robust. Public consultation comments are sought on the topic, and on situations in which Bayesian methods satisfy the core adaptive design principles, and in which the use of Bayesian methods could be considered.

A special case is the use of adaptive design elements in the context of clinical trials that use Bayesian methods to borrow external information based on an informative prior distribution, with decision criteria for the primary analysis based on the posterior distribution for the estimand of interest (i.e., a threshold on the posterior probability for efficacy). Borrowing of external data to inform inference requires a thorough scientific justification that addresses the feasibility of alternative approaches not involving borrowing (e.g., design and conduct of a fully powered trial without using external data) and supports the relevance and quality of the external data. Misspecification of the prior distribution can lead to lack of control of the probability of false positive conclusions. Ensuring that a prior accurately reflects relevant available information and addressing the potential for conflict between prior and current trial data introduce additional uncertainties that are not present when using frequentist inference with no borrowing.

For such designs, sponsors should discuss and document in the protocol the source of the external information used to generate the prior, the relevance of the external information to the trial design (e.g., whether the populations and concomitant care are sufficiently similar, and the endpoints are the same), the list of all potentially relevant sources of information, and why selected information sources were used and other potentially relevant sources were discarded. Input from clinical subject matter experts is crucial for evaluating the relevance of external information. When considering the source of external information, data from randomized controlled trials and recent data are generally preferred. Patient-level data are generally expected because they allow a thorough evaluation at the planning stage of the relevance of the external information and may facilitate strategies to address potential conflict between the prior and current trial data at the assessment stage.

Sponsors should pre-specify and justify the details of a proposed prior distribution, including the amount of borrowing from the external data, as well as the criteria for defining trial success. The prior and decision criteria should ensure the design fulfills the principles in Section 3.3, including control of the chances of false positive conclusions. The justification for the prior should include a discussion of the balance between the prior and trial data and strategies to mitigate the risk that observed trial data may conflict with the prior. There should be a sufficient amount of current trial data to support benefit-risk assessment. Simulations should be performed to evaluate the chances of erroneous conclusions, including the chances of false positive conclusions, under various scenarios of prior-data conflict. There should be a

discussion at the planning stage about the maximum amount of borrowing and the relationship between observed conflict and the degree of borrowing, including circumstances that would question the relevance of the external data and lead to no borrowing. Sensitivity analyses should also be planned to investigate the robustness of the trial conclusions against alternative reasonable choices for the prior distribution. It is also important to evaluate the current trial data with no borrowing.

5.4 Adaptive Designs in Time-to-Event Settings

There are additional considerations specific to trials in which the primary endpoint is the time to occurrence of a certain event. In such time-to-event trials, the statistical power of the trial depends on the number of events rather than the number of participants. It is therefore common for such trials to target a fixed number of events when calculating the sample size at the trial planning stage. In addition, the follow-up time of participants is often unspecified, meaning that the trial does not have a fixed observation period, and all participants are followed until a certain number of events have occurred. For trials with adaptive designs in time-to-event settings, interim analyses are therefore often planned at target numbers of events rather than target sample sizes. Furthermore, a sample size adaptation based on an interim treatment effect estimate in a time-to-event trial may entail modification of the initially planned number of events. For example, targeting a larger number of events than originally planned could be achieved by simply waiting longer for events to occur (i.e., allowing for longer follow-up times) with the originally planned number of trial participants. Alternatively, the number of trial participants could be increased, or both approaches could be applied. In considering increases in the number of trial participants relative to the number of events, sponsors should ensure that sufficient data will be available for the benefit-risk assessment (e.g., to understand longer term treatment effects and to evaluate relevant subgroups of the patient population, including those with lower background risk of the event).

Adaptive designs are most straightforward when each trial participant only takes part in one stage of the trial. If the data collected prior to an interim analysis are completely independent of the data collected afterwards, a statistical analysis combining all information can proceed in a relatively simple way. In a time-to-event setting, however, some trial participants may be enrolled and remain event-free in one stage, but may contribute an event in a later stage. Using information (e.g., on secondary endpoints) from participants who have been enrolled in the trial but not yet experienced the event of interest at an interim analysis to inform potential

adaptations violates the independence assumption and can inflate the Type I error probability (even when using adaptive test statistics). Therefore, it is important to define plans with specific methodology for maintaining the Type I error probability. One option is to fully pre-specify an adaptation rule that relies on only the primary endpoint, without the possibility of deviations from such a rule. Another option is to use special methods that involve defining stages based on the sets of participants enrolled before and after the interim analysis, while also setting in advance either a fixed follow-up time or a fixed number of events for each stage. Alternatively, rather than incorporating adaptations to the number of events, sponsors can consider a design that targets a larger number of events and includes the option to stop the trial early at an interim analysis. Similar conceptual problems and respective considerations also apply to adaptive designs with longitudinal outcomes, as using surrogate or intermediate outcome information on participants who have not completed all follow-up visits at the interim analysis can increase the Type I error probability unless appropriate analysis methods are used.

5.5 Adaptive Designs in Exploratory Trials

This guideline focuses on the use of adaptive designs in confirmatory clinical trials. If a trial may be intended to confirm efficacy and support benefit-risk assessment, it is critical that the principles in Section 3 are followed. Adaptive designs may also be used in exploratory trials early in drug development that are intended to obtain information on a wide range of aspects of treatment use (e.g., choices of dose, regimen, population, endpoints). Trials at this stage of the development program may include a larger number of adaptations to generate information that support important decisions about subsequent development phases. The principles in this guideline are also relevant in these settings to ensure the reliability and interpretability of the results and subsequent decision-making based on such trials.

Additional considerations may apply, however, for exploratory trials because independent confirmation of findings will usually follow in one or more separate trials. For example, it may be sufficient that the protocol describes general principles for trial adaptations rather than the specific adaptation rule. This may be appropriate in, for example, dose-escalation trials where model-based dose recommendations are to be considered in the context of other emerging information (e.g., about toxicities that do not qualify for a dose-limiting toxicity). In addition, it is critical that exploratory trials with an adaptive design can reliably inform the decisions they are intended to support. For example, providing a convincing basis for decision-making about the appropriate target dose to be investigated in a confirmatory trial is critical as a

suboptimal conclusion can have serious consequences for the subsequent development program. Maintaining the integrity of exploratory trials with an adaptive design is also important, but there may be additional considerations for the sponsor's role in interim decision-making. For example, monitoring of an adaptive dose-ranging trial intended to inform the adequate dose for subsequent confirmatory trials may entail multidimensional adaptation decisions that require considerable input from various disciplines within the sponsor. Sponsors should then take into account the questions a trial intends to answer and its position within the development program, as well as the tradeoffs for sponsor involvement in the monitoring process versus limitation of access to unblinded results to maintain trial integrity. Any monitoring plan should ensure the protection of trial participants' safety.

5.6 Operational Considerations

Use of an adaptive design can add challenges to the operational execution of a clinical trial and these should be addressed at the trial planning stage. For example, measures should be implemented to minimize the information that can be inferred from an interim analysis to maintain trial integrity (Section 3.5). As another example, informed consent forms should cover the possibility of adaptive changes in the trial. Participants should understand the reasons for such changes (e.g., the goal of selecting the dose with the best benefit-risk profile from among multiple doses at an interim analysis), that these changes reflect improved knowledge about the treatment under investigation, and that their rights and safety remain protected. As yet another example, the infrastructure needed for trials with an adaptive design, such as data management systems, may differ from that of trials with a non-adaptive design. Clinical trials with an adaptive design typically use an interactive voice or web randomization system to manage randomization and assignment of participants to treatment arms. Such systems should be fully integrated into clinical trial operational processes and drug supply chain mechanisms. Pre-specified algorithms should be built into the system to ensure it is capable of handling the foreseeable scenarios (e.g., a change in the treatment arms or randomization ratio) with minimum sponsor involvement. Also, adaptations to the sample size, treatment arms, or participant allocation can lead to drug supply challenges. One such challenge is lead times for manufacturing drugs, as rapid adaptations can strain drug supply chains and lead to delays in participant treatment if sufficient drug supply is not readily available. These challenges may be increased when a clinical trial with an adaptive design spans multiple countries or even regions, as drugs need to be distributed to these locations in a timely manner. Simulations may help support supply-related decisions at planning and execution stages of the trial. Finally, processes

should be established at the planning stage to ensure relevant interim data can be appropriately validated and cleaned in a timely manner to ensure quality interim data informing the adaptation decision. This may include requiring a formal interim database lock to ensure completion of data validation and cleaning activities.

6. DOCUMENTATION

6.1 Documentation Prior to Conducting a Confirmatory Trial with an Adaptive Design

Documentation is a critical part of adequate planning of a confirmatory trial and allows a rigorous evaluation of the proposed adaptive design. In addition to the information typically included in a clinical trial protocol or in other documents, where suitable, documentation should include the following:

1. A rationale for the proposed adaptive design: The rationale should include both clinical and statistical considerations, justifying the proposal to adapt in a confirmatory trial and the adequacy of the proposed trial design within the clinical development program. A discussion of advantages and limitations as compared to alternative designs (including non-adaptive designs) will help regulators evaluate the acceptability of any additional uncertainty attributable to proposed adaptive elements.

2. A description of the adaptations being proposed: This should include the aspects of the trial that may be modified, the number and timing of interim analyses, and the anticipated rule governing the adaptation decision (e.g., the formula for determining the target sample size as a function of the interim treatment effect estimate, including the minimum and maximum potential sample size, in a design with sample size adaptation). If the design involves selection of an estimand at an interim analysis (e.g., through treatment or population selection), this should include precise definitions of all candidate estimands.

3. A description of the statistical analysis methods: This should include the methods for producing interim results and guiding adaptations decisions, the statistical approach for primary and secondary analyses (e.g., for hypothesis testing and for estimating treatment effects and corresponding measures of uncertainty), and important sensitivity and supplementary analyses.

4. A description of how the adaptive design will be implemented: This should include who

will carry out interim analyses; who will be responsible for reviewing interim analysis results and making adaptation recommendations and/or decisions; and membership, roles, responsibilities, and operational aspects of any relevant committees.

5. A description of steps to maintain confidentiality of interim results and protect trial integrity, among other details of the operational execution: This should include processes for information transfer and access; who will have access to unblinded interim results; how access to unblinded interim results will be controlled, what type of information will be disseminated following adaptive decisions, from whom, and to whom; and where records about information access and dissemination will be saved.

6. A description of important operating characteristics of the design. In cases where simulations are critical for understanding operating characteristics, this should include a report that describes the objective, design, implementation, and results of the simulation study (Section 5.2).

This information should be documented and included in regulatory submissions prior to initiation of the trial, in accordance with applicable national and regional regulatory requirements and practices. The protocol should contain the core elements, including the trial objectives and corresponding estimand(s), and the principal features of the trial design, conduct, and statistical analysis, including all adaptive design elements and their rationale. Some information, such as details on operation of an IDMC and data access processes, may instead be included in a separate document such as an IDMC charter. In some cases, details of the anticipated adaptation rule should be reserved for specific documents with access restrictions, rather than the protocol, to maintain trial integrity (Section 3.5).

6.2 Documentation to Include in a Marketing Application After a Completed Confirmatory Trial with an Adaptive Design

A marketing application for a treatment that relies on a confirmatory clinical trial with an adaptive design should include sufficient documentation to allow a comprehensive review of the trial results. In addition to its typical components, a marketing application should include:

1. All prospective plans described in Section 6.1.

2. Information on how the adaptive design was implemented, including the actual number and timing of interim analyses, an evaluation of whether aspects of trial conduct (e.g.,

baseline characteristics, enrollment rate, adherence, retention) varied notably before and after the interim analysis, the results of interim analyses used for adaptation decisions, any notable heterogeneity between results from different stages of the trial, the adaptation decisions that were made, whether anticipated adaptation rules were followed, and the date of sponsor unblinding. If there was any deviation from the anticipated plan (e.g., in terms of the number or timing of interim analyses or adherence to the anticipated adaptation rule), this should include a discussion of the reasons for the deviation, any measures taken to minimize impact on trial integrity, and any other potential impact on the interpretation of trial results.

3. Any information on compliance with planned processes for data access and maintaining trial integrity, such as results of any audits and reporting of any known deviations from the processes, along with a discussion of potential implications.

4. Records of deliberations by the IDMC (e.g., all closed and open IDMC meeting minutes), including records of discussions related to any adaptation decisions.

5. Reporting of results that appropriately account for the adaptive design (e.g., appropriately adjusted estimates, confidence intervals, and p-values).