Identification of subgroups of subjects with enhanced treatment effect

Amelie Rodrigues

Supervised by Emmanuelle Boëlle-Le Corfec, Christophe Candelas and Marie-Karelle Riviere

November 28, 2016

SANOFI

Overview

Plan

Phase IV Real World Pragmatic Trial

- **Design :** Open label, randomized, two treatment arm parallel groups
- In case of superiority of new treatment not achieved on overall population : an exploratory analysis **to identify subgroups with a greater treatment effect from the new therapy** is planned

$\Rightarrow$ Subgroup Identification based on Differential Effect Search (SIDES)

Plan

Methodology

Goal : Build a collection of subgroups of potential interest where the differential treatment effect is maximized between two treatment arms

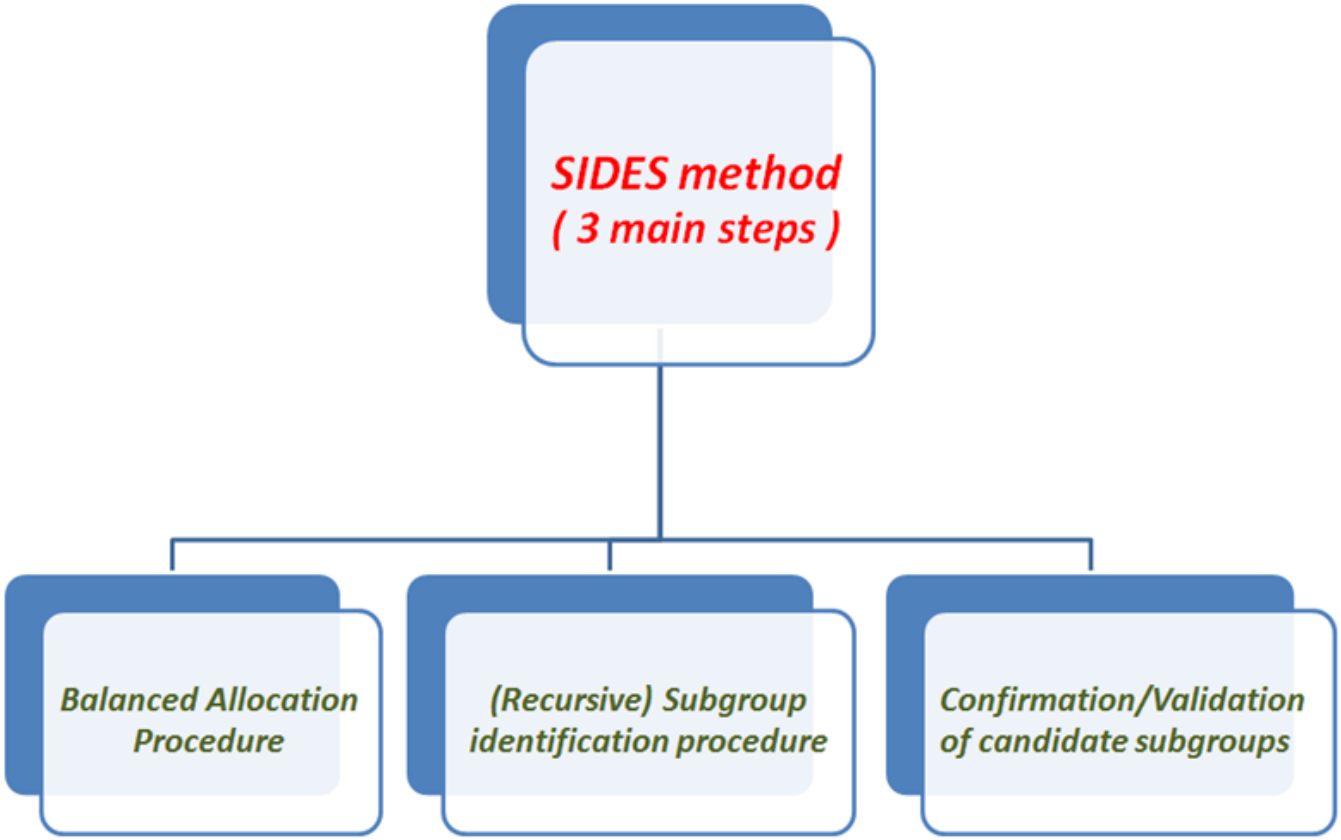


Figure: Steps of the SIDES method

Subgroup identification procedure

Three parameters to define:

Recommendations on parameters:

- **L** : maximum number of covariates defining a subgroup (the recommended value is 3)
- **S** : minimum subgroup size (the recommended value is determined based on clinical considerations)
- **M** : maximum number of best promising subgroups at each step of the algorithm (the recommended value is 5)

Splits from the parent group

Make **all possible splits** of parent group **in two subgroups** (at level 0 : start with the entire training data set)

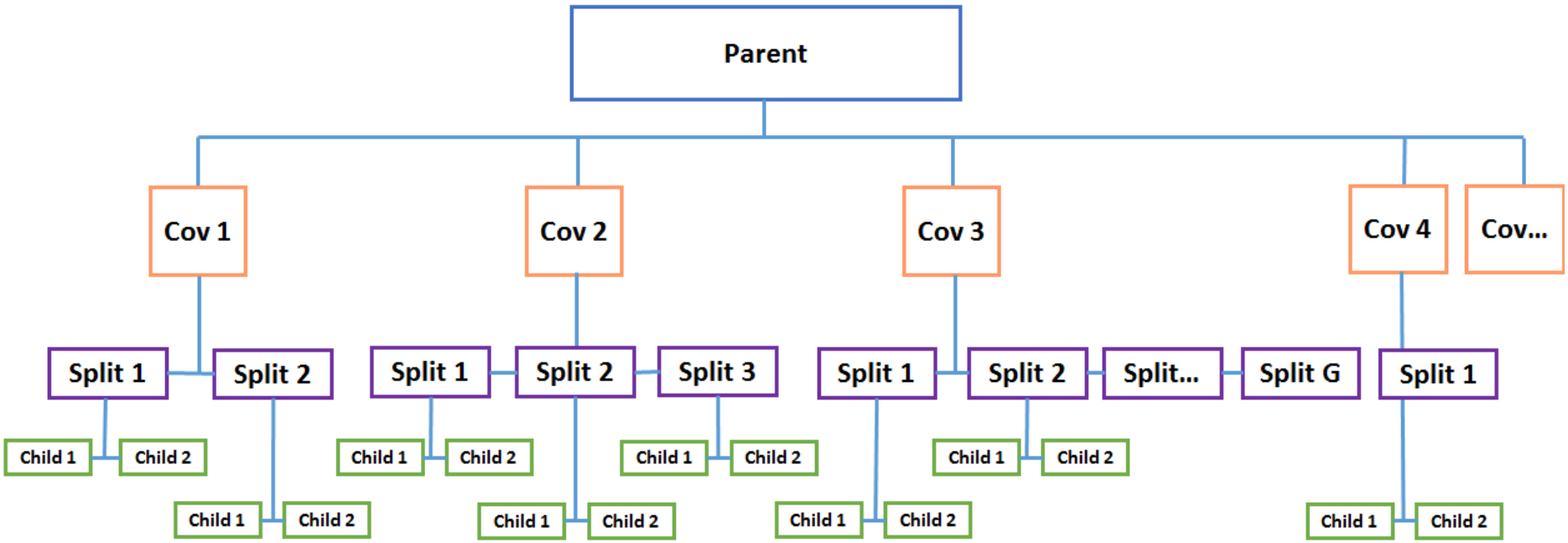
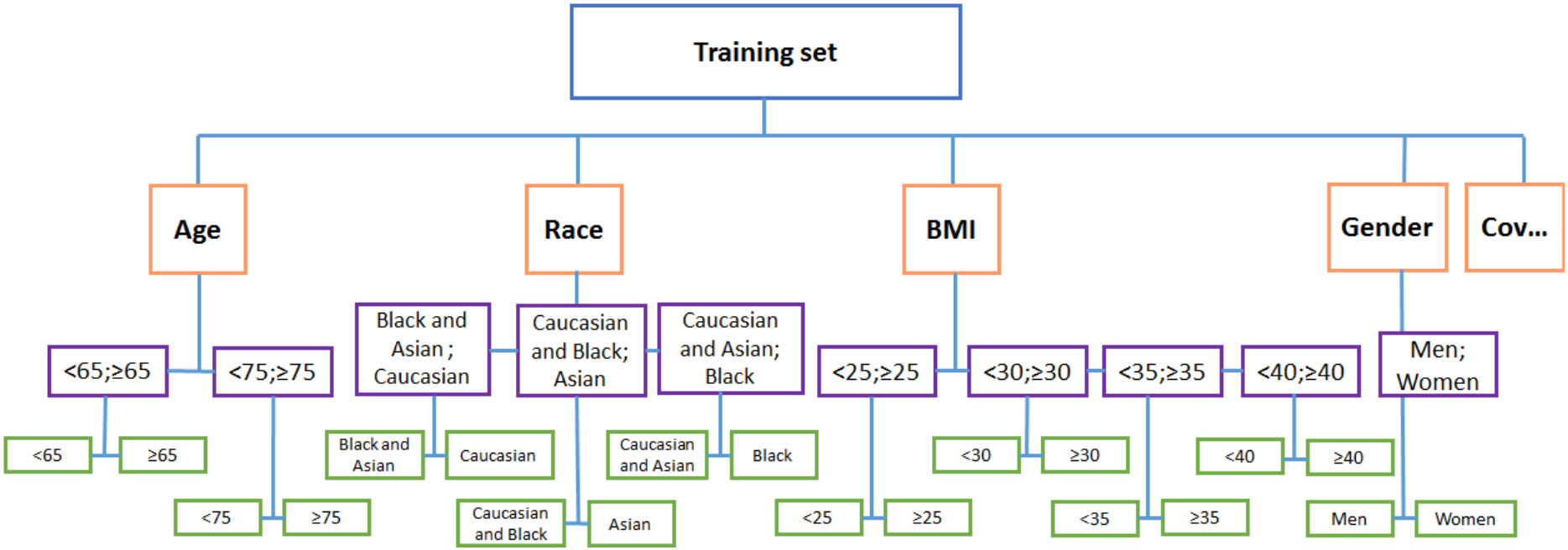


Figure: Split of Parent group

Example: Covariate with 3 levels $\{A, B, C\}$
Ordinal: $(\{A\}, \{BC\})$ and $(\{AB\}, \{C\})$
Nominal: $(\{A\}, \{BC\})$, $(\{AB\}, \{C\})$ and $(\{AC\}, \{B\})$

Example of split of the parent group



If $\#(Child_i$ with maximal treatment effect$) \leq S$: the split is discarded

Research of promising groups

$\Rightarrow$ For each split:

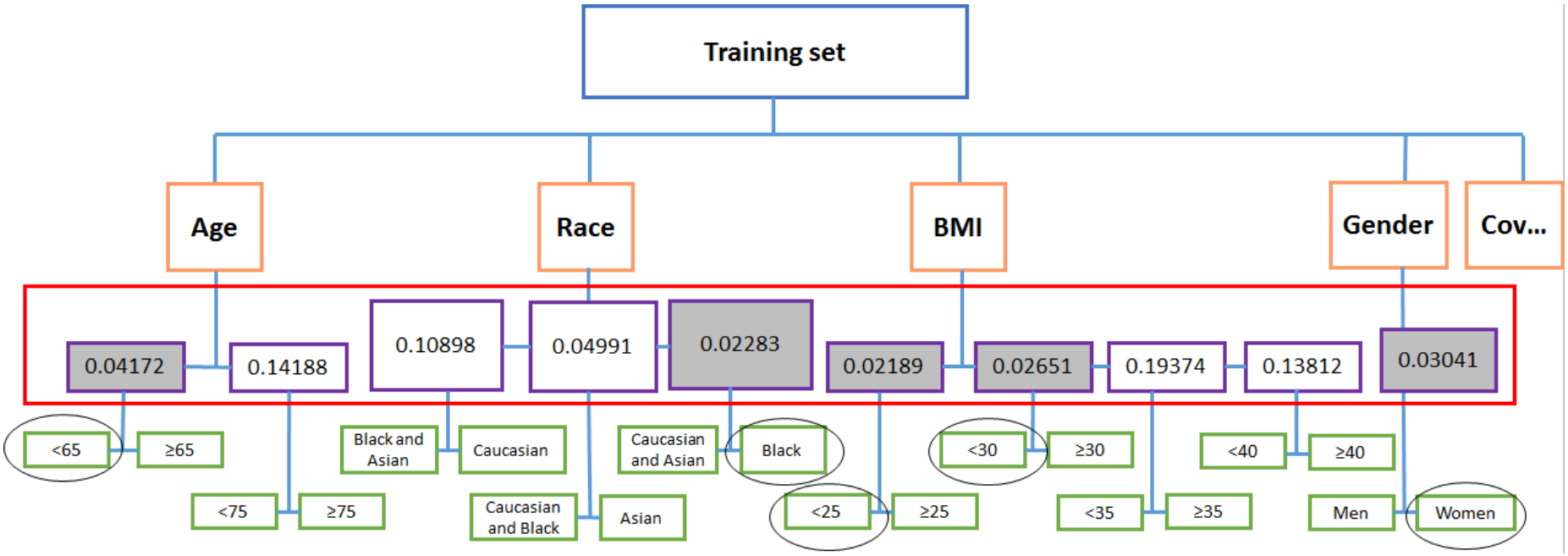
- Calculate the splitting criterion (p-value scale) $\Rightarrow$ Maximizing the differential effect between the two child subgroups :

$$p_1 = 2\left[1 - \Phi\left(\frac{|Z_{E1} - Z_{E2}|}{\sqrt{2}}\right)\right]$$

  where Z_{E1} and Z_{E2} are efficacy tests statistic in child 1 and child 2

- Adjust the p-value for covariates with more than 2 levels (more than 1 possible split) with Sidak-based multiplicity adjustment

Example of selection of the promising groups



- Select the best M splits in terms of splitting criterion
- For each split, select the subgroup child with the best treatment effect

Definition of new parent groups

Parent group : Subset of the database restricted according to covariates selected in previous levels

The continuation criterion to become a parent group:

$$P_c \leq \gamma P_p$$

where :
γ is the vector of the relative improvement parameters : $0 < \gamma_i \leq 1$
P_c is the p-value of the treatment effect in child subgroup
P_p is the p-value of the treatment effect in parent group

Identification of candidate groups

Candidate group : Subset of potential responders with enhanced treatment effect

The selection criterion to become a candidate group:

$$P_c \leq \nu$$

where ν is the adjusted significant threshold

- Controlling the overall Type I error rate in a weak sense using a resampling method
- Generating 1000 data sets under the global null hypothesis : no interaction between covariates and the treatment group
- Computing the proportion of time where at least one subgroup is wrongly returned

Example of new parent groups and candidate groups

$\Rightarrow \gamma_1 = 1$ and $\nu = 0.01$

Subgroup	Size	Splitting criterion	Treatment effect p-value*
Parent group : Training set	1500		0.06168
Split 1 by Age (<65 vs ≥65)		0.04172	
Child subgroup 1A : Age <65	480		0.04744
Child subgroup 1B : Age ≥65	1020		0.08222
Split 2 by Race ('Caucasian and Asian vs 'Black')		0.02283	
Child subgroup 2A : Race = 'Caucasian and Asian	1125		0.22920
Child subgroup 2B : Race = 'Black'	375		0.09766
Split 3 by BMI (<25 vs ≥25)		0.02189	
Child subgroup 3A : BMI <25	152		0.06888
Child subgroup 3B : BMI ≥25	1348		0.51332
Split 4 by BMI (<30 vs ≥30)		0.02651	
Child subgroup 4A : BMI <30	454		0.00174
Child subgroup 4B : BMI ≥30	1046		0.47561
Split 5 by Gender ('Men' vs 'Women')		0.03041	
Child subgroup 5A : Gender = 'Men'	702		0.16743
Child subgroup 5B : Gender = 'Women'	798		0.07587

*fictitious values

- Two new parent groups $\{Age < 65\}$ and $\{BMI < 30\}$
- One candidate group $\{BMI < 30\}$

Recursive partitioning

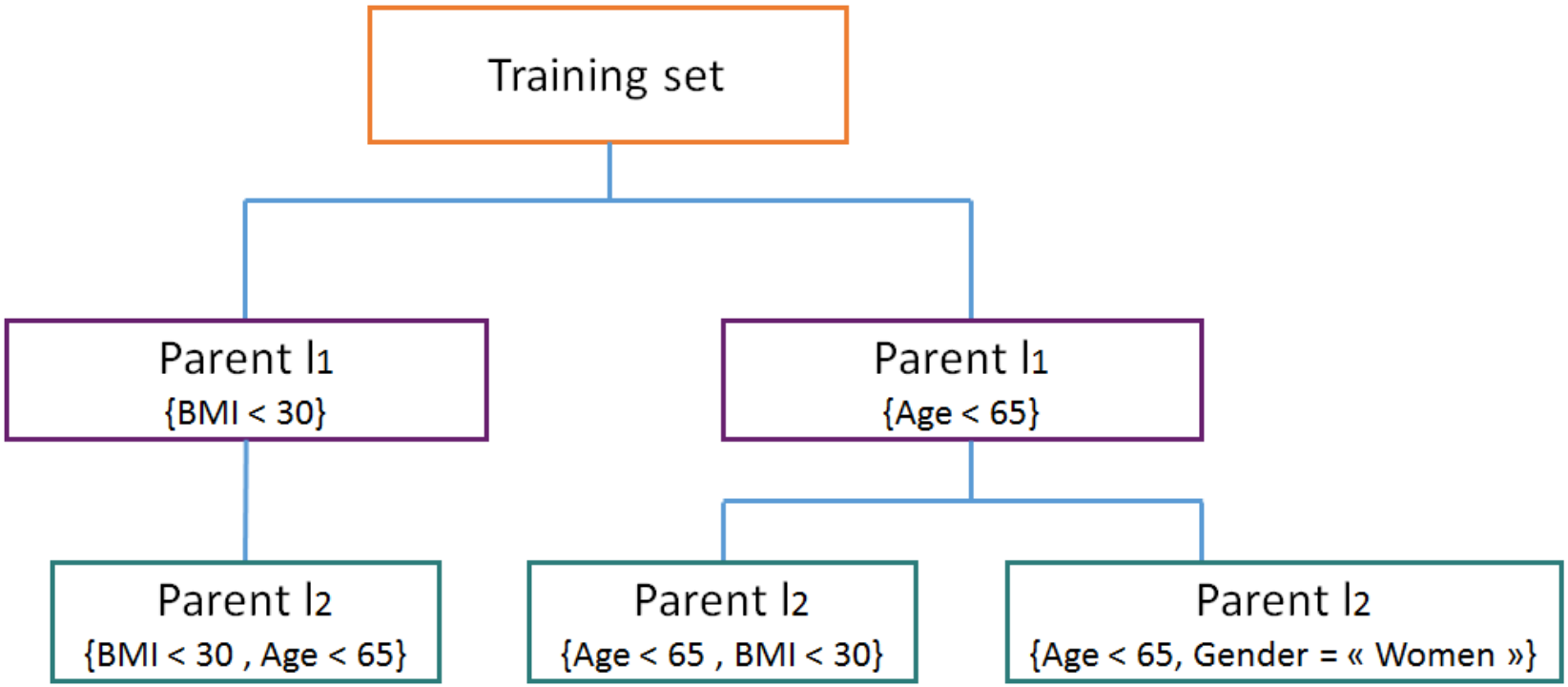


Figure: Example of different levels

The algorithm stops when :

- There is no new parent group
- Number of levels $l > L$

Plan

Simulation : Evaluation of performances

Goal : Assess the method SIDES operating characteristics under 8 scenarios

1000 Data generations / 20 cuts training-validation (0.7,0.3)

- n=1500
- Z : Binary treatment
- X=$(x_1, ..., x_{10})$: Covariates of different types
- Y : Binary response variable based on 8 scenarios

- In each scenario there are between 0 and 3 predictive covariates
- Treatment effect non significant in the overall population
- Target differential treatment effect : 20%

Example of a scenario with two predictive covariates

M_1: $P(Y = 1|X) = 0.3 + 0.25 * I_{(Z=1)} * I_{(X_1=1)} + 0.2 * I_{(Z=1)} * I_{(X_2 \leq 1)}$

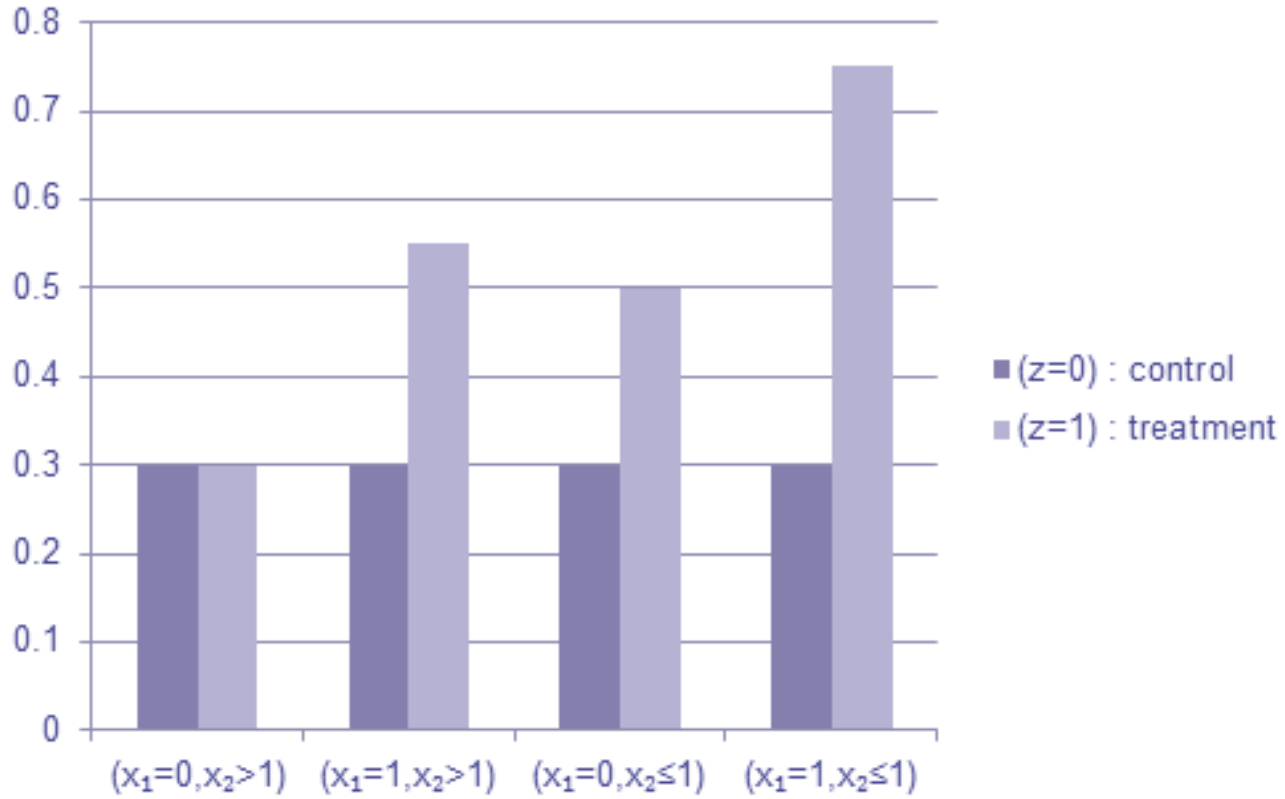


Figure: Probability of expected response under scenario one

Overall results

- The Best subgroup (PVB) : includes all existing predictive covariates
- A large subset (PVLS) : subset of the Best subgroup
- No subgroup (NS)

Table of percentage of simulations where subgroups are returned :

	$\alpha = 0.05$			$\alpha = 0.10$			no VS $\alpha = 0.10$		
	NS	PVB	PVLS	NS	PVB	PVLS	NS	PVB	PVLS
SC1	26.8	42.2	59.0	13.6	**55.1**	**73.3**	1.8	**84.6**	94.0
SC2	24.0	41.8	75.9	10.4	**55.6**	**89.6**	0.1	**87.6**	99.9
SC3	61.9	10.4	10.4	45.4	16.7	16.7	15.0	**38.7**	38.7
SC4	33.1	26.6	26.6	18.5	37.2	37.2	2.8	**63.5**	63.5
SC5	22.9	54.8	71.9	11.1	**68.5**	**83.0**	1.1	**91.7**	97.6
SC6	99.9	-	-	**99.5**	-	-	**89.7**	-	-
SC7	55.3	18.9	35.3	37.1	**29.7**	**49.0**	7.2	**58.4**	80.8
SC8	99.6	-	-	**99.1**	-	-	**89.6**	-	-

$\Rightarrow \alpha = 0.10$ without Validation Set is recommended

Performance of SIDES method

$\Rightarrow$ Advantages :

- Good performances
  - High % to select a subgroup with the highest treatment effect
  - High % stop when there are not subgroup
- Easy to interpret clinically
- Controlling the overall Type I error rate in a weak sense

$\Rightarrow$ Disadvantages :

- Highly time consuming when there are validation sets
- When there are prognostic covariates : more difficulties to return the best subgroup
- When there are too many predictive covariates : need a large database

Plan

Conclusion

- Two different methods were investigated and assessed with simulation study
- Recommend SIDES method :
  - Clinical interpretation
  - Controlling the overall Type I error rate in a weak sense
- Package of SIDES developed internally in Sanofi under R software and will be used in clinical studies
- An extension : SIDEScreen selects covariates to keep in the dataset

Thank you for your attention !

BACK-UP SLIDES

Phase IV Real World Pragmatic Trial

- **Design :** Open label, randomized, two treatment arm parallel groups
- **Population :** Patients with type 2 diabetes mellitus
- **Sample size :** 3270 patients (1635 per treatment arm)
- **Aim :** Demonstrate clinical effectiveness of a new diabetes therapy compared to a standard of care diabetes therapy
- **Primary endpoint :** Binary endpoint at 6 month
- **Sample size assumption :** expecting a modest difference between the two therapies

$\Rightarrow$ In case of superiority of new treatment not achieved on overall population : an exploratory analysis **to identify subgroups with a greater treatment effect from the new therapy** is planned
$\Rightarrow$ Two methods investigated and assessed

Balanced Allocation Procedure

In the remaining 1-c fraction of the global data set calculate :

- The proportion of patients :

$$f_{ijl}^{h} = \frac{n_{ijl} + I_{(i=h)}}{n_i + I_{(i=h)}}$$

- The imbalanced score :

$$d_{jh} = \max_i f_{ijl}^{h} - \min_i f_{ijl}^{h}$$

- The total imbalance score :

$$d_h = \sum_j d_{jh}$$

- The probabilities constructed to be inversely proportional to d_h :

$$p_h = \frac{1}{H-1}\left(1 - \frac{d_h}{\sum_{h=1}^{H} d_h}\right)$$

Splitting criterion

1. Maximizing the differential effect between the two child subgroups (i.e. identify a treatment effect that differs across subgroups):

$$p_1 = 2\left[1 - \Phi\left(\frac{|Z_{E1} - Z_{E2}|}{\sqrt{2}}\right)\right]$$

2. Maximizing the treatment effect in at least one of the two child subgroups (i.e. identify a large treatment effect relative to the overall population):

$$p_2 = 2\min(1 - \Phi(Z_{E1}), 1 - \Phi(Z_{E2}))$$

3. Combination of the two criteria :

$$p_3 = \max(p_1, p_2)$$

Sidak-based multiplicity adjustment

$$q_i = 1 - (1 - p_i)^{G^*}$$

Where $G^* = G^{1-r}$ with G the total number of possible splits for the covariate and r is the average pairwise correlation across the $\frac{G(G-1)}{2}$ test statistics

Identification of candidate groups

Candidate group : Subset of potential responders with enhanced treatment effect

The selection criterion to become a candidate group:

$$P_c \leq \nu$$

where ν is the adjusted significant threshold

- Controlling the overall Type I error rate in a weak sense using a resampling method
- Generating 1000 data sets under the global null hypothesis : no interaction between covariates and the treatment group
- Computing the proportion of time where at least one subgroup is wrongly returned

ν	0.001	0.005	**0.01**	0.02	0.03	0.04	0.05	...
Type I error	0.028	0.067	**0.1**	0.124	0.182	0.36	0.41	...

Table: Example of the selection criterion

Validation part

- All candidate subgroups are evaluated in each validation set
- Confirm the candidate group when the efficacy criterion is satisfied in every set

Method 2 : Predictive Enrichment Procedure to identify potential responders

Goal : Identify an enrichable subpopulation

$\Rightarrow$ An enrichable subpopulation : subgroup whose patients respond better to a given treatment than the rest of population

$\Rightarrow$ Three steps : Create, Evaluate and Validate this enrichable population with two data sets

Figure: Steps of the PEP method

Step 1 : Creation of a treatment difference score

- Regress two logistic models (one for each treatment group) : vectors of parameter estimates of covariates $\hat{\beta}_0$ (control group) and $\hat{\beta}_1$ (test group)
- Calculate predicted probabilities for each individual i:

$$P_{0(i)} = P(\widehat{Y = 1|X} = x_i) = \frac{e^{x_i\hat{\beta}_0}}{1 + e^{x_i\hat{\beta}_0}}$$

$$P_{1(i)} = P(\widehat{Y = 1|X} = x_i) = \frac{e^{x_i\hat{\beta}_1}}{1 + e^{x_i\hat{\beta}_1}}$$

- Calculate the treatment difference : $D_i(X = x_i) = P_1 - P_0$

Step 2 : Choosing an enrichable subpopulation

- Decile of the vector of the treatment difference D(X) are calculated

Decile	0.1	0.2	0.3	**0.4**	0.5	0.6	0.7	0.8	0.9
D(X)	0.11	0.16	0.21	0.25	0.29	0.33	0.37	0.39	0.43
Z-Stat	0.9	1.8	1.2	**3**	2.1	1	0	-1	-1.4

Table: Example of decile of D(X)

- For each decile
  - A subgroup of population is chosen : subgroup with D(X) higher than the decile
  - Estimate the treatment effect in this subgroup : Z-Stat
- The optimal threshold value q0 is determined across all subgroups where Z-test statistic is the highest
- The rank scoring of the enrichable subgroup : $Q(.) \geq 1 - q_0$

Step 3 : Validation of the selected enrichable subgroup

- Estimate predicted probabilities in the validation data set from vectors $\hat{\beta}_0$ and $\hat{\beta}_1 \Rightarrow P_1$ and P_0
- Calculate D(X) with the new values of P_1 and P_0
- Define the enrichable subpopulation from q_0 previously fixed on the training set
- Calculate the treatment effect in the enrichable subpopulation
- Validate if the enrichable subgroup responds favorably to the new therapy compared to control

Overall PEP results

	SC1	SC2	SC3	SC4	SC5	SC6	SC7	SC8
Se	0.81	0.93	0.9	0.91	0.89	NA	0.9	NA
Sp	0.93	1	0.78	0.8	0.87	NA	0.82	NA
PVP	0.81	1	0.41	0.45	0.73	NA	0.63	NA
PVN	0.93	0.94	0.97	0.98	0.96	NA	0.96	NA
ACC	0.87	0.97	0.78	0.82	0.87	NA	0.8	NA
Empty	5	30	40	0	5	100	25	100

Table: Results of simulation for the PEP method

Evaluation of PEP performances

- Advantages of PEP :
  - Good performances from simulation studies
  - Shortly time consuming

- Disadvantages of PEP :
  - When there are too many predictive covariates : seems more difficult to discard non responders
  - No control for inflated Type I error
  - Difficulties to interpret a score clinically : How to know if a new patient is included in the enrichable subgroup?

Presentation of Edition studies

Non-inferiority Phase III

- **Design :** multicenter, randomized, open label, parallel group studies
- **Pool of 3 studies :** 2260 Patients with T2DM
- **Aim :** Compare the efficacy of a new diabetes therapy with a standard of care in terms of HbA1C change between baseline and month 6
- **Results :**
  - Similar efficacy profile in terms of HbA1C change between baseline and month 6 for both basal insulins
  - Benefit of the new basal insulin in terms of hypoglycemia incidence

Application of SIDES method

Goal : Assess whether some patients subgroups may benefit more from the new basal insulin in terms of hypoglycemia applying SIDES method

Endpoint : At least one severe and/or symptomatic documented by plasma glucose hypoglycemia ($\leq$70 mg/dl) during the 6-month study period

Result : No subgroup returned when $\alpha = 0.10$ without validation set (consistent with other analysis conducted)