



European Medicines Agency

London, 18 October 2007  
Doc. Ref. CHMP/EWP/2459/02

**COMMITTEE FOR MEDICINAL PRODUCTS FOR HUMAN USE  
(CHMP)**

**REFLECTION PAPER ON METHODOLOGICAL ISSUES IN CONFIRMATORY  
CLINICAL TRIALS PLANNED WITH AN ADAPTIVE DESIGN**

<b>DRAFT AGREED BY THE EFFICACY WORKING PARTY</b>	11 January 2006
<b>ADOPTION BY CHMP FOR RELEASE FOR CONSULTATION</b>	23 March 2006
<b>END OF CONSULTATION (DEADLINE FOR COMMENTS)</b>	30 September 2006
<b>AGREED BY THE EFFICACY WORKING PARTY</b>	September 2007
<b>ADOPTION BY CHMP</b>	18 October 2007

<b>KEYWORDS</b>	Adaptive Design, Interim Analyses; Design Modifications; Randomised Clinical Trials; Confirmatory Clinical Trials; Biostatistics
-----------------	--

**REFLECTION PAPER ON METHODOLOGICAL ISSUES IN CONFIRMATORY  
CLINICAL TRIALS PLANNED WITH AN ADAPTIVE DESIGN**

**TABLE OF CONTENTS**

**EXECUTIVE SUMMARY ..... 3**

**1. INTRODUCTION ..... 3**

**2. SCOPE ..... 3**

**3. LEGAL BASIS ..... 4**

**4. MAIN REFLECTION PAPER TEXT ..... 4**

**DEFINITIONS ..... 10**

**REFERENCES ..... 10**

## **EXECUTIVE SUMMARY**

In some instances studies can be planned with a so-called adaptive design involving design modifications based on the results of an interim analysis. Such a design has the potential to speed up the process of drug development or can be used to allocate resources more efficiently without lowering scientific and regulatory standards. This is especially welcome if at the same time the basis for regulatory decision-making is improved.

However, in a clinical development plan the purpose of phase III is to confirm the findings from pre-clinical studies, tolerability studies, dose-finding and other phase II studies (CPMP/EWP/2330/99). To argue for design modifications in a phase III trial (or a late stage phase II trial supposed to be part of the confirmatory package) is then a contradiction to the confirmatory nature of such studies and will be rarely acceptable without further justification: adaptive designs should not be seen as a means to alleviate the burden of rigorous planning of clinical trials. Instead, adaptive designs would be best utilised as a tool for planning clinical trials in areas where it is necessary to cope with difficult experimental situations. In all instances the interim analysis and the type of the anticipated design modification (change of sample size, discontinuation of treatment arms, etc.) would need to be described and justified in the study protocol. Adaptations to confirmatory trials introduced without proper planning will render the trial to be considered exploratory.

Using an adaptive design implies that the statistical methods control the pre-specified type I error, that correct estimates and confidence intervals for the treatment effect are available, and that methods for the assessment of homogeneity of results from different stages are pre-planned. A thorough discussion will be required to ensure that results from different stages can be justifiably combined. The body of evidence justifying the final treatment recommendation must be discussed. The need for a change in the study design and the change itself may have implications for the clinical interpretation of the results, which deserve consideration at the planning stage.

### **1. INTRODUCTION**

Clinical trials often take years to recruit and adequately follow up patients and even with the best knowledge from a carefully planned phase II programme, there may still be uncertainty at the beginning of phase III concerning various aspects of design or analysis. There is much interest, therefore, in being able to carry out interim assessments of long running trials to ensure that the design is still appropriate to meet its objectives or that accumulating data on safety and efficacy do not indicate that the trial should be modified or even stopped.

So-called “group sequential” designs have been developed that avoid inflating the pre-specified type I error associated with the repeated testing of the treatment effect based on accumulating data. Thereby, these designs avoid increasing the probability of a false positive conclusion based on the results of the ongoing trial. Methodology has been developed further to include, for example, varying the number and timing of the interim analyses, or the rules for stopping the trial early due to efficacy or futility. Newer developments of so called “adaptive designs” allow much broader design modifications (e.g. sample-size, randomisation ratio, number of treatment arms) at an interim analysis of an ongoing trial, whilst still fully controlling the pre-specified type I error.

### **2. SCOPE**

The option to modify the design of an ongoing clinical trial in the framework of an “adaptive design” is intuitively appealing. The opportunity to correct misjudgements on the basis of data from a planned interim analysis is likely to increase the chance of the trial formally being a success (i.e. that the null-hypothesis can be rejected).

Whilst the increased flexibility that is now available may well fit the needs in early phases of drug development, their use in late phase II or confirmatory phase III trials deserves a more cautionary approach. This Reflection Paper does not discuss specific statistical methods. Rather it focuses on the opportunities for interim trial design modifications, and the prerequisites, problems and pitfalls that must be considered as soon as any kind of flexibility is introduced into a confirmatory clinical trial intended to provide evidence of efficacy.

This document first outlines some general considerations for studies incorporating interim analyses. This should be seen as a basis for an in-depth discussion of those potential obstacles to interpretation that need be foreseen if, in addition, a design modification at an interim analysis is pre-planned. As, in principle, a large variety of design modifications are possible from a statistical point of view a set of minimal requirements is outlined that must be fulfilled whenever confirmatory clinical trials are planned with an adaptive design. Later sections comment on specific design modifications that have been proposed in the relevant literature.

This document should be read in conjunction with existing regulatory guidance. A selection of relevant guidance documents is referenced in the respective section later in this document.

### **3. LEGAL BASIS**

Not applicable

### **4. MAIN REFLECTION PAPER TEXT**

#### ***4.1 Interim analyses - general considerations***

##### ***4.1.1 The importance of confidentiality of interim results***

Assessment of results from clinical trials involves, amongst other issues, a full discussion of potential sources of bias. In trials that have had interim analyses, it is possible to assess patient demography and to estimate the size of the treatment effect from the data collected before, and after, the interim analysis and check these for consistency. Substantial discrepancies with respect to the types of patients recruited and / or results obtained will raise concern: it will be difficult to interpret the conclusions from the trial if it is suspected that the observed discrepancies are a consequence of (intentional or unintentional) dissemination of the interim results. This problem is usually of even greater importance in situations where treatments cannot be fully blinded or the assessment of results incorporates some subjective element.

Although even substantial discrepancies in the estimated treatment effects could be simply due to chance, and although most sponsors would plan careful procedures to minimise the risk of communication of interim results, it would always be difficult to convincingly demonstrate that no unblinded interim results have been released. Interim analyses, therefore, always introduce the possibility of damaging the integrity of a trial. To minimise these risks, three important issues need to be considered during the planning stage of the study: (i) is there a need to perform any interim analysis? (ii) is the number of interim analyses justified? and (iii) is the information flow carefully described and controlled? In general, interim data should be provided by an independent statistician to an independent decision-making committee (see Guideline on DMCs); sponsor involvement is discouraged.

A balance has to be achieved between the needs for assessing accumulating information and the risk of damaging the integrity of the trial. Routinely breaking the blind should be avoided particularly when it can be foreseen that insufficient information will be available for stopping the study because of proven efficacy, or futility, or meaningful safety concerns of the experimental treatment.

##### ***4.1.2 Considerations about stopping trials early for efficacy***

Different interim analysis plans include different types of adjustment to the nominal significance level (i.e. the level that is chosen for the assessment of an interim result in order to maintain the desired overall significance level). The statistical theory is well developed and permits a flexible spending of the type-I- error over the course of the trial.

The choice of which plan to use can depend on which properties are considered important for a particular trial. Often it may not be acceptable to stop a trial very early, despite convincing efficacy results, because insufficient data on safety, or on secondary endpoints may be available.

Therefore, from a regulatory point of view, any interim analysis without realistic objectives should be avoided when planning the trial. In addition, decision making about stopping a trial early should also consider the fact that acceptance of study results is not only based on a statistically significant primary result. Primary efficacy data should be complimented by a careful assessment of consistency of trial

results beyond the primary variable(s), including results in important subgroups, and the adequacy of the safety database. A discussion is needed as to whether all the requisite information can be provided if the study is stopped at an interim analysis.

#### 4.1.3 *Overrunning*

In many clinical trials the primary endpoint is not observed immediately for each patient (e.g. survival or time to event data). Furthermore, in trials with a complex organisational structure, additional patients are likely to be randomised or some even followed to the primary endpoint before the results of a pre-planned interim analysis are known. If a trial is to be terminated as a result of an interim analysis it is always important to carry out an additional analysis including all of these further patients that did not contribute to the interim analysis. It may be that when this analysis is carried out, the null hypothesis can no longer be rejected and apparently decision making may depend on whether or not these so called overrunning patients are included or excluded from the analysis. In such a situation, it is accepted regulatory practice to base decision making on the final results of the trial (not the interim analysis). This is also in accordance with the intention to treat principle that all randomised patients should be analysed. Obviously, overrunning patients need to be treated and observed according to the protocol and due attention should be given to this at the planning stage of the trial.

A full discussion of the results of a trial should be based on estimates of the treatment effect rather than simply on *P*-values alone. If the estimate of the treatment effect *including* the overrunning patients is not very different from that *excluding* them, then a small increase in the *P*-value might not be regarded as a concern. An important reduction in the size of the point estimate might, in contrast, lead to reluctance to accept the overall result as “positive”, especially as, unless a trial is stopped very early, the proportion of overrunning patients will usually be sufficiently small such that the estimate of the treatment effect should not be substantially altered. In all cases, results including and excluding the overrunning patients should be presented and differences between these two analyses should be discussed.

### 4.2 *Interim analyses with design modifications*

#### 4.2.1 *Adaptation of design specifications: minimal requirements*

In general, changes to the design of an ongoing phase III trial are not recommended even though, in an adaptive design, such changes could be introduced with full control of the type 1 error. If design changes are anticipated in a confirmatory clinical trial this would require pre-planning and a clear justification from an experimental point of view: limitations of the current knowledge (e.g. at the end of phase II) should be outlined in the study protocol to justify the option to modify the design. Argumentation should be provided as to why an anticipated design modification can safely be incorporated into the study and why the interpretability of the trial is not endangered. A discussion would be required why it would not be more appropriate to resolve the outstanding issues in additional exploratory studies.

If an adaptive design is used, the number of design modifications should be limited. Phase III trials are supposed to confirm hypotheses generated in earlier trials about efficacy, and to some extent safety, of a particular drug under particular experimental conditions. The need to modify a number of design aspects, e.g. re-assess sample size, change inclusion or exclusion criteria, change dosing, treatment duration, mode of application, allow for alternative co-medications, may change the emphasis from a confirmatory trial to an hypothesis generating, or exploratory, trial.

A minimal prerequisite for statistical methods to be accepted in the regulatory setting is the control of the pre-specified type I error (for a definition see Point to Consider on multiplicity issues in clinical trials). Corresponding methods to estimate the size of the treatment effect and to provide confidence intervals with pre-specified coverage probability are required in addition to the presentation of the *P*-value. The mere presentation of *P*-values (as often done to compare interim results to pre-specified statistical “stopping” rules) is of little value for interpreting clinical benefit.

Whenever the effect of treatment on a certain endpoint can be measured on different scales, a measure for the treatment effect that is readily interpretable for clinicians should be preferred. Standardised treatment effects should not be used to overcome the requirement to estimate the treatment effect in a

situation where a change in scale or other design modifications are foreseen that would then prohibit estimation of the treatment effect.

Studies with interim analyses where there are marked differences in estimated treatment effects between different study parts or stages will be difficult to interpret. It may be unclear whether the estimated treatment effects differ just by chance, as a consequence of the intentional or unintentional communication of interim results, or for other reasons. This problem can be even greater if the study design has been changed as a result of an interim analysis.

From a regulatory point of view, whenever trials are planned to incorporate design modifications based on the results of an interim analysis, the applicant must pre-plan methods to ensure that results from different stages of the trial can be justifiably combined. The justification should include consideration of the impact of the modification on patient population, schedule of assessments, patient management and other features of the trial, which, if affected, would complicate the interpretation of the trial. In this respect, studies with adaptive designs need at least the same careful investigation of heterogeneity and justification to combine the results of different stages as is usually required for the combination of individual trials in a meta-analysis.

Depending on the nature of the design modification, the simple rejection of a global null hypothesis across all stages of the trial may not be sufficient to establish a convincing treatment effect.

Addressing such issues at the planning stage of the trial is essential to avoid post-hoc discussions about whether observed data may indicate that combining results from different stages of the trial is justified or not. Hence, trials should not be planned to make many changes along a series of small steps with limited numbers of patients at each step. Insufficient information will be available from such trials to justify the consistency of the treatment effect after a design modification.

The involvement of sponsor personnel in interim decision making remains controversial. It is appreciated that decisions in certain types of adaptive trials are more complicated to set into an algorithm for independent interpretation than, for example, a sample-size re-estimation problem or group-sequential stopping guidelines. Nevertheless sponsor involvement introduces an additional risk when the credibility of the trial results is challenged: with sponsor involvement it would be more difficult to argue that importantly different results from different stages are only due to chance.

#### *4.2.2 Sample size reassessment*

When planning late phase II or even phase III trials, considerable uncertainty may still exist about the assumptions needed for an appropriate sample size calculation. Some experimental situations exist where studies often fail because the placebo response cannot be reliably predicted. Consequently, assumptions used for determining the sample size may not reflect reality for a particular study.

The option to reassess sample size in an ongoing trial should not be seen as a substitute for careful planning. The relevance of a particular size of treatment effect should be discussed at the planning stage of the trial and not deferred to the point where interim results are already available. In general, consideration of what magnitude of treatment effect might be of clinical importance should not be influenced by trial results (interim or final).

In cases where good justification can be provided that uncertainty about the required sample size is not an indicator for insufficient research in earlier stages, sample size reassessment based on results of an ongoing trial is an option. Analysis methods that control the type I error must be pre-specified. Whenever possible, methods for blinded sample size reassessment that properly control the type I error should be used, especially if the sole aim of the interim analysis is the re-calculation of sample size. In cases where sample size needs to be reassessed based on unblinded data, sufficient justification should be made.

Whatever approach to sample size reassessment is taken, the very need for reassessment may indicate that crucial design assumptions (e.g. about response rate or variability) are not met.

The need to reassess sample size in some experimental conditions is acknowledged. However, if more than one sample size reassessment becomes necessary, concerns would be raised that the experimental conditions are fluctuating and are not fully understood.

### 4.2.3 *Change or modification of the primary end-point*

A large and steadily increasing number of indication-specific guidelines for the conduct of clinical studies investigating new treatments have been developed by regulatory authorities and learned societies. These guidelines cover principal aspects of study design such as acceptable control groups and preferred primary and secondary endpoints. Nevertheless, in some cases there may still be choice regarding the selection of the primary endpoint and there are many therapeutic areas where respective guidelines have not yet been developed. External knowledge from other studies or interim results may suggest that assumptions and expectations regarding the definition of the primary endpoint do not hold, or that other variables may be better suited to describe a treatment benefit. In such cases, adaptive designs have been proposed as a methodological framework to incorporate changes of the primary endpoint, changes in the components of a composite primary endpoint, or the definition of so called responder criteria.

A change in the primary endpoint is difficult to justify: primary endpoints are chosen to describe a clinically relevant treatment effect. It is acknowledged that, when a certain primary endpoint is chosen at the planning stage, practicality and feasibility sometimes also play important roles. Once the study is ongoing it is difficult to imagine any situation where the perception of what constitutes a relevant clinical benefit should change based on interim results, especially as primary endpoints are usually not selected to differentiate between treatment and control group.

Furthermore, in a confirmatory setting, effects must always be attributable to specific endpoints to clarify the capabilities of the drug treatment. The mere rejection of a global hypothesis combining results from different endpoints will not be sufficient as proof of efficacy.

### 4.2.4 *Discontinuing treatment arms*

In many therapeutic areas, standards of treatment and their benefits over placebo are well established. New experimental treatments might then be licensed based on a demonstration of non-inferiority to an established reference treatment. In other cases, the demonstration of assay sensitivity is difficult (e.g. regarding description of the patient population and placebo response). Response with a reference treatment may sometimes be difficult to predict. This, however, is a minimal pre-requisite for the justification of a non-inferiority margin. In such cases it will usually be necessary to include placebo as well as an active comparator in a confirmatory phase III trial.

An adaptive design, combined with a multiple testing procedure, may offer the opportunity to stop recruitment to the placebo group after an interim analysis as soon as superiority of the experimental treatment (or the reference treatment, or both) over placebo has been demonstrated. The trial might then be continued into a second stage to demonstrate an acceptable level of clinical non-inferiority between the experimental treatment and the reference treatment.

Such an approach deserves very careful planning. It is a well known that clinical trials do not recruit random samples of potential patients. It might be that different types of patients would be recruited into a two-arm trial comparing an experimental treatment with placebo, a three-arm trial including experimental, reference and placebo arms, and a two-arm, non-inferiority, trial comparing the experimental with the reference treatment (e.g. placebo-controlled trials may include a patient population with less severe disease compared to a trial including active treatments only). The treatment effect may then differ to an extent that may make the combination of results from different stages impossible where the placebo arm has been stopped after an interim analysis. Consequently, all attempts would ideally be taken to maintain the blind and to restrict knowledge about whether, and at what time, recruitment to the placebo arm has been stopped. Concealing the decision as to whether or not the placebo-group has been stopped may complicate the practical running of the study and the implications should be carefully discussed. Given these difficulties, an imbalanced randomisation favouring active treatments over placebo for the total duration of the trial may be the more advantageous approach from experimental grounds.

Similarly, in some instances, even after a carefully conducted phase II programme, some doubts about the most preferable dose for phase III may still exist. Investigators may wish to further investigate more than one dose of the experimental treatment in phase III. Early interim results may resolve some of the ambiguities with regard to efficacy or safety considerations, and recruitment may be stopped for

some doses of the experimental treatment. The second stage of such a study would then be restricted to the control treatment and the chosen dose of the experimental treatment. Again, potential implications regarding the patient population selected for inclusion should be discussed in the study protocol. The mere rejection of the global null hypothesis at the end of the trial is not usually sufficient as proof of efficacy: it is not sufficient to show that some dose of the experimental treatment is effective. In consequence, a multiple testing procedure to identify the appropriate dose should be incorporated.

In addition, suppose that a particular dose group has not formally shown efficacy with data from stage I (e.g. this dose has not formally shown superiority over placebo) and is not taken forward to stage II of the trial. In this situation only those hypotheses that have been selected for the second stage should form the basis of a claim at the end of the study, even if at this stage, for example, superiority over placebo can be demonstrated retrospectively for a dose that has not been taken forward to the second stage. Multiple testing procedures should be carefully developed and explained in the study protocol.

#### *4.2.5 Switching between superiority and non-inferiority*

Active controlled trials are of increasing importance. Whenever a non-inferiority trial is planned, discussion of the assay sensitivity of the trial is of paramount importance and justification of a non-inferiority margin needs particular consideration (for principal considerations see: Points to Consider on switching between superiority and non-inferiority; Points to Consider on choice of the non-inferiority margin).

Whenever superiority *or* non-inferiority compared to an active treatment are acceptable outcomes, it is wise to plan the study as a non-inferiority trial and to foresee in the plan how a switch to superiority could be accomplished based on the trial results. This avoids the ambiguity inherent in all post-hoc justifications of non-inferiority margins. Similarly it is generally not acceptable in the adaptive framework to change the trial objective from superiority to non-inferiority (and to justify a non-inferiority margin) after interim results are available.

If a trial has been planned to show non-inferiority and this can be established based on interim results, there may be a desire to continue the study to demonstrate, with additional patients, superiority of the experimental treatment over the active comparator. This possibility should, however, be set into perspective. Replicating the non-inferiority finding in an independent study population and combining the findings of the two studies in a meta-analysis (which may then show superiority) would usually be a preferable strategy to demonstrate both consistency of findings and superiority.

If a trial is to continue after formal proof of non-inferiority at an interim analysis then any ambiguity in interpretation of results at the final analysis should be avoided by basing final conclusions on *all* data from *all* stages of the study, even if the final confidence interval is less supportive than the one obtained from the interim analysis. If the final confidence interval is less supportive than the result at the interim analysis this might be an indicator of heterogeneity in treatment effects estimated from the two stages of the trial. Such a finding would require further investigation and discussion.

#### *4.2.6 Randomisation ratio*

The 1:1 randomisation ratio, usually applied in trials intended to demonstrate superiority of the experimental treatment over comparator, may not necessarily be the most efficient allocation in non-inferiority trials. It is always useful to increase the number of patients that are treated with the new, experimental treatment as in general the safety profile of the comparator is much better established. If, based on interim analysis results, it can be assumed that the trial will still have sufficient power but using a randomisation ratio of, say 1:2, then this may be seen as a useful option.

#### *4.2.7 Phase II / phase III combinations, applications with one pivotal trial and the independent replication of findings*

In some cases in late phase II development, the selection of doses is already well established and further investigation in phase II would be performed in precisely the same patient population and with the same endpoints that are of relevance in phase III. Similar considerations as outlined for the selection of treatment arms at an interim analysis may apply and permit the conduct of a combined



phase II/phase III trial. The study protocol should then discuss why sufficient evidence is expected from the phase II / phase III combination trial compared to the strategy with another phase II trial that is followed by a separate phase III clinical trial. Sponsors should consider carefully the value of independent replication of results and the value of waiting to confirm the detailed design of the confirmatory study until all important exploratory data are available.

In indications where two phase III clinical trials are usually expected as a basis for drug licensing, it will usually not be acceptable to argue for a combination of a phase II and a phase III trial and, at the same time, for the acceptability of an application with only this one combined phase II / phase III trial: a major prerequisite for an application with one pivotal trial in phase III has always been that a sufficient body of evidence from phase II is *already available* so that phase III can be limited to simply replicating these findings in an independent setting. It may, however, be an alternative to plan two independent phase II / phase III combination trials, although it may be logistically challenging to guarantee that the same decision rules are followed in both trials. Similarly, for indications where a single confirmatory trial might be acceptable, evidence from a Phase II / III trial only, without other supporting evidence, would not be sufficient.

Investigation of drugs for the treatment of orphan diseases is a difficult task and specific requirements apply (see Guidance on Clinical Trials in Small Populations). A single phase II / phase III combination trial may be justified if such an approach is more efficient to display the totality of available information that can be derived from a limited number of patients.

Phase II / phase III combination trials, when appropriately planned, may be used to better investigate the correlation between endpoints usually used to support dose-finding and clinical endpoints, and may, therefore, support the process of providing justification that an optimal dose-regimen for the experimental drug has been selected. They may also permit a more extensive investigation of dose-response than might otherwise be feasible.

#### 4.2.8 *Substantial changes of trial design*

In rare instances, changes may be considered necessary to an ongoing trial which have to be introduced via a formal protocol amendment. Examples are changes in the duration of treatment, to mandatory co-medications, or to the criteria for inclusion or exclusion of patients.

In many regulatory discussions the impact of such changes has been felt to be so substantial that the recommendation was to re-size the trial so that the primary analysis can be based on results of the trial when restricted to patients randomised after the change was made, even if this formally contradicts the ITT principle. As a minimal requirement the primary analysis should be stratified according to whether patients were randomised before or after the protocol amendment and homogeneity of the results should be carefully investigated and discussed.

Even though such modification of an essential design feature is possible in a study planned with an adaptive design with full control of the type 1 error, obviously the same degree of caution is needed before interpreting the overall results in such a study. In all such cases, the applicant will need to carefully argue why a combination of the results from different stages is capable of substantiating a final treatment recommendation and why the confirmatory nature of the trial is not damaged.

#### 4.2.9 *Futility stopping in late phase II or phase III clinical trials*

In some cases, studies are terminated based on informal decisions regarding futility based on results of interim analyses. Newer developments in the methodology of group sequential trials and adaptive designs allow more formal futility decisions to be made that impact on the type I error used for efficacy assessment at the interim analysis. If, however, a sponsor decides to continue a trial despite the fact that at an interim analysis suggest stopping the trial for futility, the type I error rate is usually no longer controlled.

Similarly, the overall type I error rate of the MAA (that is, after all the trials have been completed) may also no longer be controlled if a study is stopped, although pre-specified stopping criteria are not met, and a new study is started thereafter. As always, reasons for stopping a study have to be fully explained and understood.

## DEFINITIONS

A study design is called “*adaptive*” if statistical methodology allows the modification of a design element (e.g. sample-size, randomisation ratio, number of treatment arms) at an interim analysis with full control of the type I error.

The term “*difficult experimental situation*” has been used in this document to describe diseases, indications, or patient populations, where it is common knowledge that clinical trials will be difficult to perform. Examples include situations where (i) placebo response is difficult to predict, even in situations where criteria for inclusion and exclusion of patients to trials are well defined (ii) small populations or orphan diseases with constraints to the maximum amount of evidence that can be provided, and (iii) ethical constraints to experimentation.

***Confirmatory trial, confirmatory nature of a trial:*** In section 2.1.2 in ICH-E9 a confirmatory trial is defined as “an adequately controlled trial in which the hypotheses are stated in advance and evaluated. As a rule, confirmatory trials are necessary to provide firm evidence of efficacy and safety”.

## REFERENCES

[Note for Guidance on Statistical Principles for Clinical Trials \(CPMP/ICH/363/96\)](#)

[Points to Consider on Multiplicity issues in Clinical Trials \(CPMP/EWP/908/99\)](#)

[Points to Consider on Application with 1.\) Meta-analyses and 2.\) One Pivotal study \(CPMP/2330/99\)](#)

[Points to Consider on Switching between Superiority and Non-inferiority \(CPMP/EWP/482/99\)](#)

[Points to Consider on Choice of the Non-Inferiority Margin \(CPMP/EWP/2158/99\)](#)

[Guideline on Data Monitoring Committees \(CHMP/EWP/5872/03\)](#)