European Medicines Agency
<Unit>

21OCT2009

SUBMISSION OF COMMENTS ON

Guideline on missing data in confirmatory trials CPMP/EWP/1776/99 Rev. 1 Corr*

COMMENTS FROM:

Name of Organisation or individual
German Region of the International Biometric Society (DR IBS)

Please note that these comments and the identity of the sender will be published unless a specific justified objection is received.

Comments should be sent to the EMEA electronically and in word-format (not pdf).

1. GENERAL COMMENTS

Stakeholder No. *<to be completed by EMEA>*	**General Comment (if any)**	**Outcome (if applicable)** *<to be completed by EMEA>*
	It should be the objective to provide methods yielding unbiased analyses. From our point of view not every method that is conservative is automatically appropriate. The guideline recommends the use of methods that are biased in favour of the control treatment. We think that the primary analysis should provide an estimate of the treatment effect that neither favours the experimental nor the control treatment, based on a plausible assumption on the missing data mechanism. As part of the sensitivity analysis, "conservative" approaches favouring the control arm can be explored. The guideline can be interpreted as pointing out flaws of complex method and favouring simple imputation methods like LOCF instead. Current literature on this topic, clearly showing disadvantages of LOCF vs MMRM, appears to be disregarded (Siddiqui O, Hung HM, O'Neill R. *MMRM vs. LOCF: a comprehensive comparison based on simulation study and 25 NDA datasets.* J Biopharm Stat. 2009;19(2):227-46).Instead of dispelling specific statistical methods, especially sophisticated complex methods for imputation, the Guideline should clearly set out the principles that underpin the analysis of a trial with missing data. It should be recommended to apply the most sensible statistical methods adequate under the corresponding assumption irrespective of the complexity of the method.	
	Sensitivity analyses should show the influence of different assumptions underlying the different statistical analysis models. The three broad classes of missingness mechanism Missing Completely at Random (MCAR), Missing (conditionally) at Random (MAR), and Not Missing at Random (NMAR) have distinct implications for the analysis.	

Stakeholder No. *<to be completed by EMEA>*	**General Comment (if any)**	**Outcome (if applicable)** *<to be completed by EMEA>*
	In practice, one therefore has to 1) assume a reasonable class of missingness mechanism, and 2) perform an analysis which is valid for that class of missingness mechanism. Sensitivity analyses should repeat steps 1) and 2) under different assumptions in 1). In some text passages, the Guideline suggests sensitivity analyses by applying different methods, instead of making different assumptions. A large number of methods all usually rely on the same assumption of MAR. Therefore, sensitivity analysis using a variety of methods all based on the same assumption could lead to a misleading optimistic view of the robustness of the conclusions.	

2. SPECIFIC COMMENTS ON TEXT

Line No of the first line(s) affected. *<e.g. Line 20-23>*	**Stakeholder No.** *<to be completed by EMEA>*	**Comment and Rationale; proposed changes** *<if changes to the wording are suggested, they should be highlighted using "track changes">*	**Outcome** *<to be completed by EMEA>*
Lines 132-133 Lines 138-139 Lines 157-158		**Comments:** Some sentences of the guideline seem to be out of context or unclear, and, therefore, rather lead to confusion than to clarity of the subject. These are for instance lines 132-133, lines 138-139, and lines 157-158. **Proposed change (if any):** These sentences should be skipped or reworded for clarity.	
Lines 179-182 Lines 339-340 Lines 358-397		**Comments:** It would be very helpful to have some references to literature on some statements, e.g. lines 179-182, lines 339-340, and several topics discussed between line 358 and line 397. **Proposed change (if any):**	
Lines 144-153		**Comments:** The statements in lines 144-153 are rather vague and imprecise. It would be helpful to include the definitons of the different missingness mechanisms given in chapter 6.1 already in section 4.2. The statements on bias could then be based on these definitions and could be made much more precise. **Proposed change (if any):** To include the definitons of the different missingness mechanisms given in chapter 6.1 already in section 4.2.	
Lines 175-187		**Comments:** Points like adequate monitoring of the trial, investigator instructions/meetings, data cleaning processes or other possibilities to avoid missing data should be stated more clearly and detailed. **Proposed change (if any):** The cross-functional efforts needed to successfully avoid missing data should be stressed in this section. Training of investigators, adequate monitoring and data cleaning processes as well as other possibilities to avoid missing data should be described in more detail.	
Line 315		**Comments:** The statement that LOCF produces unbiased estimates under the MCAR assumption is not valid. In addition to MCAR, LOCF require the constancy	

Line No of the first line(s) affected. *<e.g. Line 20-23>*	**Stakeholder No.** *<to be completed by EMEA>*	**Comment and Rationale; proposed changes** *<if changes to the wording are suggested, they should be highlighted using "track changes">*	**Outcome** *<to be completed by EMEA>*
		assumption. Since the bias in LOCF has been shown to depend on many factors it is never known how likely it is LOCF is conservative. Therefore, even under MCAR, LOCF can produce severely biased estimates as described by Molenberghs and Kenward (Wiley, 2007). **Proposed change (if any):** This statement should be skipped or reworded accordingly.	
Line 337		**Comments:** Unconditional mean imputation should be strongly discouraged since it is no statistically valid method: beside the fact, that the imputed values are not independent from the other values, this method artificially increases the sample size and reduces the variability but does not change the mean treatment effect at all; this results in an increase of the statistical power which is not based on any additional information. **Proposed change (if any):** Methodological issues of this method should be discussed. Otherwise, statement should be skipped**.**	
Lines 374-383		**Comments:** The guideline argues against the methods which are based on the MAR assumptions as these are seen to correspond to a per-protocol analysis. The guideline favors the ad-hoc single imputation methods, implying that these are closer to the ITT principle. **Proposed change (if any):** It is unclear why this should be the case, and a justification should be given.	
Lines 398-408		**Comments:** For time-to-event data, the assumption of independent censoring is not testable, but nevertheless assumed in all of the standard analyses (Cox-regression, log-rank test, Kaplan-Meier estimates), which are generally accepted as the primary analyses. Alternative assumptions on the censoring mechanism can then be explored as part of the sensitivity analysis. We consider this a very reasonable, pragmatic approach. It is unclear why a corresponding approach could not be generally recommended for other types of data, i.e. a primary analysis based on the	

Line No of the first line(s) affected. *<e.g. Line 20-23>*	**Stakeholder No.** *<to be completed by EMEA>*	**Comment and Rationale; proposed changes** *<if changes to the wording are suggested, they should be highlighted using "track changes">*	**Outcome** *<to be completed by EMEA>*
		MAR assumption, and a sensitivity analysis exploring MNAR assumptions. **Proposed change (if any):**	
Line 453		**Comments:** To conform with the ITT principle, the guideline recommends to take measurements of dropped-out patients until the end of the study, but then suggest to down play a positive outcome when a patient was switched to an active treatment. This approach does not respect ITT, and may often be unfeasible. It is also completely unclear how the retrieved drop-out information should then be used. **Proposed change (if any):**	
Line 459		**Comments:** Worst case analysis does not comprise a meaningful sensitivity analysis for missing data. Already with rather small amounts of missing data, the result of the worst case analysis is predictable and independent of any observed data. (see Unnebrink, K. and Windeler, J.: *Sensitivity analysis by worst and best case assessment: is it really sensitive?* Drug Information Journal 1999 Vol. 33, pp.835-839). **Proposed change (if any):** Worst case analysis should be dropped as an example of a recommended sensitivity analysis.	