European Medicines Agency

SUBMISSION OF COMMENTS ON Reflection Paper on Methodological Issues in Confirmatory Clinical Trials with Flexible Design and Analysis Plan, CHMP/EWP/2459/02

COMMENTS FROM German Region of the International Biometric Society
GENERAL COMMENTS
1. The paper summarizes issues to consider when performing interim analyses in general. Much of the critique concerns group sequential designs and interim analyses rather than its adaptive extension. We propose a structure of the paper where – at least – the arguments for and against interim analyses are more explicitly divided from arguments for and against adaptive designs. Recommendations when to use group sequential and/or adaptive techniques should be given. At the moment, it is rather a mix of statements on different levels concerning issues and pitfalls when performing trials with interim analyses.
2. An overall quite negative attitude towards the use of adaptive designs is expressed in the reflection paper. Generally, interim analyses (based on unblinded data) are considered as factors that might damage the integrity of the trial and there is a strong distrust when performing these analyses. So the intention of the paper is unclear. It does not serve as a guideline, but as a paper that prevent people from using these designs. There are currently various ongoing initiatives by researchers from regulatory agencies, pharmaceutical companies and academia to develop and understand new methods for more efficient clinical development, including strategies to adapt clinical trials. The current document – if released as EMEA's official firm position - may actually have a negative impact on those scientific efforts.
3. It is stated that the reflection paper *"should be read in conjunction with regulatory guidance"*. According to ICH-E9 and ICH-E6, it is a basic requirement for trial sponsors to ensure scientific credibility of results and ethical trial conduct through trial monitoring. and interim analysis. This may well require interim analyses and measures to adapt trial designs. Interim analyses (based on unblinded data) are described as a valid and useful tool in clinical trials and principally approved. Thus it is difficult to understand why the reflection paper has strong reservations against performing interim analyses. Conflicting guidance should be avoided.
4. It is common practice in drug development trials to prevent dissemination of interim results. This is usually realized and assured by an Independent Data Monitoring Committee (IDMC) having sole access to the results. The role of an IDMC in the conduct of a trial with planned interim analyses is completely missing in the paper. It would be helpful if the paper addresses the interim data review and decision process incorporating an IDMC in general, and especially gives a thorough discussion of additional issues related to adaptive designs.
5. At several points, the necessity of justification of a specific statistical technique is mentioned. This requires a "positive list" rather than a "negative list", the latter being provided in the paper. As an example, we propose not to rule out unblinded sample size reassessment strategies, but to lay out guidelines under which circumstances these might be acceptable.

7 Westferry Circus, Canary Wharf, London, E14 4HB, UK
Tel. (44-20) 74 18 84 00 Fax (44-20) 74 18 85 95
E-mail: mail@emea.eu.int http://www.emea.eu.int

6. More detailed definitions should be given in order to clarify the discussion. Specifically, it is necessary to define the “confirmatory nature” of a design. Is adaptive = flexible? What is group sequential and what is adaptive? Note that group sequential designs are already adaptive in the total sample size.
7. This reflection paper contains a lot of relevant and important aspects that should be considered when planning to perform a clinical trial with a so-called adaptive design. However, the role of a “Reflection paper” in contrast to “Note for guidance” or “Points to consider” is unclear. Generally, it seems problematic to publish a paper by EMEA that seems to warn against the use of a specific statistical method. Although it is a so-called “reflection paper” it will be understood like a Guideline or a Points to Consider paper and hence will have an enormous influence on the acceptability of such trials.

SPECIFIC COMMENTS ON TEXT

GUIDELINE SECTION TITLE

Line no[1]. + paragraph no.	**Comment and Rationale**	**Proposed change (if applicable)**
4.1 in general	The whole chapter implies that the planning and conduct of interim analyses will in general damage the integrity of the trial. Its content is in contrast to ICH-E9, sections 3.4, 4.1 and especially 4.5.and 4.6, where interim analyses are not discouraged and guidance on their proper conduct is provided.	
4.1.1	The text of this section does not give guidance on how the confidentiality of interim results can be guaranteed, but keeps an unsubstantiated, sceptical attitude towards all measures implemented to assure confidentiality. The role of an Independent Data Monitoring Committee (IDMC) for inspection of results of unblinded interim analyses is important when considering confidentiality of interim results. This is not mentioned at all in the present document.	Give guidance which measures have to be undertaken to guarantee confidentiality (role of IDMC).
4.1.1 2nd paragraph line 5	*"To minimize these risks, two important issues need to be considered: (i) the need to perform any interim analysis, and (ii) the total number of interim analyses, should be carefully justified:"* Add "(iii) dissemination of results, i.e. who get what information when	Add (iii)
4.1.2 5th paragraph	This statement seems to be in conflict with the rational of group sequential designs. The objective of “early stopping” is to stop the study when there is convincing evidence that the trial objectives are met with sufficient information. In addition, the complete picture on a safety and efficacy is never gained in only one trial. It would be helpful to include that early stopping of a well planned trial implies that the study design questions have been sufficiently addressed.	Rephrase paragraph

[1] Where available

4.1.2 2nd paragraph line 1 and 4th paragraph line 5	The term *"fixed nominal level"* in contrast to an increasing nominal level is misleading. This should be changed to the term *"constant nominal level"*. But also in this case, the statement in parentheses *"although it is larger than in a plan where a fixed nominal level is used"* is only true for a trial with sufficient power.	Change wording from 'fixed' to 'constant'. Rewrite or delete sentence in parantheses.
4.1.3 2nd paragraph	*"However, it is well known that estimates of the size of treatment effect after terminating a trial based on an interim analysis, on average, over-estimates the true treatment difference. An important reduction in the size of the point estimate might thus lead to a reluctance to accept the overall result as "positive"."* The statistical methods dealing with bias correction should be mentioned. Additionally, over-estimation of the true treatment difference applies also to the analysis including overrunning patients.	Rephrase paragraph
4.2.1 1st paragraph line 1	*"In general, changes to the design of an ongoing phase III trial are not recommended."* It is not clear whether this refers only to unplanned design changes or also to planned design changes.	Clarify
4.2.1 1st paragraph line 5	*"The need to re-assess sample size, change inclusion or exclusion criteria, change dosing, treatment duration, mode of application, allow for co-medications, etc. typically change the emphasis from a confirmatory trial to an hypothesis generating, or exploratory, trial."* Again, this statement is in contrast to ICH-E9. Moreover, there is no rationale for such a broad statement and no justification, why e.g. sample-size re-assessment - which not even requires unblinded interim analysis and does not change any design feature – is mentioned in the same category as change of treatment duration.	Rephrase paragraph
4.2.1 3rd paragraph	*"Whenever a treatment effect with respect to a certain endpoint can be measured on different scales a measure for the treatment effect that is readily interpretable for clinicians should be preferred."* This statement is true, but it is not clear what it has to do with the topic of this paper.	Delete paragraph
4.2.2 4th paragraph line 1	*"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."* One must bear in mind that there are many areas and indications where it is not possible to have such reliable information for the endpoints needed in a phase III trail. Therefore, sample size reassessment may be a proper way. One should rather add a paragraph discussing the dissemination of information on sample size adjustment and how to avoid bias.	
4.2.3 last paragraph	*"The mere rejection of a global hypothesis combining results from different endpoints will not be sufficient as proof of efficacy. Technical prerequisites are that it will always be necessary to demonstrate that any proposed new endpoint (or proposed change of component to a composite endpoint, etc.) has been recorded and monitored with the same diligence*	

	as the originally pre-specified endpoint." Guidance is expected whether there are circumstances where a change of endpoint is acceptable. Instead, an unclear statement about technical prerequisites is given.	
4.2.4 3rd paragraph line 7	*"Consequently, all attempts should be taken to maintain the blind and restrict knowledge that recruitment to the placebo arm has been prematurely stopped.....Given these difficulties an imbalanced randomization favouring active treatments over placebo for the total duration of the trial may be the more advantageous approach from experimental grounds."* While this may be the case, ethical considerations must be taken into account. It may be more ethical to stop a placebo arm than having fewer patients on placebo for the entire trial. Ethical consequences regarding patient information and informed consent must also be taken into account. It does not seem appropriate to restrict knowledge from the patients that one of the treatment arms have been stopped prematurely.	
4.2.6	The change of the randomisation ratio seems to be one of the most accepted modification planned in an adaptive design. We do not see the big difference to other modifications seen more sceptical, since also the change of the randomisation ratio is some kind of sample size modification and, additionally, may have implications regarding the patient population selected for inclusion.	
4.2.7	This section seems to refer to inferentially seamless Phase II/III trials. It would be helpful if recommendations could be provided for seamless trials where data is not combined across phases to make inference but information from the Phase II portion is used to determine some design features of the Phase III portion (such as dose). The particular design features to be determined and criteria can be planned a-priori. This is similar to the typical case of a separate Phase II and Phase III design except that down time is eliminated.	
4.2.9.	*"… the overall type I error rate of the MAA …".* This concept appears here for the first time. So far, only the family-wise type I error rate of the trial has been a regulatory concern. The general topic of this reflection paper seems unrelated to the discussion of statistical properties of MAA evaluation and the Points-to-Consider document CPMP/EWP/2330/99 ("Application … one pivotal study") would be a better place.	Delete paragraph
4.2.9	*"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."* This implies that in a properly planned study with futility stops an increase of the nominal alpha level (allowed by design) is accepted in principle from a regulatory point of view. This should be stated explicitly.	