Application de la méthode SIDES pour identifier des sous-groupes dans un essai clinique en oncologie

Pierre Bunouf

Toulouse - 17 Mars 2023

Case study

- Randomized Phase III trial in patients with metastatic CRC in 2nd & 3rd lines.
- 441 patients randomized to Test or Control.
- Power of 90% on the *overall survival* for Test vs Control to evidence HR=0.67. The *objective response rate* was also part of the inferential procedure
- Cox proportional-hazard model to analyze OS and Logistic regression to analyze ORR adjusted on :
  - ECOG (0,1),
  - Prior use of a medication (Y/N),
  - Source of a pharmaceutical component (EU,US).

Case study

- Efficacy demonstrated and MAA in Europe

OS	ORR
HR=0.60 p=0.0001	0-R=13.7 p<0.0001

- Subgroup analysis to investigate potential sources of heterogeneity in the treatment effect
  - ✓ *18 biomarkers as candidate predictors for demography, baseline clinical status, and biological data*

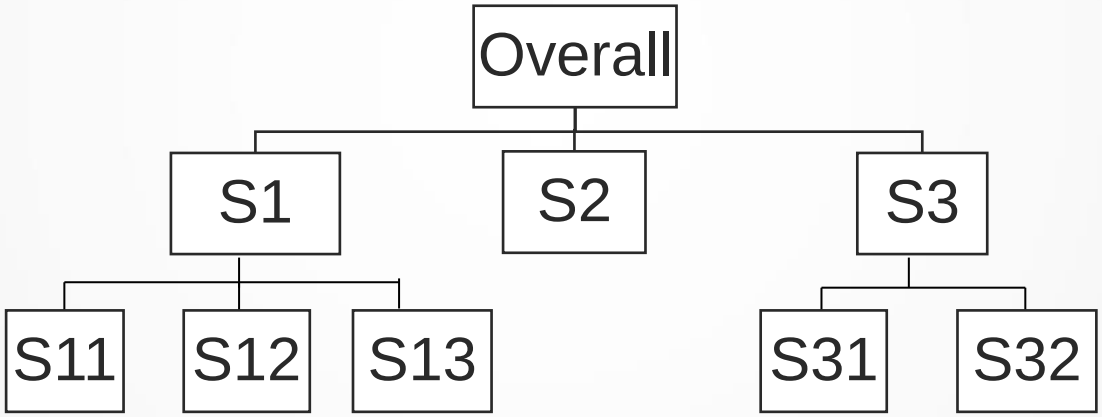
Method SIDES

- **Subgroup Identification based on Differential Effect Search**

  - Main features:
    - Recursive hierarchical procedure to obtain a classification tree
    - Identification of the promising subgroups based on an optimal cutoff value of a biomarker
    - Control of the multiplicities to limit the type 1 error

Method SIDES

- Classification tree
  - L levels and M subgroups by parent node
    - E.g., L=2, M=3



  - Recursive method where child becomes in turn parent for the next iteration

Method SIDES

- Identification of the promising subgroups
  - Effect-size estimated on the z-scale (score)
    - Continuous endpoint
      - ✓ t-test / ANCOVA
    - Binary endpoint
      - ✓ Z-test for proportion / Logistic regression
    - Time-to-event endpoint
      - ✓ Log-rank test / Cox model

Method SIDES

- Identification of the promising subgroups
  - The differential effect $D(X, c)$ is used based on a cutoff value c of a biomarker X.
    - Difference in effect size (ES) between the X values above (+) and below (-) the value c :

$$D(X, c) = 1 - \Phi\left(\frac{|ES^{+}(c) - ES^{-}(c)|}{\sqrt{2}}\right)$$

      - ✓ Φ is the CDF of N(0,1)
    - Easier to use -log $D(X, c)$

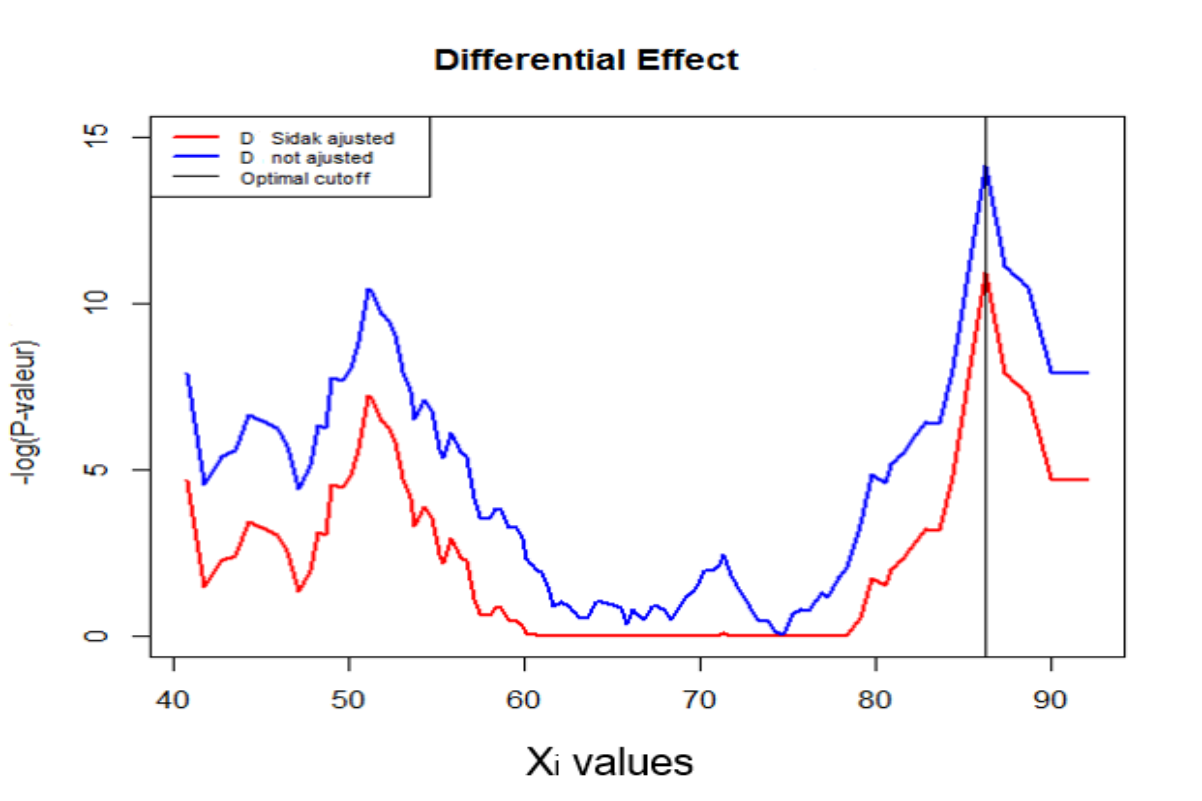
Method SIDES

- Identification of a promising subgroup
  - Define a set of cutoff values of a biomarker X_j,
  - Calculate the differential effect for each partitioning,
  - The promising subgroup is identified based on the maximum differential effect.

➢ Graphical representation of -log $D(X_j, c_j)$

Method SIDES

- Differential effect on the log p-value scale
  - Probability of identifying subgroups depends on the number of partitionings
    - Multiplicity of cutoffs controlled with Simes method



Method SIDES regular

- Determination of the classification tree
  - For each parent node, 18 candidate subgroups based on 18 biomarkers
  - SIDES regular
    - Select the M most promising subgroups based on the strongest “corrected” differential effects,
    - Prune by keeping promising subgroups according to a value of γ ($0<\gamma\leq1$) such that:
      - ✓ $p_{child} \leq \gamma\, p_{parent}$

Method SIDEScreen

- Multiplicity of candidate biomarkers
  - Selection of the most influencing biomarkers using the variable importance (VI) to reduce the subgroup space

  - SIDEScreen fixed
    - ✓ Based on the greater VI,

  - SIDEScreen adaptive
    - ✓ Based on the statistical significance of VI.

Method SIDEScreen

- VI reflects the predictive potential of the candidate biomarker
  - Determine the classification tree using SIDES regular with γ=1
  - For each biomarker X_i, sum the differential effects in the m final subgroups, such that:

$$\mathrm{VI}(X_i) = \frac{1}{m} \sum_{k=1}^{m} -\log(D(X_i, c_{ik}))$$

where c_{ik} is the optimal cutoff.

Method SIDEScreen adaptive

- Selection of biomarkers based on the statistical significance of VI
  - Distribution of VI under the hypothesis of no predictors
    - Permute randomized treatments in the overall sample,
    - Keep the maximum VI among the biomarkers and derive the distribution under hypothesis,
    - Based on a pre-defined significance level, select the predictors with significant VI value.

Method SIDES

- Multiplicity of the final subgroups for the treatment effect
  - Distribution of the treatment effect *p*-values under the hypothesis of no treatment effect
    - Permute randomized treatments in the overall sample,
    - Keep the minimum *p*-value for the treatment effect among the final subgroups and derive the distribution under H_0,
    - Adjust the observed *p*-values for final inference such that:

$$\tilde{p}_k = \frac{1}{N} \sum_{j=1}^{N} I\left(q^j \leq p_k\right)$$

Results

- Use SIDEScreen adaptive with L=M=2 and minimum subgroup size=60 to investigate the treatment effect on OS and ORR

Results

- Results of subgroup analysis

OS	ORR
Cox: type 1 error = 0.106 S1 (n=370): CEA ≤ 220.4 HR=0.524 p=0.0012 S2 (n=352): nb organs ≤ 3 HR=0.505 p=0.0012 **S3 (n=306): CEA≤220.4 & nb organs≤3** **HR=0.416 p<0.0001** **S4 (n=286): nb organs≤3 & CEA≤139.3** **HR=0.403 p<0.0001**	No subgroups identified

Results

- Results of subgroup analysis
  - The treatment effect is maximum in 65% of patients with moderate status benefit the most from Test (S4: HR=0.403 p<0.0001) …
  - … but patients with more aggressive cancer also benefit from the treatment (complementary of S4: HR= 0.595 p=0.234).
  - Treatment effect is homogeneous

Discussion

- Subgroup analysis in confirmatory clinical trials
  - Subgroup analysis could be planned systematically to demonstrate the homogeneity of treatment effect across patients' categories.
  - For a new claim, the results of retrospective subgroup analysis are regarded as hypothesis generating…
    - ✓ *Exception for orphan drug ?*

Back-up

Publication

Therapeutic Innovation & Regulatory Science
https://doi.org/10.1007/s43441-021-00329-1

DIA

ORIGINAL RESEARCH

Data-Driven Subgroup Identification in Confirmatory Clinical Trials

Pierre Bunouf[1] · Mélanie Groc[1] · Alex Dmitrienko[2] · Ilya Lipkovich[3]

Received: 29 April 2021 / Accepted: 22 July 2021

Abstract
Data-driven subgroup analysis plays an important role in clinical trials. This paper focuses on practical considerations in post-hoc subgroup investigations in the context of confirmatory clinical trials. The analysis is aimed at assessing the heterogeneity of treatment effects across the trial population and identifying patient subgroups with enhanced treatment benefit. The subgroups are defined using baseline patient characteristics, including demographic and clinical factors. Much progress has been made in the development of reliable statistical methods for subgroup investigation, including methods based on global models and recursive partitioning. The paper provides a review of principled approaches to data-driven subgroup identification and illustrates subgroup analysis strategies using a family of recursive partitioning methods known as the SIDES (subgroup identification based on differential effect search) methods. These methods are applied to a Phase III trial in patients with metastatic colorectal cancer. The paper discusses key considerations in subgroup exploration, including the role of covariate adjustment, subgroup analysis at early decision points and interpretation of subgroup search results in trials with a positive overall effect.

Keywords Confirmatory clinical trials · Data-driven subgroup analysis · Recursive partitioning method · Interim analysis · Covariate adjustment · Multiplicity adjustments

Introduction

Data-driven subgroup identification plays an important role in all Phase III clinical trials. To define the goals of data-driven or post-hoc assessments of patient subgroups, it is instructive to compare those to the goals of confirmatory subgroup analysis. Confirmatory subgroup analysis focuses on a small set of prospectively specified subsets of a trial's population. These subsets are most often defined using key demographic variables such as age and gender as well as important clinical variables such as baseline disease severity. These pre-defined subgroup analyses help characterize the homogeneity of treatment effects across the patient population of interest and support the conclusion of broad consistency with the overall trial result. By contrast, data-driven subgroup identification deals with open-ended subgroup searches aimed at identifying subsets with desirable characteristics such as an enhanced efficacy profile or improved benefit-risk profile.

The development of tailored therapies and personalized medicine relies heavily on methods for subgroup discovery based on a variety of baseline patient characteristics, most importantly biomarkers. Post-hoc subgroup assessments have been successfully applied to multiple confirmatory trials to discover new treatments or define the most relevant patient population for an existing treatment [1]. Because of this, the general topic of subgroup exploration has attracted much attention in the literature as well as regulatory guidance documents. For a general overview of statistical considerations in exploratory subgroup analysis, see Alosh et al. [2] and Ondra et al. [3]. Numerous statistical methods developed to support exploratory subgroup assessments are reviewed in Lipkovich et al. [4, 5]. In addition, post-hoc subgroup analysis strategies were discussed in recently published regulatory guidelines such as the guideline on the investigation of subgroups in confirmatory clinical trials [6] published by the European Medicines Agency.

It is important to note that exploratory subgroup analysis is often thought to be inferior to confirmatory subgroup

✉ Pierre Bunouf
pierre.bunouf@pierre-fabre.com

[1] Pierre Fabre, Toulouse, France
[2] Mediana, Carolina, PR, USA
[3] Eli Lilly, Indianapolis, IND, USA

Published online: 29 July 2021

Springer