<Date of submission>

EMA/CHMP/539146/2013
Guideline on the investigation of subgroups in confirmatory clinical trials
DRAFT

Comments from:

Name of organisation or individual
Danish Society of Biopharmaceutical Statistics (DSBS)

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

When completed, this form should be sent to the European Medicines Agency electronically, in Word format (not PDF).

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

Stakeholder number *(To be completed by the Agency)*	General comment (if any)	Outcome (if applicable) *(To be completed by the Agency)*
	Guidance on principles for subgroup analyses is welcomed. The described taxonomy of subgroup factors to be pre-specified is useful although a priori plausibility considerations may be difficult in practice in some cases. The problem of multiplicity and risk of “false positives” is recognized and this prioritization of sensitivity over specificity in the subgroup investigation is counterbalanced by the principles for credibility including a priori plausibility specification and replication. There is still a lot of uncertainty about criteria for flagging and visual inspection of Forest plots leaves a lot of room for subjective assessments. More accurate criteria with better understood operating characteristics would be welcome. Hopefully, this can stimulate statistical research into methods that can operationalize the principles.. Consider to simplify the language used, as the document may not be easy to read for non-native English speakers, due to the long length of the sentences.	

2. Specific comments on text

Line number(s) of the relevant text *(e.g. Lines 20-23)*	Stakeholder number *(To be completed by the Agency)*	Comment and rationale; proposed changes *(If changes to the wording are suggested, they should be highlighted using 'track changes')*	Outcome *(To be completed by the Agency)*
131		"…Extreme and/or pharmacologically plausible…" should be "…Extreme and/or pharmacologically *implausible*…" for the sentence to make sense	
294-297		Subgroup analyses said to not be independent and not providing mutually exclusive confirmation of findings. The subgroup findings based on levels of a single factor are in fact independent. They are not independent of the overall analysis findings based on the whole dataset or on subsets defined by other factors. This should be clarified.	
311-314		Initial analysis on the commonly used scale is advocated followed by analysis on absolute scale for those covariates/subgroups that become relevant for B/R decision making. Rather than promoting analyses on different scales which carries some statistical problems (often if model assumptions are satisfied on one scale they will not also be satisfied on another scale), analyses should be done on the statistically appropriate scale. If absolute scale is good then subgroup analyses may proceed exploring for treatment*factor interactions. If relative scale is natural then it becomes important to identify *prognostic* variables, i.e. variables that predict outcome regardless of treatment. Variables that predict worse outcome would then predict larger benefit if the treatment effect is constant on a relative	

Line number(s) of the relevant text *(e.g. Lines 20-23)*	Stakeholder number *(To be completed by the Agency)*	Comment and rationale; proposed changes *(If changes to the wording are suggested, they should be highlighted using 'track changes')*	Outcome *(To be completed by the Agency)*
		scale and would be candidates for subgroup identification. Hence, the method of subgroup identification should depend on the natural scale of the treatment effect considering the statistical model assumptions. The solution is not to analyse the same data on difference scales which would typically lead to problems with the assumptions in the statistical model on at least one of the scales.	
339		Please clarify the meaning of “utility” in this sentence	
443-446		It is argued that trial size may be increased to allow for better assessment of consistency of effect in subgroups. This issue may be better left for an integrated analysis across confirmatory studies where consistency between studies regarding a particular subgroup finding can be assessed.	
526-530		This section appears to discourage use of statistical measures of inconsistency which is in contradiction to e.g. line 263-265 where addition of treatment by factor interaction terms is advocated. We think that statistical testing for interaction can be useful in investigations of possible interaction with treatment effect even though absence of significance is not evidence of absence of an effect.	
531-533		Visual inspection of Forest plots may be useful but the criteria for flagging will invariably be very subjective and uncertain. Both the number of factors and subgroups defined by their levels must be considered and the precisions expressed by the CIs on the plots do not take multiplicity into account. This is	

Line number(s) of the relevant text *(e.g. Lines 20-23)*	Stakeholder number *(To be completed by the Agency)*	Comment and rationale; proposed changes *(If changes to the wording are suggested, they should be highlighted using 'track changes')*	Outcome *(To be completed by the Agency)*
		therefore done in an ad-hoc, not very transparent way.	
547-551		Is it wise to give specific limits for inconsistency flagging without context of number of factors and subgroups considered given that the CIs are not corrected for multiplicity? Either delete or rephrase to say that this could be an example of a rule in a particular situation	
630-634		Biological plausibility may be a fragile concept, in particular if this is not defined a priori. Consequently an assessor may make post hoc assessments of biological plausibility subject to biases from knowing the results. This could lead to erroneous indication restrictions	
688-691		Why the need for stratification of the randomization on the factor? This mainly secures the balance between treatment and stratification factor which has some efficiency implications for the overall stratified analysis. The results in a subgroup defined by a factor, not used for stratified randomization, are still valid since the allocation to treatment is statistically independent of the factor. A chance imbalance between treatments in the subgroup is only an issue of efficiency of the treatment comparison in the subgroup.	

Please add more rows if needed.