Discussion about the Proposed Different Levels of Subgroup Analysis

Claudia Schmoor and Frank Langer

German Region of the International Biometric Society (IBS-DR)

German Society for Medical Informatics, Biometry and Epidemiology (GMDS)

Germany

EMA Workshop

London, 07 November 2014

Who we are

- **German Region of the International Biometric Society (IBS-DR)**
    - About 1000 members
    - CS and FL 'Committee for Comments on EMA guidelines'

- **German Society for Medical Informatics, Biometry and Epidemiology (GMDS)**
    - About 2000 members
    - CS elected head of 'Division Medical Biometry'

- **Call for comments among all members**

- **Working group provided and consolidated comments**
    - R. Bender, F. Langer, J. Röhmel, A. Scherag, M. Schmid, C. Schmoor, S. Wagenfeil, H-H. von Zepelin

Overview

- **General comment**

- **Levels of subgroup analyses (SGA)**

- **Precision and sample size in key subgroups**

- **Role of qualitative interactions**

General comment I

- **Declared general aim of the guideline**
  - Maximizing a priori discussion and minimizing a posteriori discussion of the importance of subgroups

  - Focus on the important subgroups

  - Reduce the risk of performing abundant analyses

  - Reduce the risk of erroneous conclusions

- **This idea is highly appreciated and would really lead to a better interpretation of clinical trials' results**

General comment II

- **Guideline text: Sponsor should investigate subgroups defined by almost all baseline variables in all possible variations**

  - SGAs for full spectrum of demographic, disease and clinical characteristics (lines 422-424)

  - For completeness, review of other exploratory analyses (613-616)

  - Always absolute and relative effect on binary outcomes (311-314)

  - Always complement of each subgroup (290-291)

  - Treatment effect over full range of continuous covariate (292-293)

  - Several categories of continuous factor (278-279)

- **This contradicts the declared aim**

- **Guideline should focus on which <u>few</u> subgroups really should be investigated and which not**

Levels of subgroup analyses (SGA) I

- **Guideline proposes three levels of subgroup analyses (SGAs)**
  1) confirmatory SGAs in case of high expectation of heterogeneous treatment effects (not covered by the guideline)

  2) "exploratory key" SGAs in case of plausibility or evidence of heterogeneous treatment effects

  3) "truly exploratory" SGAs in case of plausibility of homogeneous treatment effects

- **Discrimination of factors in level 2) and 3) is a crucial point**

Levels of subgroup analyses (SGA) II

- **Clarification with better terminology to distinguish between "exploratory key" and "truly exploratory" SGAs would be very useful**

- **Scientific advice at the planning stage will be necessary for the decision which factors are considered key factors for exploratory SGAs**

- **The relevance of "truly exploratory" SGAs for factors defined in number 3 (homogeneity is plausible) is questionable**

- **Contradict guideline's aim of focusing on the important subgroups, reducing the risk of abundant analyses and erroneous conclusions**

- **Can level 3) SGAs be skipped?**

Precision and sample size in key subgroups I

- **Guideline text:**
    - For level 2) factors a formal proof of efficacy is not required in order to conclude on effects across the breadth of the trial population (lines 397-399)

    - Recommendation of an increase in sample size to ensure an effect estimate with reasonable precision to be able to substantiate therapeutic efficacy in important subgroups (lines 399-402, 443-446, 472-475)

- **Adequate power in subgroups not far from formal proof**

- **Guideline should account for implications of a sample size increase on feasibility**

- **Guideline should give an explicit statement about the expected precision of effect estimates in key subgroups**

Precision and sample size in key subgroups II

- **Statement on the expected precision:**

    - "… subgroups where the effect can also be estimated with reasonable precision (such that the width of the relevant CI is up to approximately 2x or 3x as wide as for the overall effect) …" (lines 547-549)

    - 2x or 3x CI-width could mean subgroups with 10-25% of total sample size

- **Statement on flag for inconsistency:**

    - "… flag for inconsistency would be an estimated effect that is outside the span of the CI for the overall effect such that the CIs for the subgroup and the overall effect are largely non-overlapping …" (549-551)

    - high probability that subgroup mean is outside 95%-CI for total population

    - "largely non-overlapping" CIs should be stated more precisely

- **"No formal efficacy proof in subgroups" statement should be consistently kept in mind throughout the guideline**

Role of qualitative interactions I

- **Discussion on statistical interactions should differentiate between quantitative and qualitative interactions**

- **Quantitative interaction**

    - means that treatment effect is in the same direction in subgroups, but size is different

    - simple numerical example: Effect on response rate

	Control	Experimental	Absolute difference	Relative difference
Subgroup 1	0.6	0.8	0.2	33%
Subgroup 2	0.4	0.6	0.2	50%

    - may only be present on a certain scale and may thus be sensitive to scale transformations

Role of qualitative interactions II

- **Qualitative interaction**

  – means that treatment is beneficial in one subgroup and harmful in another

  – simple numerical example: Effect on response rate

	Control	Experimental	Absolute difference	Relative difference
Subgroup 1	0.6	0.8	+0.2	+33%
Subgroup 2	0.6	0.4	-0.2	-33%

  – more relevant and can result in zero overall effect

- **Evaluation of qualitative interactions should be considered**

Summary

- **Guideline's aim highly appreciated**

- **Guideline should be more focused on which subgroups really should be investigated and which not**

- **Required precision in key subgroups should be specified**

- **Potential implications of a sample size increase on the feasibility of clinical trials should be considered**

- **Qualitative interactions should be considered**