16 January 2012
EMA/CVMP/EWP/249785/2011
Committee for Medicinal Products for Veterinary use (CVMP)

# Overview of comments received on 'Guideline on statistical principles for clinical trials for veterinary medicinal products (pharmaceuticals)' (EMA/CVMP/EWP/81976/2010)

Interested parties (organisations or individuals) that commented on the draft document as released for consultation.

| Stakeholder no. | Name of organisation or individual |
|---|---|
| 1 | Paul-Ehrlich-Institut (Germany) |
| 2 | Association of Veterinary Consultants (AVC) |
| 3 | IFAH-Europe |
| 4 | ANSES- ANMV – FRANCE |
| 5 | Pfizer Animal Health, Veterinary Medicine Research and Development |

7 Westferry Circus ● Canary Wharf ● London E14 4HB ● United Kingdom
**Telephone** +44 (0)20 7418 8400 **Facsimile** +44 (0)20 7418 8416
**E□mail**info@ema.europa.eu **Website** www.ema.europa.eu

## 1. General comments – overview

| Stakeholder no. *(See cover page)* | General comment (if any) | Outcome (if applicable) |
|---|---|---|
| 1 | While it is stated in section 2 (Scope) that this GL focus on statistical principles, the GL provides a lot of details and even examples. Probably the GL would gain more flexibility if it would focus on general principles while more detailed information (and examples) would be provided with specific appendices (e.g. on multiplicity, non-inferiority etc) that might be easier edited / updated than a GL. | Having examples included into the text facilitates readability. |
| 1 | In the Scope of the GL it should be made clear that the principles outlined in this GL might be overruled in case a study has to follow other guidance such as the European Pharmacopoeia. | Disagreed.<br>This is already stated in the Introduction. |
| 2 | The association of veterinary consultants (AVC) appreciates the opportunity to review and comment the "Guideline on statistical principles for veterinary trials".<br>The new guideline provides clarification and further guidance on how the statistical components of a clinical trial should be addressed. This will have a positive impact on the quality of the data generated and analysed during the trials. However, as the needs and requirements (including the statistical elements) for conducting clinical trials increase in complexity and resource requirements, AVC would welcome within the guideline some guidance of statistical requirements for the registration veterinary medicines for MUMs. Furthermore, we believe that, to promote the development of veterinary medicines, the requirements described in the current guideline should be more realistic. A flexible approach is needed to address the specific challenges of the development of veterinary therapeutics where differences between animal species, treatment approaches and limited market size are not comparable to the ICH statistical requirements on which the current guideline is inspired. | To have a more "realistic" or "flexible" guideline would mean that one possibly waives statistical principles. Also for veterinary medicinal products one should be able to e.g. distinguish real effects from just random ones.<br><br>Reduced dossier requirements for MUMs products are usually decided on a case-by-case basis, hence, it would be very difficult to provide some general guidance. However, a clinical field study to demonstrate efficacy and safety of a given MUMs product should be conducted in a way that the results are statistically valid and clinically relevant, which means, that no reduced requirements for field studies would be applicable. No change considered necessary. |
| 2 | The terms "protocol amendment, deviation and violation" are used | Agreed. |

| Stakeholder no. *(See cover page)* | General comment (if any) | Outcome (if applicable) |
|---|---|---|
| | throughout the guideline. Even though these terms are described in VICH-GCP (except for "violation") it would be useful to include them in the Definitions section. | Definitions have been included in the Glossary. |
| 3 | IFAH-Europe acknowledges that the scope of the GL is now limited to pharma (non-immunologicals) products as this was not specified in the previous version. This resolves some of the issues with regards to the description provided for studies that were not always relevant for immunologicals. Nevertheless, it is suggested that this scope should be more clearly stated and clarified to avoid any misinterpretation.<br>It is noted that most of the IFAH-Europe comments on the concept paper for the revision of this guideline have not been taken into account. The revised guidelines are certainly scientifically accurate, but they increase the complexity of the clinical studies and may hinder innovation especially when dealing with novel therapeutic areas.<br>Although IFAH-Europe commended in its comments on the Concept paper that the revised guideline should not be "basically similar to its counterpart in human medicine", the document currently under consultation is mostly a copy of the human guideline with few minor changes. The only section specific to the vet medicine refers to the experimental unit, though no guidance is provided for group treatment such as pen housed animals or fish medicine.<br>Since the guideline used ICH documents as references (see line 94), it is important to consider that for human products, dose determination and dose confirmation trials are closer to 'field studies' from the vet sector than e.g. an in-house dose confirmation study (DCS) in a known experimental model (e.g. bovine respiratory disease DCS for an antibiotics or canine worm infestation DCS for a dewormer).<br>**Flexibility and specificities of the vet sector should be maintained to avoid unrealistic requirements** that may create unnecessary use of animals, costly studies, increased development | The CVMP's Immunologicals Working Party (IWP) discussed this and concluded that inclusion of immunological products into the current guideline would probably take considerable time, and would significantly delay the publication of the version for pharmaceutical products only. IWP therefore agreed NOT to do anything to the EWP guideline for the time being, so that EWP and CVMP could go ahead with adoption (December EWP, Jan CVMP).<br><br>IWP would independently review the final version of the guideline, and then decide whether the guideline should be further amended/revised to add immunological specific information (either directly into the guideline or as an annex), or if a new IWP statistics guideline would need to be prepared.<br><br>Less scientifically accuracy implies less reliable study results. |

| Stakeholder no. *(See cover page)* | General comment (if any) | Outcome (if applicable) |
|---|---|---|
| | time, etc.<br>The complexity of the clinical studies described in the guideline has increased, especially when dealing with novel therapeutic areas. In such cases, validated primary end points are not available from the literature and must be selected following extensive pilot studies. The validation of the new end points, which requires to measure intra-assessor and inter-assessor validity of rating scales, may be also very complex and long, if ever feasible. Also the complexity of the blinding procedure has increased: keeping a study blinded at all levels until the data base is locked is difficult in small scale trials with limited personnel. The study personnel would have to be substantially increased to fully protect the blinding. Flexibility is essential.<br>Further clarification is required on the following aspects:<br>- the guidance reflects the current state-of-the-art for statistical analysis in clinical context, but should not be regarded as a mandatory lists of items to be met. In particular, it is understood that the wording "clinical trials" may refer to dose-determination, dose-confirmation as well as field trials. Please specify throughout the guideline which requirements apply to field studies and which apply to laboratory studies.<br>Sponsors have the flexibility to report information in a separate statistical report, even though the guideline frequently discusses what is to be "reported". A clear description of which portions from the statistical report must be included in the main study report would be helpful. | |
| 4 | The ANSES- ANMV thanks the EWP for this very informative and clear paper. The guideline is generally supported, and we have only one specific comment. | |

## 2. Specific comments on text

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| 0000 Title | 3 | A simple improvement for unequivocally clarifying the scope of this guideline would be to extend its title as follows: "*Guideline on statistical principles for ~~veterinary~~ clinical trials <u>with veterinary pharmaceutical products</u> (EMA/CVMP/EWP/81976/2010)*" | As the scope of the guideline does not include immunologicals, the title remains as it is. |
| 0080 | 2 | Comment: The guideline refers only to investigational veterinary pharmaceutical products<br>Proposed change (if any): Please clarify whether or not the guideline applies only to pharmaceuticals or pharmaceutical and biological (vaccines) trials | The title clearly indicates that the current guideline is for veterinary pharmaceutical products only. |
| 0098 | 3 | Comment: please make clear that the GL only applies to pharma products. It is suggested that a new guideline for immunological studies is developed. Without such document there is a risk that the present guideline will be the reference document for assessors.<br>Proposed change: *"This guideline is intended to give direction to sponsors in the design, conduct, analysis, and evaluation of clinical trials of an investigational veterinary <u>pharmaceutical</u> product in the context of its overall preclinical and clinical development. Clinical studies of immunological products do not fall within the scope of this guideline."* | See above. |
| 0115 | 2 | Comment: The guideline refers to the need of an appropriately qualified and experienced statistician to perform the analysis (see also the definition of the trial statistician on line 1294)<br>Proposed change (if any): We believe that a specific statistical degree should not be required to perform the statistical analysis. Although there are university specific degrees on statistics, it is | Not accepted.<br>The definition of statistician in the glossary is considered sufficiently wide. |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | also part of any study in mathematics as well as for many other scientific diploma. A qualified scientist with experience in statistical analysis should be sufficient as long as the expertise is justified by experience in this area. | |
| 0126 | 2 | Comment: Unless the amendments affect the statistical sections of the protocol, we consider that the statistician's approval is not necessary.<br>Proposed change (if any): "the protocol and subsequent amendments affecting the statistical sections should be approved…." | Not accepted.<br>Any amendment of the protocol may have an impact on statistics and possibly require a change in statistical methods, therefore, it should be reviewed and signed also by the statistician. |
| 0126-128 | 3 | Comment: The text says the statistician and responsible personnel should 'approve' the protocol. Please clarify whether it should be 'approved' or 'signed' and add a reference to the statistical report for the cases that this is issued separately from the study report. | It should be signed (this includes approval) |
| 0134-148 | 3 | Comment: The definition of "clinical trial" is confusing. One might interpret from this section that the GL applies also to exploratory trials, whereas for the latter a more flexible approach may be needed so that changes can be made once results are accumulating.<br>Proposed change: please clarify that the statistics principles do not apply to the three categories of studies, in order to avoid that defining good quality of statistical principles creates too stringent (unrealistic) requirements in early phases of development.<br>Please add after line 157 (as per line 88): *"Exploratory trials are usually non-GCP and therefore do not fall within the scope of this guideline."* | Accepted.<br>The following text is added: "Exploratory trials are often non-GCP and, therefore, are not in the main focus ~~do not fall within the scope~~ of this guideline. Nevertheless, their conduct should have quality, be ethical and pre-planned." at the end of the section. |
| 0135 – 138 | 1 | Comment: The definition of a study according to the | Accepted. |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | understanding of the GL is linked to hypothesis testing. However, in some situations studies (e.g. exploratory studies or safety studies) might focus on estimation, not on testing hypotheses.<br>Proposed change: A study is a single scientific experiment conducted in a target species. ~~to test at least one hypothesis relevant to the proposed effectiveness claim(s) or to in-use safety in the target animal for a veterinary medicinal product under investigation.~~ | Text is changed accordingly, and definition moved to Glossary. |
| 0147 – 148 | 1 | Comment: The term 'composite trial' is quite unusual. Usually every confirmatory trial has an explorative component, but is still a confirmatory trial.<br>Proposed change: Depending on the aim of the trial, it can be classed in one of the following ~~three~~ two categories: confirmatory or exploratory~~, or composite~~ trial. | Accepted.<br>Section on composite trials is deleted and text changed accordingly. |
| 0148 | 2 | Comment: confirmatory, exploratory and composite trial<br>Proposed change (if any): Suggest to change the order to exploratory, confirmatory and composite trial to be consistent with subsequent sections | See above. |
| 0149 – 182 | 1 | Comment: Several of the bullet points in lines 164 to 176 to characterise a confirmatory trial also apply to explorative trials (e.g. an agreed protocol).<br>Proposed change: Re-arrange order of bullet points like<br>• standard requirements (applying for all kind of trials),<br>• specific aspects of exploratory trials,<br>• specific aspects / requirements for confirmatory trials. | Not accepted.<br>All of the bullet points are considered valid for confirmatory trials. Exploratory trials are not within the main focus of the guideline (see above), therefore, no amendments relevant for exploratory trials are deemed necessary. |
| 0160 – 161 | 1 | Comment: The reference to CVMP and EP is not related to the (general) definition of a confirmatory trial.<br>Proposed change: | Accepted.<br>Information to consider other guidance documents is provided in the Introduction. Sentence deleted. |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | ~~For some specific product studies the design may be subject to other guidance such as that provided by the CVMP and the European Pharmacopoeia.~~ | |
| 0160-161 | 3 | Comment: *"For some specific product studies the design..."*<br>It would be helpful to clearly state for which specific products the CVMP and/or the European Pharmacopoeia provide particular guidance on study design.<br>Proposed change: *"For some specific products (i.e. immunologicals, anthelmintics and ectoparasiticides) the studies design ~~may~~ will be subject to other guidance such as that provided by the CVMP and the European Pharmacopoeia."* | Not accepted.<br>See above, the sentence has been deleted. |
| 0175 | 2 | Comment: Confirmatory trials use validated and clinically relevant parameters.<br>Proposed change (if any): The word validated implies a validation process which is not always achievable for some parameters (mainly clinical ones). We suggest the following text: "validated, accepted, well established or recognised" | Partly accepted.<br>The wording has been modified: "…use validated, or well established, recognised parameters that are clinically relevant" |
| 0175 & 208 | 3 | Comment: *"use validated and clinically relevant parameters."* (line 175)<br>*"The variable should be reliable and validated and derived..."* (line (208)<br>It is not clear what is meant exactly with 'validated'. The validation of a laboratory test is not the same as the validation of a clinical variable. A variable or score can be used routinely in clinical studies and be described in peer reviewed publications without being formally validated.<br>Proposed change: *"use valid~~ated~~ and clinically relevant parameters"*;<br>*"The variable should be reliable and valid~~ated~~ and derived..."* | See comment above.<br>There may be validated methods available such as VAS scoring system, other clinical variables may be used that are well-established (valid) based on previous studies or relevant published data. |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| 0176 | 1 | Comment: Confirmatory studies do not necessarily produce robust results but should be planned to allow for robust conclusions. | Accepted.<br>Text changed accordingly. |
| 0178 | 5 | Comment: We think the issue of replication of studies applies more to field safety and efficacy studies than laboratory efficacy studies. | Partly accepted<br>Text changed to: "*Usually, the weight of evidence from a single confirmatory trial is sufficient, and there is no need for replication of the results. But if there are weaknesses with respect to internal or external validity, clinical relevance, statistical significance, data quality, or internal consistency, a second confirmatory trial should be performed.*" |
| 0178-182 | 3 | Comment: The statement *'in general, the results need to be replicated'* is not clear. This is far from being a general case, and applies in fact only to some specific therapeutic areas (e.g. parasiticides clearly mentions 2 DC per claim) but exact similar studies are seldom repeated during a product development, both for animal welfare and budget reasons. One confirmatory study is sufficient if provides enough evidence of efficacy and safety.<br>Proposed change: Please remove the paragraph. | See above. |
| 0183 | 1 | Comment:<br>From the definition provided, a composite trial is a confirmatory trial. Thus, a specific mentioning of this category of trial is not necessary.<br>Proposed change:<br>Delete line 183 | Accepted.<br>Section deleted |
| 0183 – 188 | 1 | Comment:<br>In section 4.1 the GL mentions 'composite trials' as a separate kind of clinical trial. This category is quite unusual in the | Accepted.<br>Section deleted.<br>Added "In confirmatory trials, the option may exist to use the |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | statistical literature and should not be used in the context of this GL in order to avoid confusion. Instead it should be mentioned under 'Confirmatory trial' that confirmatory trials might also serve exploratory purposes. | data for further exploratory analyses, which may serve to explain and support the trial findings and to suggest further hypotheses for research. The protocol should make a clear distinction between those aspects of the trial which are confirmatory, and those which are exploratory." |
| 0184 | 1 | Comment:<br>The statement in line 184 is also included in lines 185 - 188<br>Proposed change:<br>Delete line 184 | Accepted.<br>See also above |
| 0208 – 212 | 1 | Comment:<br>Lines 210-212 repeat the contents in lines 208 – 209<br>Proposed change: Merge text | Accepted.<br>Text changed accordingly. |
| 0218 – 219 | 1 | Comment:<br>The wording (… after unblinding) implies that a redefinition of the primary variable might only be a problem in blinded trials. But redefinition of the primary variable in knowledge of the data is a problem in any kind of trial.<br>Propose change:<br>Redefinition of the primary variable in knowledge of the data ~~after unblinding~~ will almost always be unacceptable, since the biases this introduces are difficult to assess. | Accepted.<br>Text changed accordingly. |
| 0220 – 223 | 5 | Comment: Interpretation of secondary variables is always compromised by the absence of a priori estimations of type 1 and type 2 errors, especially P-values. | Point noted and agreed, however, any further guidance on this would go beyond the scope. |
| 0220-223 | 2 | Comment: In relation to secondary variables, the guideline states that "an explanation of their relative importance and roles in interpretation of trial results" is needed.<br>Proposed change (if any): it would help to have some concrete recommendations and guidance as to what is expected. | See above. |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| 0227 | 2 | Comment: The description on how to deal with missing values should be described not only for composite variables but also for other variables described later on in the text.<br>Proposed change (if any): We suggest to move this reference to the introduction of primary and secondary variables (Lines 204-223) | Not accepted.<br>Missing values are a special problem at composite variables (what to do if one component of the composite variable is missing?). In general, missing values are dealt with elsewhere (Section 7.3.) |
| 0233 | 5 | Comment: Rating scales should be avoided if possible.<br>The practice of consolidating clinical signs into a unifying score to simplify analysis and interpretation should also be avoided—if for no other reason, because of the loss of information. | Partly accepted.<br>It is agreed that the use of rating scales as primary parameters should be avoided whenever possible because of the disadvantages addressed in detail in this paragraph. However there may be situations where no other possibilities exist, thus, the wording should be "open". |
| 0234 – 238 | 1 | Comment:<br>There should be a separation between the GL text and the definition section<br>Proposed change:<br>The definitions provided in the text, e.g. *'Content validity (the extent to which the variable measures what it is supposed to measure)'* should be put into the Definition section | Accepted.<br>Definitions moved to the glossary |
| 0234-238 | 3 | Comment: The inter-assessor and intra-assessor validity of some clinical parameters may be very complex if not impossible in a number of cases (e.g. behavioural signs, pruritus, etc.).<br>Proposed change: "*... should be addressed <u>when possible, and omission justified</u>.*" | Not accepted.<br>Intra- and inter assessor validity is considered a prerequisite for the reliability of the results, therefore, no change recommended. |
| 0243 – 246 | 5 | Comment: We agree with these cautions and would prefer even stronger language regarding performing "arithmetics" and treating "ordered categorical data as if it were continuous. This is a practice that too often used which is totally incorrect. It implies linearity and equal spacing which simply not the case for | Accepted.<br>Text modified accordingly. |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | ordered categorical data that use "verbal descriptions" to place subjects in categories.<br>Proposed change (if any): Use other specific examples and stronger language on not summing, multiplying, calculating means or using other such statistics meant for continuous data on ordered categorical data. | |
| 0243-246 | 3 | *"It should be noted that… […]… should be appropriate for this type of data."*<br>Comment: These two sentences do not provide any useful information for the study design. In many instances ratings cannot be avoided; in all cases statistical methods should be appropriate.<br>Proposed change:<br>Please delete these 2 sentences. | Not accepted.<br>Clearly, methods always should be appropriate. But here this should be emphasized because experience has shown that applicants often do chose inappropriate methods when dealing with rating scales. |
| 0263 – 266 | 1 | Comment:<br>It should be mentioned that surrogate variables have to be validated in order to allow for confirmatory conclusions. | Accepted.<br>Text amended accordingly. |
| 0268 | 1 | Comment:<br>The heading 'categorised variables' is misleading as lines 269 to 276 deal with dichotomous variables (a special case of categorised variables)<br>Proposed change:<br>Replace 'Categorised variable' by 'Dichotomised variable'. | Not accepted.<br>The examples are dichotomous but the text refers to arbitrary categorical variables. |
| 0279 | 1 | Comment:<br>Time to event data might be of interest not only in case of a long-term treatment.<br>Proposed change:<br>….where time-to-event data ~~in long-term treatments~~ are of | Accepted.<br>Text changed accordingly. |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | interest. | |
| 0280 | 1 | Comment: In order to preserve the randomisation and to avoid bias, in randomised trials time to event data should be measured from the time of randomisation (not from the time of treatment start).<br>Proposed change:<br>…. time-to-event data could be the time span from randomisation ~~treatment~~ to death, | Accepted.<br>Text changed accordingly. |
| 0287 | 1 | Comment:<br>A p-value (or significance) is not an appropriate measure to estimate or describe precision, as it depends also on systematic components (e.g. treatment differences).<br>Proposed change:<br>Random errors lead to low precision – they can be kept small by increasing sample sizes or at least their size can be estimated by ~~presenting significances or~~ confidence intervals | Accepted.<br>Text changed accordingly. |
| 0293 | 5 | Comment: typo<br>Proposed change (if any): 'enrolement' should be 'enrolment' | Accepted.<br>Text changed accordingly. |
| 0306 – 311 | 5 | Comment: More clarity - Blocking is a commonly used term used in experimental design and ultimately can be modelled as a "random" effect in subsequent analyses.<br>Proposed change (if any): Examples of "adjustment techniques" such as using "blocking" in an Randomized Complete Block Design for a response gradient in one direction, or using Latin Square Designs when one has a response gradient in two directions, could be given for clarity. | Not accepted.<br>The text is considered clear enough. Since there are a number of methods for adjustment it should be left to the investigator to choose an appropriate one. No change deemed necessary. |
| 0312 and 347 | 3 | The section numbers are missing (4.3.1 and 4.3.2). | Accepted.<br>Section numbers included. |
| 0318 | 5 | Comment: 'allocation concealment': We presume this means | Accepted. |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | blinding or masking.<br>Proposed change (if any): more clarity | "Allocation concealment" is not the same as "blinding". More clarity might be necessary. The following text has been inserted (line318):<br>"*Randomisation and allocation concealment (i.e. keeping investigators and animal owners unaware of upcoming treatment assignments – without allocation concealment randomisation might become corrupted; note that allocation concealment is possible in each randomized study, blinded as well as not-blinded ones) help to avoid possible bias …"* |
| 0324 | 1 | Comment:<br>It is not clear what the term 'systematic randomisation' means as in case of a systematic treatment allocation there is no random element (and thus – strictly speaking – usual statistical methods are not applicable).<br>Proposed change:<br>… are simple randomisation~~, systematic randomisation~~, stratified randomisation and block randomisation. | Accepted.<br>Text changed accordingly. |
| 0324 | 5 | Comment: Is 'systematic randomisation' actually systematic selection? If so, it is not a form of randomisation<br>Proposed change (if any): Remove systematic randomization as it is not a form of randomization by definition or give the term more clarity | Accepted<br>See above. |
| 0350-361 | 2 | Comment: the concept of double blinding applies to human studies where both investigator and patient are blinded. In veterinary trials the concept of double and single blinding is confusing and according the definition described in this paragraph the difference relates to total binding (all agents involved in the study) versus partial blinding (only some individuals blinded) | Not accepted.<br>"Single" and "double blinding" are the *termini technici* and detailed explanation of what it means are provided in the text. |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | Proposed change (if any): Change double and single blinding by total and partial blinding | |
| 0372 | 2 | Comment: Breaking the blind for a single study animal is not an option for studies where group treatments are administered (e.g. water or in-feed medication). | Point noted.<br>However, no change considered necessary. |
| 0401-403 | 2 | Comment: These additional features which complicate the analysis would also apply to other designs and not only to the Parallel group design. | Accepted.<br>Here it is meant that although parallel designs are less complex than other designs, they might become complicated. No change necessary. |
| 0423 – 442 | 1 | Comment:<br>The whole section consists mainly of 2 examples, it might be worthwhile to shorten this section (e.g. be providing only 1 example).<br>Proposed change: provide only 1 example | Accepted.<br>Text changed as follows (line 428):<br>*… the levels may be application of the respective treatment or of placebo, or application of distinct doses in a dose finding study for a combination product) …*<br>Lines 434-440 deleted. |
| 0443-475 | 3 | Comment: The use of random effect for the centre in multicentre trials should be mentioned as an alternative strategy.<br>Proposed change: Please add *"When centers are considered as fixed or random effects, a treatment by center interaction should be explored."* | Not accepted.<br>See comment to lines 465-475. |
| 0459 – 460 | 1 | Comment: clarification<br>Proposal:<br>Furthermore, the usual sample size and power calculations depend upon the assumption that treatment effects and variances do not differ between centres. ~~the differences between the compared treatments in the centres are unbiased estimates of the same quantity.~~ | Accepted.<br>Text amended accordingly |
| 0465 – 475 | 5 | Comment: Treating centres as fixed effects will compromise | |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | extrapolation of results. It inappropriately moves the denominator mean square of the hypothesis test to "residual" and can radically inflate F statistics which in turn causes P values to be unrealistically small. It is hard to imagine a case in drug/vaccine development where centres could be considered fixed unless one were truly interested in comparing one centre to another (a characteristic associated with fixed effects). From an inference point of view, we just want to estimate the variance of the centre by treatment interaction as it is a random effect. If this variance is large, it will manifest in the F test of the treatment effect and make statistical significance more difficult to obtain for the broad inference. Devoting so much language to a fixed centre by treatment interaction and the size of centres seems unnecessary as in the vast majority of cases, centres and the interaction will be random effects allowing for a broad inference of the results. In line 465, the statement about a centre by treatment is also confusing because we always want to generalize results to the population of inference. One should always fit centre and centre by treatment as a random effect in the statistical model. This gives the broad inference necessary to generalize to the population. The definition you give for a Fixed Effect would strongly suggest the centres have to be random. (i.e. *Explanatory variables, such as treatment group or gender, in which all levels of the factor about which inferences are to be drawn from the results of the measured clinical variable, are included in the experimental design and analysis.)* Proposed change (if any): Make it very clear that centre and the centre by treatment interaction should be treated as random effects in almost, if not all cases for drug/vaccine development | Accepted.<br>Text has been modified. |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | studies | |
| 0467 | 3 | Proposed change: Please correct typo at *"...by graphical display of the results of individual centres or~~f~~ by analytical..."* | Accepted. |
| 0467 | 5 | Comment: Typo<br>Proposed change (if any): Remove floating 'of' | Accepted. |
| 0484 | 5 | Comment: Typo<br>Proposed change (if any): 'is' should be 'are' | Accepted. |
| 0484-485 | 4 | The last sentence " One generally accepted exception from this rule is the use of 90% confidence intervals in bioequivalence studies " could be misunderstood :<br>It is agreed that the consumer risk is 5 %. However, it could be suggested through this underlined sentence that for bioequivalence studies a 10 % risk is admitted….<br>The bioequivalence interval is the result of two unilateral T tests; for each test, the type 1 error is 5 %, we also build a 1-2α confidence interval, and for a bioequivalence study the controlled statistical risk is also of 5 %. | Accepted.<br>The statement on bioequivalence studies has been deleted here. |
| 0486-502 | 3 | Comments: The one-sided approach should be addressed in this section | Partly agreed.<br>One-sided approach is ok but with 2.5% significance level. Text changed accordingly. |
| 0486-502 | 3 | Comment: The primary analysis set (full set) should be introduced here (Further details are given in section 7.2) | Accepted.<br>Text amended accordingly |
| 0492 – 497 | 1 | Comment:<br>According lines 492 – 494 demonstrating a dose-response relationship is sufficient to show efficacy in a superiority trial; | Accepted.<br>The following text is amended:<br>"A successful superiority study shows a statistically significant |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | according lines 496 – 497 a successful superiority study is characterised by a statistically significant difference between test and comparator. However, these 2 statements are not consistent as the statistical proof of a dose-response relationship does not require a statistically significant difference between one dosage group and a comparator.<br>Proposal:<br>~~A successful superiority study shows a statistically significant difference between the test and the control group.~~ The clinical relevance of the observed effects ~~this difference (in particular if superiority to placebo was demonstrated) and~~ the additional benefit in relation to possible adverse effects should be discussed. | difference between the test and the control group or a statistically significant dose-response relationship." |
| 0494 – 495 | 1 | Comment:<br>The reference to the roles of the different analysis sets (7.2.3) is not fully comprehensible in this context<br>Proposed change:<br>~~(see section 7.2.3)~~ | Partly accepted (see also above).<br>Add in line 494 just before the reference to 7.2.3: *Superiority should be demonstrated for the full analysis set.* |
| 0499 | 2 | Comment: the suggestion of placebo controlled trials for non-serious illnesses is welcome. However, some guidance as to what is considered a serious illness would be welcome. | Accepted.<br>It is difficult to define "serious illness". From an animal welfare perspective it is considered unethical to prevent seriously diseased animals from treatment (and to prolong animal suffering), if a therapeutic alternative is available.<br>Text slightly amended. |
| 0499-502 | 3 | Comments: There are other reasons why a placebo controlled could be problematic; e.g. for zoonotic diseases or more pragmatically, for enrolment issues. Besides, it is unclear what should be regarded as serious illness in the veterinary medicine, knowing that 'unserious' illness in the animal may have serious | See above. |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | consequence for human health, whether owner or consumers.<br>Proposed change: *"~~For serious illnesses,~~ when an appropriate positive control exists, a placebo-controlled trial may be considered unethical, impracticable, or may be an issue for public health.* In that case the scientifically sound use of the active control should be considered~~. The appropriateness of placebo-control vs. active control must be considered~~ on a study-by-study basis." | |
| 0512 | 1 | Comment: The problem is not only to interchange type I and II errors. The problem is that one does not control the type II error of a statistical test (as the true alternative is unknown).<br>Proposed change:<br>… type II errors and in general one does not control the type II error. | Accepted.<br>Text amended accordingly. |
| 0521-523 | 3 | Comment: The undertaking of clinical equivalence trials should not be regarded as mandatory and this should be reflected in the text.<br>Proposed change: *"In some situations, clinical equivalence trials ~~are~~ may be also undertaken for other regulatory reasons, e.g. demonstrating the clinical equivalence…"* | Accepted.<br>Text changed as suggested. |
| 0539 | 3 | Comment: Please add the statement *"The choice of equivalence margins requires clinical justification."* (lines 528-529) also in the Non-inferiority trials section. | Not accepted.<br>Information already included |
| 0539 – 577 | 1 | Comment:<br>This section deals with non-inferiority trials. Thus, the term 'non-inferiority margin' instead of 'equivalence margin' should be used throughout this section in order to avoid any confusion with the section on equivalence trials.<br>Proposed change: | Accepted.<br>Text changed accordingly. |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | Change 'equivalence margin' to 'non-inferiority margin' in this section. | |
| 0541 | 1 | Comment:<br>In a non-inferiority trial one has to accept that the test product might be (not clinically relevant) worse than the reference product. If any inferiority is considered not acceptable, superiority has to be proven. Thus, the wording 'not worse' is not correct.<br>Proposed change:<br>… investigational product is not <u>clinically relevant</u> worse …' | Accepted.<br>Text changed (.. not worse within a predefined non-inferiority margin..) |
| 0552 – 553 | 1 | Comment:<br>The statement in lines 552 – 553 about providing point estimates is true also for superiority trials. | Not accepted.<br>In superiority studies, p-values are considered in the first instance, but not confidence intervals and point estimates. |
| 0558-559 | 3 | *"Ideally, active control equivalence or non-inferiority trials may also incorporate a placebo..."*<br>Comment: The statement infers that authorities can expect that a placebo should be included in all studies. This will not be possible in some therapeutic areas for welfare and practical reasons. The inclusion of a placebo should be mentioned only as a possibility.<br>Proposed change: *"~~Ideally,~~ active control equivalence or non-inferiority trials may also incorporate..."* | Not accepted.<br>"Ideally" means when there are no sound reasons for not including a third arm. |
| 0563 – 564 | 1 | Comment:<br>The terms 'internal validity' and 'external validity' should be explained in the definition section | Accepted.<br>Definitions added in the glossary |
| 0565-568 | 3 | Comment: The choice of a suitable active comparator is unclear – what is a *'well designed and well documented superiority trial(s)'*? How are applicants supposed to know in detail whether e.g. competitors products were actually tested and if it was in | Partly accepted.<br>Usually, information on comparator products can be found in published assessment reports and relevant scientific literature.<br>Text slightly changed |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | well-designed superiority trials when they are the only possible comparators?<br>Proposed change: *"A suitable active comparator would be a ~~widely used~~ therapy registered in the EU ~~whose efficacy in~~ with the relevant indication ~~has been clearly established and quantified in well designed and well documented superiority trial(s)~~ and which can be reliably expected to exhibit similar efficacy…"* | |
| 0569 | 1 | Comment:<br>The term 'assay sensitivity' should be explained in the definition section | Accepted.<br>Definition added in the glossary |
| 0569 | 3 | Comment: Importance of "assay sensitivity" is defined and understood. However, no guidance is given on determination of assay sensitivity. | Point noted.<br>However, it appears difficult to recommend a method for determination since the assay sensitivity depends on the parameter to be monitored in a certain clinical condition. |
| 0574-577 | 3 | Comment: It should be pointed here that no primary analysis set can be *a priori* defined and that the full and the per-protocol analysis set are considered co-primary (further details are given in section 7.2.3). | Accepted.<br>Text amended |
| 0579 / 580 | 1 | Proposal:<br>Equivalence / Non-inferiority margin | Accepted.<br>Changed accordingly. |
| 0580-581 | 3 | Comment: A '*clinically acceptable*' difference between treatments is highly subjective and cannot be measured with certainty. How can applicants make sure that the difference they consider clinically acceptable is also agreed by the authorities? | There is no general answer. Whether or not a difference between test and reference product would be assessed as clinically relevant depends on the use of a given product, the indication and target animal species, and the size of effects. In case of uncertainties applicants should seek scientific advice by the competent authority. |
| 0601 – 602 | 5 | Comment: Why would odds be used to compare rates which are different things? Is this just technically easier or more | Fixed differences of proportions might be problematic with proportions near 100%. E.g. with a success rate of 95% a |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | satisfactory mathematically? | margin of 15% points would result in a failure rate of 20% compared to 5%. This problem does not occur with odds ratios. Therefore, this method is recommended as a suitable alternative to fixed differences. |
| 0604-622 | 3 | Comment: Please introduce a reference to CPMP/EWP/482/99 (*Points to consider on switching between superiority and non-inferiority)*. | Accepted.<br>Reference added in the reference list. |
| 0605- 608 | 1 | Comment:<br>Depending on the definition of non-inferiority there might be situations where a confidence interval entirely above the non-inferiority margin (and above 0) does not indicate superiority but inferiority (e.g. in case of death rates).<br>Proposed change: rephrase | Accepted.<br>Text amended |
| 0623 | 2 | Comment:<br>It should be made clear that two groups of designs are commonly used: the multiple comparisons approach and the model-based approach. Recently, designs combining the two approaches have been developed. Will all three approaches be equally considered?<br>Adaptive designs may be used to determine a first-intent dose for the clinical trial. For some therapeutic classes, the dose cannot be determined on healthy animals housed in standard conditions and no experimental model has been validated to reproduce the disease. Furthermore, it sometimes impossible to find a dose that works for all animals and after administering the starting dose, the ultimate goal of the practitioner might be to increase or decrease the previous dose until the end-point (eg glycaemia) falls within a range of values considered as safe. A few adaptive designs have been developed to deal with these | The current text only refers to principles in dose response designs without providing specific guidance for certain types of study design. How to design a dose-response trial is very much up to the sponsor. Detailed guidance is not required on this point. |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | particular dose-finding studies. Their use should be discussed. | |
| 0623-635 | 3 | Comment: Please note that for some therapeutic areas, specific guidance documents exist for the design of studies. Unless specific information are sought which requires statistical input, these guidance documents should be followed with priority. Statistical analyses are not always required in these documents and the text should reflect this.<br>Proposed change: The dose response studies should be excluded from the scope of this guideline, unless they are used for confirmatory studies. | Not accepted.<br>Corresponding information is already included in the Introduction. |
| 0636 – 644 | 1 | Comment:<br>The section title is misleading. This section does not deal with a specific design but with a specific kind of endpoint. Furthermore, time to event data might not only be of interest for long-term treatments but also for treatments that are given once.<br>Proposed change: Merge with the section 'Time to event variables' (lines 278 pp) and delete section 5.2.4 | Accepted.<br>See above |
| 0645 | 2 | Comment: Group sequential designs<br>Proposed change (if any): Instead of using a group sequential design allowing for early termination of the trial in case of success or futility, one might want to use an adaptive design size in case the effect size is lower than expected, though still being clinically relevant. Would this approach be acceptable? | Accepted.<br>If described approach was predefined, it is acceptable. However, no text change considered necessary. |
| 0645-656 | 3 | Comment: In some instances studies can be planned with a so-called adaptive design involving design modifications based on the results of an interim analysis. Such a design has the potential to speed up the process of drug development or can be used to allocate resources more efficiently without lowering scientific and regulatory standards. This is especially welcome if | Point noted<br>One could use "adaptive" instead of "sequential" and provide some more information. However, adaptive designs relay on interim analyses, and there is already a section on this issue. No change considered necessary. |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | at the same time the basis for regulatory decision-making is improved. The need for a change in the study design and the change itself may have implications for the clinical interpretation of the results, which deserve consideration at the planning stage.<br>Adaptive study designs become increasingly common in human clinical trials. It would be good if this guideline would also take these new developments into consideration. Please refer to the CHMP 'Reflection Paper on methodological issues in confirmatory clinical trials planned with an adaptive design' (CHMP/EWP/2459/02) and the FDA draft guidance on 'Adaptive Design Clinical Trials for Drugs and Biologics'. | |
| 0651 – 656 | 1 | Comment:<br>These lines deal more with interim analyses then with design aspects<br>Proposed change:<br>Move to section 6.3 (interim analysis) | Accepted.<br>Text has been moved. |
| 0657-669 | 3 | Comment: More flexibility should be given in the use of experimental unit or observation unit for sample size calculation and analysis. If efficacy is observed in individual animals, but treatment is a pen treatment, it makes sense to base the analysis on the observation unit and use the experimental unit merely as an additional effect in the analysis.<br>In case the experimental unit is a pen, the clinical condition cannot be always measured on an individual basis (i.e. diarrhoea in a pen of pigs and feed medication). If the experimental unit is a pen or a house, the power of the test should be relaxed as well as the alpha significance level (i.e. P=0.1), since the size of the study may be enormous and | Not accepted.<br>Introducing the notion of "observation unit" is considered not necessary (see also below).<br><br>Relaxation of requirements for group treatment is not accepted. |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | practically impossible to set up. | |
| 0667 | 2 | Comment: This sentence refers to the Observation unit as defined on line 1240.<br>Proposed change (if any): To avoid confusion between experimental unit (see definition online 1172) and observation unit (defined on line 1240), the following could be added to the end of line 667: "… should be done at individual animal level which is defined as the "Observation unit"." | Not accepted.<br>Introducing the notion of "observation unit" is considered not necessary, since it is not referred to anywhere else in this guideline. |
| 0668 – 669 | 1 | Comment:<br>The experimental unit is not only important for sample size calculation but also for the statistical analysis model (in an ideal world the analysis model would serve as a starting point for sample size calculation).<br>Proposed change:<br>The experimental unit should be clearly specified in the protocol, since it is essential for the statistical analysis model and thus ~~to~~ the sample size calculation. | Accepted.<br>Text changed accordingly. |
| 0700 – 702 | 1 | Comment:<br>This sentence is hard to understand and (in some way) a repetition of the former statement in lines 698 - 699.<br>Proposal:<br>~~When the hypotheses to be tested are well written (i.e. in a way that the null hypothesis is the one to be rejected), it is not useful for guidelines to impose any specific value for the type II error.~~ | Accepted.<br>Sentence deleted |
| 0715 – 716 | 1 | Comment:<br>There is no section 6.3.1<br>Proposed change:<br>~~(see Section 6.3.1)~~. | Partly accepted.<br>Change the reference to 6.3. |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| 0756 -757 | 5 | Comment: This is an interesting point. Some people claim that a meta-analysis is either an analysis of pooled data, or an analysis of summary data from different studies.<br>This discussion clearly embraces both as meta-analysis. A meta analysis is an analysis of summary data from each study. An analysis using raw data from each study is a "multi-study" analysis. For a multi-study analysis, each study should have been conducted almost exactly the same way. In such a case, you need to model the random effect of the study sites and the site by treatment interaction.<br>Proposed change (if any): We propose you make the distinction between "multi-study" and "meta" analyses more clear. | Not accepted.<br>The term "multi-study analysis" is no statistical standard term. The term Meta-analysis is used for both types of studies. See for example the definition in ICH E9. No change of the text considered necessary. |
| 0760 – 767 | 1 | Comment:<br>Lines 761 – 762 state '.. changes may be appropriate, …, as a consequence of an interim analysis (see Section 6.3)…'<br>However, in line 764 it is stated 'changes should be made without breaking the blind…'.<br>As section 6.3 deals with unblinded interim analyses there is a contradiction between these 2 statements in lines 761 – 762 and 764. | Accepted.<br>Text has been modified. |
| 0769-771 | 3 | Comment: The monitoring function is typically separate from the statistician function. *"In order to protect the power of the trial"*, or a similar statement should also appear in guidelines typically read by monitors as well. | Point noted. However, no change necessary for this guideline. |
| 0772 – 811 | 1 | Comment:<br>Lines 775 - 785 give the (wrong) impression that careful pre-planning is not necessary as nearly everything might be changed based on an unblinded interim analysis. In this respect the bullet point in lines 784 – 785 seems the most critical one as | Accepted.<br>The possibility of changing the design due to interim analyses has been deleted. |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | it allows changing a superiority hypothesis into a non-inferiority hypothesis based on study data. However, such data driven analyses strategies should be avoided. | |
| 0799-800 | 3 | Comment: Please note that the establishment of an independent Data Monitoring Committee may apply to large studies sponsored by human pharmaceutical companies but are not feasible in smaller studies in vet companies.<br>Proposed change: Please relax the definition so that any personnel not directly involved in the study may monitor the safety data. | Accepted.<br>Text changed to<br>"Blinding can best be warranted if these analyses are performed by person(s) that is/are not directly involved in the study; the independence of this/these person(s) has to be made plausible." |
| 0810 – 840 | 1 | Comment:<br>Bullet points dealing with similar issues should be combined, e.g. 'hypotheses to be tested, ….' and 'justification of the use of one-sided tests' etc | Accepted.<br>Order of bullet points changed. |
| 0816 – 841 | 5 | Comment: This is justified, but would require considerably more input into the data analysis section than is seen in current protocols. | Point noted. No change necessary. |
| 0817 | 1 | Comment: typo?<br>Proposal<br>The~~is~~ …. | Accepted. |
| 0836 | 1 | Comment:<br>According lines 1009 – 1014 the GL does not deal with Bayesion methods, thus it is not clear why Bayesian estimates are mentioned here<br>Proposal:<br>• ~~Bayesian estimates~~ | Accepted.<br>Deleted. |
| 0849 – 864 | 5 | Comment: We do need to do a better job of accounting for all animals in reports, and identifying which ones are flagged from analyses—and why. | No change necessary. |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| 0881 – 888 | 1 | Comment:<br>The statements in these lines are only true if one can assume that e.g. the failure of entry criteria is completely at random (which is hardly possible to prove).<br>Proposal:<br>Delete lines 881 - 88 | Accepted.<br>Text deleted |
| 0910 – 911 | 1 | Comment:<br>Introduce an additional section '7.2.3 Safety Dataset', moving lines 9108-911 into such a section.<br>Proposal:<br>7.2.3 Safety Dataset<br>All animals that received at least one dose of study medication should comprise the safety dataset. They should be included into the analysis of safety variables according to the treatment actually received. | Accepted.<br>New section included as suggested. |
| 0927 | 3 | Comment: Does the last sentence refer to equivalence trial as well?<br>Proposed change: if it is the case, please adapt *"In non-inferiority and equivalence trials..."*. | Accepted.<br>Text amended accordingly. |
| 0931 | 1 | Comment:<br>According to the numbering system 7.2.4 'Comparison of baseline values' is a subitem to 7.2 'Analysis sets'. However this is not logical.<br>Proposal:<br>Change 7.2.4 to 7.3 (Cave: this imposes further changes in the section numbering) | Accepted.<br>Numbering changed |
| 0940 – 941 | 1 | Comment:<br>This sentence does not deal with the comparison of baseline values (but with the choice of the endpoint or the analysis | Accepted.<br>Sentence deleted. |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | method for the endpoint).<br>Proposal:<br>delete | |
| 0941 | 3 | Comment: Please correct: 7.~~6~~7.1 | Numbering has changed. |
| 0942 | 5 | Comment: This is an important issue with serological results and the ascribing of default values for values outside the limits of detection. This should only really be done with some sort of sensitivity analysis of outcome.<br>On another note, as long as data is missing at random, mixed model analyses should account for the missing data. That is the reason one estimates least squares means as they are unbiased estimates when you have missing data at random. PROC MIXED in SAS can handle this type of missing data quite nicely.<br>Proposed change (if any): Make it clear that mixed model methods take into account "missing at random" data when statistical model based estimates are obtained…like least squares means | Not accepted.<br>It is in the responsibility of the sponsor how to treat missing data statistically. Providing more specific guidance would be too prescriptive and go beyond the scope of the guideline. |
| 0944– 945 | 1 | Proposal:<br>Hence, every effort should be undertaken to ~~fulfil all the requirements of the protocol concerning the collection and management of data.~~ avoid missing data. | Accepted.<br>Text changed accordingly. |
| 0948 – 951 | 1 | Comment:<br>These lines deal with extreme values (outliers) while the lines above and below deal with missing values.<br>Proposal:<br>Shift these lines below line 963. | Accepted.<br>Text moved |
| 1015 – 1026 | 5 | Comment: We find this discussion of multiplicity with multiple primary variables a little odd considering the ready acceptance | Not accepted.<br>This section is important and should be kept as is. |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | to analyse declared secondary variables which would suffer the same problem.<br>Proposed change (if any): more clarity | Secondary variables play a different role. They serve as confirmation and support for the results in the primary variables and as such no adjustment for multiplicity is needed. On the other hand if direct claims (e.g. in the SPC) are to be made based on the secondary variables, multiplicity has to be controlled for these as well. |
| 1016 – 1017 | 1 | Comment: strengthen the wording<br>Proposal:<br>When multiplicity is present, the usual frequentist approach to the analysis of clinical trial data ~~may~~ requires an adjustment to the type I error. | Accepted.<br>Text changed accordingly |
| 1020 | 1 | Comment: strengthen the wording<br>Proposal:<br>…reduce multiplicity are always ~~sometimes~~ preferable when available, such as the identification of the key… | Accepted. |
| 1041 – 1042 | 1 | Comment:<br>The term 'protocol defined covariates' might be misleading. It should be made clear that in case of time varying covariates, only treatment independent covariates measured after randomisation and pre-specified in the study protocol should be used.<br>Proposed change:<br>This does not include protocol-defined treatment independent covariates (e.g. ambient temperature) measured after randomisation. | Accepted.<br>Text changed as suggested. |
| 1043-1046 | 3 | Comments: Does this paragraph mean that if ANCOVA is used to analyse a variable and the results are included in the statistical and/or final report, it is also required to provide ANOVA results? Please clarify. | Unadjusted results should always be reported. |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| 1058 | 1 | Comment: typo?<br>Proposal:<br>…subgroups (see 4.3.1) should be considered. Issues relating to multiplicity (see 7.6~~5~~) and statistical | Accepted.<br>(however, it is another section now, 7.7) |
| 1078 | 1 | Comment:<br>The text references to section 5.7. However there is no such section<br>Proposed change: delete/correct reference | Accepted.<br>Sentence and reference deleted (see below). |
| 1078 | 3 | Comments: Section 5.7 does not exist; is it a reference to section 4.3? Nevertheless, the added value of this statement is unclear since this is covered by GCP principles.<br>Proposed change: Please remove the sentence *"it may be helpful… section 5.7"*. | Accepted.<br>Sentence deleted. |
| 1081 – 1087 | 5 | Comment: Current EU guidelines suggest sample sizes of around 8 as a minimum for safety studies. This will only detect the most common adverse events. | Accepted.<br>However, the current guideline recommends at least 8 animals.<br>No change necessary. |
| 1082 | 1 | Comment:<br>The addition in brackets in obvious (as this is a basic requirement for studies<br>Proposal:<br>~~(both for pharmaceuticals and biologicals)~~ | Accepted.<br>Text changed accordingly |
| 1082-1083 | 3 | Comments: This sentence can only relate to field study. Safety studies in the veterinary sector are assessed in details through TAS – which is GLP, and therefore not falling within the scope of this guideline. Biologicals are now also out of scope of the guideline.<br>Proposed change: *"Safety variables ~~(both~~ for pharmaceuticals ~~and biologicals)~~ collected during field studies ~~are~~ may be* | See above.<br>Text changed. |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | *evaluated, where appropriate, according to the same statistical principles as clinical efficacy endpoints."* | |
| 1083 – 1085 | 1 | Comment:<br>This statement is true with regard to any measurable safety variable (e.g. laboratory or vital parameter) but not for any kind of adverse event | Accepted.<br>Text amended accordingly. |
| 1083-1085 | 3 | Comments: The sentence mentions *"the need to refer to normal ranges"*. This may be not always feasible: what is the reference for instance, for abortions in a bovine herd? Literature references? Data from the breeder?<br>Proposed change: In order to allow for some flexibility, please change as follows: *"One additional requirement is the need to refer to ~~normal~~ appropriate ranges for safety variables ~~when~~ for a relevant interpretation~~ing~~ of the results of any statistical analysis."* | Accepted.<br>Text amended ("appropriate reference ranges") |
| 1089 – 1128 | 1 | Comment:<br>It should be explicitly made clear that the assumptions for sample size calculation as well as the statistical methods used for analysis are briefly described in the study report. | Accepted.<br>Text amended accordingly. |
| 1093-1096 | 3 | *"In particular, a reviewer should be able to check a statistical procedure by taking the raw data, applying the statistical method and software to arrive at the same conclusions presented in the report."*<br>Comments: This seems already covered in lines 1089 to 1091. The underlying rationale for this additional sentence is unclear. If studies are conducted under GCP and a normal QA process is conducted, the stat analysis conducted and provided in the stat report should be reliable and complete enough. Besides, datasets and/or program files are not routinely submitted | Accepted.<br>Sentence deleted. |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | electronically.<br>Proposed change: please delete the sentence. | |
| 1100 | 1 | Comment:<br>Not only the experimental units but all animals entering a trial (e.g. in case a litter is the experimental unit for randomisation) should be accounted for in the report | Accepted.<br>Text amended accordingly. |
| 1109 – 1112 | 5 | Comment: We agree on the use of more graphic output and table, However, descriptive statistics are purely that and have no basis for inference. Be careful how you use the term descriptive statistics. Some people may think this means you don't need a statistical model that partitions the variability in factors based on the design. We think you mean more graphs and tables, but descriptive statistics are just summaries of the data and are not model based. One should use " statistical model based statistics" that produce least squares means, standard errors and confidence intervals. These are necessary for proper inference.<br>Proposed change (if any): Don't use the term "descriptive statistics". Just speak in terms of more graphs and tables that describe the statistical output. | Partly accepted.<br>Text amended, change "descriptive presentation" to "graphical presentation". |
| 1113-1116 | 3 | Comments: If *"Additional and perhaps complex statistical analysis"* result from the emergence of *"new questions based on the observed data"*, they should be presented as explanatory analysis. | Accepted.<br>This is evident from the text. |
| 1126-1128 | 3 | Comment: Please clarify whether the report should be 'approved' or 'signed'. If a separate statistical report is written and signed by the statistician, the final study report does not need to be signed by the statistician. | Not accepted.<br>The study report should be signed by the statistician even if there is a separate statistics report, because it may contain conclusions drawn from statistical results, and by signing the |

| Line no. | Stakeholder no. | Comment and rationale; proposed changes | Outcome |
|---|---|---|---|
| | | Proposed change: *"... and should ~~approve~~ sign the final statistical report or the final report in case no separated statistical report is provided."* | report the statistician confirms that from a statistical point of view these conclusions are plausible. |
| 1129 | 2 | Comment: No definition of the null hypothesis is included<br>Proposed change (if any): We suggest to include a definition of "null hypothesis" given that it is referred throughout the guideline | Accepted.<br>Definition added |
| 1176 and 1255 | 5 | Comment: We agree with these definitions. Your definition of a fixed effect provides evidence of the confusion one has when one reads earlier in the document about "Fixed" Centres and your discussion of the Fixed Centre by Treatment interaction. Unless you are only interested in particular centres, they would have to be considered Random effects.<br>Proposed change (if any): It should be very clear that in almost all cases, centre and centre by treatment interactions should be a random effects in the statistical model. | Accepted.<br>See above. |