European Medicines Agency
<*Unit*>

22 October 2009

SUBMISSION OF COMMENTS ON

GUIDELINE ON MISSING DATA IN CONFIRMATORY CLINICAL TRIALS

CPMP/EWP/1776/99 Rev. 1 Corr

DRAFT **COMMENTS FROM:**

Name of Organisation or individual
European Federation of Statisticians in the Pharmaceutical Industry (EFSPI) Contact person: Pierre Verweij (pierre.verweij@spcorp.com)

1. GENERAL COMMENTS

Stakeholder No. *<to be completed by EMEA>*	**General Comment**	**Outcome (if applicable)** *<to be completed by EMEA>*
	This draft guideline is considered an improvement over the previous guideline and raises many important points on missing data which are fully supported: • The need for pre-specification of statistical analyses with missing data (including sensitivity analyses). • Inclusion of graphical summaries of drop-out patterns (including discussion) • The fact that it is difficult to elucidate whether the relationship between missing values and the unobserved outcome variable is completely absent.	
	However we disagree with a number of suggestions made in the draft guideline. In other places the documents lacks the clarity in a number of topics which are important for analysis of missing data: 1. The guideline generally seems to fail to distinguish between "assumptions" and "methods". The guideline can and should be written based on "assumptions", which will then provide a solid framework for a technically appropriate and logically coherent guideline. Methods can be discussed in terms of their properties relative to these assumptions. Different methods may share the same assumptions, and the distinction between them does not add supporting value in a sensitivity analysis; for example,	

Stakeholder No. *<to be completed by EMEA>*	General Comment	Outcome (if applicable) *<to be completed by EMEA>*
	performing a different method with the same MAR assumption will not provide additional support. A clear advantage of focusing the guideline on "assumptions" and not "methods" is that in contrast to assumptions typically made, methods will certainly evolve and develop over a relatively short period of time. Placing all discussions relating to methods and how they apply to the underlying assumptions in appendices may make updating the document easier in the future. 2. In some settings early withdrawal may be important and appropriately regarded as a treatment failure; for example, trials where the outcome is success/failure and withdrawal is classified as a failure. In other settings, missing data are a "nuisance" which is to be accommodated in the analysis; for example, when there is a requirement to have measurements on each trial participant at particular time points. The guideline seems to be written in the context of the latter scenario, with examples relating to the former scenario juxtaposed. This distinction needs greater clarification in the guideline. 3. The guideline appears to recommend the use of methods that are **biased** in favour of the control treatment. We find that the primary task of a statistician should be to provide an *unbiased* analysis. To favour analyses that tend to work for the null hypotheses drives sponsors of clinical trials to artificially increase the sample size in order to avoid the risk of being not conservative enough in the eyes of a regulator. Such behaviour would increase the costs of drug development without tangible benefit. To create the impression that every method is acceptable as long as the applicant can justify that it is conservative in the	

Stakeholder No. <to be completed by EMEA>	General Comment	Outcome (if applicable) <to be completed by EMEA>
	study at hand does not pave the way for innovations. Furthermore, 'bias' depends on is the definition of the 'true' treatment effect. For example, in most settings, LOCF and MMRM provide estimates of different treatment effects. 4. The guidance on acceptable primary analysis is neither appropriate nor helpful. While in survival analysis it is accepted practice to consider an analysis assuming random censorship (i.e., MCAR) primary, a similar approach is not acceptable for other data. In fact, even an analysis valid under the more general MAR assumption does not seem to meet the expectations of this guideline. The important question on what constitutes an acceptable primary analysis is therefore left unanswered. 5. The guideline seems to implicitly promote single imputation in favour of more formal model-based procedures. If this was not the intention, this could be mentioned more explicitly. Otherwise this would not be in line with the general consensus within the statistical community as –for example– reflected in issue 43(4) of the DIA Journal (2009): "All discussants agree that we should be extremely careful with simplistic methods, and arguably abandon them completely. This includes, in particular, LOCF" [Page 447]; "We do not believe that highly conservative statistical approaches, such as those deliberately penalize only the experimental treatment, should necessarily play a main inferential role, because the overriding objective should be to characterize true effects of treatment regimens as accurately as possible" [Page 474]. 6. The list of sensitivity analyses envisaged in Chapter 7 suffers	

Stakeholder No. *<to be completed by EMEA>*	**General Comment**	**Outcome (if applicable)** *<to be completed by EMEA>*
	from three main problems. First, it is not made clear what constitutes a sensitivity analysis. Second, not all the analyses proposed are indeed sensitivity analyses in the accepted sense. For example, they are not all consistent in the analysis goal, and some are merely alternative analyses made under the same assumptions. Third, the range of possibilities presented points to sensitivity analyses that are too extensive in nature, rewarding diligence with a greater opportunity of undermining the principal analysis. A coherent approach to sensitivity analysis is needed, making clear the importance of (a) the consistency of the analysis goal, which must match the original trial aims, (b) transparency in the assumptions made and varied, and (c) a realistic and appropriate extent of the analysis**.**	
	Additional points to consider in the revision are: • The guideline seems to focus only on analysis of efficacy, and not on safety. It would be helpful to clarify whether the guideline is as applicable for safety endpoints, e.g. lab safety values. • It would be helpful to include some specific mention of the need to consider the potential impact of missing data when performing meta-analysis. The same principles of seeking appropriate estimates which are not inappropriately biased and not spuriously precise apply, and it should also be noted that the choice of methods will impact weighting across trials as well as the estimate from each trial. • Consistency in terminology with other guidance documents, especially ICH E9, should be improved. The text should include a	

Stakeholder No. *<to be completed by EMEA>*	**General Comment**	**Outcome (if applicable)** *<to be completed by EMEA>*
	reference to E9 when mentioning, for example, the ITT principle. Also, *internal* consistency of terminology could be improved. For example: 'active treatment', 'test treatment' and 'experimental treatment' are used interchangeably. Finally, a glossary of terms would be helpful.	

2. SPECIFIC COMMENTS ON TEXT

Line No of the first line(s) affected	**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>*
032		**Comments:** Any analysis in which there are missing data rely on untestable assumptions. This should be discussed early in the document. **Proposed change:** Insert after 1st sentence "When there are missing data, all approaches to analysis rely on untestable assumptions".	
033-035		**Comments:** The use of the terms "method" and "approaches" is inappropriate (see general comments) and should be replaced by "assumptions". This needs to be consistently applied throughout the document. **Proposed change:** "There is no universally applicable set of assumptions concerning missing data, and **therefore** approaches **based on different assumptions will generally** lead …."	
048-051		**Comments:** It is impossible to establish the absence of bias... See comment on line 32. **Proposed change:** Delete sentence beginning "Hence …".	
051-054		**Comments:** The statement is self contradictory (see general comments). Whether a method is appropriately conservative or not depends on the assumptions made, including those about the missing data mechanism (e.g. MAR or MCAR). **Proposed change:** "The justification of selecting a particular method should **be based on the goal of the analysis, assumptions and the way in which the chosen method of analysis achieves that goal in the light of these assumptions**."	

Line No of the first line(s) affected	**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>*
064		**Comments:** See general comments. **Proposed change:** Add additional "We also make a clear distinction between those settings where early withdrawal may be important and appropriately regarded as a treatment failure and those in which missing data are a "nuisance" to be accommodate in the analysis. Only in the latter case does withdrawal lead to what is regarded as missing data in the present context."	
067-076		**Comments:** This paragraph is logically incoherent. Missing data do not violate anything. **Proposed change:** The paragraph needs to be re-considered from a technical perspective and to make clear the scientific points the authors wish to communicate.	
067-068		**Comments:** Missing *data* do not violate the strict ITT principle. Some types of *analyses* violate that principle (e.g. complete case analysis) or an *as treated* analysis. Also the suggestion is made that according to the ITT principle missing data should not occur since all outcomes are to be collected even if the protocol is violated. According to ICH E9, the ITT principle implies that no data that are collected should be excluded from the analyses. The proposal to observe patients until the planned end of study and take measurements, regardless what treatment patients may receive during that time, may be challenging to be followed through in practice there is no generally accepted approach how to use the data in the (primary) analysis. When, for example, a patient on placebo drops out and switches to active treatment the treatment difference may be downplayed. This approach would not respect ITT either. **Proposed change:** Reword as missing data do not violate ITT and clarify the	

Line No of the first line(s) affected	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>*
		comments regarding the ITT principle accordingly.	
083-084		**Comments:** "It should be noted that the strategy employed to handle missing values might in itself provide a source of bias and that there is no universal best approach for all situations." This statement is non trivial and requires explanation. In what cases does a strategy to handle missing values create bias? **Proposed change:** Please add an example about how the choice of strategy could introduce bias and provide guidance.	
100-109		**Comments:** This section is labelled 'Scope', but much of the content does not seem to be on the scope of the guideline, but is a discussion on content. **Proposed change:** This section needs to be shorter, consider removing the parts on methodology..	
105-106		**Comments:** The method chosen should aim to provide a conservative estimate… It would be more correct to aim for an unbiased estimate. **Proposed change:** Replace sentence with "**Given the assumptions made,** the method chosen should aim to provide an **unbiased** estimate…." and delete the phrase "In other words" in the subsequent sentence.	
108		**Comments:** **Proposed change:** Replace "methods" with assumptions"	
132-133		**Comments:** **Proposed change:** Change to "If values for missing data are **multiply** imputed or modelled …… ITT principle**, while single imputation will typically lead to an overestimation of precision**".	

Line No of the first line(s) affected	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>*
134-158		**Comments:** The section on bias is unclear. **Proposed change:** Re-word as per Carpenter and Kenward; for example • In principle missing values will not be expected to lead to bias in naive treatment effect estimates if they are not related to any inference we wish to draw about the treatment effect (e.g. some observations may be missing due to equipment failures in the clinic) • Conversely, if the unmeasured observation is related to the real value of the outcome (e.g. the unobserved measurements have a higher proportion of poor outcomes), this will lead to bias in simple treatment effect estimates even if the missing values are not apparently related to treatment (i.e. missing values are observed equally frequently in all treatment arms) • Missing observations will lead to bias in simple treatment effect estimates if they are related to both the treatment and the unobserved outcome variable (e.g. missing values are more likely in one treatment arm because it is not as effective)	
154-155		**Comments:** Self contradictory for reasons given above in comment on lines 51-54. **Proposed change:** Please re-word.	
155-156		**Comments:** Here the argument is made that because we cannot exclude MNAR we should adopt a conservative approach (often an approach for which we know there is a bias – favouring the null hypothesis). Hence, solutions for which there *may* be a bias are put on equal footing with approaches for which we know there *will* be a bias. **Proposed change:** We suggest that methodological progress concerning principles such as MAR and MNAR is referred to. Sensitivity analyses should address such	

Line No of the first line(s) affected	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>*
		topics and should be interpreted within that framework	
157-158		**Comments:** While it is an excellent idea to address the case of non-inferiority trials, the current statement offers very little guidance on what will be different in the case of a non-inferiority trial. **Proposed change:** Add a paragraph on the impact of missing data in non-inferiority trials.	
175-187		**Comments:** Some other factors avoiding missing data could be mentioned. **Proposed change:** After 'by favouring design' add: '**and monitoring-specific recommendations**'**.**	
183		**Comments:** In some cases the collection of outcome data after withdrawal may not be relevant. **Proposed change: 'w**here possible **and relevant'**	
192-194		**Comments:** Section 5.2 seems to suggest that assumptions should be selected that will result in a realistic, plausible, unbiased treatment effect, which we support. However, throughout the document this seems to be contradicted, and the need for a "conservative estimate" is stated in a number of places. **Proposed change:** Delete the term "conservative", and utilise expressions such as "plausible, realistic, unbiased estimate given the assumptions". Where this is not achievable, an estimate that does not favour the new treatment to an important degree (taken from section 2 - Scope) should be utilised.	

Line No of the first line(s) affected	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>*
195		**Comments:** an indication of the acceptable amount of missing data is highly recommended. It seems self-contradictory with line 164 (where it is stated that there is no rule for the maximum number acceptable). **Proposed change:** Delete 'and acceptable'.	
198-200		**Comments:** "..., and thirdly because the uncertainty in interpreting the results **introduced** increases …" **Proposed change:** "..., and thirdly because the uncertainty **introduced** in interpreting the results increases …", or delete "**introduced**"	
204-206		**Comments:** "This section must include a detailed description of the selected methods and a justification of why the methods to be applied are expected to be an appropriate way of summarising the efficacy results of the study and to result in an absence of bias in favour of experimental treatment." **Proposed change:** "This section must include a detailed description of the selected methods and a justification of why the methods to be applied are expected to be an appropriate way of summarising the efficacy results of the study and to **allow assessment without an important degree of** bias in favour of experimental treatment."	
208-210		**Comments:** It is not possible to ensure the selected method has any particular properties, yet alone conservative ones, without first stating the underlying assumptions (see general comments). **Proposed change:** See proposed change for comment on lines 105-106.	
213-214		**Comments:** Under what circumstances could the process of imputation or modelling be relevant to the baseline variables? Common practice is that if baseline values of relevant variables in the model are missing, the subjects will be	

Line No of the first line(s) affected	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>*
		excluded from the analyses. **Proposed change:** An example of a realistic setting where imputation of baseline values would be helpful.	
223		**Comments:** Please clarify how graphical summaries (e.g. Kaplan-Meier plots) can identify reason for dropout. **Proposed change:** Change "These graphical summaries should identify the reason for dropout." to "These graphical summaries **could** identify the recorded reason for dropout."	
226-227		**Comments:** It is unclear how data presentation can help to determine the contribution of each patient to the statistical analysis. **Change:** Clarify or delete sentence.	
238		**Comments:** "at least one analysis which gives …" **Proposed change:** "at least one **such** analysis gives …"	
270-273		**Comments: "**For example, when a patient drops out due to lack of efficacy reflected by a series of poor efficacy outcomes that have been observed, the appropriate value to assign to the subsequent efficacy endpoint for this patient can be calculated using the observed data." **Proposed change: "**For example, when a patient drops out due to lack of efficacy reflected by a series of poor efficacy outcomes that have been observed, **it would be appropriate to impute poor efficacy outcomes subsequently for this patient.**"	
288		**Comments:** The consequence of this sentence would be to always assume MNAR. Is this really the way forward?	

Line No of the first line(s) affected	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>*
289-290		**Comments:** Self contradictory for reasons given above in comment on lines 51-54. **Proposed change:** Please re-word.	
291-293		**Comments:** "Therefore the method chosen should not depend primarily on the properties of the method under the MAR or MCAR assumptions but on whether it is considered to provide an appropriately conservative estimate in the circumstances of the trial under consideration." **Proposed change:** "Therefore **justifications for** method**s** chosen should not depend primarily on the properties of the method**s** under the MAR or MCAR assumptions but on **the expected magnitude and direction of bias** in the circumstances of the trial under consideration."	
294-414		**Comments:** A more logical framework would be to have the main body of the guideline to focus on a discussion of "assumptions" relating to missing data, not methods. Comments on specific methods could be placed in an appendix (see publication by PSI Missing Data Working Group as an example, which can be made available on request). **Proposed change:** Place methods discussion in Appendix, and outline how they apply in the light of the assumptions being made. Appropriate references to technical statement may be beneficial.	
311		**Comments:** "response collection is interrupted after one point, …" **Proposed change:** "response collection is **prematurely terminated**, …"	

Line No of the first line(s) affected	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>*
315-316		**Comments:** Sentence is technically incorrect: LOCF does not always produce unbiased results under MCAR. Even under MCAR, LOCF can produce severely biased estimates as described in Molenberghs and Kenward (Wiley, 2007). **Proposed change:** Delete sentence.	
317-332		**Comments:** A longish list of methods is provided based on examples (instead of methodology), but no real recommendations are made. Also, the discussion of BOCF is somewhat puzzling. It seems that the method is recommended in some cases, even though here there is no reference to the method being 'appropriately conservative'. Why would BOCF be more appropriate than LOCF in the example? **Proposed change:** Use methodology rather than examples to introduce the methods.	
324-328		**Comments:** LOCF seems to be acceptable if conservative; however, the example is not entirely convincing. Even if there is differential drop out (more in the active group) and earlier drop-out, but the drop-outs have low scores (for example, the compound works quite well but in those where it works it is not tolerated that well leading to drop-out on relatively good conditions) the LOCF method may not necessarily be conservative. More importantly, it typically will not be so straightforward and not easily judged whether a method is conservative or not. In fact, it will probably would take a less biased method (e.g., MMRM or MI) to be able to judge what direction the bias of LOCF takes. **Proposed change:** Do not refer to LOCF as a conservative method. In fact, remove all suggestions that LOCF is necessarily conservative.	

Line No of the first line(s) affected	**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>*
328		**Comments:** "Establishing a treatment effect based on a primary analysis which is clearly conservative represents compelling evidence of efficacy from a **statistical** perspective." If a primary analysis is clearly conservative but very biased, it does not represent compelling evidence of efficacy from a **statistical** perspective. **Proposed change:** "…from a **regulatory** perspective"	
339		Comments: Hot deck imputation usually requires large samples. Proposed change: Please explain how you envision these methods can be used in the settings of this guideline.	
341-347		**Comments:** This part is rather speculative. It is unlikely that the reason for withdrawal will be predictive for the outcome to be imputed, without any reference to a criterion driven by actual data collected to decide on this. It sounds like a personal preference not founded in methodology. **Proposed change:** Take out this paragraph.	

Line No of the first line(s) affected	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>*
374-378		**Comments:** "The methods above are unbiased under the MAR assumption and can be thought of as aiming to estimate the treatment effect that would have been observed if all patients had continued on treatment for the full study duration. Therefore, for effective treatments these methods have the potential to overestimate the size of the treatment effect likely to be seen in practice and hence to introduce bias in favour of experimental treatment in some circumstances." **Proposed change:** "Under the MAR assumption, the methods above provide an unbiased estimate of the treatment effect that would have been observed if all patients had continued on treatment for the full study duration. **However, since it is likely that some data may be MNAR,** these methods have the potential to overestimate the size of the treatment effect likely to be seen in practice and hence to introduce bias in favour of experimental treatment in some circumstances."	
377-379		**Comments:** It is suggested that methods based on MAR overestimate treatment effects 'in practice'. But if the MAR assumption is valid, MAR is 'the practice' and these estimates are unbiased. It is also suggested that estimates from an MMRM are similar to those from a complete case (CC) analysis. This is untrue because (1) an MMRM includes all patients (and all available data points as per ITT principle) whereas the complete CC analysis includes only the subset of patients who have all data points and (2) MMRM is based on MAR whereas the CC analysis is based on MCAR. **Proposed change:** Reconsider the critical approach towards MAR-based methods.	
383-387		**Comments:** "The appropriateness of these methods will be judged by the same standards as for any other approach to missing data (i.e., absence of important bias in favour of the experimental treatment) but in the light of the concern above the use of only these methods to investigate the efficacy of a medicinal product in a regulatory submission will only be sufficient if missing data are negligible."	

Line No of the first line(s) affected	**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>*
		Proposed change: Add how to establish 'bias in favour of the experimental treatment'. Or delete these lines.	
391-392		**Comments:** The mentioned methods will not always be relevant. **Proposed change:** Change to "The potential impact of MNAR should be discussed"	
391-397		**Comments:** To advocate complex models such as pattern mixture models seems strange. The contrast with models such as MMRM is also remarkable. MMRM and MI models are well understood in the literature. The proposed models are however much more complex, involving assumptions above and beyond those of regular MMRM models. The interpretation and value of these models becomes therefore also even more complex. Currently the research on these models is not so extensive that the properties and value of such models is sufficiently established. **Proposed change:** Be less prescriptive. Or write 'One possibility is (..)'	
393		**Comments:** What is a "combined strategy incorporating several methods for handling missingness"? **Proposed change**: Clarify what is meant with the combined strategy.	
400-401		**Comments:** **Proposed change:** Replace "missing outcome" with "censoring".	
410		**Comments:** Responder analysis could be interpreted in many ways. **Proposed change:** Add cross-reference to later definition of a responder analysis (currently lines 455-458)	

Line No of the first line(s) affected	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>*
415-418		**Comments:**. 'Sensitivity analyses can be defined as a set of analyses where the missing data are handled in a different way in each analysis. This will show the influence of different methods of handling missing data on the study results.' **Proposed change:** Replace with "Sensitivity analysis is defined as a set of analyses in which the underlying assumptions are varied with the same analysis goal. This will show the influence of different assumptions on the study results and can therefore help to justify the set assumptions that underlie the primary analysis."	
429		**Comments:** "Conversely, whilst not all sensitivity analyses must necessarily give statistically significant results, …" **Proposed change:** "Conversely, **for a study with statistically significant primary analysis**, whilst not all sensitivity analyses must necessarily give statistically significant results, …"	
437		**Comments:** "Compare the results of the full set analysis to those of the complete case analysis." With missing data, there is no definitive full set analysis available. **Proposed change:** Delete this bullet	
437-462		**Comments:** Bullets relate to specific method and may be better positioned in an appendix for reason given above (294 – 414, pages 8-10). **Proposed change:** Place bullets in appendix.	
455-458		**Comments:** This sensitivity analysis addresses different clinical questions of interest. Subsequently, the approach is not a valid sensitivity analysis. **Proposed change:** Remove bullet point	