Sequential graphical approach for a phase 3 study with multiple primary endpoints, experimental treatment arms and populations

SMAC 2023

Laure Charbonnier, Sanofi, Biostatistics & Programming, Chilly-Mazarin, France

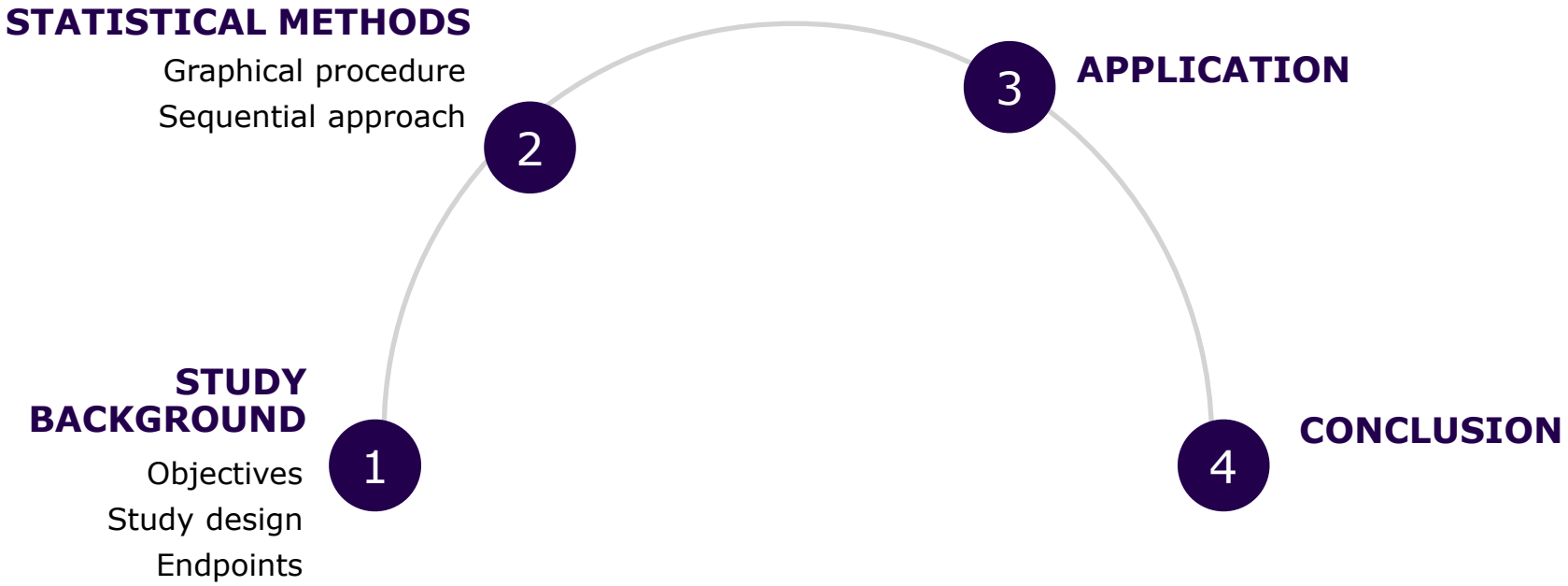
17 . Mar . 2023

Disclosures

Laure Charbonnier is a Sanofi employee and may hold shares and/or stock options in the company

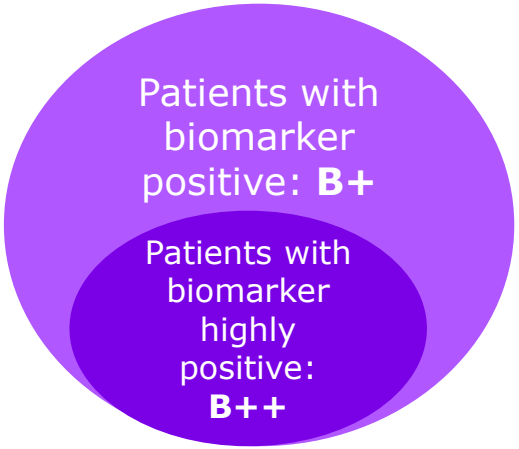
Research and analyses were supported by Sanofi

Agenda


STATISTICAL METHODS
Graphical procedure
Sequential approach
2
3
APPLICATION
STUDY BACKGROUND
1
Objectives
Study design
Endpoints
4
CONCLUSION

Complex phase 3 to address multiple objectives

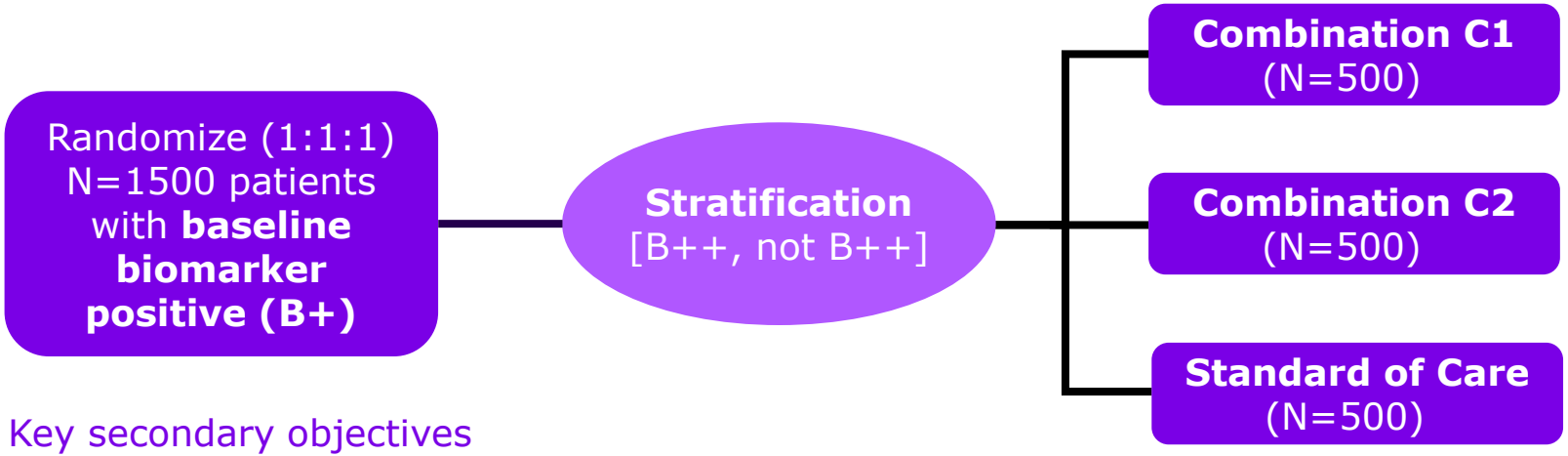
- Two populations



- Two experimental arms to be compared to a Standard of Care
  - Add-on approach: combination C1
  - Substitution approach: combination C2
- Two primary endpoints
  - Progression Free Survival based on RECIST 1.1
  - Overall Survival

Study design: multicenter international randomized controlled phase 3 study in biomarker stratified population

- Primary objectives: To demonstrate an improvement in PFS or OS with C1 vs SOC in B++



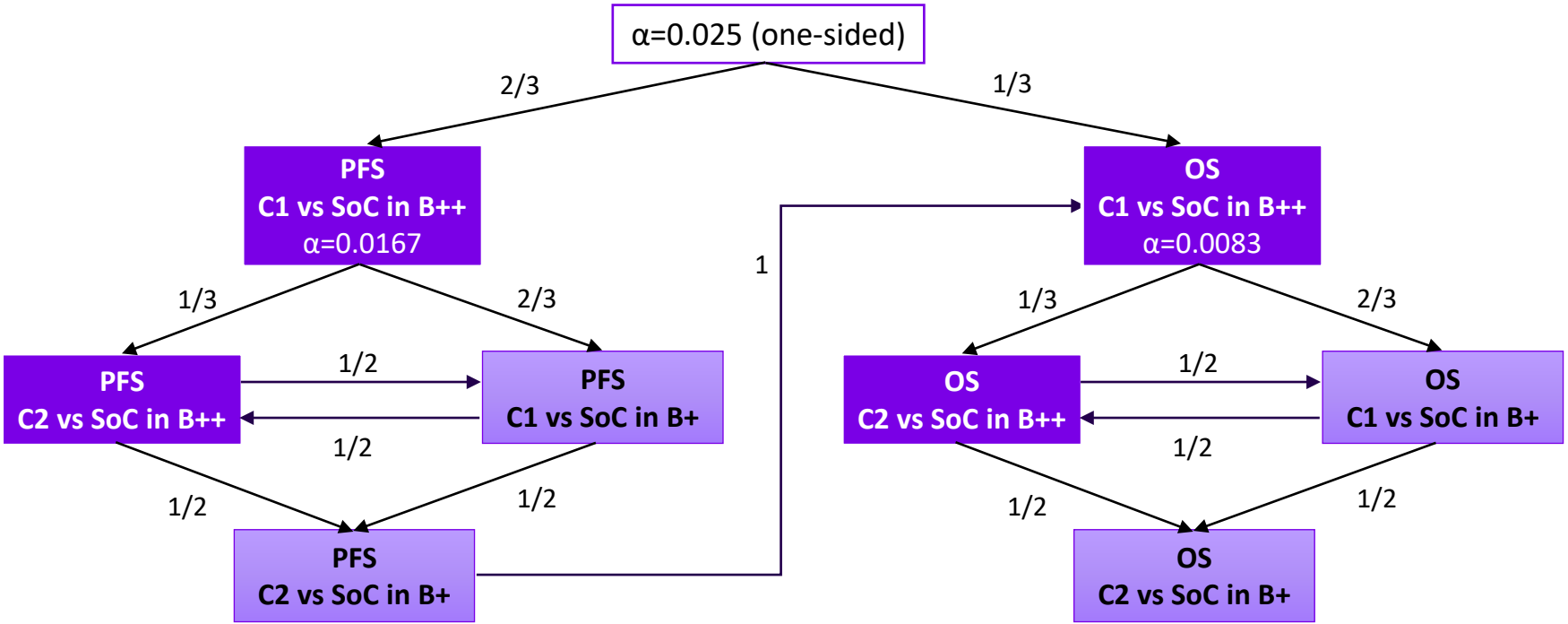
- Key secondary objectives
  - To demonstrate an improvement in PFS or OS with C1 versus SOC in B+
  - To demonstrate an improvement in PFS or OS with C2 versus SOC in B++
  - To demonstrate an improvement in PFS or OS with C2 versus SOC in B+

Two primary efficacy endpoints: PFS and OS

- Progression Free Survival (PFS), is defined as the time from the date of randomization to the date of the first documented disease progression or death due to any cause, whichever comes first
- Overall Survival (OS) is defined as defined as the time from the date of randomization to the date of death due to any cause
- Clinical assumptions: Hazard Ratio (HR)

	PFS	OS
C1 vs. SOC in B++	HR = 0.65 (8 mo → 12.3 mo)	HR = 0.70 (20 mo → 28.6 mo)
C1 vs. SOC in B+	HR = 0.69 (8 mo → 11.6 mo)	HR = 0.74 (20 mo → 27.0 mo)
C2 vs. SOC in B++	HR = 0.69 (8 mo → 11.6 mo)	HR = 0.74 (20 mo → 27.0 mo)
C2 vs. SOC in B+	HR = 0.74 (8 mo → 10.8 mo)	HR = 0.79 (20 mo → 25.3 mo)

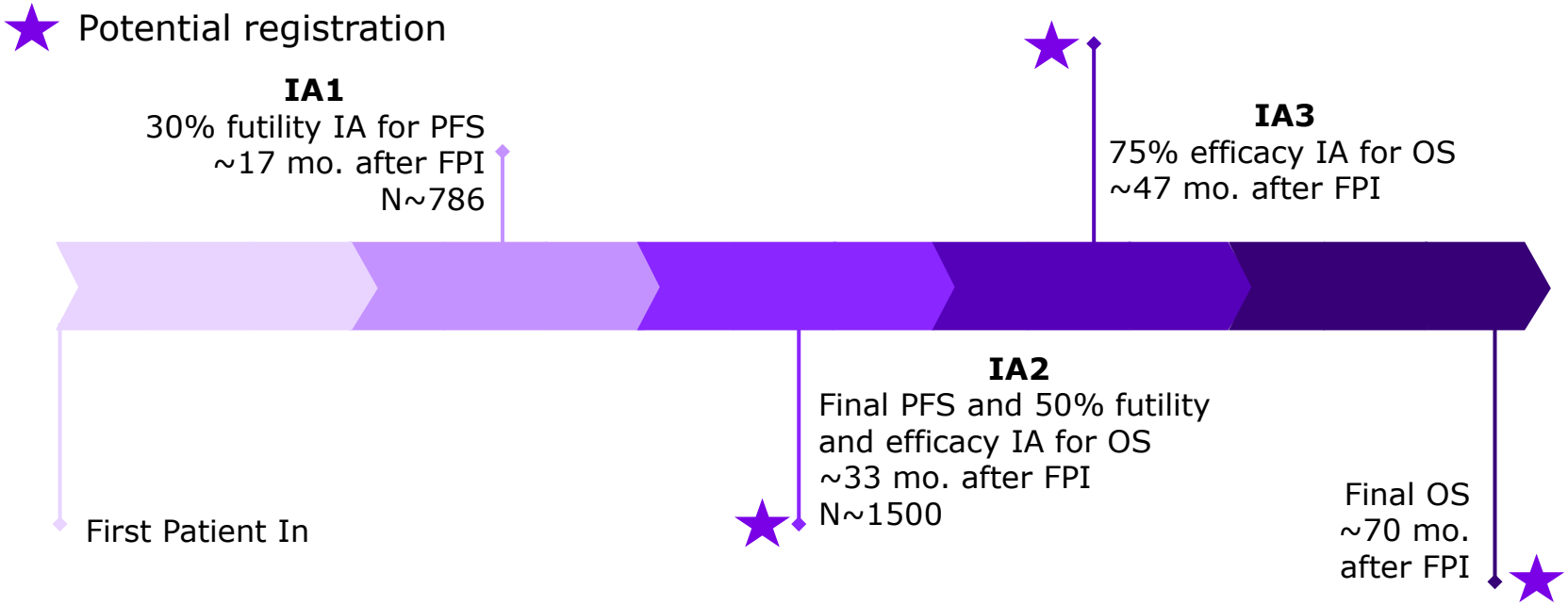
Graphical procedure (1)



(1) Bretz 2009

Multiple interim analyses are planned to check for futility or demonstration of early outstanding efficacy

Timelines are driven by the comparison of C1 vs. SoC in the B++ population



PFS: Progression Free Survival, OS: Overall Survival, IA: Interim Analysis, SoC: Standard of Care

Sequential approach (2)

- For the comparison of C1 vs. SOC in B++ on OS: alpha-spending function according to Lan-DeMets O'Brien-Fleming (OBF)
- For other hypotheses: information fraction cannot be precisely determined as the analysis timeline is driven by the number of OS events from C1 vs. SoC in the B++ population,
- To control the overall type-I error rate due to repeated testing
  - Each hypothesis has its prespecified α level to be spent at each analysis
  - Corresponding sequence of group sequential p-values (GP) is computed
- General idea for GP computation
  - At each IA (not including the final look), if the test statistics crosses the efficacy boundary, then the p-value is defined as the actual value
  - Otherwise, the p-value is set to one to be conservative as we should not make a claim
- The graphical procedure is then applied at each analysis time on the group sequential p-values

(2) Luo and Quan 2023

Sequential approach: Notations

- PFS hypotheses: H1 (C1 vs SoC in B++), H2 (C1 vs SoC in B+), H3 (C2 vs SoC in B++), H4 (C2 vs SoC in B+)
  - No interim analysis is planned for efficacy
- OS Hypotheses: H5 (C1 vs SoC in B++), H6 (C1 vs SoC in B+), H7 (C2 vs SoC in B++), H8 (C2 vs SoC in B+)
- We assume all the hypotheses are one-sided and smaller test statistics indicate stronger evidence in rejecting the hypotheses
- For H5, X_{52}, X_{53} and X_{54} are the log-rank test statistics at IA2, IA3 and Final Analysis (FA)
- For H5, OBF type of alpha spending function spent at IA2, IA3 and FA
  - Cumulative alpha levels based on OBF: α_{52}, α_{53} and α_{54}
  - Critical values: c_{52}, c_{53} and c_{54}
  - Number of OS events: D_{52}, D_{53} and D_{54}
- Use similar notations for H6, H7, H8

Sequential approach: Group-Sequential P-values

Computation of Group-Sequential P-values for H5

- $GP_{52} = I(X_{52}<=c_{52})\Phi(X_{52}) + I(X_{52}>c_{52})$
- $GP_{53} = I(X_{52}<=c_{52})\Phi(X_{52}) + I(X_{52}>c_{52}, X_{53}<=c_{53})\{\alpha_{52}+F_{53}(X_{53})\} + I(X_{52}>c_{52}, X_{53}>c_{53})$
  - If the efficacy boundary has been crossed at IA2, then $GP_{53}=GP_{52}$
  - GP_{53} should not be smaller than α_{52} as this part of α has already been spent
- $GP_{54} = I(X_{52}<=c_{52})\Phi(X_{52}) + I(X_{52}>c_{52}, X_{53}<=c_{53})\{\alpha_{52}+F_{53}(X_{53})\} + I(X_{52}>c_{52}, X_{53}>c_{53})\{\alpha_{53}+F_{54}(X_{54})\}$

Where "I" is the indicator function,

"Φ" is the distribution function of standard normal distribution

$F_{53}(x)=Pr(Z_3<=x, Z_2>c_{52})$, $F_{54}(x)=Pr(Z_4<=x, Z_2>c_{52}, Z_3>c_{53})$

Z_2, Z_3 and Z_4 follows multivariate normal distribution with mean zero, variance 1 and correlations

$corr(Z_2, Z_3) = sqrt(D_{52}/D_{53})$, $corr(Z_2, Z_4) = sqrt(D_{52}/D_{54})$, $corr(Z_3, Z_4) = sqrt(D_{53}/D_{54})$

Note: $GP_{52} \geq GP_{53} \geq GP_{54}$

Sequential approach: Group-Sequential P-values

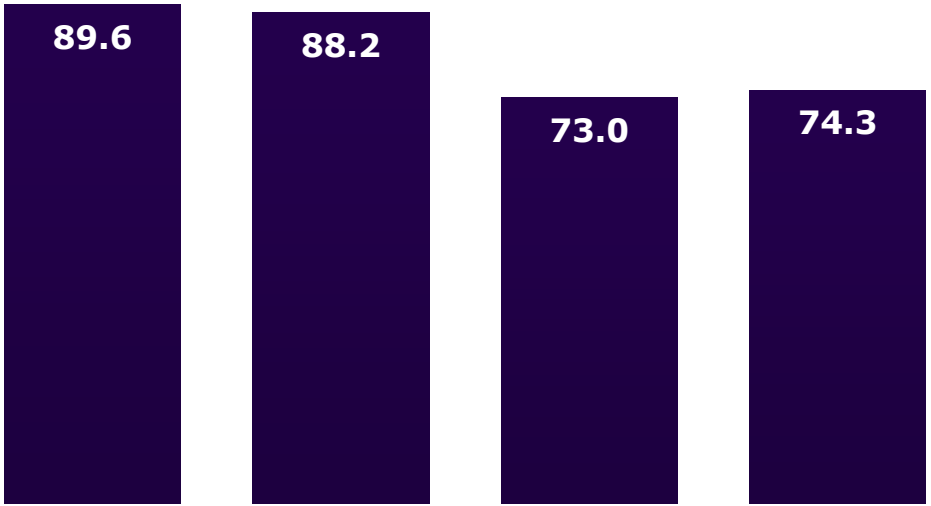
Computation of Group-Sequential P-values for H6 (same logic applies for H7 and H8)

- Use same alpha level as for H5: α_{52}, α_{53} and α_{54}
- Determine critical values c_{62}, c_{63} and c_{64} according to the following algorithm
  - Compute c_{62} such that $\Phi(c_{62})=\alpha_{52}$
  - After observing D_{62} and D_{63}, compute c_{63} such that $F_{63}(c_{63})=\alpha_{53}-\alpha_{52}$
    - let Z_2 and Z_3 follow a bivariate normal distribution with means 0, variances 1 and correlation $corr(Z_2, Z_3) = sqrt(D_{62}/D_{63})$, define function $F_{63}(x)=Pr(Z_3<=x, Z_2>c_{62})$
  - After observing D_{62}, D_{63} and D_{64}, compute c_{64} such that $F_{64}(c_{64})= \alpha_{54}-\alpha_{53}$
    - let Z_2, Z_3 and Z_4 follow a multivariate normal distribution with means zero, variances 1 and correlations $corr(Z_2, Z_3) = sqrt(D_{62}/D_{63})$, $corr(Z_2, Z_4) = sqrt(D_{62}/D_{64})$ and $corr(Z_3, Z_4) = sqrt(D_{63}/D_{64})$, define function $F_{64}(x)=Pr(Z_4<=x,\ Z_2>c_{62}, Z_3>c_{63})$
- Use same logic as for H5, to compute Group-Sequential P-values for H6

Sequential approach: Group-Sequential P-values

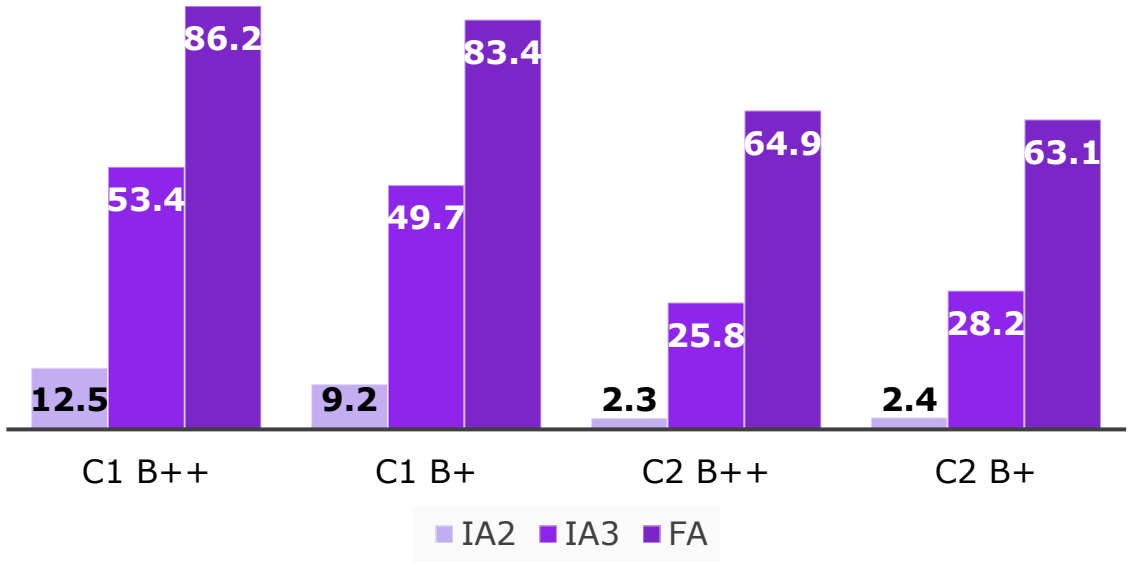
- P-value at each analysis point is compared with full α available for testing the hypothesis
  - To reject the null hypothesis: (1) $GP_{52} \leq \alpha$ or (2) $GP_{52} > \alpha$ and $GP_{53} \leq \alpha$ or (3) $GP_{52} > \alpha$ and $GP_{53} > \alpha$ and $GP_{54} \leq \alpha$
  - Equivalent to the conventional rejection criteria in the group-sequential trial: (1') $X_{52} \leq c_{52}$, or (2') $X_{52} > c_{52}$ and $X_{53} \leq c_{53}$ or (3') $X_{52} > c_{52}$ and $X_{53} > c_{53}$ and $X_{54} \leq c_{54}$
  - Therefore, the GPs are just transformations of the data preserving the information for rejection decision
- The procedure prevents the type-I error rate inflation due to multiple hypotheses and multiple analyses
  - Because the GPs computed from earlier analyses are not smaller than those from later analyses by definition
  - and each of the p-values at final analysis follows a uniform distribution
- Benefit of defining the GPs
  - Most of the multiplicity comparison procedures are based on a set of p-values instead of the test statistics
  - p-value at each analysis point is compared with full α available for testing the hypothesis, instead of the stage-specific α

Power results for PFS for each comparison



	C1 B++	C1 B+	C2 B++	C2 B+
Number of events:	286	581	290	592

Power results for OS for each comparison and analysis



Number of events:	189 | 284 | 378	385 | 577 | 764	192 | 289 | 383	393 | 588 | 775

Conclusion

- The sequential graphical approach can be used as multiplicity adjustment procedure for complex studies with multiple objectives and interim analyses
  - Offers flexibility for this phase 3 case study with two primary endpoints, two experimental treatment arms and two populations
  - Possible extension to other multiple comparison procedures (Hochberg, Hommel)
- Strategy to target a potential registration in the overall population of patient with positive baseline biomarker (B+) or only in the subgroup with high expression of the biomarker (B++)
  - Graphical procedure is designed to address first the primary objectives in B++ subgroup
  - Possibility to demonstrate also an effect in the overall population B+
    - Consistency of the results for the subset of patients moderately positive (not B++) will also need to be observed
- FDA feedback: the proposed approach for controlling overall study-wise Type I error rate appears reasonable

References

(1) Bretz, F., W. Mauer, W. Brannath, and M. Posch. 2009. “A graphical approach to sequentially rejective multiple test procedures.” Statistics in Medicine 28 (4): 586–604.
(2) Luo, X, Quan, H. 2023. “Some multiplicity adjustment procedures for clinical trials with sequential design and multiple endpoints”. Statistics in Biopharmaceutical Research. To appear

Thank you

sanofi