ACRO
ASSOCIATION OF CLINICAL RESEARCH ORGANIZATIONS

March 9, 2017

Division of Dockets Management (HFA–305)
Food and Drug Administration
5630 Fishers Lane, Rm. 1061
Rockville, MD 20852.

**Re: Multiple Endpoints in Clinical Trials; Draft Guidance for Industry**
**Docket No. FDA–2016–D–4460**

Dear Sir/Madam:

The Association of Clinical Research Organizations (ACRO) represents the world's leading, global clinical research organizations (CROs). Our member companies provide a wide range of specialized services across the entire spectrum of development for new drugs, biologics and medical devices – from discovery, pre-clinical, proof of concept and first-in-man studies through post-approval and pharmacovigilance research. With more than 130,000 employees engaged in research activities around the world (including 57,000 in Europe), ACRO advances clinical outsourcing to improve the quality, efficiency and safety of biomedical research. Each year, ACRO member companies conduct more than 7,000 clinical trials involving 1.3 million research participants in over 100 countries. On average, each of our member companies works with more than 700 research sponsors annually.

ACRO welcomes and supports the draft guidance issued by FDA on Multiple Endpoints in Clinical Trials, as it addresses several contemporary statistical methods being used in clinical trials. It includes graphical models for constructing testing procedures, which is powerful. It also contains robust and useful guidance on non-parametric procedures of testing multiple hypotheses in clinical trials. However, it is a long and fairly technical guidance document, and ACRO proposes several recommendations in order to increase the usefulness of the document and provide greater clarity, in order to make it easier for users to follow:

- ACRO recommends that the scope of the guidance is more clearly defined than at present. Currently, the Background and Scope section begins with the statement that "Failure to account for multiplicity when there are several clinical endpoints evaluated in a study can lead to false conclusions regarding the effects of the drug." This, together with the title of the document, implies that the scope is the management of multiple endpoints. Lines 241 – 243 also state that "The primary goal of this guidance is to provide recommendations for designing studies that control the chances of erroneously concluding that a treatment is effective with respect to a particular endpoint." However, the final paragraph of the section (lines 345 – 347) states that "Both the issues and methods that apply to multiple endpoints also apply to other sources of multiplicity, including multiple doses, time points, or study population subgroups." This last statement is more consistent with the ICH E9 guideline on Statistical Principles for Clinical Trials, which requires measures to adjust the Type I error for any form of multiplicity (endpoints, treatments, repeated evaluation over time, and interim analysis). However, ACRO notes that lines 182 – 184 of the draft FDA guidance confirm that "When multiple endpoints are examined at an interim analysis, the appropriate adjustments can become complex; discussion of this issue is outside the scope of this guidance." Also, ACRO notes that,

according to lines 324 - 325, certain types of safety analyses are also outside the scope of the guidance. Rather than distributing these references throughout the relatively long section on Scope and Background, ACRO recommends that the final FDA guidance should include separate sections on (1) Scope and (2) Background in order to bring these statements together and provide a clear statement of the multiplicity issues that are considered to be in and/or out of scope. This may also require changing the title of the guidance to, for instance, Multiplicity Issues in Clinical Trials, and ACRO would support such a change if the scope of the guidance warrants it.

- As noted above, the treatment of multiple endpoints examined at an interim analysis is stated (lines 182 – 184) to be outside the scope of this guidance, and ACRO agrees with FDA that this is a complex issue. However, ACRO believes that sponsors would welcome FDA guidance on this important topic and recommends either inclusion in the present guidance or an acknowledgement that FDA intends to address this topic in a separate guidance document. At the very least, ACRO recommends that this statement in the current draft guidance be accompanied by a statement that clinical trial sponsors intending to perform an interim analysis of multiple endpoints should discuss their proposed methodology in advance with the relevant FDA review division.

- While noting (lines 237 – 246) that "The emphasis of this guidance is not on the confidence interval, but rather on the test of a hypothesis" the draft guidance further notes that, in certain situations, "it is critical to ensure that the confidence intervals appropriately reflect multiplicity of hypothesis tests." This, however, is not addressed further in the draft guidance. ACRO considers that this is an important omission and recommends that a greater discussion on the need for simultaneous confidence intervals corresponding to multiple test procedures be included in the final guidance.

- The draft guidance frequently reads as a basic statistics primer with regard to the need to control Type I and Type II errors arising from multiplicity and a description of some of the available techniques to do so. This is very informative, but it is not always clear exactly what FDA is recommending in terms of the application of statistical methodology, and ACRO considers that the value of the guidance would be greatly improved by the addition of clear recommendations, where possible. For instance, section IV.C states that "The Holm (section IV.C.2) and Hochberg (section IV.C.3) methods are more powerful than the Bonferroni method for primary endpoints and are therefore preferable in many cases. However, for reasons detailed in sections IV.C.2-3, sponsors may still wish to use the Bonferroni method for primary endpoints in order to maximize power for secondary endpoints or because the assumptions of the Hochberg method are not justified." This still leaves open the question of the situations in which it is appropriate for the Bonferroni method to be used. The guidance could be expressed more clearly in terms of recommending that sponsors use the Holm or Hochberg methods unless there is a need to maximize power for secondary endpoints or the assumptions of the Hochberg method are not justified, in which case the Bonferroni method should be used.

- The discussion of common statistical methods for addressing multiple endpoint-related multiplicity problems in section IV.C of the draft guidance focuses on the Bonferroni, Holm and Hochberg methods. However, several other methods exist for maintaining Type I error rate control and are not

discussed. The overall impression given by the document (until the Conclusion) is that the Bonferroni, Holm and Hochberg methods are the only techniques acceptable to FDA. ACRO does not believe that it is necessary to address all of the available options in detail, but recommends that the guidance should point out that other options are available and will be acceptable to FDA when used in appropriate circumstances. Additionally, ACRO notes that the draft guidance provides little information on the use of methodology in parametric procedures. ACRO considers that this is an important omission and recommends that a greater discussion of parametric procedures be included in the final guidance document.

- One method that is not addressed in the draft guidance is that of false discovery rate (FDR) control, which has become increasingly standard practice in pharmacogenomic studies. It is now generally recognized that when a study involves a large number of tests, the FDR error measure is a more useful approach to determining a significance cut-off, as the "family-wise error rate" (FWE) approach is too stringent. Recently, some authors (e.g., Glickman et al, Journal of Clinical Epidemiology 67 (2014) 850e857; Benjamini & Cohen, Biostatistics 2017, 18 (1): 91-104) have also advocated the applicability of this approach to clinical research. Consequently, ACRO recommends that the final guidance should address the role of false discovery rate control, and clarify whether pharmacogenomics testing is in or out of scope of the guidance. In this context, ACRO notes that FDA's final Guidance For Industry – Clinical Pharmacogenomics: Premarket Evaluation in Early-Phase Clinical Studies and Recommendations for Labeling (January 2013) recommends that that these studies should be controlled "for the multiplicity and the risk of incorrectly identifying associations in genomic data when many searches are performed (e.g., using Bonferroni correction, false discovery rate, or other method)." However, this final guidance does not include any recommendation on how the FDR should be controlled. Also, the present draft guidance under discussion does not make clear whether the information provided on use of the Bonferroni correction is applicable to pharmacogenomic studies, and ACRO recommends that this is made clear in the final version.

- With regard to the weighted Bonferroni test, the draft guidance states that the "weights are prespecified in the design of the trial, taking into consideration the clinical importance of the endpoints, the likelihood of success, or other factors" (lines 1035 – 1038). However, it is not clear from this statement how the weighting should be calculated on the basis of these factors, and ACRO recommends that this be explained in the final guidance.

- Reference to the impact on sample size of the power (alpha level) specified for the test of hypothesis for each individual endpoint is made throughout the document. While recognizing that it is not possible to give definitive guidance that will cover all circumstances, ACRO recommends that the value of the guidance would be increased by including some general guidance on sample size and power calculations, including those to allow for multiple comparisons in correlated endpoints. Additionally, while recognizing that FDA will not recommend specific statistical software packages, ACRO recommends that it would be helpful if the final guidance were to define the desirable features of appropriate software for these calculations.

ACRO
ASSOCIATION OF CLINICAL RESEARCH ORGANIZATIONS

- ACRO considers that the key recommendation in the draft guidance is stated in the Conclusion section at the very end of the document (lines 1613 - 1617): "There are many strategies and/or choices of methods that may be used, as appropriate, as described in this guidance. Each of these methods has advantages and disadvantages and the selection of suitable strategies and methods is a challenge to be addressed at the study-planning stage. Statistical expertise should be enlisted to help choose the most appropriate approach." ACRO agrees strongly with this statement and recommends, to ensure clarity, that it is presented early in the final guidance document and highlighted as the guidance's key recommendation.

- The draft guidance focuses solely on the statistical issues associated with the analysis of multiplicity in clinical trials. However, in the case of multiple endpoints, these analyses are valid only if the respective endpoints are assessed consistently in all trial subjects. Sponsors should avoid the application of different instruments or techniques for the measurement of any given endpoint. ACRO therefore recommends that the final guidance should include a clear statement of this principle, especially with regard to multi-regional clinical trials where regional differences in assessment techniques may well occur.

**Conclusion – consideration of stakeholder workshop**

The draft guidance provides important and detailed descriptions of various strategies for grouping and ordering endpoints for analysis and for applying some well-recognized statistical methods for managing multiplicity within a study in order to control the chance of making erroneous conclusions about a drug's effects. However, because of both the importance and challenges of managing multiplicity, ACRO asks the FDA to consider convening a stakeholder workshop on this issue before finalization of the draft guidance document. The in-person interaction of a stakeholder workshop might help facilitate moving beyond the general principles ACRO has raised in its comment to a more granular discussion of these important guidelines for managing multiplicity.

ACRO thanks the Agency for the opportunity to provide comments on this draft guidance on Multiple Endpoints in Clinical Trials. Please do contact ACRO if we can provide additional details or answer any questions.

Respectfully submitted,

Karen A. Noonan
Vice President, Global Regulatory Policy
knoonan@acrohealth.org