Exploratory subgroup analysis: Post-hoc subgroup identification in clinical trials

Alex Dmitrienko (Quintiles)
Ilya Lipkovich (Quintiles)

EMA Expert Workshop 2014
Outline

Exploratory subgroup analysis
Guideline-driven and principle-driven approaches

Key principles of subgroup identification
Analytic subgroup search procedures, complexity control, adjustment for selection bias, biomarker screening, reproducibility assessment

Case study
Phase III development program in patients with nosocomial pneumonia
Exploratory subgroup analysis
Subgroup analysis

Subgroup analysis approaches

Several classification schemes proposed in clinical trial literature (Varadhan et al., 2013; Lipkovich and Dmitrienko, 2014b)

Simplified classification scheme

Confirmatory subgroup analysis relies on a small set of well defined patient subgroups

Exploratory subgroup analysis focuses on a large set of loosely defined patient subgroups
Applications of exploratory subgroup analysis

Scenario 1 (positive trial)
Assess consistency of treatment effects across key subgroups

Scenario 2 (positive trial)
Analyze subgroups in a post-hoc manner to (1) exclude a subgroup due to lack of efficacy or (2) focus on a subgroup without safety issues
Add a subgroup with enhanced treatment effect

Scenario 3 (negative trial)
Discover subgroups with enhanced efficacy profile
Applications of exploratory subgroup analysis

Scenario 1 (positive trial)
Consistency assessment

Scenario 2 (positive trial)
Post-hoc subgroup identification

Scenario 3 (negative trial)
Post-hoc subgroup identification
Post-hoc subgroup identification

Guideline-driven approaches
Multiple sets of guidelines attempt to improve credibility of exploratory subgroup analysis
Checklist with 25 rules (Brookes et al., 2001), checklist with 21 rules (Rothwell, 2005), checklist with 11 rules (Sun et al., 2010)
Main rule: Proceed with caution

Principle-driven approaches
Subgroup identification ought to be based on specific operationalizable principles
Post-hoc subgroup identification

Key idea
Utilize recent developments in machine learning and data mining to pre-specify a **subgroup exploration strategy**

Principles of subgroup identification
Define an analytic subgroup search procedure
Control complexity of search space
Perform reliable inferences in selected subgroups
Key principles of subgroup identification

Analytic subgroup search procedures

Haphazard/unplanned subgroup exploration leads to spurious results

Tools used in subgroup search algorithms

Recursive partitioning algorithms with pre-specified rules for subgroup generation to select the most relevant subgroups (e.g., partitioning rules based on maximum differential treatment effect)
Key principles of subgroup identification

Complexity control

Unconstrained (greedy) subgroup search creates a very large search space, which hinders the assessment of clinical relevance

Tools for reducing the size of search space

Efficient subgroup pruning rules to choose child subgroups in recursive partitioning algorithms
Key principles of subgroup identification

Reliable inferences and interpretation

Unadjusted treatment effects in subgroups are misleading due to “optimism bias”

Tools for performing reliable inferences

Resampling- or cross-validation-based adjustments (p-value adjustment and “honest” treatment effect estimates) to perform reliable inferences in subgroups
Subgroup identification methods

Global outcome modeling
Virtual Twins method (Foster et al., 2011), Bayesian subgroup search (Xu et al., 2014)

Global treatment effect modeling
CART-based (Classification And Regression Trees) methods, e.g., Interaction Trees method (Su et al., 2009)

Local modeling
Responder Identification method (Kehl and Ulm, 2006), SIDES method (Lipkovich et al., 2011)
Local modeling

Subgroup Identification based on Differential Effect Search (SIDES)

Recursive partitioning-based subgroup identification method which provides a multivariate assessment of biomarkers, employs complexity control and accounts for selection bias

SIDEScreen method

Extension of original SIDES method with efficient biomarker screening for complex settings, e.g., > 100 biomarkers (Lipkovich and Dmitrienko, 2014a)
Case study
Clinical trial database
Total sample size: 1289 patients
Primary endpoint: All-cause mortality at 28 days
Overall outcome: Slightly negative treatment effect in overall patient population

Exploratory objective
Identify biomarkers that help predict positive treatment response

Reference
Dmitrienko et al. (2014)
Main challenge

Candidate set included 26 biomarkers (mostly demographic and clinical variables)

Large set of candidate biomarkers created a vast search space

SIDES-based subgroup search

Aggressive pruning rules to reduce the search space

Biomarker screens to filter out non-informative (noise) biomarkers and focus on best predictors of treatment response
Greedy subgroup search (390 subgroups)

Black dot: Overall patient population
Red dots: Patient subgroups
Efficient subgroup search (3 subgroups)

Black dot: Overall patient population
Red dots: Patient subgroups
Selected patient subgroup
Serum creatinine clearance $> 67$ mL/min
Sample size: 352 patients
Raw treatment effect $p$-value: $p = 0.0077$

Adjustment for selection bias
Adjusted treatment effect $p$-values were computed using a resampling-based method
Efficient subgroup search: Lower multiplicity burden due to reduced search space
Additional important considerations

Adjustment for optimism bias
Cross-validation to derive honest (bias-adjusted) estimates of treatment effects in selected patient subgroups

Reproducibility assessment
“Learn and confirm” method to assess the likelihood of replicating results in another clinical trial
Summary
Summary

Principled-based approach to post-hoc subgroup identification

Analytic subgroup search procedures for examining all relevant patient subgroups to find subsets of overall population with desirable characteristics

Statistical methods

Multiple methods have been developed with available software implementation

Web site: http://biopharmnet.com/wiki/Subgroup_Analysis
References


References


References

