Eric Derobert - Julie Perez

9/7/2011

Emerging Topics in Pharmaceutical Statistics
Basel Biometric Society Seminar

A Parametrized Strategy of Gatekeeping, Keeping Untouched the Probability of Having at Least One Significant Result

Analysis of Primary and Secondary Endpoints in a Many-To-One Comparison of Two Treatments with a Control

SANOFI

Presentation summary

- Multiplicity issue with a Many-to-One Comparison of Two Treatments with a Control, when a Primary and a Secondary Endpoint are analyzed
- A Gatekeeping Procedure which do not decrease the Probability of Success (i.e. at least a Significant Result)
- Tuning up the Gatekeeping Procedure
- Modeling the Gatekeeping Parameter with a Logistic Regression
- Conclusion and Possible Extension

SANOFI

Multiplicity issue with a Many-to-One Comparison of Two Treatments with a Control, when a Primary and a Secondary Endpoint are analyzed

SANOFI

A common multiplicity problem (1/3)

- Let's consider two treatments compared to a control, with a primary and a secondary endpoints.

Primary	H_1 (Treatment 1)	H_2 (Treatment 2)
Secondary	H_3 (Treatment 1)	H_4 (Treatment 2)

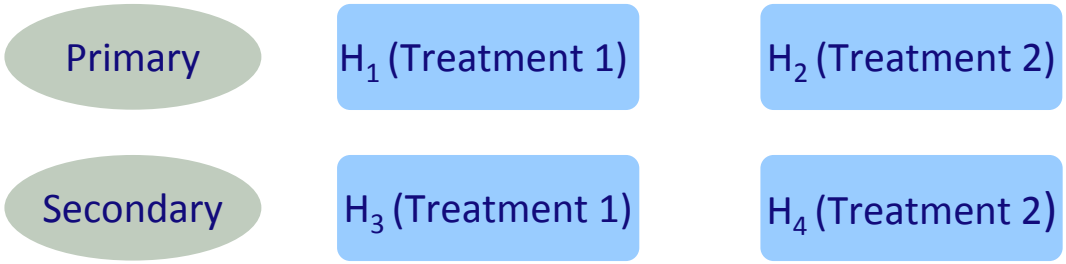
- Treatments 1 and 2 can be different drugs or two doses (or galenic forms…) of the same drug

SANOFI

A common multiplicity problem (2/3)

- It is well known that if

  a closed simultaneous test of the two treatments on the primary endpoint performed at a level α,



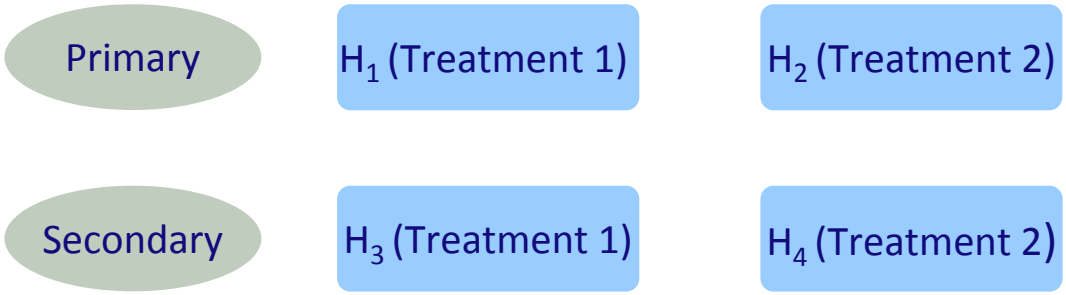
  is followed by the"natural" hierarchical rule: **the null hypothesis on the secondary endpoint for one treatment is tested as soon as the hypothesis on the primary endpoint has been rejected for the same treatment**

  controls multiplicity in a **weak** sense

- (from expected correlations between endpoints, from the expected dose-response curve (if Treatments 1 and 2 are two doses of the same drug), it could be argued that **weak control is ok**; but **this is not the point of view of Health Authorities**)

SANOFI

A common multiplicity problem (3/3)



- Dunnett test is performed on H_1 and H_2 at the global one-sided level 0.025.
- If H_2 is the only false null hypothesis and if H_2 is rejected at the Dunnett nominal level

  (α_D=0.0135 – balanced groups), then
  - H_1 is tested at the level 0.025 (Dunnett closed testing procedure)
  - H_4 is tested at some supplementary level
- Then the probability of rejecting H_1 or H_4 (true null hypotheses) is greater than 0.025

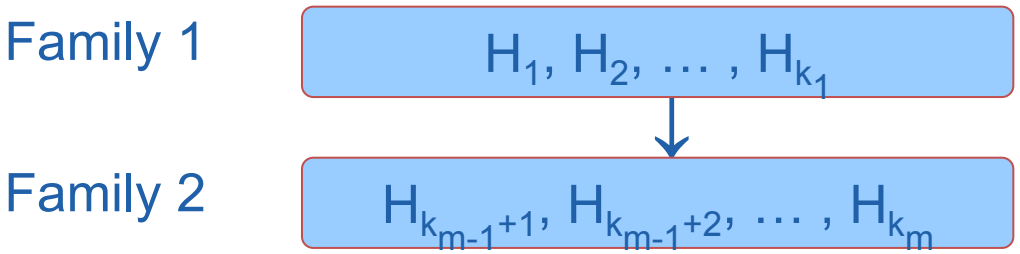
SANOFI

A Gatekeeping Procedure which do not decrease the Probability of Success (i.e. at least a Significant Result)

SANOFI

Using a Gatekeeping Procedure (1/2)

- **Gatekeeping:** A procedure which permits to deal with the multiplicity issue using hierarchical testing procedure **(with strong control)**
- Testing m (here m=2) families of null hypotheses
  - Family 1 serves as a gatekeeper for Family 2
  - The gatekeeper family 1 is tested; with parallel gatekeeping, at least one significant result must be observed in one family (ie at least one hypothesis must be rejected in the family) to proceed to the next family)
  - The family 2 is tested only if the gatekeeper has been successfully passed

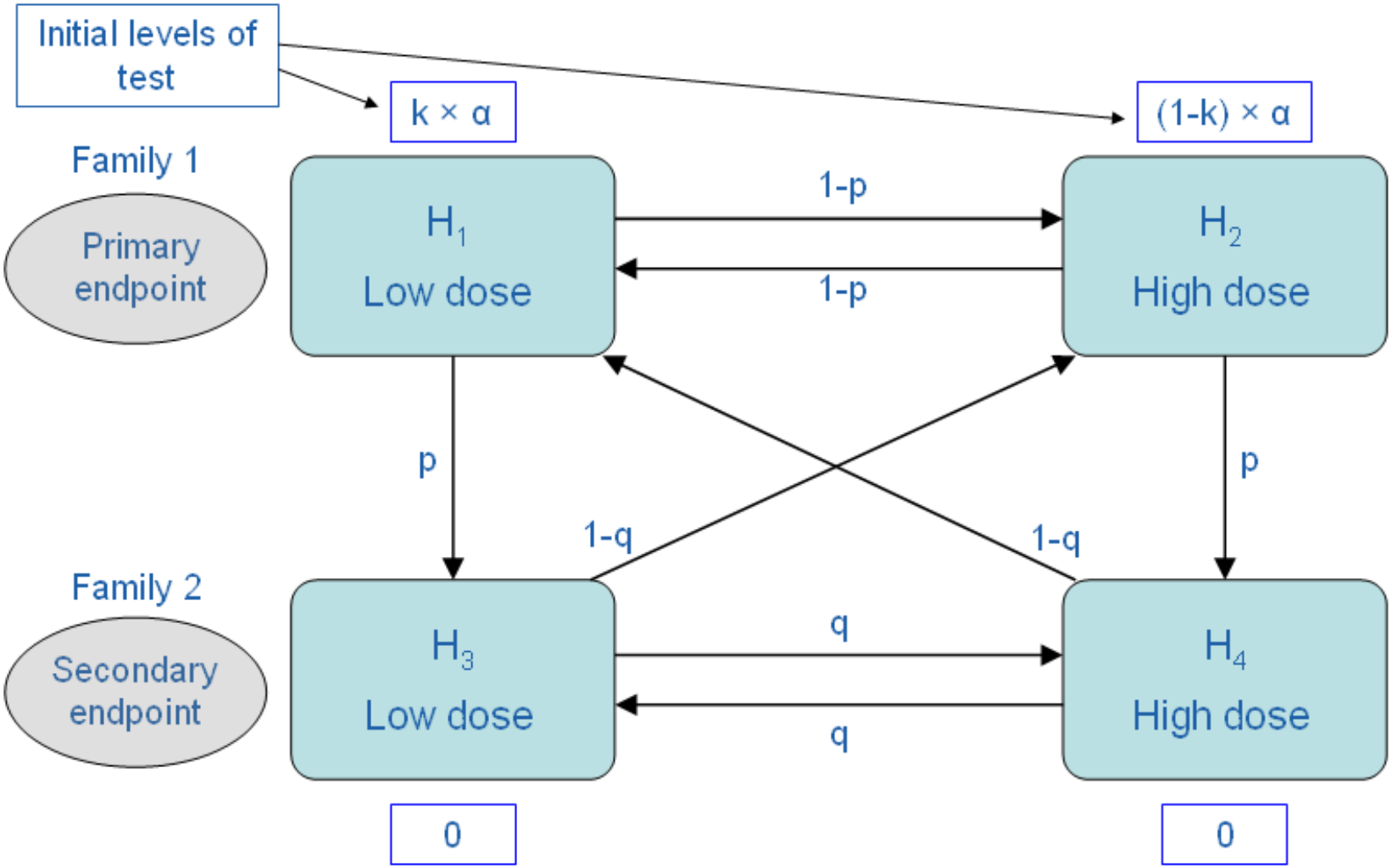


Using a Gatekeeping Procedure (2/2)

- On this two-family testing problem
- With 2 null hypotheses in each family
  - Testing two treatments (e.g. Low dose and High dose) versus a control
  - On a primary endpoint (family 1) and a secondary endpoint (family 2)
- With a Gatekeeping procedure
  - We test the secondary endpoint only if some significant results have been obtained on the primary endpoint
- Using a graphical approach (Bretz *et al*., 2009: "*A graphical approach to sequentially rejective multiple test procedures*")
- From this graphical approach, we will define an associated closed testing strategy, and freely choose related test procedures

SANOFI

Gatekeeping scenario depending on 3 parameters



SANOFI

Two Constraints for our Gatekeeping strategy

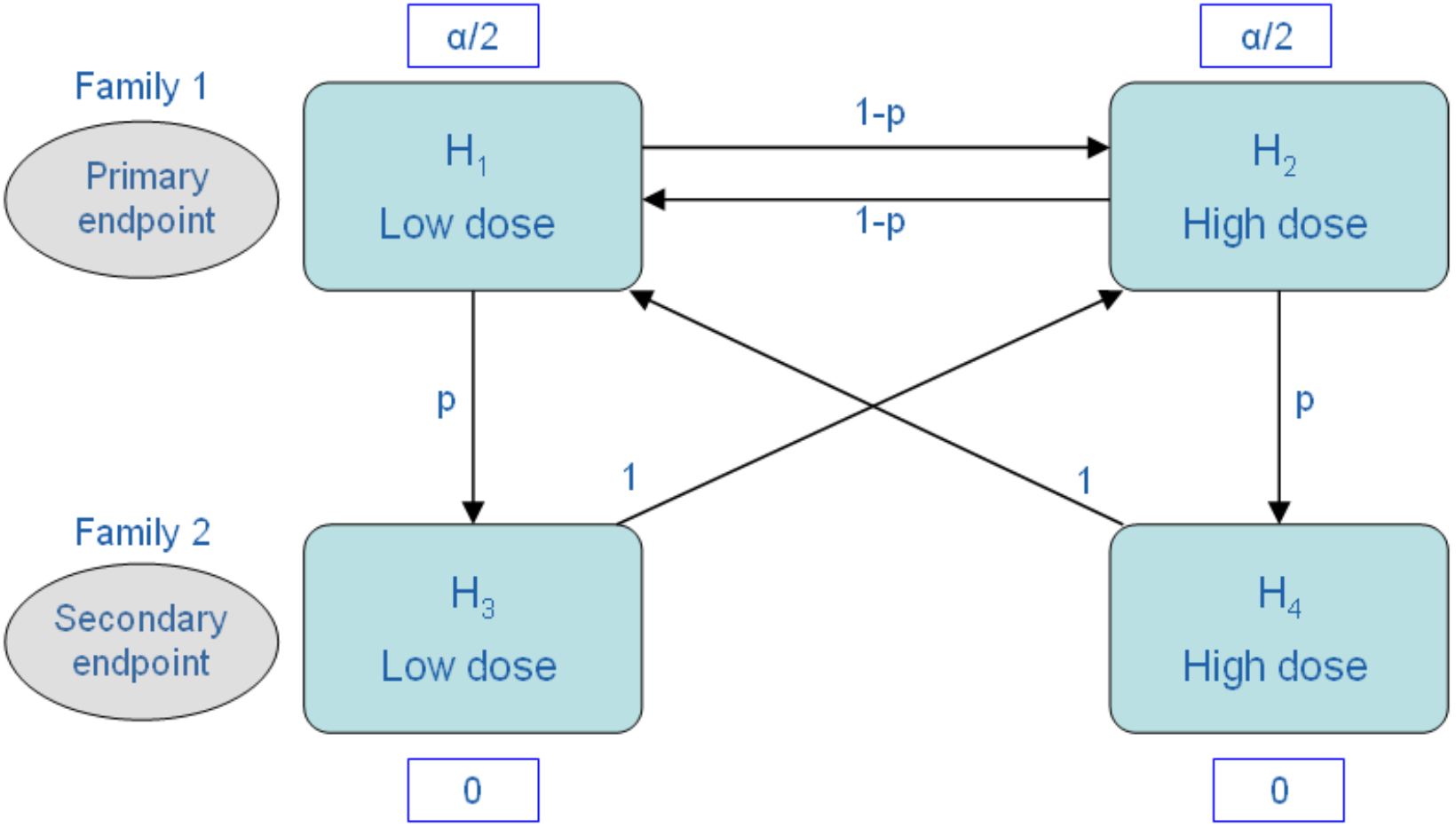
- N°1: Constraint CPE: The probability of rejecting at least one null hypothesis on the primary endpoint remains unchanged
  - Some gatekeeping strategies do not satisfy this condition
    - Therefore Gatekeeping reduces the power of the test
    - Main reason why Gatekeeping is not so often used

- N°2: Constraint CSE: The rejection of a null hypothesis for one dose on the secondary endpoint without having rejected the hypothesis on the primary endpoint is useless

SANOFI

Simplified scenario

- k = ½, q = 0 (CSE is met) -> only depends on p ($0 \leq p \leq 1$)



Link between the closed testing procedure and the graphical approach

Intersection hypotheses	Respective weighting for each hypothesis				Performed test
	H_1	H_2	H_3	H_4	
$H_1 \cap H_2 \cap H_3 \cap H_4$	$\frac{\alpha}{2}$	$\frac{\alpha}{2}$	0	0	DUNNETT
$H_1 \cap H_2 \cap H_3$	$\frac{\alpha}{2}$	$\frac{\alpha}{2}$	0	0	DUNNETT
$H_1 \cap H_2 \cap H_4$	$\frac{\alpha}{2}$	$\frac{\alpha}{2}$	0	0	DUNNETT
$H_1 \cap H_2$	$\frac{\alpha}{2}$	$\frac{\alpha}{2}$	0	0	DUNNETT
$H_1 \cap H_3 \cap H_4$	$\alpha - p \times \frac{\alpha}{2}$	0	0	$p \times \frac{\alpha}{2}$	Weighted SIMES
$H_1 \cap H_3$	α	0	0	0	SINGLE TEST
$H_1 \cap H_4$	$\alpha - p \times \frac{\alpha}{2}$	0	0	$p \times \frac{\alpha}{2}$	Weighted SIMES
H_1	α	0	0	0	SINGLE TEST
$H_2 \cap H_3 \cap H_4$	0	$\alpha - p \times \frac{\alpha}{2}$	$p \times \frac{\alpha}{2}$	0	Weighted SIMES
$H_2 \cap H_3$	0	$\alpha - p \times \frac{\alpha}{2}$	$p \times \frac{\alpha}{2}$	0	Weighted SIMES
$H_2 \cap H_4$	0	α	0	0	SINGLE TEST
H_2	0	α	0	0	SINGLE TEST
$H_3 \cap H_4$	0	0	$\frac{\alpha}{2}$	$\frac{\alpha}{2}$	DUNNETT
H_3	0	0	α	0	SINGLE TEST
H_4	0	0	0	α	SINGLE TEST

SANOFI

How to satisfy the Constraint CPE ?

- Constraint CPE: Using Dunnett test for hypotheses concerning a common endpoint and Simes test for hypotheses on different endpoints
  - Condition: if min(p1,p2) is significant with the Dunnett test at the nominal level $\alpha_{D)}$, then min(p1,p2) is always significant with the Gatekeeping procedure
  - i.e. if $\min(p1, p2) \leq \alpha_D$ then $\min(p1, p2) \leq \alpha - p \times \alpha / 2$
  - Finally, this is possible provided that $\alpha_D \leq \alpha - p \times \alpha / 2 \iff p \leq 2 \times (\alpha - \alpha_D) / \alpha$
  - Here is the definition of the value $\max(p)=p_{Dunnett}$: $p_{Dunnett} = 2 \times (\alpha - \alpha_D) / \alpha$
  - If $\alpha=0.025$ (one-sided), then, $\alpha_D = 0.0134787$ and $p_{Dunnett} = 0.9217057$ (for balanced groups): $(0 \leq p \leq 0.9217057)$
  - Using only weighted Simes tests cannot really fulfill CPE: if $H1 \cap H2$ is rejected with $p1= \alpha$ and $p2=\alpha$, $H1 \cap H3$ (as well as $H2 \cap H4$) is always rejected only if $\alpha \leq \alpha - p \times \alpha / 2$, i.e. p=0 (serial gatekeeping)

SANOFI

How it works with extreme values of “p” ?

- When p = 0
- Only serial gatekeeping is performed
  - The secondary endpoint can only be tested if and only if the significance of the primary criterion is proved for both treatments.
  - Dunnett closed testing procedure performed on H1 and H2
  - If both H1 and H2 have been rejected, then Dunnett closed testing procedure is performed on H3 and H4.

- When $p=p_{Dunnett}$
- Let’s assume $p_1<p_2$ (wlog) and $p_1<\alpha_D$
  - H1 is rejected with Dunnett test
- Case 1: $p_2<\alpha_D$
  - H2 is rejected with Dunnett test too
  - Dunnett closed testing procedure on H3, H4
- Case 2: $\alpha_D<p_2<\alpha$
  - H2 and H3 are tested with weighted Simes
  - The result only depends on p_3
    - If $p_3> \alpha$: Stop (only H1 is rejected)
    - If $p_3< \alpha$: H2 can be rejected too. Then, Dunnett closed testing procedure on H3 and H4.
- Case 3: $p_2> \alpha$
  - H2 is not rejected
  - H3 is tested at a level
    - $w_3= p_{Dunnett}\times\alpha/2= \alpha - \alpha_D$

SANOFI

Tuning up the Gatekeeping Procedure

SANOFI

Gatekeeping and empirical score

- Choice of p

- Definition of an empirical score to maximize
  - Reflecting the level of success of the clinical trial
  - Used to assess the performance of the gatekeeping procedure

- Examples of empirical scores:
  - Score=1 if H_1 only (or H_2) is rejected
  - Score=1+d if H_1 and H_2 are rejected
  - Score=1+s if (H_1 and H_3) or (H_2 and H_4) are rejected

Empirical score to maximize

Rejected hypotheses	Empirical score	Empirical score with an additive model
None	0	0
H_1	1	1
H_2	1	1
H_3	0	0
H_4	0	0
H_1 and H_2	$1+d$	$1+d$
H_1 and H_3	$1+s$	$1+s$
H_2 and H_4	$1+s$	$1+s$
H_1 and H_4	1	1
H_2 and H_3	1	1
H_3 and H_4	0	0
H_1, H_2 and H_3	$1+t$	$1+d+s$
H_1, H_2 and H_4	$1+t$	$1+d+s$
H_1, H_3 and H_4	$1+s$	$1+s$
H_2, H_3 and H_4	$1+s$	$1+s$
H_1, H_2, H_3 and H_4	$1+g$	$1+d+2\times s$

Score = (h1×h2)×(1+d) + (h1+h2)×(1-h1×h2) + (h1×h3) ×s + (h2×h4) ×s

SANOFI

The different parameters of the clinical trial involved in the procedure 1/

- The priority ratio RP = d/s :
  - Illustrates the relative importance given to the rejection of hypotheses
    - For both doses on the primary endpoint (d)
    - For only one dose on both the primary and the secondary endpoints (s)
- For the time being:
  - FWER = 2.5% (gatekeeping)
  - Sample size depends on α=2.5% and β=10% (naive)

SANOFI

The different parameters of the clinical trial involved in the procedure 2/

- Correlation coefficient between primary and secondary endpoints
  - Chosen values: $\rho=0$ or $\rho=+0.5$ or $\rho=+0.8$ $\rho=+0.9$ or $\rho=+1$

- Randomization ratio:
  - $r = 1$ → scenario with balanced groups
  - $r = \sqrt{(2)}$ → scenario minimizing the variance of pairwise differences versus placebo
    - $N_{tot} = n_0 + n_1 + n_2 = (2+r).n_1$ with $n_1=n_2$ and $n_0= r.n_1$

- Effect-sizes of treatments:
  - Target effect-size Δ (true for at least one treatment, primary endpoint)
  - True effect-sizes δ's (used as a ratio of Δ)
    - $\delta = \Delta$ or $⅔.\Delta$ or $⅓.\Delta$ or 0, for the other treatment on the primary endpoint
    - δ = between 0 and 2Δ, for each treatment on the secondary endpoint

SANOFI

Optimization of the average empirical score

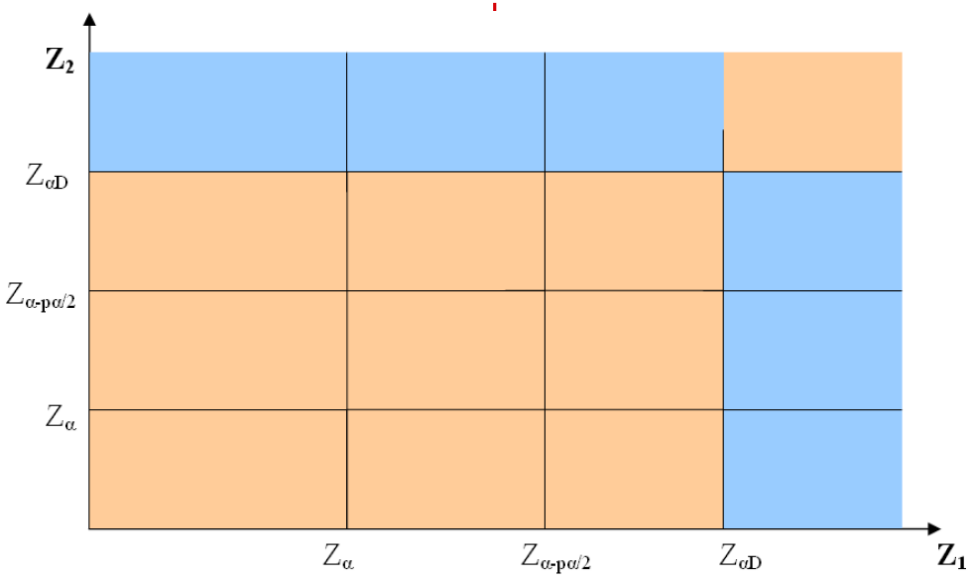
- We are in R^4, with four statistics Z1, Z2, Z3, Z4
- 2 algorithms to compute the empirical score depending on all the parameters involved
  - Bretz and Genz' s SAS macro
  - Joe's article
    - Joe's method a little less precise than Bretz-Genz's method
    - But less computational time, and precise enough in terms of difference of score
- Method based on the computation of multi-normal rectangle or orthant probabilities
- Parameters :
  - Priority ratio
  - Randomization ratio
  - Correlation coefficient
  - Effect-sizes
    - For the second dose on the primary endpoint
    - For the first dose on the secondary endpoint
    - Proportionality of the dose-responses

SANOFI

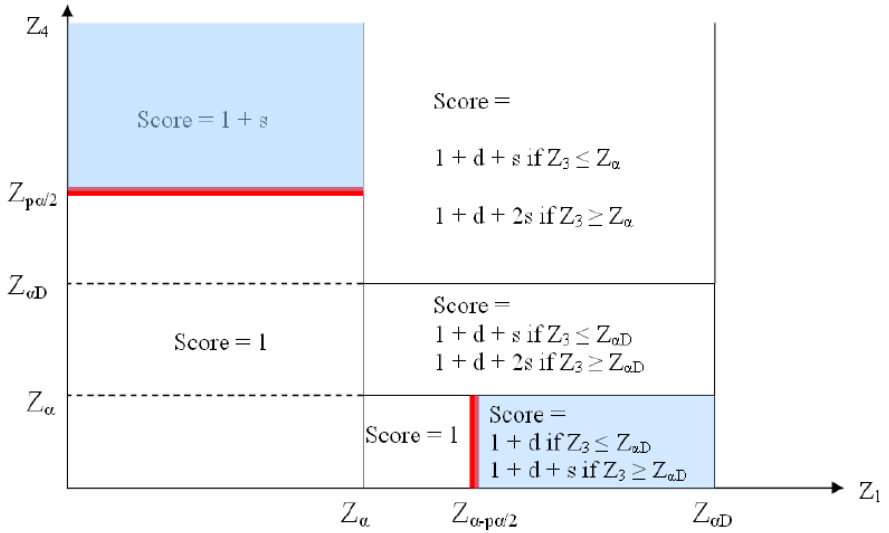
Computation of the empirical score in each case

- Computation of the “p-dependent” part (in blue) of the score :
- The empirical score only depends on “p” for some precise cases and values of the four statistics associated to each hypothesis

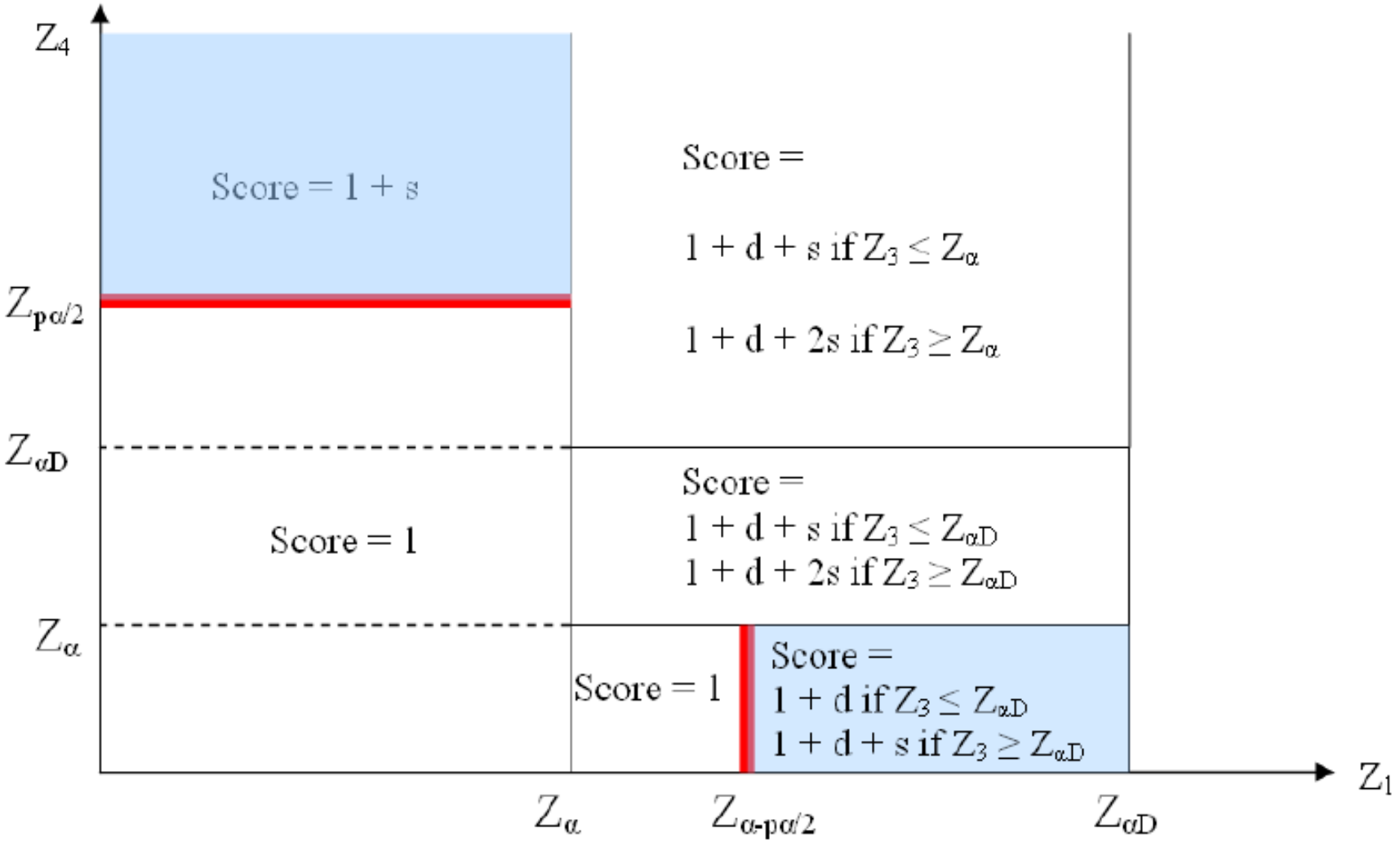
(Z2 vs Z1)



(Z4 vs Z1)

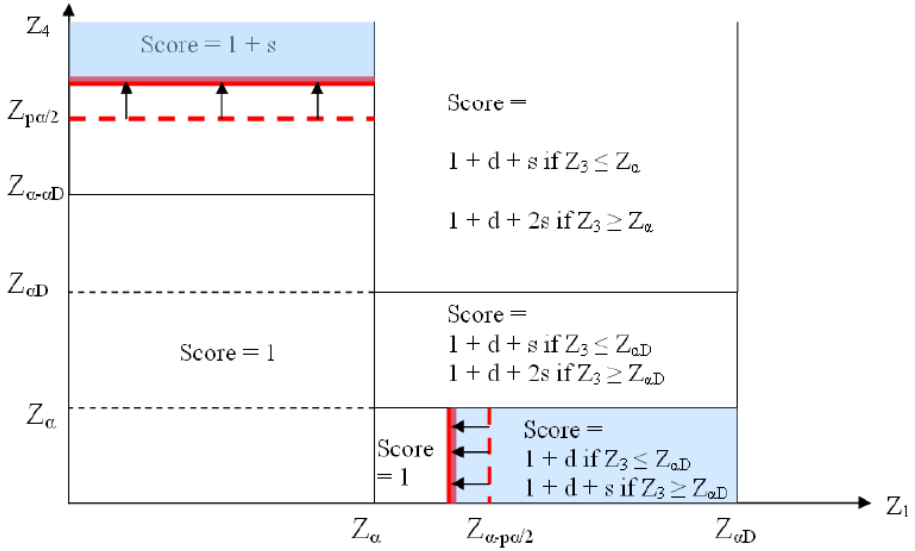


Computation of the empirical score in each case

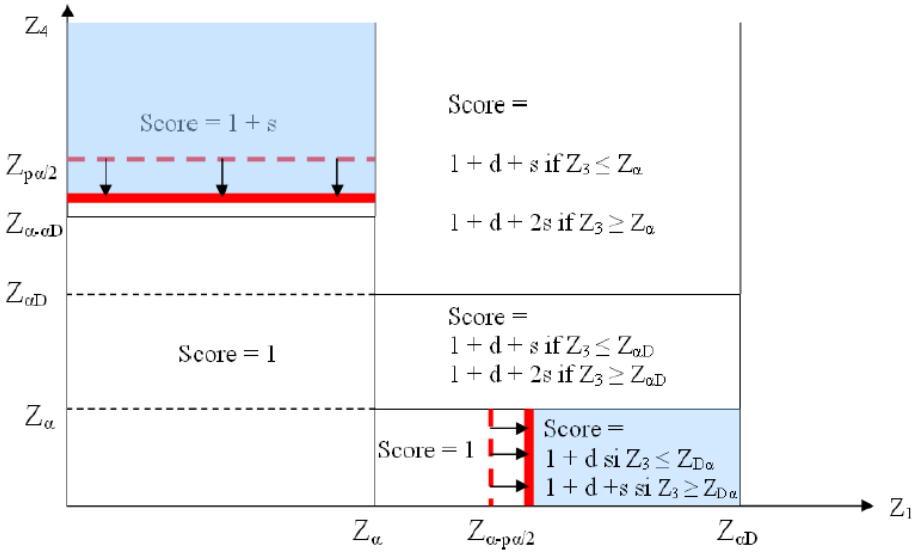


Evolution of each score area depending on “p”

- When p decreases



When p increases



- Opposite evolutions of the area “Score=1+d” and of the area “Score=1+s” depending on “p”
  - Therefore: important influence of the subjective priority ratio =d/s

SANOFI

Results obtained with Joe's method (1/2)

- Gatekeeping parametrized with a single parameter "p"
- For each combination of values for the parameters involved in the clinical trial:
  - Effect-sizes, correlation, randomization ratio, priority ratio
- Joe's method computes the optimal value of "p"
  - To maximize the value of the empirical score we defined

SANOFI

Results obtained with Joe's method (2/2)

Parameters of studied cases				Computation of "p"
$\delta_1(D2)$	$\delta_2(D1)$	ρ	RP	Joe method
1	0	0	0.1	0
$\frac{1}{3}$	0	0.8	0.5	0
1	$\frac{1}{3}$	0.5	1	0
1	$\frac{1}{3}$	0	2	0
1	0	0	9	0
1	$\frac{2}{3}$	0.9	2	0.11
$\frac{1}{3}$	$\frac{2}{3}$	0	9	0.35
$\frac{2}{3}$	$\frac{2}{3}$	0.5	2	0.52
$\frac{2}{3}$	$\frac{2}{3}$	0	1	0.74
$\frac{2}{3}$	2	1	0.1	0.9217
$\frac{2}{3}$	$\frac{2}{3}$	0	0.5	0.9217
$\frac{2}{3}$	2	0.5	1	0.9217
$\frac{2}{3}$	2	0.8	9	0.9217

- But this can only be processed individually
- Necessity to systematize the process
  - With a model estimating the value of "p" depending on all the other parameters

SANOFI

Modeling the Gatekeeping parameter

With a logistic regression model
maximizing the empirical score

SANOFI

Modeling the Gatekeeping parameter

- Idea of a logistic regression
  - Logistic modeling of $\ln [p/(p_{Dunnett}-p)]$
  - In many cases, “p” is either equal to 0 or $p_{Dunnett}$
  - Intermediary values of “p” are not so frequent
  - $\ln [p/(p_{Dunnett}-p)] =$

    $L (FWER, \alpha, \beta, \rho, RP, r, \delta_1(D2), \delta_2(D1))$

    where L is a linear function

SANOFI

Modeling the Gatekeeping parameter

- Weighting of each case, based on the range of scores for all values of “p”
  - idea: the empirical score should be adequately estimated rather than the gatekeeping parameter itself; then underweight cases where p has few influence on the results
- Additive model
  - Important influence of the priority ratio (d/s)
  - Antagonistic influences of the two effect-sizes
    - Important value for δ_1(D2) → No gatekeeping
    - Important value for δ_2(D1) → Gatekeeping
  - Very small influence of the correlation coefficient
- Most important terms of interaction (but finally ignored)
  - Interactions between the correlation coefficient and both effect-sizes

SANOFI

Modeling the Gatekeeping parameter

- FWER = 2.5%, (α, β) = (0.025,0.10)
- ln (p/($p_{Dunnett}$-p) = L (ρ, RP, r, δ_1(D2), δ_2(D1)) where L is a linear function
- Influence of the randomization ratio only through the value of $p_{Dunnett}$

$$p_{opt} = p_{Dunnett}(r) \times \left(1 - \frac{1}{\exp(L_1) + 1}\right)$$

$$\text{with } L_1 = k + a \times \ln\left(\frac{d}{s}\right) + c \times \arcsin(\rho) + d \times \delta_1(Dose2) + e \times \delta_2(Dose1)$$

Parameters	a	c	d	e	k
Estimation	−2.4304	0.6782	−12.0316	17.7415	−1.1636

- For r=1, $p_{Dunnett}$ = 0.9217057
- For r=√(2), $p_{Dunnett}$ = 0.9448982

SANOFI

Influence of the Type I and Type II errors

- Two Type-I errors actually in the Gatekeeping procedure
  - The level of risk α (associated with the level of power 1- β) on which are based the computation of the sample size and effect-sizes

$$(u_\alpha + u_\beta) = \sqrt{\frac{r \times N_{tot} \times \Delta_1^2}{(1+r) \times (2+r)}} = \sqrt{\frac{r \times N_{tot}}{(1+r) \times (2+r)}} \times \Delta_1$$

  - The FWER level α_g used for the Gatekeeping procedure

Level α	Level β	Power $1-\beta$	u_α	u_β	$u_\alpha + u_\beta$
1.25%	5%	95%	2.2414027	1.6448536	3.8862564
2.5%	10%	90%	1.9599640	1.2815516	3.2415156
5%	20%	80%	1.6448536	0.8416212	2.4864749
10%	30%	70%	1.2815516	0.5244005	1.8059521
20%	40%	60%	0.8416212	0.2533471	1.0949683

α_g	u_{α_g}
50%	0
25%	0.674490
10%	1.251552
2.5%	1.959964
0.5%	2.575829

SANOFI

Model depending of the risk levels α, β, α_g

- Now, the sample size levels α and β and the FWER α_g can take more values, and are added in the model

$$p_{opt} = p_{Dunnett}(r) \times \left(1 - \frac{1}{\exp(L_8) + 1}\right)$$

$$\text{with } L_8 = k + a \times \ln\left(\frac{d}{s}\right) + c \times \arcsin(\rho) + d \times \delta_1(Dose2) + e \times \delta_2(Dose1) + n \times (u_\alpha + u_\beta) + g \times u_{\alpha_g}$$

Parameters	a	c	d	e	g	n	k
Estimation	−1.7827	4.5095	−12.4617	4.7110	4.0700	−2.6272	15.2455

- Of course, this model is not as good as the previous one for the classical values of (α, β) = (0.025,0.10) and α_g = 0.025

SANOFI

Some practical examples

- To compare the model with fixed levels $(\alpha, \beta) = (0.025, 0.10)$ and $\alpha_g = 0.025$ and the one depending on α, β, α_g

Parameters of studied cases				Standard range	Computation of "p"	Estimations of "p"	
$\delta_1(D2)$	$\delta_2(D1)$	ρ	RP	Weighting	Joe method	Basic model	Complete model
1	0	0	0.1	0.0480	0	0.0177	0.9201
$\frac{1}{3}$	0	0.8	0.5	0.1291	0	0.0499	0.9217
1	$\frac{1}{3}$	0.5	1	0.2175	0	6.11×10^{-4}	0.9198
1	$\frac{1}{3}$	0	2	0.2945	0	1.18×10^{-4}	0.8567
1	0	0	9	0.5160	0	8.21×10^{-9}	0.1457
1	$\frac{2}{3}$	0.9	2	0.1075	0.11	0.0493	0.9216
$\frac{1}{3}$	$\frac{2}{3}$	0	9	0.1096	0.35	0.2609	0.9216
$\frac{2}{3}$	$\frac{2}{3}$	0.5	2	0.1877	0.52	0.6787	0.9217
$\frac{2}{3}$	$\frac{2}{3}$	0	1	0.3862	0.74	0.8006	0.9216
$\frac{2}{3}$	2	1	0.1	2.2766	0.9217	0.9217	0.9217
$\frac{1}{3}$	$\frac{2}{3}$	0	0.5	1.5666	0.9217	0.9212	0.9217
$\frac{2}{3}$	2	0.5	1	1.1874	0.9217	0.9217	0.9217
$\frac{2}{3}$	2	0.8	9	0.2454	0.9217	0.9217	0.9217

SANOFI

Conclusions and possible extension

SANOFI

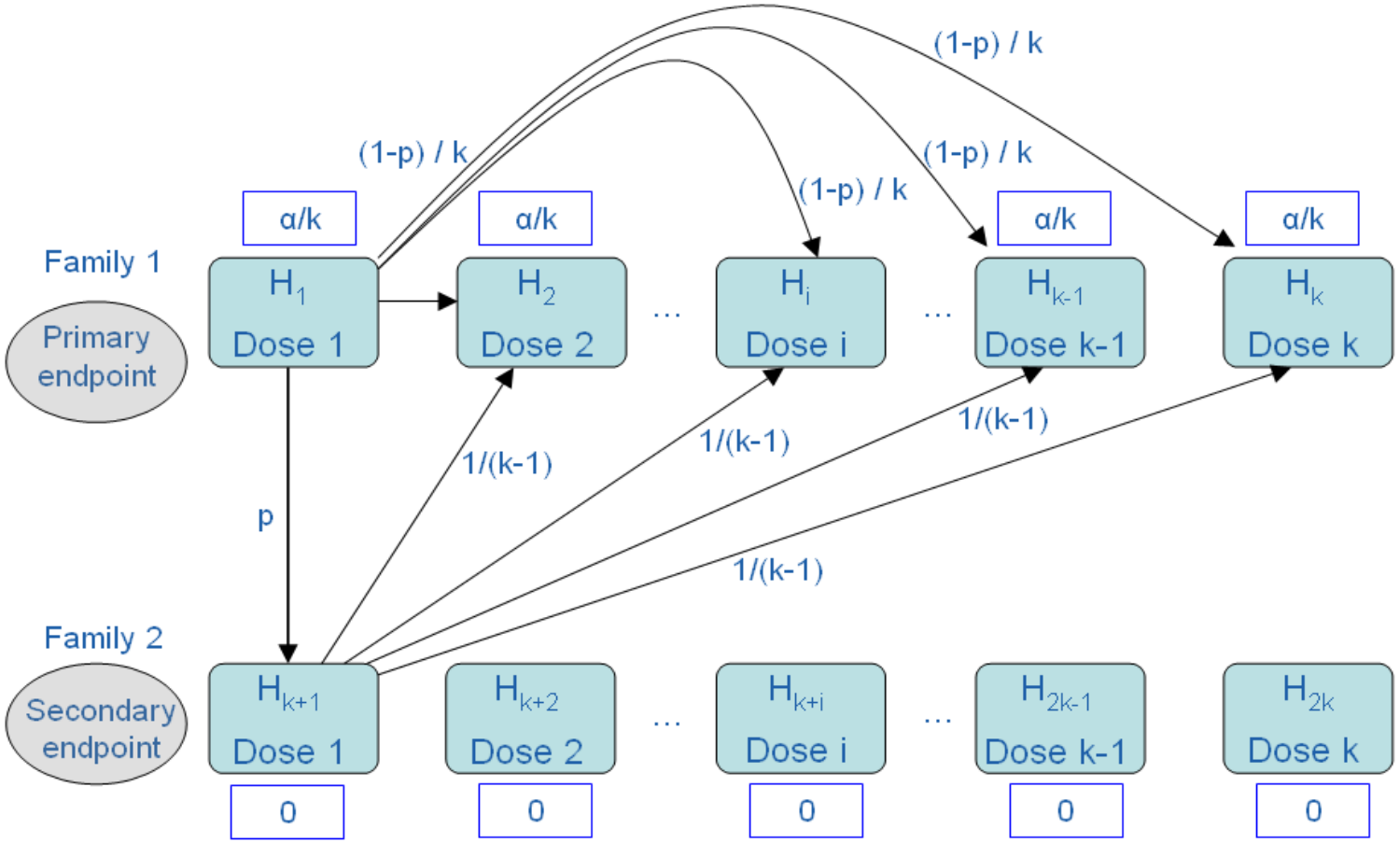
Conclusions and possible extension 1/3

- Parametrization of the gatekeeping strategy which do not decrease the probability of at least one rejection
  - For a “many-to-one” comparison of 2 treatments to a control, with one primary and one secondary endpoints
  - With only one single gatekeeping parameter “p”
- Obtain a logistic regression model to compute the optimal value of the gatekeeping parameter “p”
  - Depending on the clinical parameters
    - Randomization ratio
    - Correlation between endpoints
    - Effect-sizes
    - Subjective priority ratio
    - And possibly the levels of risks

SANOFI

Conclusions and possible extension 2/3

- Possible extension to more than 2 treatments?
  - For k treatments, the gatekeeping graph taking into account the CSE constraint is as follows depends again of a single parameter p:



Conclusions and possible extension 3/3

- Possible extension to more than 2 treatments?
  - With k treatments, constraint CSE is met with $0 \leq p \leq p_{Dunnett}$ where $p_{Dunnett} = k - k(k-1)\alpha_{Dk}/FWER$
  - If FWER=0.025:

Value of k	Value of $p_{Dunnett}$ if $r = 1$	Value of $p_{Dunnett}$ if $r = \sqrt{k}$
$k = 2$	0.9217057	0.9448982
$k = 3$	0.7409823	0.8569344
$k = 4$	0.4923294	0.7671374
$k = 5$	0.1907518	0.6808121
$k = 6$	-0.154998 i.e. no possible gatekeeping	0.5988504
$k = 10$	/	0.3091347
$k = 13$	/	0.1214795
$k = 15$	/	0.0063697
$k = 16$	/	-0.048743 i.e. no possible gatekeeping

  - Of course the optimization step for finding p becomes more and more difficult: now, we would work in R^{2k}…

SANOFI

Backup

SANOFI

Some definitions about multiplicity

- Classical unadjusted tests lead to Inflation of Type-I error (=rejection of at least one true null hypothesis).
- A testing procedure provides a **strong control** of the overall Type-I error when it controls the Type-I error for any combination of true or false individual null hypotheses.
- A testing procedure provides a **weak control** of the overall Type-I error when it controls the Type-I error for the combination of all individual null hypotheses, i.e. for the combination of order k (= intersection of all individual null hypotheses).
- Weak control is much less stringent than strong control.
- The **closed testing principle** allows the rejection of any one of these individual hypotheses, say H_i, if all possible intersection hypotheses involving H_i can be rejected by using valid local level α tests. It controls the overall Type-I error for all the k hypotheses at level α in the **strong** sense.

SANOFI

Different kinds of Gatekeeping procedures

- Serial Gatekeeping
  - All hypotheses must be rejected within one family in order to proceed to the next family in the sequence
- Parallel Gatekeeping
  - At least one significant result must be observed in one family (ie at least one hypothesis must be rejected in the family) to proceed to the next family
- Closure based tree Gatekeeping
  - Combining both serial and parallel methods

SANOFI

Weighting

$$Score(p_{opt}) - Score(p_{est}) \approx \frac{dScore}{dp}(p_{opt}) \times (p_{opt} - p_{est}) + \frac{1}{2} \times \frac{d^2Score}{dp^2}(p_{opt}) \times (p_{opt} - p_{est})^2$$

In our gatekeeping context, we must focus on optimizing the score and not the gatekeeping parameter itself, since in some cases, the score hardly depends on "p". Then, it comes down to minimize the gap between the maximal score obtained for the optimal gatekeeping parameter and the score obtained for the estimated gatekeeping parameter.

Minimizing $\sum Score(p_{opt}) - Score(p_{ext})$ can be seen as minimizing

$$\sum \frac{(Score(p_{opt}) - Score(p_{ext}))}{(p_{opt} - p_{est})^2} \times (p_{opt} - p_{est})^2$$

with a weighting least squared regression on the parameter "p".

In this case, weighting is then defined with the following formula:

$$Weight = \frac{Score(p_{opt}) - Score(p_{est})}{(p_{opt} - p_{est})^2} \approx \frac{1}{p_{opt} - p_{est}} \times \frac{dScore}{dp}(p_{opt}) + \frac{d^2Score}{dp^2}(p_{opt})$$

If p_{opt} leads to the best score, then $\frac{dScore}{dp}(p_{opt}) = 0$ and the weight becomes: $Weight = \frac{d^2Score}{dp^2}(p_{opt})$

This results tends to show that if the empirical score could be expressed with an analytical function, a natural possible weighting would be the second order derivative. A second order derivative corresponds to the local range for a point neighborhood. However, the empirical score is not expressed with an analytical function but is only computed numerically. Nevertheless, by analogy with an analytical function, a possible weighting is using the global range of empirical score. This comes from the fact that:

$$\forall \epsilon > 0, f''(x) \approx \frac{f'(x+\epsilon) - f'(x)}{\epsilon} \approx \frac{f(x+\epsilon) - f(x)}{\epsilon^2} - \frac{f(x) - f(x-\epsilon)}{\epsilon^2}$$

$Weight = Score(p_{max}) - Score(p_{min})$ with p_{max} the value of p for which the score is maximum and p_{min} the value of p for which the score is minimum.

SANOFI

Modeling the Gatekeeping parameter

- Now with interactions in the model
- Most important terms of interaction:
  - Between the correlation coefficient and the effect-sizes

$$p_{max} = p_{Dunnett} \times \left(1 - \frac{1}{\exp(L) + 1}\right)$$

with

$$L = k + a \times \ln\left(\frac{d}{s}\right) + c \times \arcsin(\rho) + d \times \delta_1(Dose2) + e \times \delta_2(Dose1)$$

$$+cd \times \arcsin(\rho) \times \delta_1(Dose2) + ce \times \arcsin(\rho) \times \delta_2(Dose1)$$

  - Parameters:

Parameters	a	c	d	e	k	cd	ce
Estimation	−2.8905	−0.8063	−10.5999	13.4193	0.00896	−4.4696	8.9513

- But adding these terms of interaction does not improve the quality of the model that much *-> really necessary ?*

SANOFI