Open discussion

(testing strategies)

Biomarker +/- Q (Collette EORTC)

- 2-arm superiority phase III study.
- Total population: stat.sign. and clinically relevant effect d
- Biomarker M
  - M+ subgroup: d1>d (not stat.sign. or stat.sign)
  - M- group: d2 with 0<d2<d1<d (not stat.sign. and small)
- some misclassifications (M surrogate, assay imperfect)
- What is the decision of EMA?
  - A) Do you approve the drug for the whole group?
  - B) Do you approve only in the M+ subgroup?
  - C) Other?

- Note that
  - B true M+ misdiagnosed as M- excluded from receiving an effective drug,
  - A could give pernicious signal: the worse the quality of their companion diagnostics, the better their chance to get an authorization for the whole group.

Hedging bets Q1 (Russek-Cohen)

- Effect in full population or only in a subgroup
- Proposal: alpha=.04 (full set) and =0.01 (subgroup)
- pros and cons?

Decide on analysis after (Sakov)

- If not clear after phase 1-2 whether whether keeping as a continuous variable or categorizing of subgroup analysis: pre-specified analysis might not be the scientific best one

- Acceptable to decide about the details of the analysis based on the data and only specifying up front the sub-groups?
- Acceptable to describe up front the considerations to decide without the need to pre-specify the analysis itself
  - say what kind of diagnostic tools will be used and what we expect to see?

Formal testing needed? (Rosenkranz)

- In what situations formal testing of hypotheses about specific subgroups needed?
  - the total population,
  - the target subgroups, their complements
  - how could a testing strategy look like?
- How to deal with conflicts (e.g., significant result in total population and subgroup, but not in complement)?
- What are pros and cons of enrolling
  - all comers,
  - enriched samples with patients of target subgroup, or
  - only patients from target subgroup?

Panel discussion: new methods

New methods 1

- What requirements to accept new methods? (Radlmaier & Fletcher)
- EMA supportive of
  - Bayesian modelling and/or the
  - formal incorporation of prior belief in a subgroup effect in the statistical analysis (assuming the strength of prior belief has been agreed up front)? (Radlmaier & Fletcher)
- When adaptive designs useful for identification and analysis of subgroups? Specific challenges in adaptive designs/ adaptive licensing (e.g. timing subgroup analysis)? (Rozenberg, Radlmaier & Fletcher)
- How can exploratory subgroup analysis (subgroup search) methods be used within guideline framework? (Dmitrienko)

New methods 2

- When covariate adjustments preferable to subgroup analyses?
- Why no distinction superiority and non-inferiority?
- When using graphical tools / techniques to assess treatment effect patterns, what patterns would be concerning to EMA?
- Is EMA considering any specific technical approaches for subgroups, e.g. a new metric to define consistency?
- EMA has preference for specific statistical methods ?
- EMA has any preferences for missing data handling in subgroup analyses, and would there be any specific sensitivity analyses expected?

Panel discussion exploratory subgroups

Exploratory subgroups

- What situations require exploratory subgroup analyses?
  - When pre-specification needed (Rozenkranz)
  - When justified as post-hoc analyses? (Rozenkranz)
- When subgroups emerge after design of trial during phase 3
  - What expectations does EMA have on how sponsors can incorporate subgroups (Radlmaier & Fletcher)
    Especially, what is the regulatory position wrt subgroup analyses when performed at end of trial (Criscuolo)
  - How much external evidence needed for biological rationale for a subgroup, especially in case of advances in clinical knowledge of the disease? (Radlmaier & Fletcher)
  - Which opportunities to consult regulators on subgroups emerging during phase 3? (Radlmaier & Fletcher)

Panel discussion overall goal

Overall goal

- What is the overall goal of subgroup analyses
  – demonstrate consistency across subgroups
  or
  – establish if there is evidence suggesting the treatment effect is inconsistent?
- Is multiplicity a problem if having observed an overall effect, subgroups analyses are ‘exploring’ the treatment effect in the population?

Panel discussion
pooled analyses

Pooled analyses

- What is the value and role of pooled analyses for subgroups (predefined or not),
  from reviewer's perspective,
  e.g. when investigating differential treatment effects?
  (Radlmaier & Fletcher, Satkov, Sauerbrei & Royston)

Panel discussion
Sample size

Sample size

- Lines 399-402 size subgroups so large to give sufficiently precise estimates. Lines 407-408 recruitment must reflect epidemiology of disease: inconsistent? (Smoor & Langer)
- Requirements on the number and size expl. subgroups?
  - Advice on “sufficient number of patients”, “reasonable precision” and “substantiation of therapeutic effect”? (Smoor & Langer)
  - Minimal level of efficacy expected for key subgroups? (Radlmaier & Fletcher)
  - When sufficient power needed for ‘important’ subgroups? (Radlmaier & Fletcher, Rozenkranz)
  - What is expected for in rare diseases or rare events? (Radlmaier & Fletcher)

Panel discussion
b/r labeling

b/r labeling

- How establish ground rules for: results exploratory . subgr. analyses -> labelling /reimbursement? (Rozenkranz)
- When results subgroups ‘sufficient’ to support label? (Radlmaier & Fletcher)
- Can non pre-specified subgroups inform product label? (Radlmaier & Fletcher)
- Which level of evidence required for a sponsor to restrict its label based on a subgroup analysis? For example, is the level of evidence required to restrict a label less than the level of evidence required to expand a label? (Radlmaier & Fletcher)
- Possible for a sponsor to pre-define all the relevant endpoints to be considered for the B-R assessment of subgroup? (Radlmaier & Fletcher)

Panel discussion
Classification

classification

- What analyses for investigating the potential for mis-classification (e.g. via cut-point) and impact? (Rozenkranz, Radlmaier & Fletcher)
- Really need to range cut-of if established or biologically meaningfull? given risk of increasing type I error (Smoor & Langer)
- When will the popular but dangerous practice of dichotomization be replaced by more suitable approaches using the full information from the data? (Sauerbrei & Royston)
- When subgroup based on a 'cut-point' (e.g. biomarker variable): are there any preferences for how the clinical utility of the cut-point? (Radlmaier & Fletcher)

Panel discussion consistency

consistency

- Criteria
  - Which would be used to conclude inconsistent trt effect?
  - And for consistence?
  - When would results from exploratory analyses call into question interpretation of the overall trial?
- What is the value of conducting subgroup analyses in secondary endpoints?
- flag for inconsistency (for subgroups with CI 2-3x wider than overall CI) is: point estimate outside overall CI and CI subgroup largely non-overlapping with overall CI.
  - What is meant by largely non-overlapping (given that point estimate will be 50% of times outside CI)

Panel discussion interaction tests

Interaction tests

- What does EMA feel is the role of treatment by subgroup interaction tests? (Radlmaier & Fletcher)
- Especially the utility and choice of alpha (as interaction test generally have low power, and do not explicitly determine if a particular subgroup ought to receive the investigational therapy (Russek-Cohen)