European Medicines Agency

London, 21 June 2007
Doc. Ref. EMEA/CHMP/EWP/260830/2007

OVERVIEW OF COMMENTS RECEIVED ON THE DRAFT GUIDELINE ON REPORTING THE RESULTS OF POPULATION PHARMACOKINETIC ANALYSES

Table 1: Organisations that commented on the draft Guideline as released for consultation

	Name of Organisation or individual	Country
1	EFPIA	
2	Bristol -Myers Squibb	
3	EUROPEAN GENERIC MEDICINES ASSOCIATION (EGA)	
4	GLAXOSMITHKLINE	
5	Daiichi Sankyo Pharma Development	
6	Next Level Solutions	
7	Metrum Research Group	
8	Novartis	

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

Table 2: Discussion of comments

GENERAL COMMENTS	**Outcome**
COMMENTS FROM EFPIA	
The Guideline captures the required and recommended formats of a population pharmacokinetic analysis report well. This is a timely action from the EMEA to ensure a high quality population analysis that facilitates model-based drug development. However, there are a number of general points that require clarification	
Specifically, there seem to be no clear distinction between what should be reported in the *Data / Results* sections on the one hand and in the *Method* section on the other hand (e.g. the last paragraph in *4.2.5 Methods*; subsection *Covariate model* clearly deals with results on the other hand model evaluation is only described in the results part). Our suggestion would be that under *Methods* a pure and complete description of the methods applied should be given. This includes description of the study(ies), the sampling and data collection, the data handling (including description of how incomplete data and outliers should be handled; currently proposed to be described under Data), the model development and evaluation and a description of the changes/modifications that took place in comparison to the planned analysis. Under *Data* and *Results* then the results of all the methods described under *Methods* should be given (e.g. in the Data section: "X patients had incomplete data which were replaced by X"; in the Results section: "The forward inclusion is documented in the run record runs X to X. Only X was found statistically significant." "The posterior predictive check as described in section X showed…"). With this suggestion *Methods* should be located before *Data* and *Results* in the report.	Accepted. Some restructuring of sections 4.2.4, 4.2.5 and 4.2.6 have been made. Also, nomenclature has been moved from the end and now is section 4.1. Hence, further sections have been renumbered.
Additionally, the guideline is sometimes very detailed e.g. see 4.2.4 below or the description which lines should be included in the GOFs (4.2.6). Our suggestion would be to add a general statement as "The general recommendations of the guideline might be appropriate for most analyses however in particular cases they can and should be adjusted."	Accepted with some rewording.
It would also be helpful to add a statement that "that it is not necessary to append documents that are already included in other part of the submitted documentation (as study protocols, analytical reports etc.)".	Accepted.
Finally, it is not clear whether the analyses datasets should also be submitted electronically. In this event the format, e.g. flat ASCII, should be specified.	In section 4.2.4, the following has been added: *Electronic files of the analysis datasets should be provided as comma separated Excel files (.csv) and ASCII text files.*

COMMENTS FROM <u>Bristol -Myers Squibb</u>	
This Guidance provides a welcome complement to the FDA Guidance on Population Pharmacokinetics, and should lead to greater uniformity in the information presented in population PK and exposure-response analysis reports. Although the FDA Guidance does include a brief section on the reporting of results, the EMEA guidance provides greater detail, and addresses issues (such as the impact of drop-outs) that have become more evident since the issuance of the FDA Guidance in 1999. Importantly, EMEA recognizes that every population PK model depends on decisions made by the developer, and clarifies that the guiding principle for reporting is: "It is vital that every assumption and decision made during model development is made clear for the assessor." This guiding principle should remove ambiguity regarding the level of detail that is expected in reporting these analyses.	
COMMENTS FROM: EUROPEAN GENERIC MEDICINES ASSOCIATION (EGA)	
The guideline concerns population PK analysis which can form part of the documentation for a New Chemical Entity (NCE) Marketing Authorisation Application. Therefore the SCOPE section should be consistent with the Introduction (Background) section and make it clear that the guideline is applicable to NCE applications.	See comment under Scope below.
COMMENTS FROM: Metrum Research Group	
This guideline provides thoughtful and valuable insight regarding expected content of population (POP) pharmacokinetic (PK) analysis reports for European regulatory submission. The emphasis on the assessment of clinical relevance of covariate effects is particularly helpful.	
The stated focus of the guideline is not to provide guidance on how to conduct a population PK analysis but rather to provide guidance on how to present results of POP PK analyses. In general, this is a useful scope for such a document, but by providing specific details for what should be presented, the guideline does set expectations for how the analysis should be conducted. The concern is that these specific details are not inclusive of alternative analysis methods that are currently in practice or may evolve in the future. A more useful approach might be to identify general points to consider, that are independent of method, when presenting scientific support for POP PK analysis conclusions.	In the Scope it is clearly stated that this guideline is written in a nomenclature that is applicable to NONMEM and that it is assumed that the reader can generalize the points to the software used in their particular analysis. If the approach suggested (i.e. to identify general points to consider) is used, the guideline would not fulfil its purpose to provide sufficient guidance on the level of detail of reports for population PK analyses. However, the scope has been revised to include information that this is an evolving science, that the guidance is written in accordance to current knowledge and that the reader is expected in the future to apply additional knowledge gained.

COMMENTS FROM: Novartis	
Generally, the guideline is very good and well-written, but additional details could make it even more useful. To that end, we have provided general and specific comments that may be considered for the final version.	
The guideline uses the words “guideline” and “guidance” interchangeably. However, the two words have different meanings. As such, one word should be used consistently throughout the document.	No change needed. The word “guideline” is used, when referring to the document. The word “guidance” in the sense “provides recommendations”.
In addition to the formal report (which is clearly defined), the definition of type and format of additional files required for submission is not clear. Thus, the guideline should concisely describe them.	Details regarding provision of electronic files have been added both for data files, model and output files.
Although this document is focusing on how to “report” the results of population PK analysis, it is worthwhile including a few sentences stressing that a well planned and conducted pop PK study is a prerequisite before writing the report.	This is true. However, it rather concerns how to conduct population PK analyses which is out of the scope of this guideline.
Relating to the above suggestion, is there any guidance document (from EMEA) on how to design and conduct pop PK trials? Probably not. There is one such guidance from FDA. One should recommend to read it along with the EMEA reporting document.	There is no specific EMEA guidance on how to design or conduct pop PK trials. A reference to the FDA guideline has been included.
The term “a secondary evaluation” appears several times. What does it mean? Less important, or detailed evaluation?	This refers to an assessment of the conducted analysis and the applicant’s conclusions from this analysis by regulatory authorities. This has been clarified in the guideline.
We suggest to use “population PK” throughout, i.e., not separating the two words.	No change has been made.
Convergence problems of non-linear mixed effects methods and their consequences are not addressed.	This is out of the scope of the guideline as it concerns how to conduct analyses.
Guideline states that final model should be re-run with outlier data points. However, such an endeavour could prove difficult as convergence may not be achieved. The guideline should recognize this difficulty.	Accepted. The guideline has been revised.
In agreement with the referenced paper by Wade et al, the guideline does not address population kinetics as a valid approach for addressing drug-drug interaction. It is deliberately omitted. We agree with this approach. It could be said.	This is out of the scope of the guideline as it more relates to how to conduct analyses.

SPECIFIC COMMENTS ON TEXT		
EXECUTIVE SUMMARY		
Line no. + paragraph no.	**Comment and Rationale**	**Outcome**
COMMENTS FROM: Novartis		
Executive Summary	What is a “secondary evaluation”?	This refers to an assessment of the conducted analysis and the applicant’s conclusions from this analysis by regulatory authorities. This has been clarified in the revised guideline.

SPECIFIC COMMENTS ON TEXT		
1 INTRODUCTION		
Line no. + paragraph no.	**Comment and Rationale**	**Outcome**
COMMENTS FROM: Novartis		
1. Introduction	Population pharmacokinetics is not defined. Provide a formal definition of population pharmacokinetics. The text to the right is the definition used in the FDA guideline, and is a quote from Aarons, L., "Population Pharmacokinetics: Theory and Practice," *Br J Clin Pharmacol*1991; 32:669-670. Add the following sentence as the second one in 1.: **“Population pharmacokinetics is the study of the sources and correlates of variability in drug concentrations among individuals who are the target patient population receiving clinically relevant doses of a drug of interest.”**	A definition has been added.

SPECIFIC COMMENTS ON TEXT		
1 LEGAL BASIS		
Line no. + paragraph no.	**Comment and Rationale**	**Outcome**
COMMENTS FROM: Novartis		
3. Legal Basis	Legal Basis, 2nd paragraph, It's not clear what is recommended for reading	It refers to the Directive 2001/83/EC amended in Directive 2003/63/EC which can be found at http://ec.europa.eu/enterprise/pharmaceuticals/eudralex/homev1.htm , Directive 2003/63/EC

SPECIFIC COMMENTS ON TEXT		
2 SCOPE		
Line no.[1] + paragraph no.	**Comment and Rationale**	**Outcome**
COMMENTS FROM: EUROPEAN GENERIC MEDICINES ASSOCIATION (EGA)		
2. Scope 1st line	The scope is applications for New Chemical Entities (as indicated in the Introduction), this should be made clear in the opening sentence of the Scope paragraph. Suggest adding the wording in bold. The aim of this guideline is to detail what the European regulatory assessors look for in a population report and the main components to be included in a report of a population PK analysis **when an applicant chooses to include this data in an application for a NCE.**	Not accepted. As clearly stated in the Introduction, the guideline does not only concern NCEs, but rather reports of all population PK analyses submitted with applications to EU regulatory authorities.

[1] Where applicable

COMMENTS FROM: Metrum Research Group		
Paragraph 1 Lines 2-5	See general comments. Proposed change: "The guideline does not provide guidance on how to conduct a population PK analysis, but rather provides guidance on points to consider when presenting the results from such an analysis, in order to provide a level of detail which will enable a secondary evaluation."	Accepted.

SPECIFIC COMMENTS ON TEXT

4. MAIN GUIDELINE TEXT

Line no. + para no.	Comment and Rationale	Outcome
COMMENTS FROM EFPIA		
Lines 1-2, 4.2.1 Lines 2 to 4, 4.2.2 P 4/9	This section is to describe the summary of a population PK analysis. The term "study" may mistakenly regarded as the population PK studies that were used in the analysis. Replace: "study" in paragraphs 4.2.1 and 4.2.2 with "***analysis***".	Accepted.
Line 1, 4.2.3 P 4/9	Related to the comment above, it is not clear whether the objectives of the clinical study used in the population analysis need to be included in this section Remove "study and" for the first line to read: "***The objectives of the analysis should be stated***"!	Accepted.
Lines 2-3, 4.2.4 P 4/9	Listing how many data points per subject were used may not add values to the assessors, especially when quite a few studies are used in the analysis. Replace: "The report should further state how many subjects and how many data points per subject have been analyzed. Information regarding	Accepted with some rewording.

	number of samples per visit should be given" with "*The report should further state how many total subjects and observed concentrations were used. Information regarding nominal number of PK samples per subjects per visit may be included in each study that was used in the analysis.*"	
Lines 3 to 5, 4.2.4 P 4/9	These instructions are very detailed and concrete, sometimes other plots or tables might be more appropriate. A separation by treatment group is not mentioned. Replace: "case there is a large range of number of samples per subject, a histogram of the distribution of the number of samples per subject and visit should be included. A histogram *of the distribution of sample times should be provided*" with "*The distribution of samples should be presented in an appropriate way that might include tables or graphics presenting the number of samples per subject / treatment group / visit / time interval etc.*"	Accepted with some rewording.
Lines 8-9, 4.2.4 P 4/9	It is not clear what is meant by "data input checking procedures" for the case of data transformation. Please clarify the meaning of "data input checking" in this context	The sentence did not refer to data transformation. After additional reconsideration, the sentence is considered redundant and has been removed.
Lines 5-6, 4.2.4 last paragraph P 5/9	Add "relevant", to keep listings concise and to be consistent with what is requested for subjects removed from the analysis. Add "relevant" for the sentence to read: "Outliers should be specified in a separate appendix to the report, with all ***relevant*** data available."	Accepted.
Line 4, 4.2.5 P 5/9	Information on bioanalytical methods are not relevant to PK population analysis on the other hand limit of quantification of individual chemical moiety would seem appropriate if a parent compound and metabolites are analysed. Change the sentence "This section should also include information regarding the bioanalytical methods used and the limit of quantification for each method." to read "***This section should also include information on the limit of quantification for each chemical moiety analysed***"	The type of bioanalytical method and Limit of quantification for all analytes should be stated. This has been clarified.
Lines 5-6, 4.2.5,	Since it is well accepted that parametric analysis with FOCE is the estimation method of choice if convergence can be reached in a timely	Not accepted. Population pharmacokinetics is an evolving science. We do not want to restrict the guideline unnecessarily to current NONMEM

P 5/9	manner, it is not clear what kind of justification would be required when FOCE with INTER is used. Insert "***If FOCE with INTER is used, no justification is necessary.***	practice.
Lines 11-12, 4.2.5 P 5/9	Reference 2 states that "Even when the factors governing the actual significance levels are known, it may be difficult to predict the impact of the approximation in a specific data set." Therefore, it is not clear how the influence of study designs on the actual significance level should be taken into account. Please clarify.	Even if it may be difficult to predict the impact of the approximation in a specific data set, a higher significance level is needed when using e.g. FO than FOCE with INTER. This should be taken into account in the analysis. After revision of this section the guideline states: *With respect to the model selection criteria, it is recommended to justify the statistical criteria used based on the fact that the actual statistical significance level obtained from the LRT (ΔOFV in NONMEM) could be markedly different from the nominal. For example, depending on the estimation method used (FO, FOCE with or without INTER) the number of subjects, number of samples per subject, residual error magnitude etc. may influence the actual significance level [2, 3].*
Lines 2-3, 4.2.5 Covariate model p 6/9	For categorical covariates it would also be appropriate to present correlations by correlation coefficients. Add a sentence mentioning that the graphical analysis of correlation between covariates could be used for excluding covariates to be tested (i.e. body weight and body mass index). Rephrase the sentence to read: "*The correlation between covariates should be presented* ***graphically or by appropriate correlation coefficients***. ***A graphical analysis can be used to exclude covariates to be tested: if two covariates are highly correlated only the most statistically significant should be tested."***	The proposal is considered too detailed. Also, it is out of the scope of this guidance to state how the population analysis should be conducted. The guideline has been revised to state that correlation between covariates should be presented without further details on how.
Lines 4, 4.2.5 Covariate model p 6/9	A school of thoughts prefers to do only the backward deletion after identifying full covariate models, in order to avoid accidental omission of an important covariate(s) during model development because of the order of adding covariates in a forward addition way. Add "and/or" in the sentence to read: "*The criteria for covariate selection (forward and/**or** backward) should be presented*"	This paragraph has been revised based on this comment and other comments.
Lines 4 to 6, 4.2.5	For the choice of covariates <u>to be tested</u> in covariate model building, clinical relevance is a good criterion. However, use of clinical relevance	This paragraph has been revised based on this comment and other comments.

Covariate model P 6/9	in selection of a covariate <u>in the process of</u> covariate model building is not appropriate because clinical relevance is a subjective matter. It is recommended that clinical relevance of a covariate that was selected based on statistical significance be tested after forward and backward covariate searches are over. Rephrase the second paragraph to read: "***The criteria for covariate selection (forward and/or backward) should be presented. It is recommended to use clinical relevance in the choice of covariates to be tested for covariate search. In the process of covariate model building, statistical significance and improvement in explaining between-subject variability and/or residual variability should be used in accepting the covariate as a significant covariate. After the covariate search with both forward inclusion and backward deletion is over, clinical relevance of each significant covariate should be discussed.*****"**	
Line 1, 4.2.6, Basic model P 6/9	Add a sentence on the log transformation and its justification at the very beginning of the paragraph. Proposed first sentence to read: "***The justification for a log transformation of the data should be presented.***"	This information is provided in the *Data* section.
Line 3, 4.2.6, Basic model P 6/9	It is difficult, if not impossible, to estimate intra-subject and residual variability from routinely collected PK samples. Delete "intra-subject" for the sentence to read: "***The forms of inter-subject and residual variability models should also be presented and supported by appropriate graphics.***"	Given the presence of observations from more than one occasion it is possible to estimate inter-occasion variability, which may be regarded as intra-subject in some cases. If modelled, then this should also be described. The sentence has been reworded.
Line 7, 4.2.6, Basic model P 6/9	The mandatory presentation of GOFs for all key models during basic model development will make the reports rather voluminous. Replace "should be" by "***might be***" Or alternatively, re-phrase the sentence to read: "***The complete model development should be described in the run record which should also contain a column for the evaluation of the runs/models. If it deemed necessary in addition GOFs of key models may be included.***"	GOF plots should be presented for the base model, and when relevant for key stages during model development. The sentence has been reworded.

Bullet points, 4.2.6,Basic model P 6/9	Sometimes trend lines could also be misleading. Replace "should be" by "***might be***" in all bullet points mentioning a trend line.	We are aware that trend lines in GOF plots sometimes could be misleading, but they are still useful.
Line 29, 4.2.6, Basic model P 6/9	The trend line used in e.g. DV vs PRED plots may also be a smooth line. Re-phrase to read: "*The trend line used in e.g. DV vs PRED plots should preferably be a linear regression line* ***or a smooth line****,…."*	Agreed. The sentence is superfluous and has been removed.
Lines 1 to 3, 4.2.6, Covariate model, 3rd paragraph P 7/9	This paragraph does not provide how sponsors come up with candidate covariate that is to be claimed to have no effect. Re-phrase the first sentence to read: "***Based on understanding of pharmacokinetics, pharmacology and physiology, if a variable is identified as a covariate candidate, a formal covariate model building with the candidate should be performed. The confidence interval for the estimated effect should be provided."***	This paragraph has been revised based on this comment and other comments.
Lines 12-13, 4.2.6, Final model P 7/9	In order to avoid unnecessary duplicates of the same plots, it would be better to refer to the same figures, if the final model is identical to the base model. Replace the sentence "These plots should also be provided in case the final model does not include covariates (and is identical to the basic model)." with "***In case the final model does not include covariates (and is identical to the basic model), the same GOF plots can be referred in the report and appropriate titles indicating that the plots are for both basic and final models should be stated.***"	Accepted.
COMMENTS FROM Bristol -Myers Squibb		
Section 4.2.1 (Summary) Page 4	*Comment:* It would be helpful to include additional detail on information that should be reported in the summary. *Rationale:* A summary usually consists of objectives, study design, data analyzed, methods, results and conclusions.	Accepted.

	Suggest replacing the following: "It should include sufficient information on the context of the study and the main findings and conclusions of the population PK study." *With the following*: "It should include sufficient information on the context of the study, **objectives, study design, data (number of subjects and samples), methods, results,** and the main findings and conclusions of the population PK study."	
Section 4.2.4 (Data) Paragraph 1 Page 4	*Comment:* Plots of data in a log-linear scale may not be applicable to data that do not vary over orders of magnitude. *Rationale:* Plots of raw data in log-linear scale may not be informative when the data vary by less than two orders of magnitude. Although log-linear plots are usually informative for PK data, the same is not always true for PD endpoints. For example, it would not be informative to plot blood pressure on a log-linear scale. Clarification is needed even though this guidance focuses on PK responses, as the Scope of the guidance states that the principles are equally applicable to PK/PD analyses. *Suggest replacing the following:* "Plots of the raw data are very useful, and should be provided on linear and log-linear scale." *With the following*: "Plots of raw data are very useful and should be provided on a linear-linear scale, **and plots should also be provided on a log-linear scale if the response variable changes by more than 2 orders of magnitude.**"	Partly accepted. There is no need to be so specific. The sentence has been revised to: *Plots of the raw data are very useful, and should be provided on linear scale and usually also on log-linear scale.*
Section 4.2.5 (Basic (or structural) model)	*Comment:* "Basic model" is not standard terminology in population PK analysis. *Rationale:*	Accepted

	"Base model" is the standard term used to denote the best model without covariate effects. *Suggest replacing* "**basic model**" *with* "**base model**" *throughout the document*	
Section 4.2.5 (Covariate model) Page 6	*Comment:* "The covariates to be tested should have been pre-specified in the analysis plan" is too restrictive. *Rationale:* This would preclude the incorporation of potentially important covariate effects that were not recognized prior to the present analysis. As population PK and exposure-response analysis are often exploratory, incorporation of covariates that were not pre-specified should be allowed, especially if inclusion of these covariates provide insight into the sources of variability in model parameters. *Suggest replacing the following:* "The covariates to be tested should be pre-specified in the analysis plan." *With the following:* "The covariates to be tested should **preferably** be pre-specified in the analysis plan. "	The covariates to be tested should be pre-specified in the analysis plan. If the covariates to be tested are changed during analysis this can be reported as a deviation. This paragraph has been revised based on this comment and other comments.
Section 4.2.5 (Covariate model) Page 6 Page 7	*Comment:* The guidance should not appear to endorse a particular method for covariate selection. Second paragraph states "... criteria for covariate selection (forward and backward) should be presented." Third paragraph states "... (both forward inclusion and backwards deletion)." These statements imply that only the specified forward inclusion/backward elimination methodology be used, and is not in alignment with the "scope" of this guideline which is to provide guidance on how to present the results from a population PK analysis.	Accepted. This paragraph has been revised based on this comment and other comments.

	Rationale: Forward inclusion/backward elimination is not the only reasonable method for screening covariates. Other methods include only backward elimination, or combined estimation of all pre-specified covariate effects. *Suggest adding the qualifier "***e.g.***" before text in the guidance that refers to the forward inclusion and backward elimination covariate selection method.* *Change text in second paragraph to:* "... criteria for covariate selection (**e.g.** forward and backward) should be presented." *Change text in third paragraph to:* "... (**e.g.** both forward inclusion and backwards deletion)."	
COMMENTS FROM GLAXOSMITHKLINE		
Section 4.1 (Analysis Plan), bullet 7	A capital letter should be removed as follows: "(e.g. ~~O~~objective function value…)"	Accepted.
Section 4.2.5 (Methods)	Abbreviations should be written in full the first time they are mentioned in the document.	The nomenclature section has been moved to the beginning of the guideline.
Section 4.2.6 (Results - Basic Model)	It would be useful to describe for each goodness of fit plot what they are intended for, i.e. what they should demonstrate and if some plots are more required than others. The term 'depending on situation' in the paragraph above the list is vague and giving examples of the most typical analyses would also be useful.	This is out of the scope of this guideline. As stated in the Scope, the guideline does not provide guidance on how to conduct a population PK analysis, but rather provides guidance on points to consider when presenting the results from such an analysis, in order to provide a level of detail which will enable a secondary evaluation. It is assumed that the reader is familiar with the use of different GOF plots.
Section 4.2.6 (Results - Final Model; last sentence on page 7):	In the following sentence it is proposed that the word 'validation' is replaced with 'evaluation': "...a posterior predictive check or external ~~validation~~ evaluation with..." This change is based on the first sentence of the paragraph.	Accepted.

COMMENTS FROM Next Level Solutions		
Section 4.2.5	Overall quite good. My primary concern is the apparent focus on step wise regression for model selection. As in the text in the "covariate selection" section at the top of page 6 (reproduced below) "The criteria for covariate selection (forward and backward) should be presented. It is recommended to use both statistical significance and clinical relevance (only effects larger than a certain pre-defined magnitude) in the process of covariate model building. The covariate model building steps (both forward inclusion and backwards deletion) to illustrate covariates that are included in the final model and those that were tested but were not retained in the final model should be clearly presented in the run record. The criteria on which the decision was based, e.g. objective function values, should be outlined as well." Step wise regression has significant weaknesses, both in sensitivity and specificity. Further, step wise regression is much less relevant when using Bayesian methods such as MCMC. (References can be provided). The possibility of other methods of model selection (least angle regression [http://www-stat.stanford.edu/~hastie/Papers/LARS/LeastAngle_2002.pdf] and machine learning methods [Dr Robert Bies has demonstrated the advantages of machine learning methods for pop pk model selection, http://www.page-meeting.org/default.asp?abstract=405]). It occurs to me that a guideline should at least acknowledge the possibility of other model selection algorithms.	This paragraph has been revised based on this comment and other comments.
COMMENTS FROM Daiichi Sankyo Pharma Development		
Paragraph 2 4.2.5 Methods	A run record is a good idea to explain the steps of model building and used in literature. However, the format and information is different according to the author. Please clarify the essential items for reports.	The information has been slightly revised. *The run record should describe the changes from the previous model and the decisions taken and could include a brief, but interpretable, description of the run, the objective function value and information whether the model converged successfully. Preferably, the run record should also include parameter estimates (for key runs) and, when needed, a comment about the run.*

Paragraph 2 4.2.6 Results	It is true that GOF plots are useful to demonstrate that the model is well described. However, some GOF plots are not necessary to show the model validity. For example, Graphs of observed data versus (individual) predicted data are essential to describe the goodness of fit. However, we do not think graphs regarding eta are always necessary. Please clarify the mandatory plots.	The need for graphs regarding eta depends on the situation. In case covariates have been included in the model, these graphs may be useful to evaluate if there are any remaining trends in the data. However, as stated in the first paragraph (and in reference 7) the value of different GOF plots depends on the situation; type of data; rich or sparse, type of estimation method etc.
4.2.5 Methods Paragraph 1 Line 7	Assessment of statistical significance in POP PK modeling is problematic, complex and practically unnecessary for making inferences about the clinical relevance of modeling results. Current analysis methods, such as the Genetic Algorithm and Full Model methods do require objective criteria for model selection, but do not require the determination of statistical significance. Proposed change: "Model selection criteria should have been defined prospectively in the analysis plan."	This paragraph has been revised based on this comment and other comments. Also, as stated in the scope this guideline is written in a nomenclature that is applicable to NONMEM and that it is assumed that the reader can generalize the points to the software used in their particular analysis. In the scope it has also been clarified that the general recommendations of the guideline might be appropriate for most analyses however in particular cases they can be adjusted.
4.2.5 Methods Paragraph 6 Line 1	This section implies that stepwise hypothesis testing is required. Also see comment above. Proposed change: "The rationale and criteria for covariate selection should be presented. It is recommended to use criteria based on both goodness of fit and clinical relevance (focused on assessment of magnitude of covariate effect) in the process of covariate model building."	Partly agreed. This paragraph has been revised based on this comment and other comments.
4.2.5 Methods Paragraph 7 Line 1	Same as above. Also consider moving this paragraph to the RESULTS section. Proposed change: "The results of covariate modeling steps should be presented in sufficient detail to support the final covariate model and reproduce the results (e.g. plots and run record)."	The paragraph has been moved to the results section as suggested. The paragraph has been revised based on other comments.
COMMENTS FROM: Novartis		
4.1 Analysis Plan	Number of patients with PK samples? Add the following sentences: **"If the pop PK study was planned/designed, the number of patients**	This paragraph has been revised based on this comment and other comments.

	should be presented. If applicable, the number of patients needed to discern an effect (e.g. of a covariate on a PK parameter) should be briefly justified (using an appropriate statistical method or based on extensive experience and ethical concerns, etc.)"	
4.1 Analysis Plan	Levels of statistical significance should be mentioned in the plan Add one bullet point to the list: **". Levels of statistical significance"** Remove the respective sentence from 4.2.5.: "The level of statistical significance should have been defined prospectively in the analysis plan."	This proposed text has been added in "criteria to be used for selection of models during model building and covariate selection (e.g. objective function value, ***level of statistical significance,*** goodness of fit plots, …..) Section 4.2.5 has been rewritten.
4.2.4 Data	The "Data" section might better be called "Experimental Methods" Rename 4.2.4. to "Experimental Methods" Histograms, number of subjects, etc. , belong in the first part of the results. This section should focus on the study designs (treatments, inclusion criteria, visit schedules, protocol sampling times) and assay methods for the related clinical study protocols. To reflect this change, a corresponding descriptive section title is suggested. Rules for outliers belong in section 4.2.5.	Not agreed. Agreed. The Data, Methods and Results sections have been restructured based on this and other comments.
4.2.4 Data	Information on specificity, sensitivity and accuracy in measurements of the drug(s) and their metabolites are not mentioned. Should list those that might affect data analysis, such as, minimum level of detection/quantification Add another bullet point requesting such information (see below)	No change needed. It is already stated in the methods section that information regarding bioanalysis method and LoQ should be given.
4.2.4 Data	Last line of 1st paragraph: reasons for dropouts should be presented, especially drug-related Add reasons for dropouts (see below)	The Data, Methods and Results sections have been revised based on this and other comments.
4.2.4 Data	The section 4.2.4. is rather specific on how to summarize the data; one could consider tightening it, e.g. as a tick list. Replace the current text by a list, e.g. as follows: **The report should briefly summarize the features of the data which are relevant to the pop PK analysis, in appropriate tabular and graphical form, including:**	The Data, Methods and Results sections have been revised based on this and other comments.

	• **Studies included for analysis, and their key design features** • **Data used for validation of the model** • **Assay properties, specifically limits of quantification** • **Number of visits with PK sampling, numbers of samples per dosing interval** • **Sampling times** • **Raw data (plots)** • **Dropouts: summary, as appropriate, e.g. number, reasons, timing relative to PK sampling** **Missing data: summary, as appropriate, specifically missing dosing or sampling times, covariates**	
4.2.5 Methods	Before deciding a more complex pop PK model, some preliminary exploratory analyses of concentration data may be encouraged as starting point. Suggest to employ the simplest models as starting point in the analysis plan	It is out of the scope of the guideline to state how analyses should be conducted.
4.2.5 Methods	The term "Methods" is too general and can encompass study design and data collection processes. To make the sub-title more specific, the section should be renamed "Data Analysis Methods" Rename 4.2.5. to **"Data Analysis Methods"**	The Data, Methods and Results sections have been restructured based on this and other comments.
4.2.5 Methods	Adapt the first paragraph within 4.2.5, according to the changes proposed above The beginning of the first paragraph should be modified as follows: **"The methods section should describe the methods used and should include the same components as the analysis plan (even if there is some repetition). If, during the analysis, any deviations from the analysis plan occur, then these should be clearly described in the methods section of the report. This section should also include information regarding the handling of missing data, outliers, and values outside the limits of quantification. The choice of parametric or non-parametric analysis …"**	The Data, Methods and Results sections have been revised based on this and other comments.

4.2.5 Methods	The guideline appears to be biased towards NONMEM Guideline should be more general to accommodate other parametric approaches, such as iterative two stage approaches. In this section, an additional point should be made about nonparametric and/or Bayesian approaches (e.g. Winbugs?) Rephrase the respective part, e.g. as follows: **"When using NONMEM, the actual significance level obtained from the LRT (□OFV in NONMEM) could be markedly different from the nominal. Depending on which estimation method used (FO, FOCE with or without INTERACTION) the number of subjects, number of samples per subject, residual error magnitude etc. may influence the actual significance level, which should be taken into account [2, 3].** **If other approaches / software are used (e.g. Bayesian / BUGS), analogous considerations should be taken into account.**	In the Scope it is clearly stated that this guideline is written in a nomenclature that is applicable to NONMEM and that it is assumed that the reader can generalize the points to the software used in their particular analysis. The section has been revised based on this and other comments and now reflects use of other approaches/softwares.
4.2.5 Methods	Lack of distinction between pharmacological and statistical modeling. The "Basic Model" sub-section should include the choice and justification for population parameter distribution and error models used. Insert the sentence: **"Choices for population parameter distribution and error models used should be described and justified"**	Agreed. The methods section has been updated based on this and other comments.
4.2.5 Methods	Covariate selection should be more than statistical exercise. Clinical relevance may also play a role. Proposition to add the following: **"Choice of covariates to be tested will be made using biological/pharmacological/clinical plausibility and/or a graphical exploration of potential covariates".** The recognition of this covariate selection approach should be made explicit in the text.	The covariate section and the analysis plan section have been revised based on this comment and other comments.
4.2.5 Methods Line 12	Version, operating system, compiler used, and level of bug fixes should be stated: In addition to stating the "software and version used", we think that in addition the operating system, compiler, and the level of bug fixes for the software should be clearly stated (in particular for	Not agreed. We do not find that this level of detailed is necessary.

	NONMEM) Replace the sentence on software version by: "**The exact software specifications should be stated, including the version and level of bug fixes. Details of hardware, operating system and compiler should be provided, either as a description or through a reference to the related QA documents"**	
4.2.7 Discussion	In the discussion section, in addition, some wording may be needed to avoid/minimize biased data interpretation Proposition to add the following paragraph: **"Model selection and interpretation of the results of a Pop PK analysis require a fundamental appreciation and integration of multidisciplinary principles (such as e.g., physiology, pharmacology, biochemistry, statistics). The relevant aspects of these areas must be considered in reaching any conclusions regarding the particular data analysed, to avoid biased interpretation of the data."**	This is considered sufficiently covered by other sections of the guideline, e.g. where justification is requested for methods used for model building and covariate selection.

DEFINITIONS		
Line no. + para no.	**Comment and Rationale**	**Outcome**
COMMENTS FROM GLAXOSMITHKLINE		
Definitions section	There is a typo under FOCE; 'method .The' should be replaced by 'method. The'.	Accepted.
Definitions section	We propose to add "epsilon" to the definitions section.	As the word epsilon is not used in the guideline, addition to the definition list is not needed.
COMMENTS FROM: Novartis		
Definitions	Definitions/abbreviations of some commonly used PK terms are provided towards the end of the guideline, however, some terms are considered as standard and are not formally defined. Make the abbreviations list more comprehensive: Definitions should be provided for terms such as BUGS, PK, AUC, NONMEM, PD, LOQ, QA(=Quality assurance) etc.	The list of "definitions" is actually a list explaining the specific population PK nomenclature used in the guideline. NONMEM has been added to the list of nomenclature. There is no need for adding the other abbreviations. When relevant these have been clarified in the text.

REFERENCES		
Line no. + para no.	**Comment and Rationale**	**Outcome**
COMMENTS FROM: Novartis		
Reference	Other regulatory references An important reference that is currently omitted is FDA's Guidance on population PK. FDA's guidance provides in depth details on study design, execution, data handling, and analysis of a population PK study. As such, referencing such a document may prove useful for many modelers performing Pop PK.	A reference to the FDA guideline has been included.
Reference	Technical references The guideline provides some technical references. It would be useful to add some others, for instance for evaluation/qualification/validation (very useful to guide novice modelers).	Additional references have been added.
Reference	Ordering The order of the references does not conform to the sequence in the text (1,2,3, 7, 4, 5, 6, 8,9). This may be un-important	Agreed. This has been revised.
Céline Dartois, Université de Médecine de Lyon sud		
	Firstly, I think that concerning the data, people should report method they used to transform raw data in NONMEM format, automatically in splus, sas by a script (which is recorded and can be modified and evaluated) or in excel by clicking, deleting, modifying….without any proof…	It is already stated that data transformation should be described and justified. Other information related to the construction of the data ser, as exemplified in the comment, is considered out of the scope of the guideline.
	Secondly, I think that more precise details should be required on clinical data than those described in the analysis plan. I am thinking about treatment characteristics (the different arms, therapeutic windows if applicable, dosage level and range, administration route, other treatments (which can be used as covariates or can be modelled), ect…) I think it is very important as we can model a subcutaneous administration by an oral one or a bolus IV of 5 minutes by an infusion but we have to justify this sort of decision.	This is sufficiently reflected in the revised data section. It is already stated that the choice of structural model should be justified.

	Thirdly, I think that EMEA should recommend only to use last versions of the softwares. Numerous bugs of NONMEM have been identified and I think it is illogical to accept results form old versions (with bugs….) but it is perhaps not the subject of your guidance.	The guideline states that software and version used should be stated. This is sufficient.
	Finally, in the evaluation paragraph, you mentioned bootstrap and jacknifes techniques. I do not understand how they can be used to evaluate a model as the same level than VPC or PPC. I think it would be necessary to be more precise and make the difference between, methods which can be used to evaluate objective of the models (VPC, PPC) by comparing **observed data** and predictions (for descriptive models) or by comparing **observed data** and simulations (for predictive models), and methods which can be used to evaluate properties of the model (bootstrap and jacknifes..) like robustness and sensibility. Moreover, I think that metrics people use for this evaluation should be stated. I think it is important as number of them are biased like the error of prediction.	The method evaluation section has been revised taking these comments into account.