



A Systematic Approach for Post Hoc Subgroup Analyses With Applications in Clinical Case Studies

Christoph Muysers
Bodo Kirsch

PSI / DIA webinar • October 2019



Further co-authors of paper:
Dmitrienko A, Kulmann H, Lippert S,
Schmelter T, Schulz A, Mentenich N,
Schmitz H, Schaefers M, Meinhardt G,
Keil T, Roll S



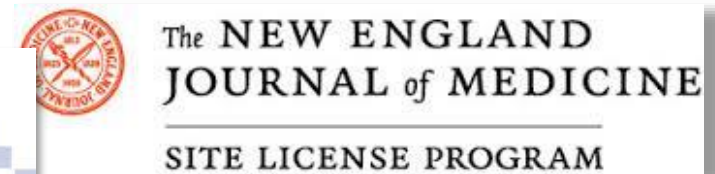
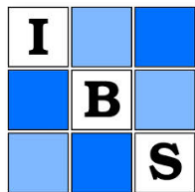


The next 20 minutes

- // The demand for subgroup analyses
- // The 'Subgroup Explorer' (tool)
- // The 'Subgroup Screening' (procedure)

Subgroup Analyses

medically important -- regulatory requirement -- many stakeholders



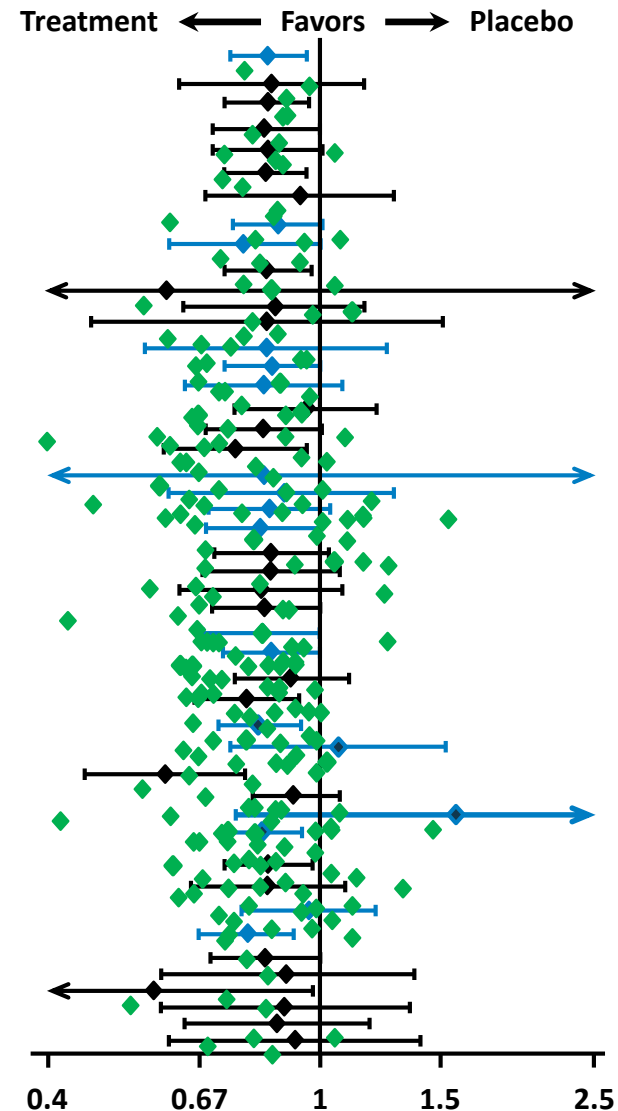
Even 'Important' only Subgroups can be Overwhelming

... factors used in stratification randomisation // factors with some biological plausibility or external evidence where heterogeneous response might be hypothesised // at least demographic factors, including genomic factors, related to the mechanism of action pharmacology // in addition, careful consideration should be given to other factors that might plausibly be predictive for different response to treatment such as stage, severity or phenotype of disease, use of concomitant medications and possibly region, country, or centre // truly exploratory analyses should be planned for the spectrum of demographic, disease and clinical characteristics, including those factors a particular factor there is good argumentation why homogeneity of response to treatment is plausible // analysis of the complement subset should also be displayed // review of other exploratory analyses // exploration of interactions and effects in subgroups on different scales // analyses of continuous variables using different cut-offs should routinely be performed ... *

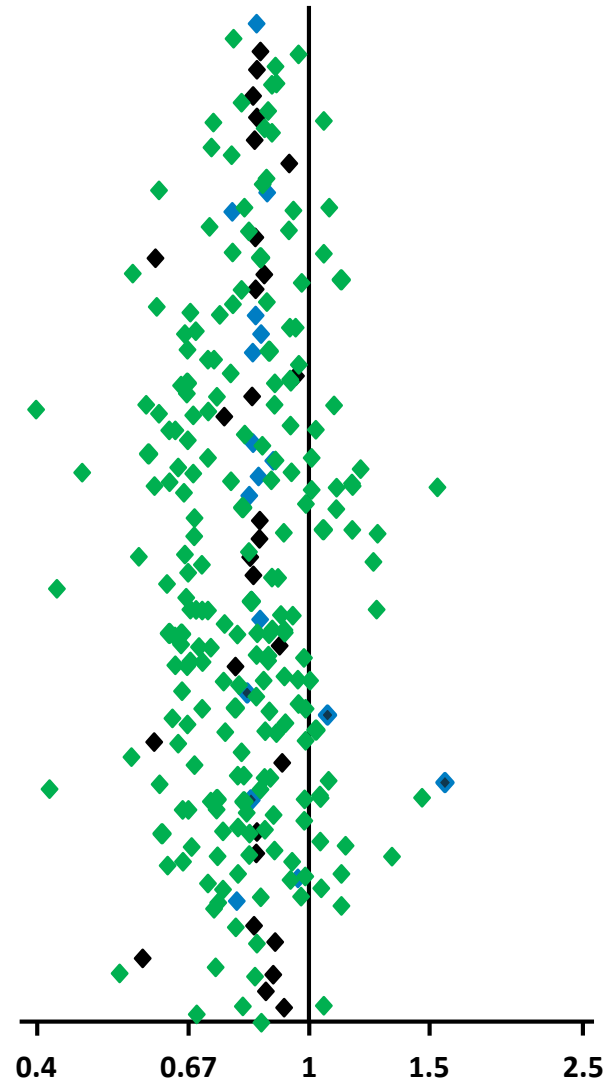
* EMA 2019, Guideline on the investigation of subgroups in confirmatory clinical trials

From Forest Plots to In-depth Subgroup Screening

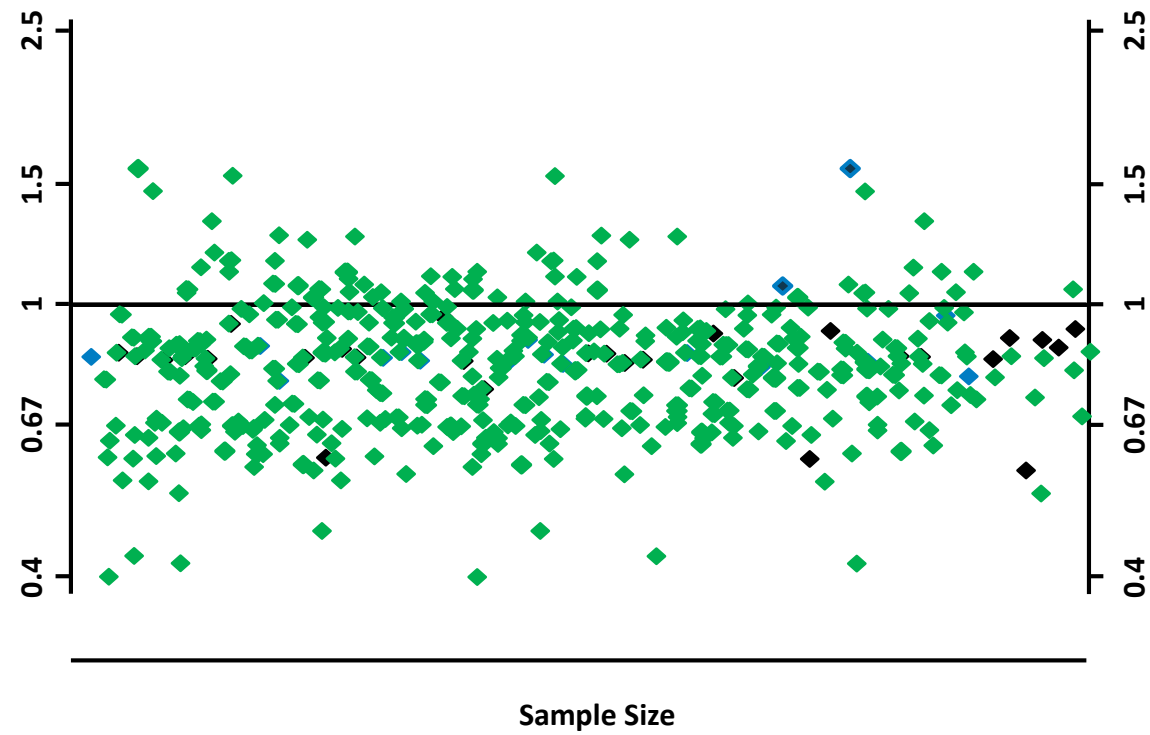
- Overall
- Age X Sex
- Age [yrs] (<55; ≥55; <65; ≥65; <75; ≥75)
- Age X Race
- Age X Weight [kg]
- Age X BMI [kg/m²]
- Sex (Male; Female)
- Sex X Race
- Sex X Weight [kg]
- Race (White; Black; Asian; Other)
- Race X Weight [kg]
- Race X BMI [kg/m²]
- Weight [kg] (<60; ≥60-<90; ≥90 kg)
- Weight X BMI [kg/m²]
- BMI [kg/m²] (<25; ≥25-<30; ≥30)
- Age X CrCl [mL/min]
- Age X CrCl [mL/min] X CrCl [mL/min]
- Sex X CrCl [mL/min]
- CrCl [mL/min] (<30; ≥30-<50; ≥50-<80; >80)
- Age X Race X CrCl [mL/min]
- Sex X Index Event
- BMI X Index Event
- Index Event (STEMI; NSTEMI; Unstable angina; NSTEMI+Unstable Angina) or MI (yes, no)
- Age X Prior MI (yes, no)
- Sex X Prior MI (yes, no)
- Prior MI (yes, no)
- Age X Prior MI (yes, no) X Index Event (yes, no)
- Age X Prior MI (yes, no) X Index Event (yes, no)
- Weight X Prior MI (yes, no) X Index Event (yes, no)
- PCI for Index Event (yes, no)
- Age X Elevated troponin (yes, no)
- Age X Elevated troponin (yes, no) X Index Event (yes, no)
- Elevated Cardiac Biomarker (yes, no)
- Age X Congestive Heart Failure (yes, no)
- Age X Congestive Heart Failure (yes, no) X Stroke/TIA (yes, no)
- Congestive Heart Failure (yes, no)
- Age X Prior Ischemic Stroke/TIA (yes, no)
- Age X Prior Ischemic Stroke/TIA (yes, no) X Hypertension/Hypercholesterolemia (yes, no)
- Prior Ischemic Stroke/TIA (yes, no)
- Age X Hypertension/Hypercholesterolemia (yes, no)
- Hypertension/Hypercholesterolemia (yes, no)
- Age X Diabetes (yes, no)
- Age X Diabetes (yes, no) X Region
- Diabetes (yes, no)
- Age X Region
- Region (East Europe; Western Europe; North America; South America; Asia; Others)
- Weight X Region
- Diabetes X Region



From Forest Plots to In-depth Subgroup Screening



From Forest Plots to In-depth Subgroup Screening





Subgroup Explorer and the Exploration Hub

<https://cran.r-project.org/web/packages/subscreen/>



explore factor level combinations



exploration of all factors at a glance



compare subgroups for two endpoints





Feature: Importance Tab

Machine learning based prioritization of analyzed factors

information about variable importance based on a random forest algorithm

Variable Options | Importance Tab

Display Options | Colour Options

Importance Value Option ?

- No Importance Value
- Use Variable Importance Values
- Use Ranking of Variable Importance Values

Choose number of Variables which are most important ?

1 3 7

Sorting order: ?

- Increase
- Decrease

Used/colored importance variables ?

bilig
albuming
sex

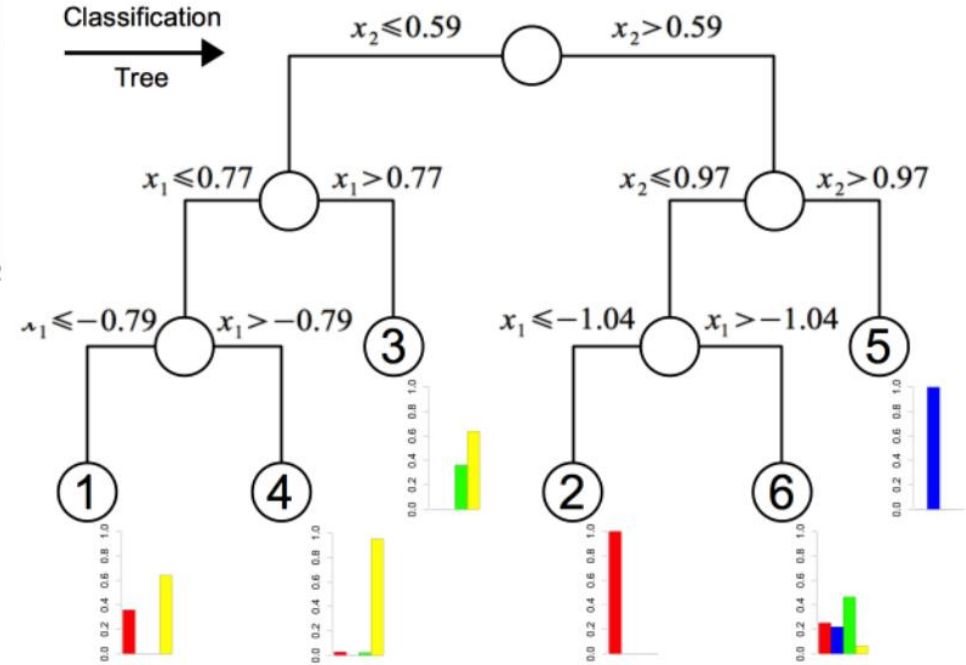
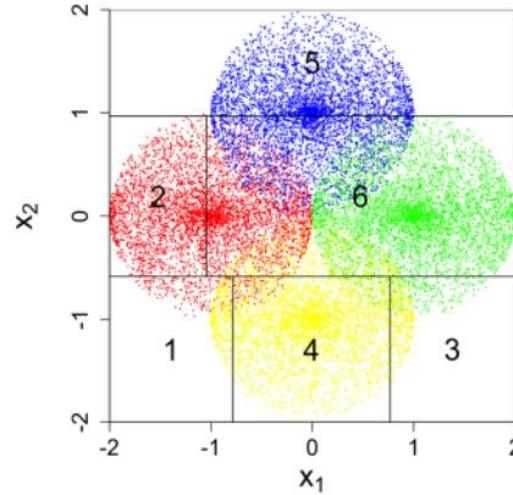


Illustration taken from Random Forest R-package

<https://kogalur.github.io/randomForestSRC/theory.html>

Variable Options

Importance Tab

Display Options

Colour Options

Target variable

Mean_changeTS_w52

Reference variable

Number.of.Subjects

Subgroup Filter

no selection

Subgroup level(s)

1

3

Plot Type

linear logarithmic

Y Range

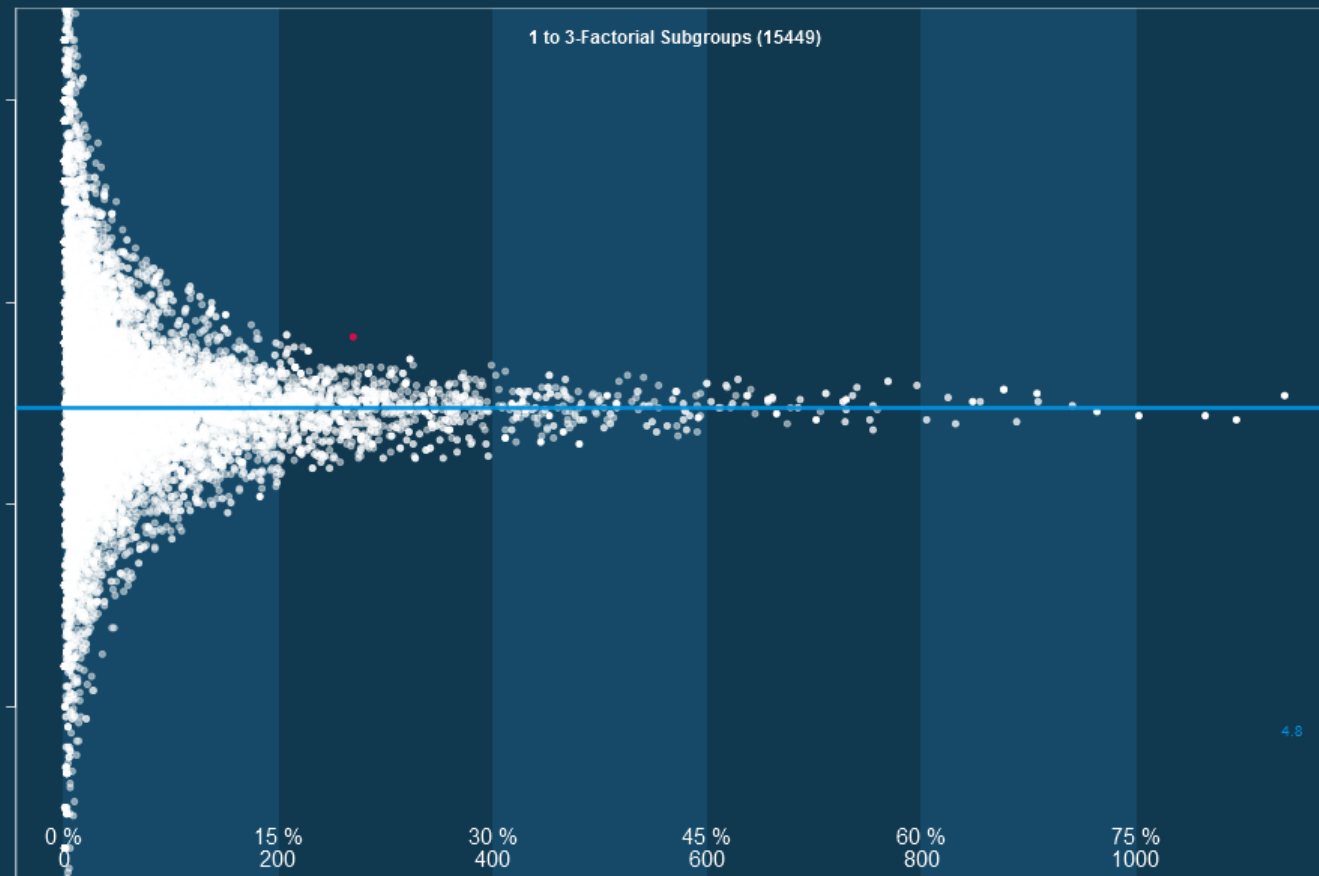
-40

-17

23

50

Help text



Selected Subgroups

Filtered Subgroups

Parent Subgroups

Memorized Subgroups

Copy

Print

Show 6 rows

Column visibility

Table of Selected Subgroups

Memorize	SGID	Number.of.Subjects	Mean_changeTS_w52	nfactors	PerProtocol	EQ5D_score
All	All	All	All	All	All	All
Memorize	452	270	8.3	2	Y	Worst QoL

Showing 1 to 1 of 1 entries

Previous

1

Next

Variable Options

Importance Tab

Display Options

Colour Options

Target variable

Mean_changeTS_w52

Reference variable

Number.of.Subjects

Subgroup Filter

EQ5D_score

Choose a value

Worst QoL

Subgroup level(s)



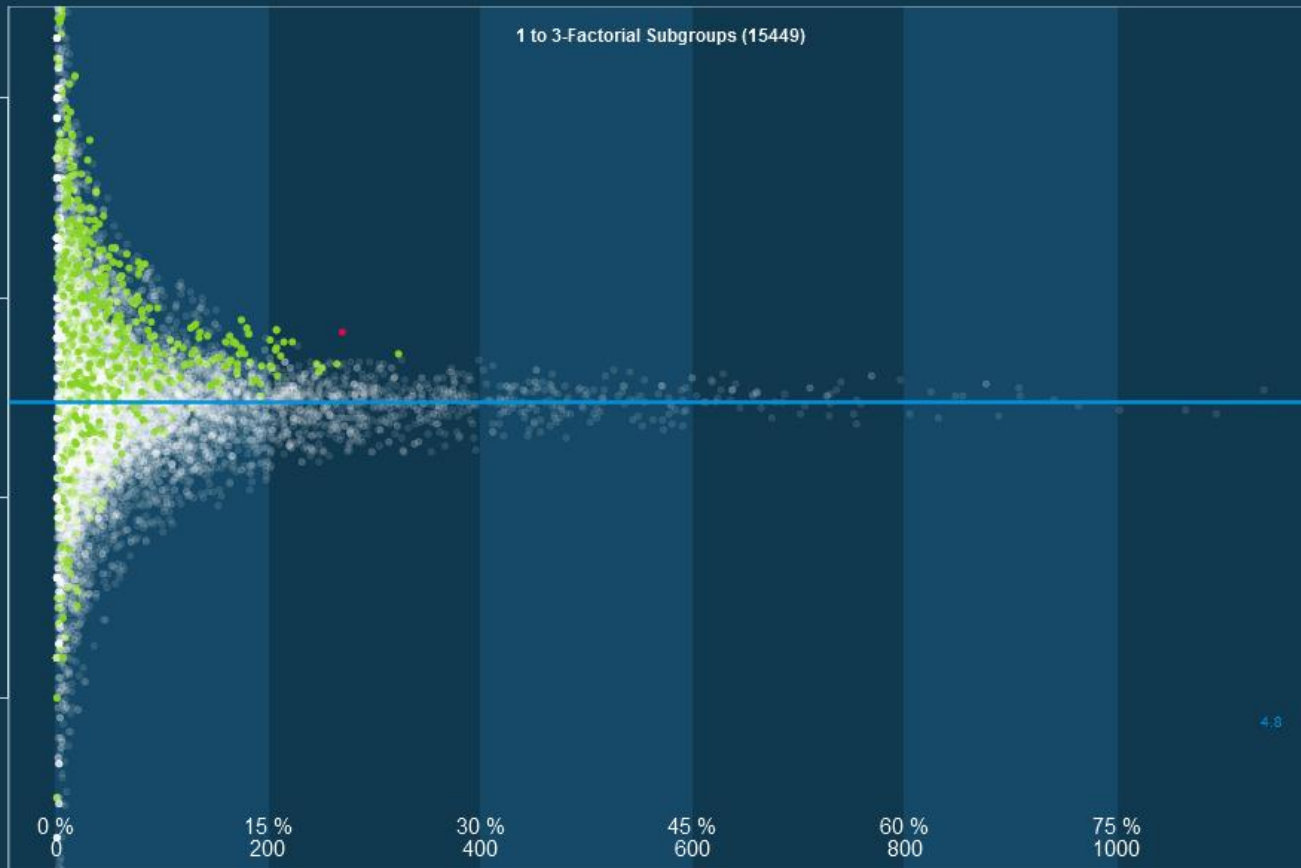
Plot Type

linear logarithmic

Y Range



Help text



Selected Subgroups

Filtered Subgroups

Parent Subgroups

Memorized Subgroups

Copy

Print

Show 8 rows

Column visibility

Table of Selected Subgroups

Memorize	SGID	Number.of.Subjects	Mean_changeTS_w52	nfactors	PerProtocol	EQ5D_score
All	All	All	All	All	All	All
Memorize	452	270	8.3	2	Y	Worst QoL

Variable Options

Importance Tab

Display Options

Colour Options

Target variable

Mean_changeTS_w52

Reference variable

Number.of.Subjects

Subgroup Filter

EQ5D_score

Choose a value

Best QoL

Subgroup level(s)



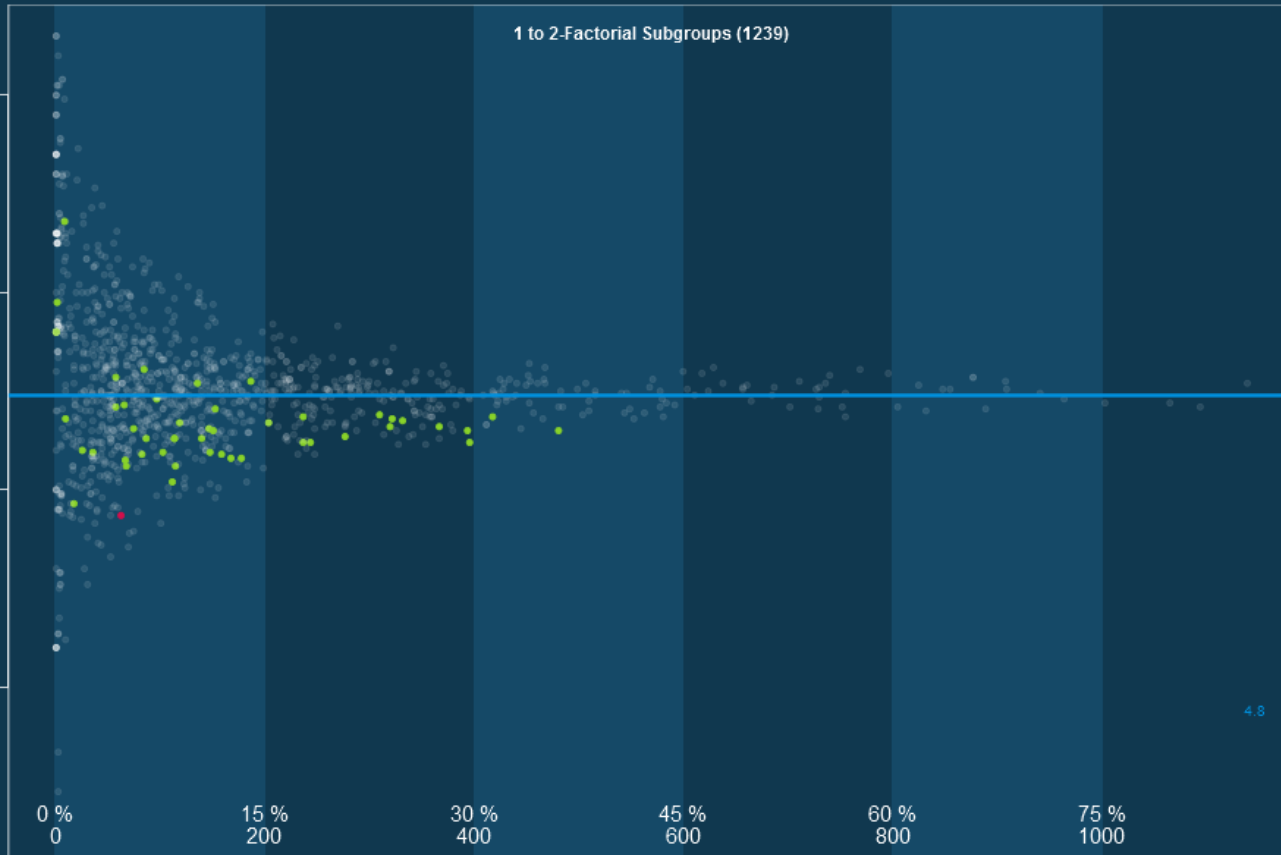
Plot Type

linear logarithmic

Y Range



Help text:



Selected Subgroups

Filtered Subgroups

Parent Subgroups

Memorized Subgroups

Copy

Print

Show 6 rows

Column visibility

Table of Selected Subgroups

Memorize	SGID	Number.of.Subjects	Mean_changeTS_w52	nfactors	PerProtocol	EQ5D_score
All	All	All	All	All	All	All
Memorize	445	63	-1.3	2	N	Best QoL


First Target variable ?

Mean_changeTS_w52

Plot Type (Compare Plot: y-axis / Bubble Plot: x-axis) ?

linear logarithmic

Range (Compare Plot: y-axis / Bubble Plot: x-axis) ?



A horizontal slider with a blue track. The left handle is at -15 and the right handle is at 18. The scale ranges from -40 to 50 with major ticks every 10 units.


Second Target variable ?

Mean_changeTS_w12

Plot Type (Compare Plot: y-axis / Bubble Plot: y-axis) ?

linear logarithmic

Y Range (Compare Plot: y-axis / Bubble Plot: y-axis) ?



A horizontal slider with a blue track. The left handle is at -14.4 and the right handle is at 18. The scale ranges from -40 to 40 with major ticks every 8 units.

Reference variable ?

Number.of.Subjects

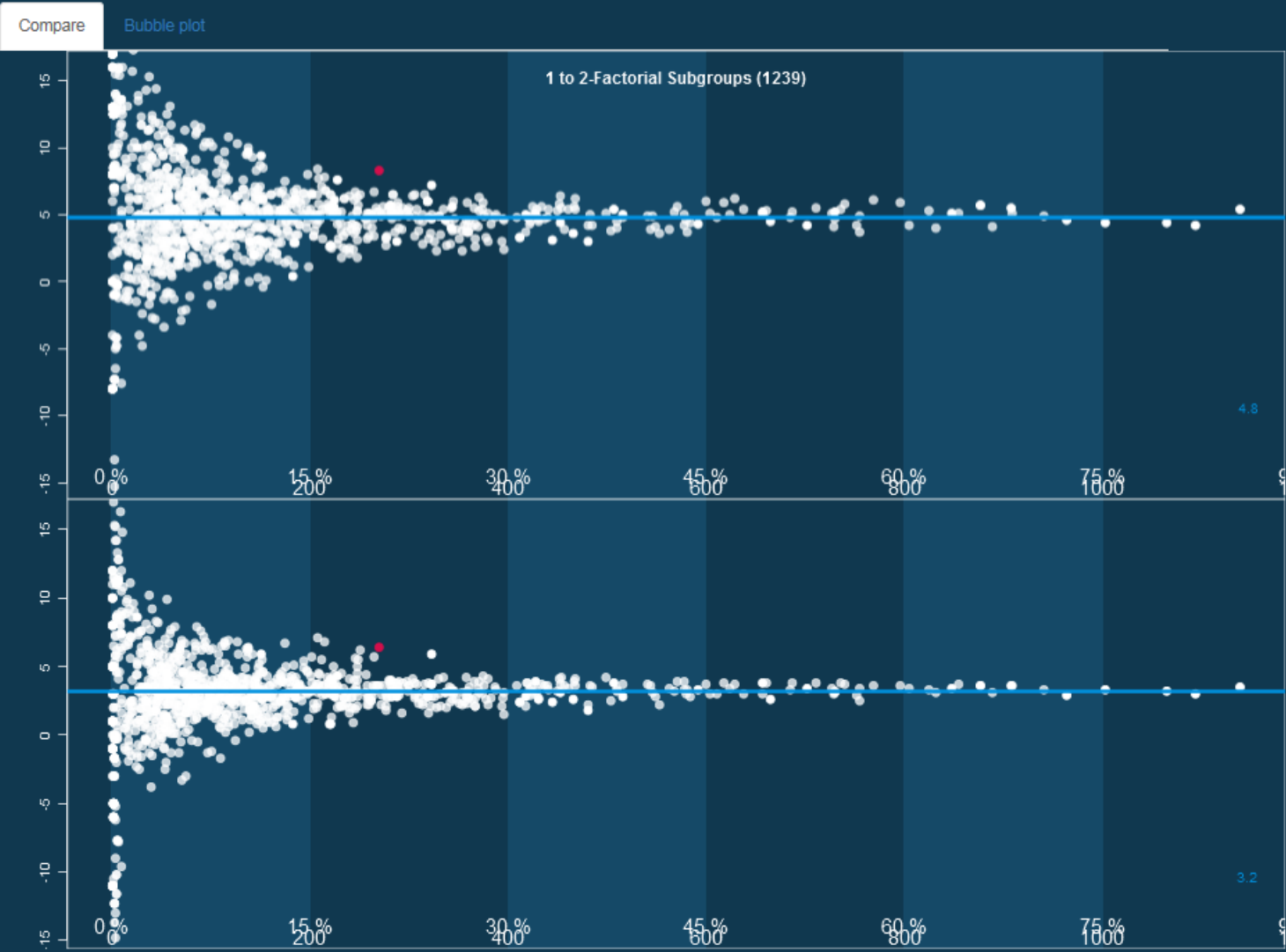
Subgroup Filter ?

no selection

Subgroup level(s)



A horizontal slider with a blue track. The left handle is at 1 and the right handle is at 2. The scale ranges from 1 to 3 with major ticks at 1, 2, and 3.



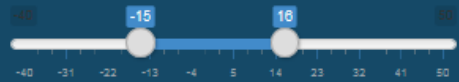
First Target variable

Mean_changeTS_w52

Plot Type (Compare Plot: y-axis / Bubble Plot: x-axis)

 linear logarithmic

Range (Compare Plot: y-axis / Bubble Plot: x-axis)



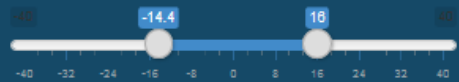
Second Target variable

Mean_changeTS_w12

Plot Type (Compare Plot: y-axis / Bubble Plot: y-axis)

 linear logarithmic

Y Range (Compare Plot: y-axis / Bubble Plot: y-axis)



Reference variable

Number.of.Subjects

Subgroup Filter

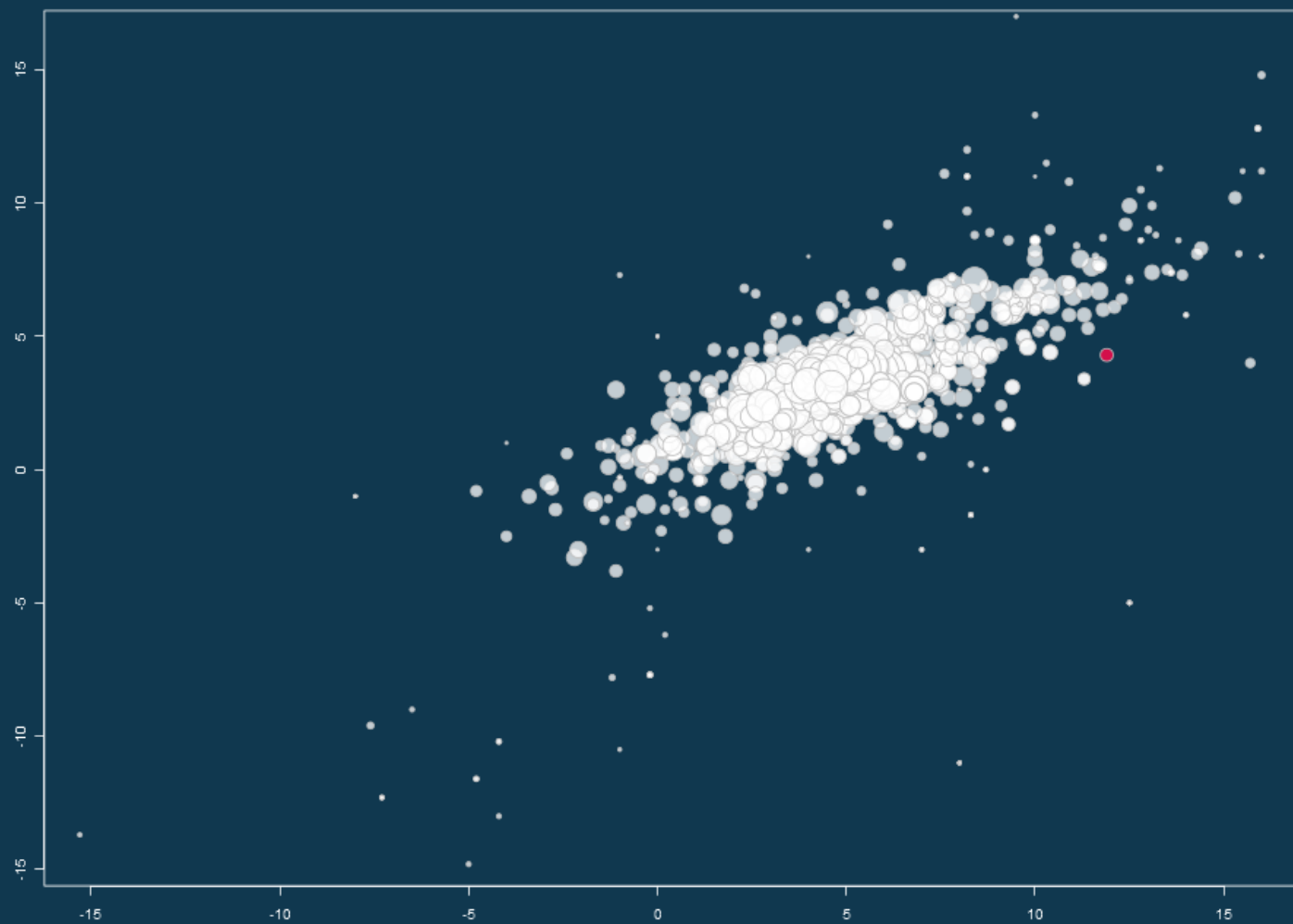
no selection

Subgroup level(s)



Compare

Bubble plot



First subgroup variable (x)

EQ5D_score

Second subgroup variable (y)

Region

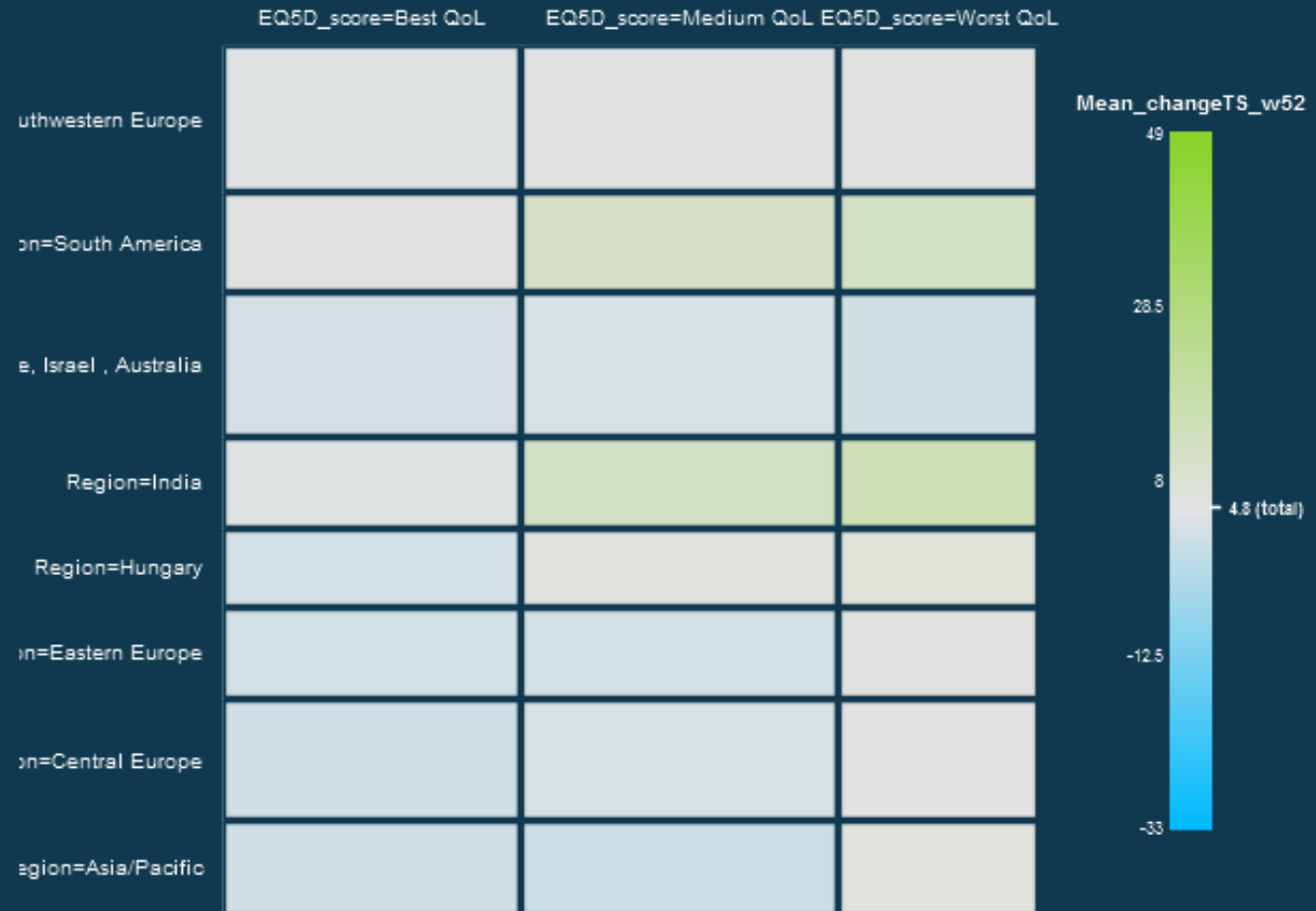
Third subgroup variable (y2)

no selection

Reference variable (color)

Mean_changeTS_w52

Plot Type

 linear logarithmic

More information in the article

Biostatistics: Original Research

A Systematic Approach for Post Hoc Subgroup Analyses With Applications in Clinical Case Studies

Christoph Muysers, MSc¹, Alex Dmitrienko, PhD²,
Hermann Kulmann, PhD¹, Bodo Kirsch, MSc¹, Susanne Lippert, MSc¹,
Thomas Schmelter, PhD¹, Anke Schulz, MSc¹, Nicole Mentenich, MSc¹,
Heinz Schmitz, MD, PhD³, Matthias Schaefer, MD, PhD³,
Gerold Meinhardt, MD, PhD⁴, Thomas Keil, MD, PhD⁵,
and Stephanie Roll, PhD⁵

Abstract

Background: The analysis of subgroups in clinical trials is essential to assess differences in treatment effects for distinct patient clusters, that is, to detect patients with greater treatment benefit or patients where the treatment seems to be ineffective. **Methods:** The software application *subscreen* (R package) has been developed to analyze the population of clinical trials in minute detail. The aim was to efficiently calculate point estimates (eg, hazard ratios) for multiple subgroups to identify groups that potentially differ from the overall trial result. The approach intentionally avoids inferential statistics such as *P* values or confidence intervals but intends to encourage discussions enriched with external evidence (eg, from other studies) about the exploratory results, which can be accompanied by further statistical methods in subsequent analyses. The *subscreen* application was applied to 2 clinical study data sets and used in a simulation study to demonstrate its usefulness. **Results:** The visualization of numerous combined subgroups illustrates the homogeneity or heterogeneity of potentially all subgroup estimates with the overall result. **With this, the application leads to more targeted planning of future trials. Conclusion:** This described approach supports the

DIA

Therapeutic Innovation
& Regulatory Science
1-12
© The Author(s) 2019
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/2168479019853782
tirs.sagepub.com

THERAPEUTIC
INNOVATION
& REGULATORY
SCIENCE

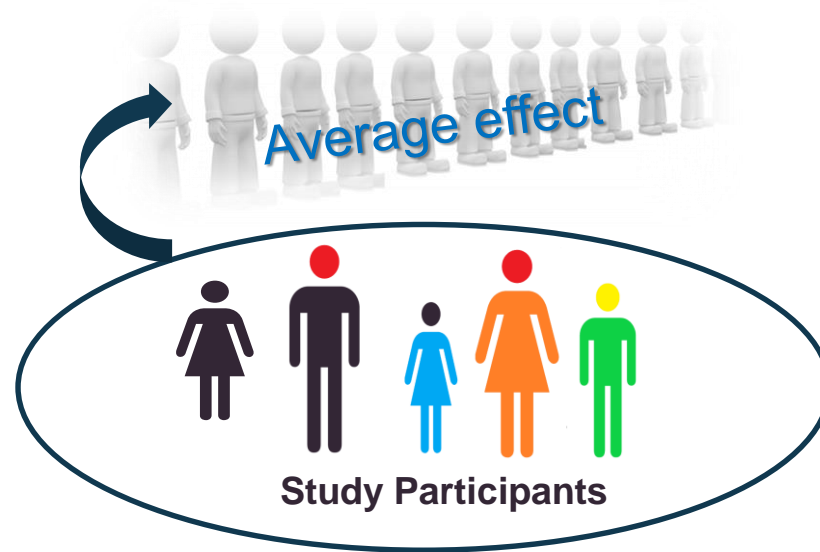
And on CRAN



subscreen (Vers 2.0.1):
Systematic Screening of
Study Data for Subgroup Effects



In the past we draw a single conclusion based on a heterogenous study population





Imagine a tool (such as 'subscreen') that allows efficient subgroup analyses for many stakeholders interactively

Sponsors



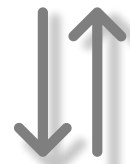
dedicated effect estimation



Health Authorities



Contributing in- and external Databases



HTAs



Investigator Researcher





Contact



e-mail: Bodo.Kirsch@bayer.com

R-PACKAGE: <https://cran.r-project.org/web/packages/subscreen/index.html>



Thank you!





at a glance

exemplary data

Select endpoint of interest

Filter by subgroup factor of interest

Number of non-empty subgroups

Click on a subgroup for tabulation

The reference line shows the overall results

Each subgroup is represented as single dot

Choose scale type and range

Number of subgroup factor combinations

Subscreen Explorer | Subscreen Comparer | Subscreen Mosaic

Variable Options | Importance Tab | Display Options | Colour Options

Target variable: hr

Reference variable: N.of.subjects

Subgroup Filter: no selection

Subgroup level(s): 1, 3

Plot Type: linear (selected), logarithmic

Y Range: 0.3 to 1.18

Help text: [Icons]

1 to 3-Factorial Subgroups (251)

0.52

0 % 8 % 16 % 25 % 33 % 41 % 50 % 58 % 66 % 75 % 83 %

0 100 200 300 400 500 600 700 800 900 1000

Selected Subgroups | Filtered Subgroups | Parent Subgroups | Memorized Subgroups

Copy | Print | Show 6 rows | Column visibility

Table of Selected Subgroups

Memorize	SGID	N.of.subjects	hr	nfactors	bi11
All	All	All	All	All	All
Memorize	2	200	0.706432308925214	1	B

Showing 1 to 1 of 1 entries

Previous | 1 | Next