Federal Institute
for Drugs
and Medical Devices

(Regulatory) views on Biomarker defined Subgroups

Norbert Benda

Disclaimer:
Views expressed in this presentation are the author's personal views and not necessarily the views of BfArM

Biomarker defined subgroups

- **Using (genetic) biomarkers to define subgroups of patients with**
  - improved efficacy
  - improved tolerability
  - improved benefit/risk
- **Stratification according to biomarker defined patient characteristics**
  - stratified medicine = precision medicine

- **Biomarker to select patients that are likely to respond to treatment**

  ≠

  - Biomarker as a surrogate to measure response to treatment

Personalized/individualized/stratified medicine

- **Unique therapies**
  - e.g. implants using rapid prototyping, (stem) cell therapy
  - complex /expensive therapies impeding large clinical trials
- **Stratification according to specific patient characteristics**
  - e.g. biomarker defined subpopulation
- **Individualized regimen**
  - dose adjustment by age/weight/renal function
  - individual dose titration
  - etc.

Stratified therapies: Examples

- **Cetuximab**
  - treatment of colorectal cancer in patients with wild-type K-ras mutation
- **Trastuzumab**
  - treatment of HER-2-positive breast cancer
- **Gefitinib**
  - treatment of NSCLC in patients with EGFR mutation

Stratified therapies

- **Example: Gefitinib (IRESSA)**
  - IPASS Study: Gefitinib vs Paclitaxel
    - PFS
      - BM+: HR = 0.482, 95% ci (0.362; 0.642)
      - BM-: HR = 2.853 , 95% ci (2.048; 3.975)
    - ORR
      - BM+: OR = 2.751, 95% ci (1.646; 4.596)
      - BM-: OR = 0.036, 95% ci (0.005; 0.273)
    - OS
      - BM+: HR = 0.776, 95% ci (0.500; 1.202)
      - BM-: HR = 1.384 , 95% ci (0.915; 2.092)

Biomarker used for stratified therapies

- **Conventional development:**
  - looking for a safe and effective treatment in a given population/indication
- **Stratified medicine**
  - looking for a treatment and a population where this treatment is safe and effective
  - given a broader population:
    - looking for a subgroup in which benefit is more favorable than in the complementary group

      = Looking for positive treatment x subgroup interaction

      = Looking for a treatment and a predictive biomarker
- **Development: Exploration and confirmation**

Research project on biomarker defined populations University Medical Centre Göttingen – BfArM (2016 - 2019)

1. **empirical investigation of evidence on subgroup effects**
2. **comparing exploratory statistical methods for subgroup identification**
3. **method assessment based on regulatory criteria**
4. **method development**
   - modelling between-study heterogeneity
5. **assessment of regulatory consequences of between-study heterogeneity**
6. **combining exploratory and confirmatory subgroup identification in clinical development**
   - using adaptive enrichment designs and basket trials.
7. **updated comprehensive biomarker classification**
8. **systematic assessment of European SmPCs and the FDA drug labels**

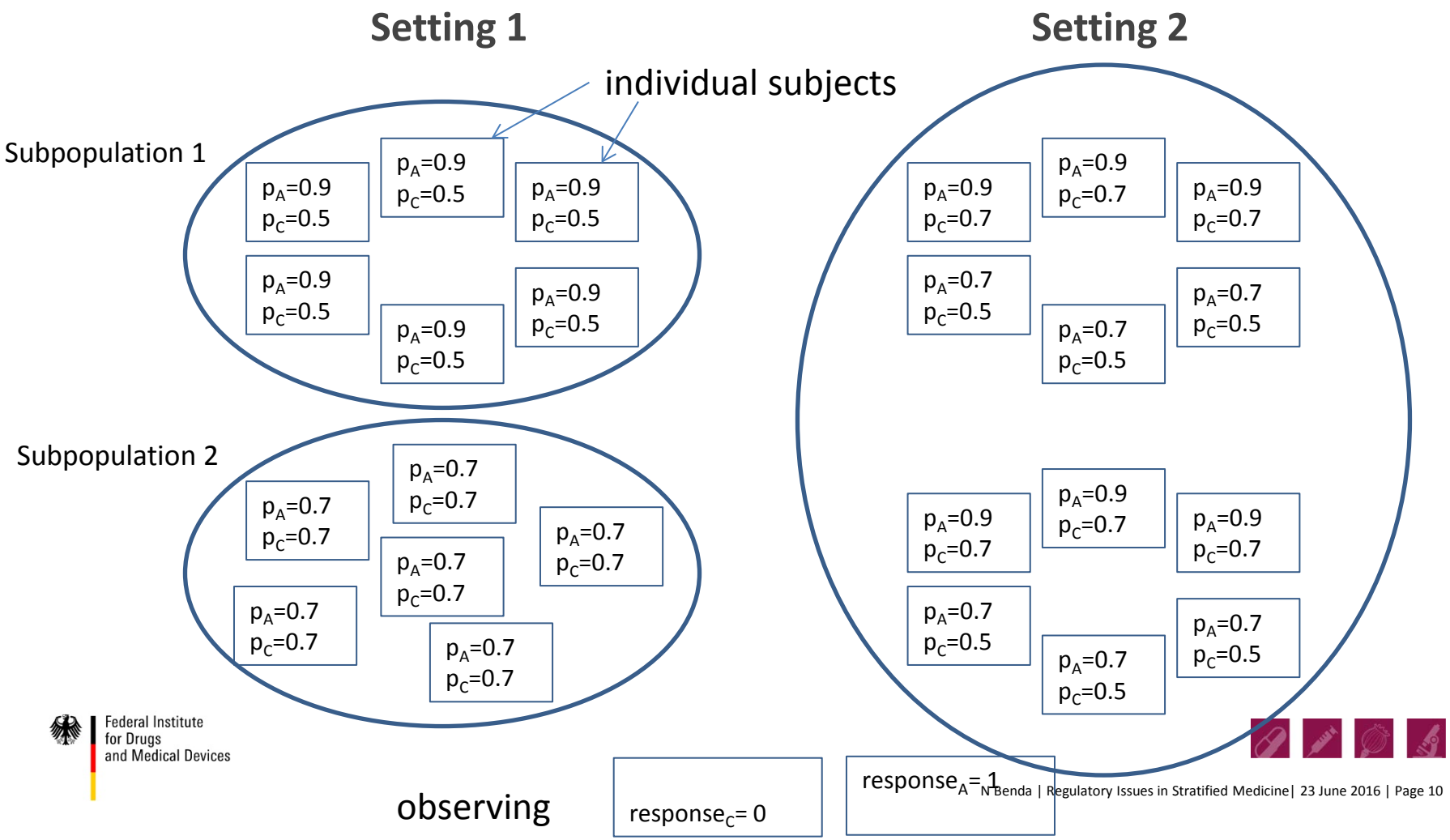
Stratified therapies: Exploration

- **Looking for most promising interaction**
  - predictive biomarker (BM)
  - inconsistency between subgroups
- **in-vitro / clinical randomize-all studies**
- **Positive interaction re.**
  - efficacy
  - tolerability
- **Questions/issues**
  - optimized strategy may consider multiple biomarkers
  - repeatability of the diagnostic tool / adjudication process
  - interaction may relate to a surrogate endpoint
  - relevant interaction size
  - positive interaction in efficacy but negative interaction in tolerability?

Stratified therapies

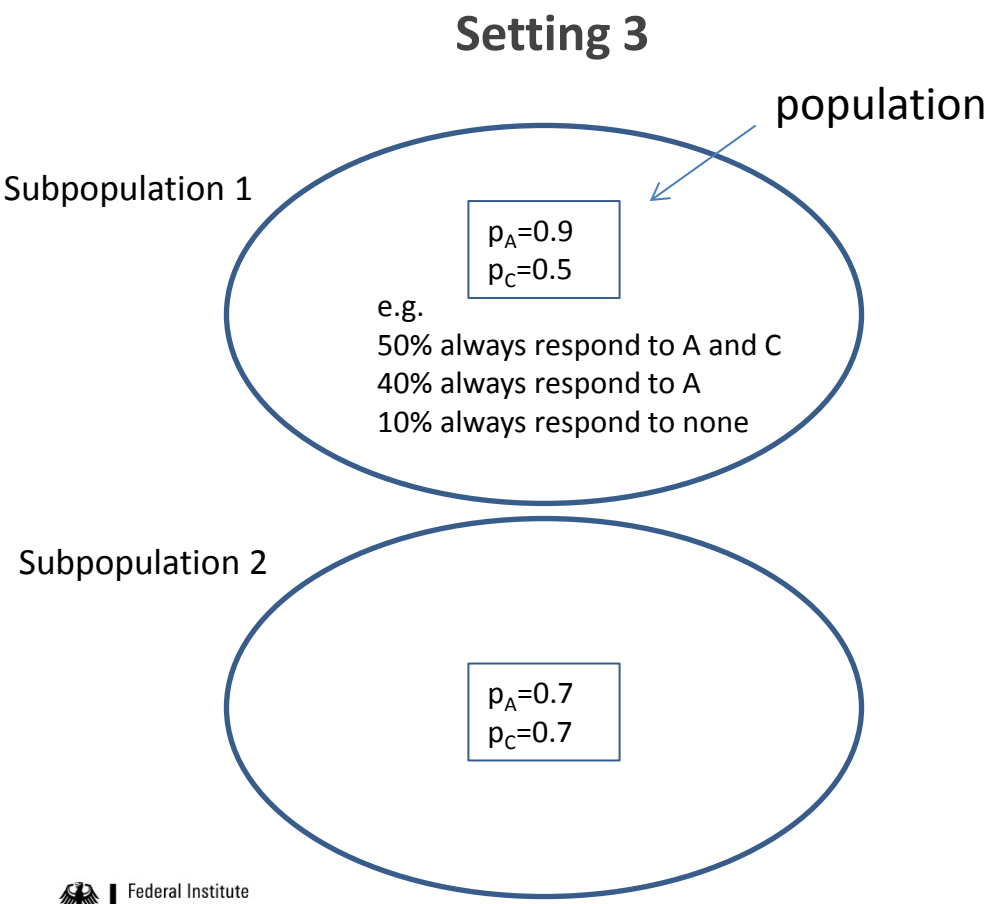
- **Treatment x subgroup interaction**
  - implies treatment x subject interaction
    - treatment effect varies across subject
  - may be difficult to verify w/o within-subject comparison/cross-over
  - interaction tests w.r.t. subgroups often lack power
- **S. Senn** (*Mastering variation: variance components and personalised medicine, SiM 2015*):
  - "Thus, I am not claiming that elements of individual response can hardly ever be identified. I am claiming that the effort necessary, whether in design or analysis, is rarely made .. "
  - "In short, the business of personalising medicine is likely to be difficult. We already know that it has turned out to be much more difficult than many thought it would be."

Any subject-by-treatment interaction?



Any subject-by-treatment interaction?

Or ?

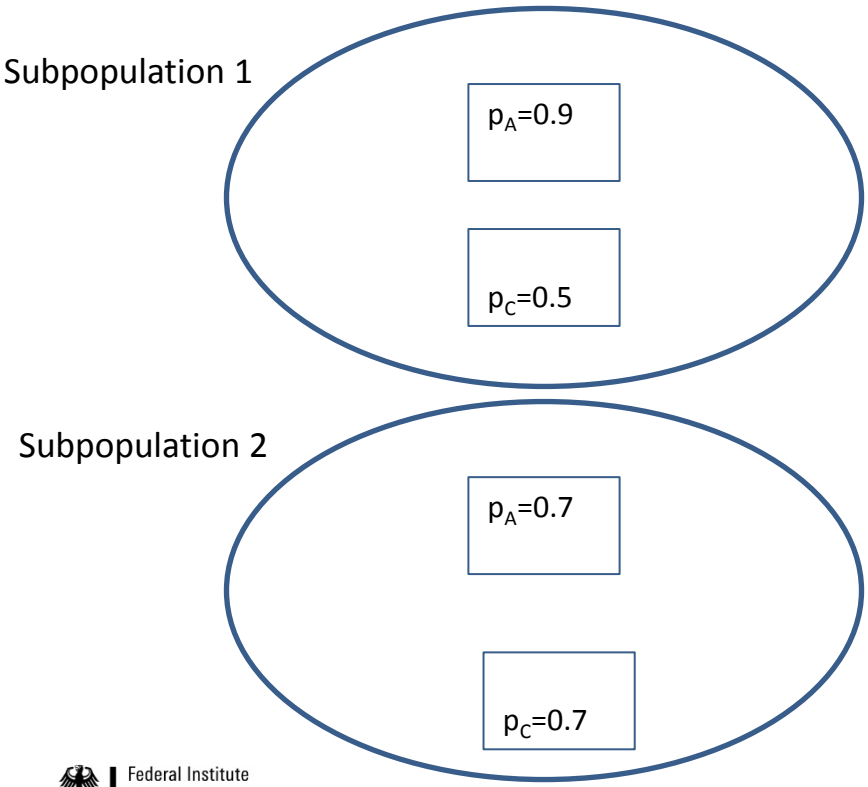


- **between subject variability**
  - vs
- **within subject variability**

Any subject-by-treatment interaction?

Observe

Setting 1



covariate (biomarker) < c

biomarker > c

lots of biomarker options:
chance finding ?

Stratified therapies: Exploration

- **Success may relate to multiple biomarker**
- **Example: Rosuvastatin**
  - cardiovascular disease prevention
  - stratification according to
    - hs-CRP (high sensitive C-reactive protein)
    - LDL cholesterol
  - risk ratio in low-LDL subjects
    - RR = 0.88 (low hs-CRP)
    - RR = 0.47 (high hs-CRP)
  - risk ratio in high-LDL subjects
    - RR = 0.42 (low hs-CRP)
    - RR = 0.72 (high hs-CRP)

Stratified therapies: Confirmation

- **Regulatory requirement**
  - confirm efficacy in subgroup (BM+) in an independent Phase III trial with proper type-1 error control
  - show positive benefit risk in BM+
  - plausibility for a reduced efficacy in BM–
- **Study design options**
  - study in BM+ only
    - (some) other data in BM– needed
  - stratification in BM+ and BM–
  - adaptive design that decides at interim for BM+ or all

Stratified therapies: Confirmation

Study in BM + only

- **Issues**
  - population size
  - information on BM–
- **Population size**
  - weaker requirements depending on medical need
    - increased model assumptions / type-1 error
- **Information on BM–**
  - usefulness of the biomarker
  - justification to exclude BM–
    - usually no confirmatory proof of effect irrelevance in BM–

Stratified therapies: Confirmation

Adaptive design to decide on BM+

- **Interim analysis to decide on subgroup or all**
  - fully pre-specified BM subgroup
  - two null hypotheses
    - no effect in all
    - no effect in BM +
  - multiplicity adjustment required
    - p-value combination test allows for free decision rule
      - decision rule may use external information
      - Bayesian rules could be applied (e.g. Brannath et al SiM 2009)
  - some information on BM– generated
    - usefulness of the biomarker

Stratified therapies: Confirmation

Possible adaptive designs

- **Predefined subgroup to be decided on at interim**
  - no subgroup definition or refinement at interim to limit the number of hypotheses to be tested
  - use of all data with adequate multiplicity adjustment
- **Adaptive signature design**
  - adjust for full population vs (any) subpopulation
  - if full population is unsuccessful
    - use first stage to define subpopulation
    - use second stage to confirm
- **Biomarker adaptive threshold design**
  - Adjust for full population vs biomarker defined subpopulation with any threshold b of biomarker score B
  - If full population is unsuccessful
  - use resampling methods to analyse $Z^* = \max\{Z(b)\}$ for test statistic Z

Stratified therapies: Confirmation

Stratified design BM + and BM -

- **Full information on different effect sizes**
  - usefulness of biomarker tested
  - exclusion or inclusion of BM – justified
- **Borrowing strength from BM –**
  - safety may be concluded from total population
    - biological plausibility required
  - efficacy may be extrapolated based on covariates
    - but difficult to justify when fundamental difference assumed

Role of subgroups in pivotal trials in general

- **Internal consistency**
  - assessing homogeneity or heterogeneity
    - evaluating interaction subgroup x treatment
- **Predefined confirmatory subgroup analysis**
  - designed to assess efficacy in subgroup
- **Exclusion of subgroups in successful trials**
  - optimizing the study population

Stratified therapies: Issues

- **Assessment of interaction / different effect sizes**
  - interaction test less informative
    - lack of power
    - scale dependency
  - in general, descriptive assessment

  scale dependency:
    - equal treatment effects on risk difference means different effect sizes on odds ratios and vice versa
    - decision on multiplicative or additive model may not be well justified

Stratified therapies: Issues

Issues in stratified design

- **Potential selection bias of treatment effect**
  - when deciding on BM+ or all
  - adaptive design alleviates selection bias
- **Success in BM+ but no clear interaction**
  - exclusion of BM - may not be informative and may be challenged
  - proof of irrelevance in BM – would require large sample sizes

Summary (1)

- **Biomarker defined subgroups → Stratified therapies**
  - main focus of the current discussion
  - based on predictive biomarkers
  - requires the identification of relevant interaction biomarker/subgroup x therapy
  - repeatability of the adjudication process paramount
  - may not be restricted to one biomarker only
    - discrimination procedures related to multiple biomarkers could be optimized

Summary (2)

- **Promise of stratified therapies depends on**
  - size of the interaction
  - further restriction appears less promising
    - residual variability limits the precision of precision medicine
- **In general more evidence required to support selection**
  - generally weak evidence on the usefulness of the selection
  - blurred by different sources of variability
- **Confirmatory strategies based on**
  - pivotal study in subgroup only
  - adaptive design to decide on subgroup or all
  - stratified design
- **borrowing information from BM– to be further justified**
  - e.g. similar safety profile ?