Politecnico di Torino

Dual criteria approach to early Accelerated Approval in time-to-event group sequential trials via historical borrowing

Marco Ratta, Donia Skanji, Mauro Gasparini, Valentine Barboux, Gaëlle Saint-Hilary, Pavel Mozgunov

SERVIER
moved by you

Regulatory Context: FDA's Accelerated Approval

- The urgency of delivering novel effective treatments against life-threatening diseases has brought various health authorities to allow for accelerated approvals **(AA)**

- FDA Accelerated Approval Draft Guidance (March 2023)[1]
  - treatment efficacy on a **surrogate endpoint** is assessed to give an **Accelerated Approval (AA)**
  - treatment efficacy on the **primary endpoint** is assessed to give **Full Approval (FA)**

- One RCT monitoring surrogate endpoint → AA
- One RCT monitoring primary endpoint → FA

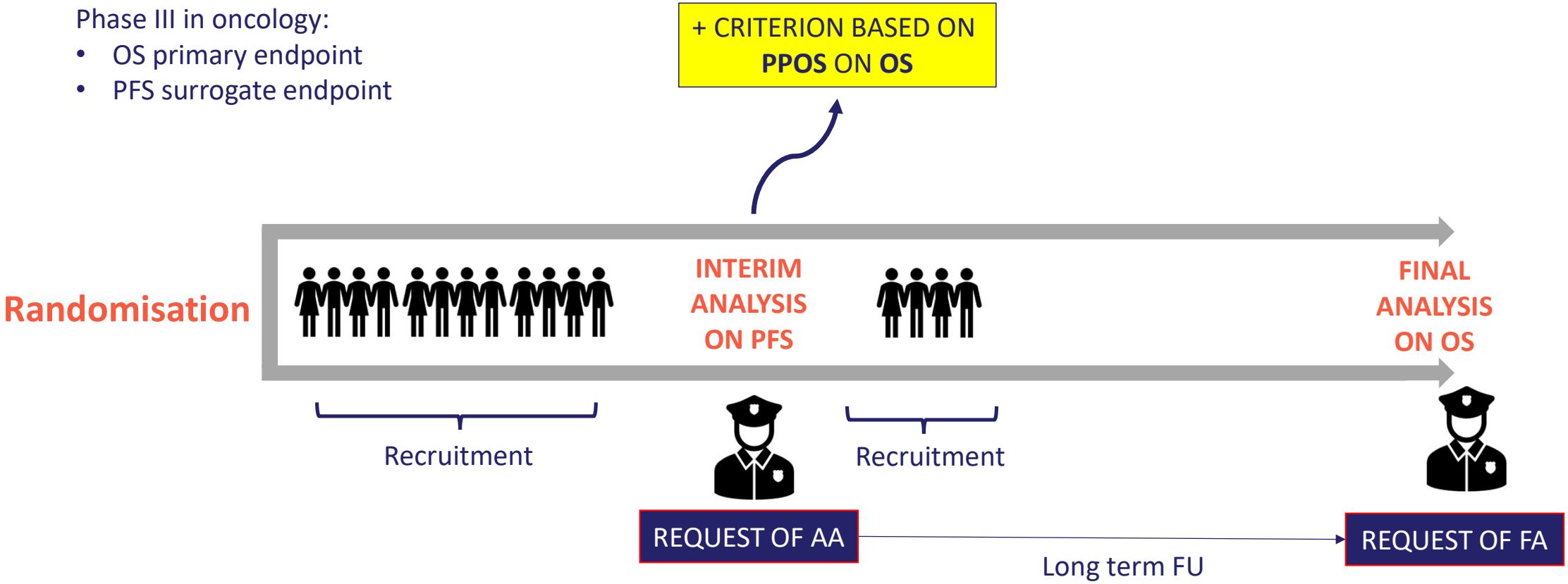
- One unified Group Sequential RCT where both surrogate & primary endpoints are monitored. An early Interim Analysis (IA) is planned for AA

1) «Clinical Trial Considerations to Support Accelerated Approval of Oncology Therapeutics – Guidance for Industry» https://www.fda.gov/media/166431/download

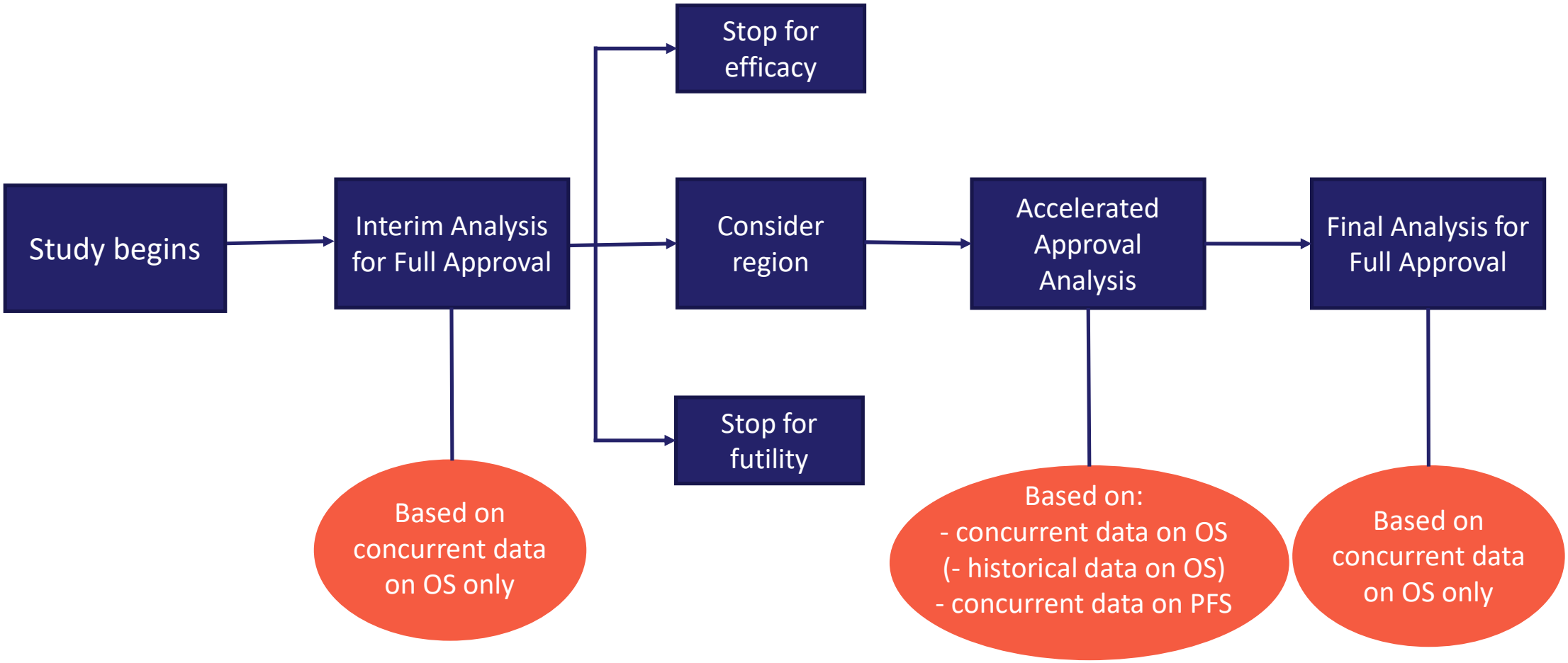
Our proposal

Phase III in oncology:

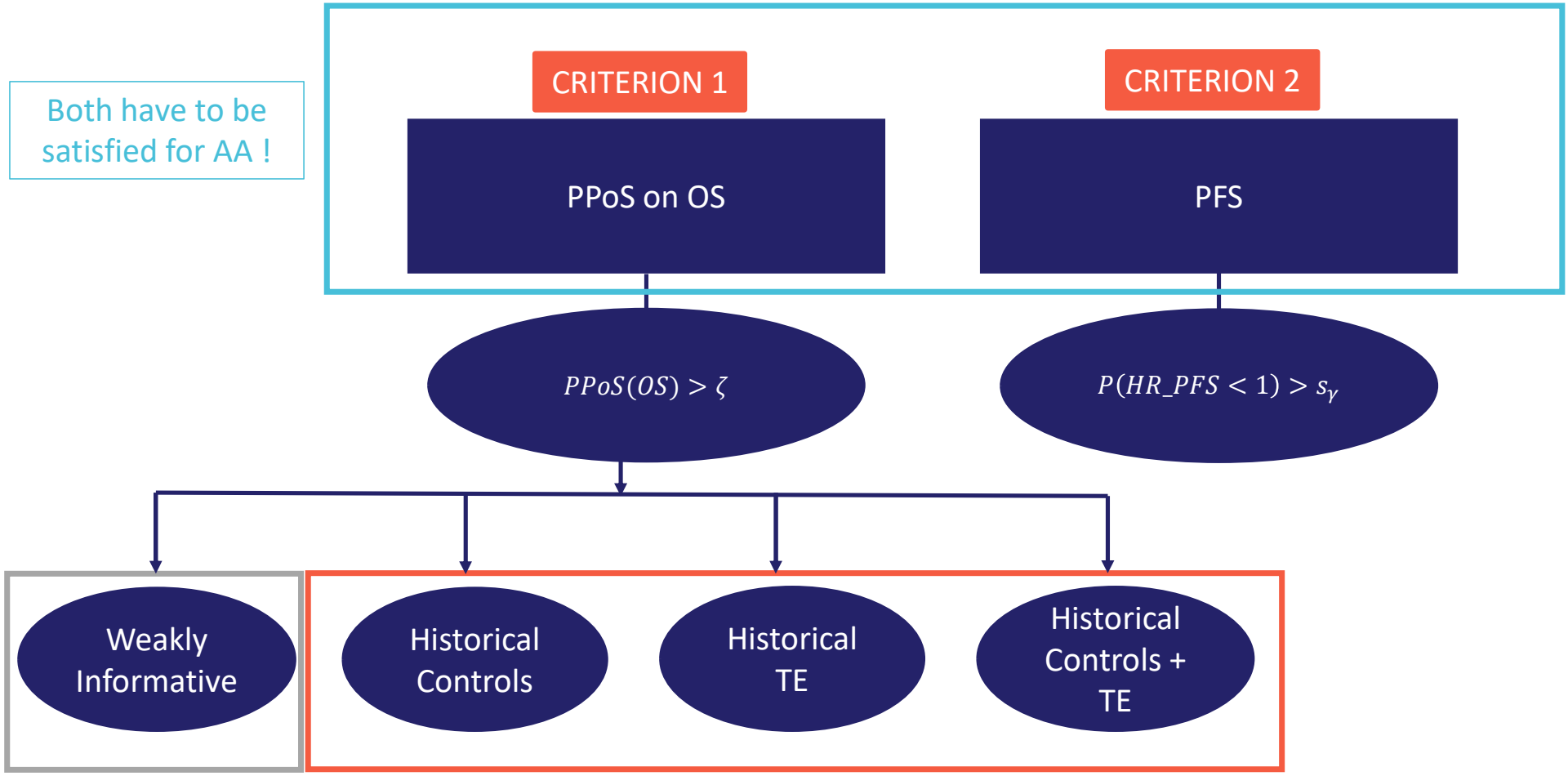
- OS primary endpoint
- PFS surrogate endpoint



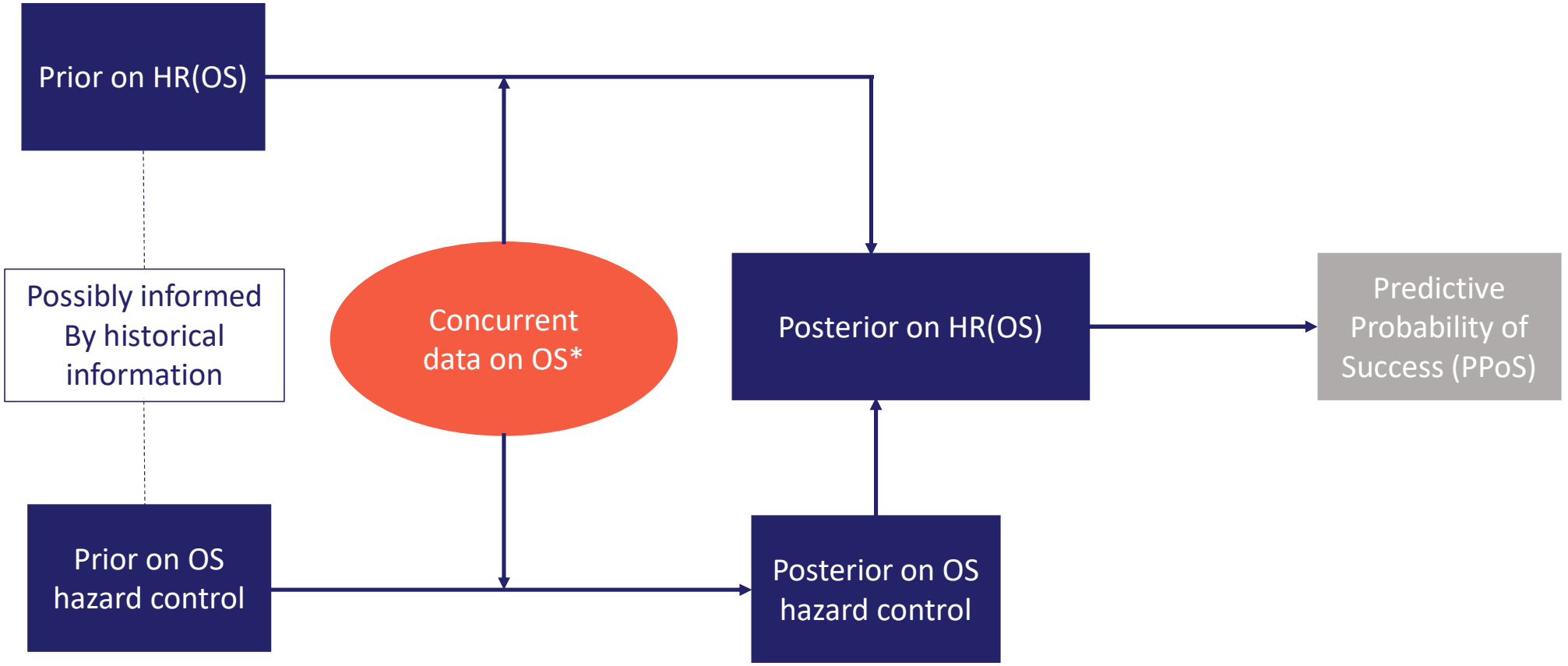
Trial design



Accelerated Approval Analysis – Dual Criteria approach



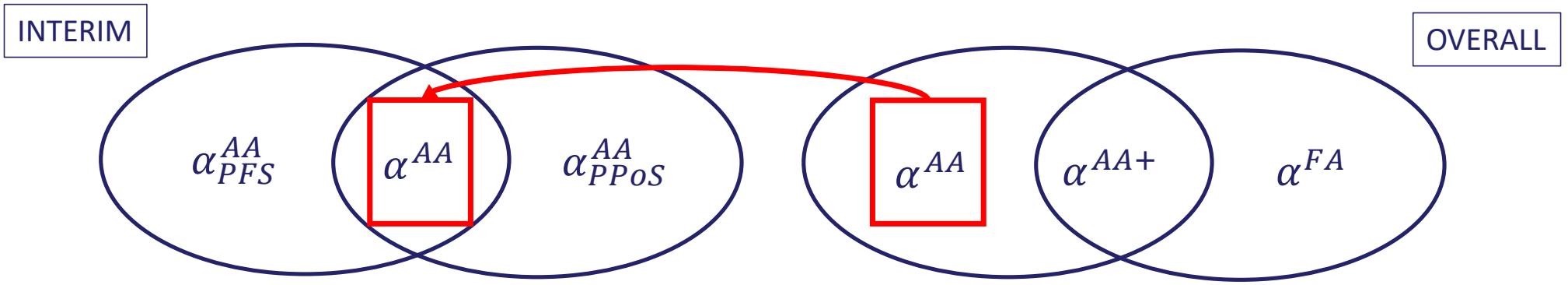
PPoS on OS at Interim



* Data are number of control events of *current* study, total exposure time on control of *current* study, number of treatment events of *current* study, total exposure time on treatment of *current* study

Type I Error

- The treatment has multiple chance to enter the market, as a consequence a multiplicity adjustment is needed in order to avoid type I error inflation.



$$\alpha = \alpha^{FA} + \alpha^{AA} - \alpha^{AA+FA} < \alpha^{FA} + \alpha_{PFS}^{AA}$$

Probability to obtain either accelerated or full approval

Probability to obtain accelerated approval

Probability to obtain full approval

Probability to obtain both accelerated and full approval

In our case:

α^{FA} = 1.25%

α_{PFS}^{AA}= 1.25%

α < 2.5%

Motivating example
(phase III in mCRC)

- OS: median control = 8.5 months, target HR = 0.71
- PFS: median control= 2.1 months, target HR =0.525
- Control Borrowing from 3 historical studies (median control = 8.5 months)
- 1:1 randomization
- 17 months recruitment, 20 months FU
- Interim at IF=0.2
- 90% OS Power

- 500 patients
- Final Analysis after 424 OS events

FULL APPROVAL CRITERION

$P(HR_{OS} < 1) > 0.9875$

Under the OS null scenario

$\alpha^{FA} < 0.0125$

ACCELERATED APPROVAL DUAL CRITERION

- $P(HR_{PFS} < 1) > 0.9875$

Under the PFS null scenario

$\alpha_{PFS}^{AA} < 0.0125$

$\alpha < 0.025$

Under the PFS/OS null scenario

- $PPoS(OS) > 0.9$

$\alpha < 0.025$

Under an «assurance scenario»

SERVIER

Case study

	INTERIM ANALYSIS		FINAL ANALYSIS	
	# Control	# Treatment	# Control	# Treatment
PFS	93	76	-	-
OS	47	37	218	206

Approach	AA criterion	Historical Controls	Historical PFS/OS Relationship	Interim Analysis			Final Analysis
				$\%P(\gamma < 1)$	%PPoS	AA request	$\%P(\theta < 1)$
CURRENT	single	X	X	99.6	X	yes	95.2
NO HIST	dual	no	no	99.6	69.6	no	95.2
HIST CON	dual	yes	no	99.6	61.6	no	95.2
HIST TE	dual	no	yes	99.6	49.7	no	95.2
BOTH	dual	yes	yes	99.6	44.0	no	95.2

INCORRECT ACCELERATED APPROVAL !

Dual criteria helped in avoiding incorrect Accelerated Approval !

SERVIER

Discussion

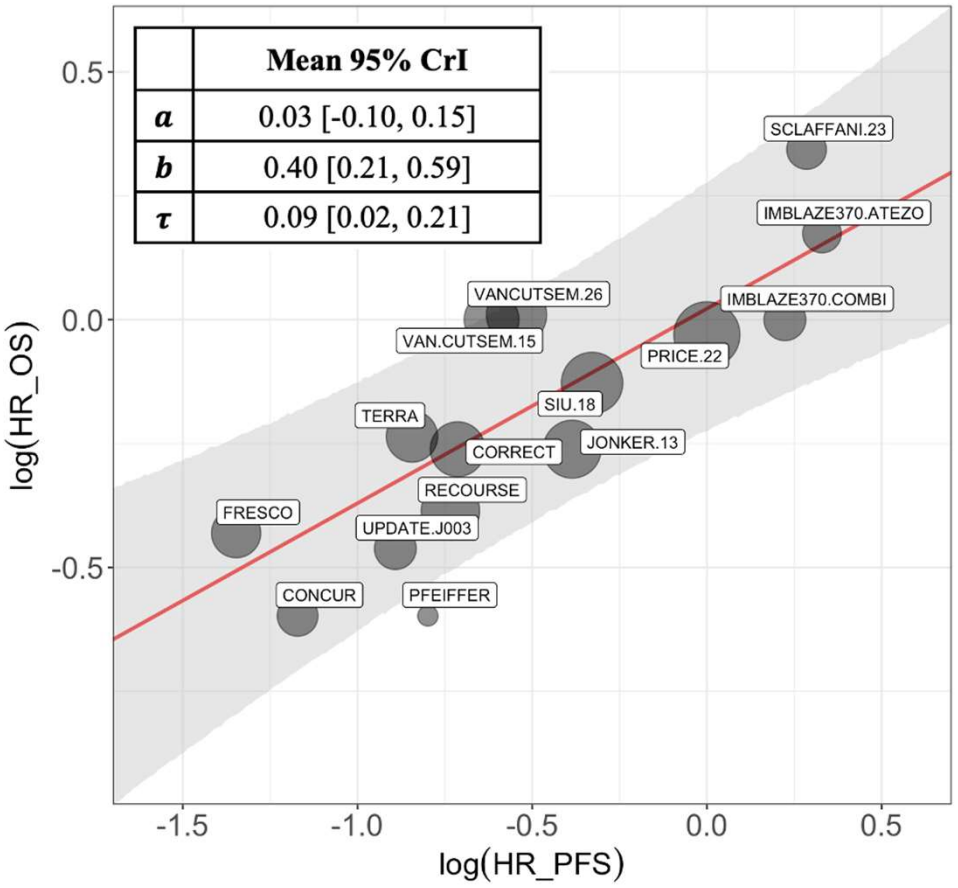
An **extensive simulation study** has been performed to test the proposed approach, showing

- Dual criteria approach substancially limits substancially the number of incorrect Accelerated Approval requests
- Historical borrowing may be helpful in informing PPoS for Accelerated Approval if there is fair accordance between current and historical data
- A conservative choice of the decision thresholds may help in maintaing the global type I error low even in case of prior-data conflict

Clinical Development and R&D CMO

Backup Slides

PFS/OS Historical relationship- Surrogate prior



$$\begin{pmatrix}\hat{\theta}_k \\ \hat{\gamma}_k\end{pmatrix}\Bigg|\begin{pmatrix}\theta_k \\ \gamma_k\end{pmatrix} \sim N\left(\begin{pmatrix}a + b\gamma_k \\ \gamma_k\end{pmatrix}, \begin{pmatrix}\sigma_k^2 + \tau^2 & \rho_k\sigma_k\delta_k \\ \rho_k\sigma_k\delta_k & \delta_k^2\end{pmatrix}\right)$$

$$\theta|a,b,\tau \sim N(a + b\gamma_{IA}, \tau^2)$$

$$\pi_\theta^S(\cdot) = \int f_{\theta|a,b,\tau}(\cdot)\, f_{a,b,\tau}(x,y,z)d(x,y,z)$$

SURROGATE PRIOR (Saint-Hilary, 2019)

γ_{IA} : Posterior distribution for HR(PFS) at the interim analysis

$\hat{\theta}_k, \hat{\gamma}_k$: point estimates of treat. Effect for OS and PFS in the k[th] historical study

σ_k, δ_k: standard errors for treat. Effect OS and PFS in the k[th] historical study

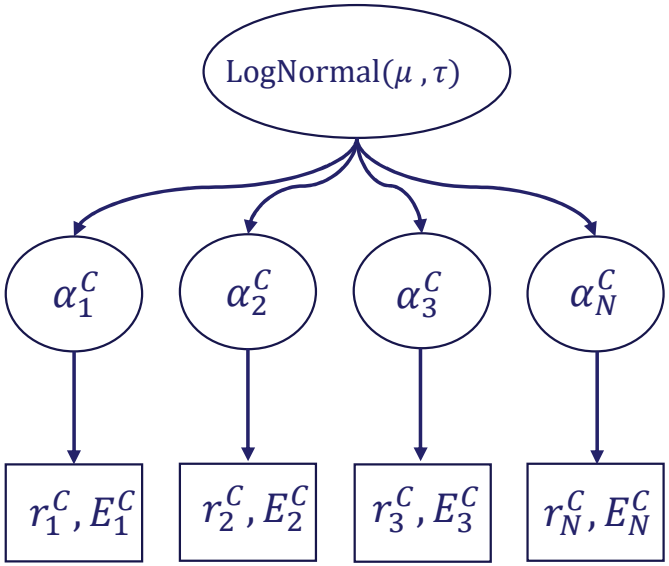
a, b, τ: regression parameters

θ: true treatment effect on OS in the current study

δ_p: standard error for PFS in the current study

Historical controls borrowing

Suppose control data are available for N studies.
The available informations for each study are the total number of events under control arm and the total exposure time (i.e. the sum of the individual times patients remain in the study).
A prior distribution for the control Hazard Rate on OS is assessed via MAP approach using the methodology below



HIERARCHICAL MODEL

$$r_i^C \sim \text{Poisson}(\alpha_i^C E_i^C)$$
$$\log(\alpha_i^C) \sim \text{Normal}(\mu , \tau)$$
$$\mu \sim \text{Normal}(0, 10)$$
$$\tau \sim \text{Half Normal}(0.5)$$

Prior distribution for the current hazard control

$$\pi_{\alpha_*^C} \sim \text{LogNormal}(\mu^{post} , \tau^{post})$$

r_i^C: number of control events of i^{th} study
E_i^C: Total exposure time on control of i^{th} study
α_i^C: Control hazard of i^{th} study
μ: True Control hazard
τ: Between study variability
r_*^C: number of control events of *current* study
E_*^C: Total exposure time on control of *current* study
α_*^C: Control hazard of *current* study

Simulation Plan – Metrics

%AA: Probability of passing AA analysis

%FA: Probability of passing FA analysis

α^G: Probability of passing either AA or FA analysis under the null hypothesis HR(OS)=1

CR: Probability of passing FA analysis conditional to having passed AA under the alternative hypothesis HR(OS)=0.71

Base approach

Alternative Hypothesis – Base Approach

PFS/OS Conflict	Scenario ID	HR(PFS)	HR(OS)	Control Median(OS)	%AA	CR	%FA (POWER)
Minor	Scenario 1B	0.39	0.71	8.5 months	50.5	97.4	89.7
Design Assumptions	Scenario 2B	0.525	0.71	8.5 months	50.0	97.4	89.7
Large	Scenario 3B	0.75	0.71	8.5 months	22.5	96.4	89.7
Extreme	Scenario 4B	1	0.71	8.5 months	0.6	>99.9	89.7

- Under minor PFS/OS conflict and under the **design assumptions** (scenario 2) there is a 50% AA, while a sensible reduction in %AA is observed in presence of large or extreme prior-data conflict (scenarios 3-4) due to the PFS testing criterion
- Conversion Rate (CR) is close to 100% in all scenarios, meaning there is a high correlation between AA and FA
- Power is maintained around 90% (not impacted by PFS)

Null Hypothesis – Base Approach

Conflict	Scenario ID	HR(PFS)	HR(OS)	Control Median(OS)	%AA	%FA	α^G
Design Assumptions	Scenario 5B	1	1	8.5 months	0.1	1.8	1.9
Large prior-data conflict	Scenario 6B	0.525	1	8.5 months	2.4	1.8	3.7

- Under the **design assumptions** (scenario 5) the global type I error is maintained below 2.5%
- When there is a large prior-data conflict with respect to PFS/OS historical relationship we have an increase in the global type I error rate, which is however maintained fairly low
- The contribution of the %FA on the global Type I error does not depend on HR(PFS), in fact the FA analysis is uniquely driven by OS concurrent data

Combined approach

Alternative Hypothesis – Combined Approach

- In case of minor PFS/OS prior data conflict using the surrogate prior makes %AA increase
- In case of inferior current controls, there is a decrease in %AA (due to underestimation of prior hazard control parameter)
- In case of superior current controls there is an increase in %AA (due to underestimation of prior hazard control parameter)

PFS/OS conflict	Scenario ID	HR(PFS)	HR(OS)	Control Median(OS)	%AA	CR	%FA (POWER)
Minor	Scenario 1B	0.39	0.71	8.5 months	72.4	96.4	89.7
Design Assumptions	Scenario 2B	0.525	0.71	8.5 months	58.0	97.4	89.7
Large	Scenario 3B	0.75	0.71	8.5 months	20.2	97.5	89.7
Extreme	Scenario 4B	1	0.71	8.5 months	0.5	>99.9	89.7
Minor	Scenario 1A	0.39	0.71	7 months	59.8	97.0	89.3
Design Assumptions	Scenario 2A	0.525	0.71	7 months	46.4	97.6	89.3
Large	Scenario 3A	0.75	0.71	7 months	14.5	97.9	89.3
Extreme	Scenario 4A	1	0.71	7 months	0.5	>99.9	89.3
Minor	Scenario 1C	0.39	0.71	10 months	78.4	96.2	89.4
Design Assumptions	Scenario 2C	0.525	0.71	10 months	65.8	97.0	89.4
Large	Scenario 3C	0.75	0.71	10 months	44.6	97.6	89.4
Extreme	Scenario 4C	1	0.71	10 months	0.6	>99.9	89.4

Scenarios A: Inferior current controls with respect to historical trials
Scenarios B: Comparable current controls with respect to historical trials
Scenarios C: Superior current controls with respect to historical trials

Null Hypothesis – Combined Approach

PFS/OS conflict	Scenario ID	HR(PFS)	HR(OS)	Control Median(OS)	%AA	%FA	α^G
Design Assumptions	Scenario 5B	1	1	8.5 months	<0.1	1.8	1.8
Large	Scenario 6B	0.525	1	8.5 months	1.7	1.8	3.1
Design Assumptions	Scenario 5A	1	1	7 months	<0.1	1.5	1.5
Large	Scenario 6A	0.525	1	7 months	1.2	1.5	2.5
Design Assumptions	Scenario 5C	1	1	10 months	<0.1	1.7	1.7
Large	Scenario 6C	0.525	1	10 months	2.3	1.7	3.5

Scenarios A: Inferior current controls with respect to historical trials
Scenarios B: Comparable current controls with respect to historical trials
Scenarios C: Superior current controls with respect to historical trials

- Under the **design assumptions** the global type I error is maintained
- In case of inferior current controls, there is a decrease in %AA (due to underestimation of prior hazard control parameter) → lower α^G
- In case of superior current controls there is an increase in %AA (due to underestimation of prior hazard control parameter) → higher α^G

Summary

Conflict	Scenario ID	HR(PFS)	HR(OS)	Control Median(OS)	%AA (base)	%AA (historical)	%AA (surrogate)	%AA (combined)
Minor	Scenario 1B	0.39	0.71	8.5 months	50.5	54.3	64.2	72.4
Design Assumptions	Scenario 2B	0.525	0.71	8.5 months	50.0	53.8	49.6	58.0
Large	Scenario 3B	0.75	0.71	8.5 months	22.5	24.4	16.8	20.2
Extreme	Scenario 4B	1	0.71	8.5 months	0.6	0.6	0.5	0.5

Conflict	Scenario ID	HR(PFS)	HR(OS)	Control Median(OS)	α^G (base)	α^G (historical)	α^G (surrogate)	α^G (combined)
Design Assumptions	Scenario 5B	1	1	8.5 months	1.9	1.9	1.8	1.8
Large	Scenario 6B	0.525	1	8.5 months	3.7	2.7	2.9	3.1

Under the **design assumptions** for median OS in current controls Combined approach seems the most promising one, in fact:

- improves the %AA when there is good accordance between current data and PFS/OS relationship
- Fairly maintains the %AA when there is major drift between current data and PFS/OS relationship
- Fairly maintains the α^G at any level of accordance between current data and PFS/OS relationship

Historical approach

Alternative Hypothesis – Historical Approach

- When current there is no drift with between current and historical controls (scenarios 1B-2B-3B-4B), there is an increase in %AA to Base Approach
- If current control is inferior to historical controls (scenarios 1A-2A-3A-4A) there is a reduction in %AA for any HR(PFS) with respect to Base Approach, due to the underestimation of the prior hazard control
- If current control is superior to historical controls (scenarios 1A-2A-3A-4A) there is an increase in %AA for any HR(PFS) with respect to Base Approach, due to the overestimation of the prior hazard control
- Conversion Rate is close to 100% in all scenarios, meaning there is a high correlation between AA and FA
- Power is maintained around 90% (not impacted by PFS)

	HR(PFS)	HR(OS)	Control Median(OS)	%AA	CR	%FA (POWER)
Scenario 1B	0.39	0.71	8.5 months	54.3	97.3	89.7
Scenario 2B	0.525	0.71	8.5 months	53.8	97.2	89.7
Scenario 3B	0.75	0.71	8.5 months	24.4	96.3	89.7
Scenario 4B	1	0.71	8.5 months	0.6	>99.9	89.7
Scenario 1A	0.39	0.71	7 months	40.8	98.0	89.3
Scenario 2A	0.525	0.71	7 months	39.4	98.0	89.3
Scenario 3A	0.75	0.71	7 months	16.5	97.0	89.3
Scenario 4A	1	0.71	7 months	0.5	>99.9	89.3
Scenario 1C	0.39	0.71	10 months	64.4	96.6	89.4
Scenario 2C	0.525	0.71	10 months	63.9	96.6	89.4
Scenario 3C	0.75	0.71	10 months	29.5	97.3	89.4
Scenario 4C	1	0.71	10 months	0.6	>99.9	89.4

Null Hypothesis – Historical Approach

	HR(PFS)	HR(OS)	Control Median(OS)	%AA	%FA	α^G
Scenario 5B	1	1	8.5 months	0.1	1.8	1.8
Scenario 6B	0.525	1	8.5 months	1.4	1.8	2.7
Scenario 5A	1	1	7 months	<0.1	1.5	1.5
Scenario 6A	0.525	1	7 months	1.2	1.5	2.5
Scenario 5C	1	1	10 months	<0.1	1.7	1.7
Scenario 6C	0.525	1	10 months	2.6	1.7	3.7

- When there is no drift with between current and historical controls (scenarios 5B-6B), global type I error is close to Base Approach (<2.5%!!)
- In case of inferior current controls (scenarios 5A-6A) we have a decrease in %AA and accordingly a decrease in the global type I error
- In case of superior current controls (scenarios 5C-6C) we have an increase in %AA and accordingly an increase in the global type I error, which is however fairly maintained

Surrogate approach

Alternative Hypothesis – Surrogate Approach

	HR(PFS)	HR(OS)	Control Median(OS)	%AA	CR	%FA (POWER)
Scenario 1B	0.39	0.71	8.5 months	64.2	97.4	89.7
Scenario 2B	0.525	0.71	8.5 months	49.6	98.2	89.7
Scenario 3B	0.75	0.71	8.5 months	16.8	98.8	89.7
Scenario 4B	1	0.71	8.5 months	0.5	>99.9	89.7

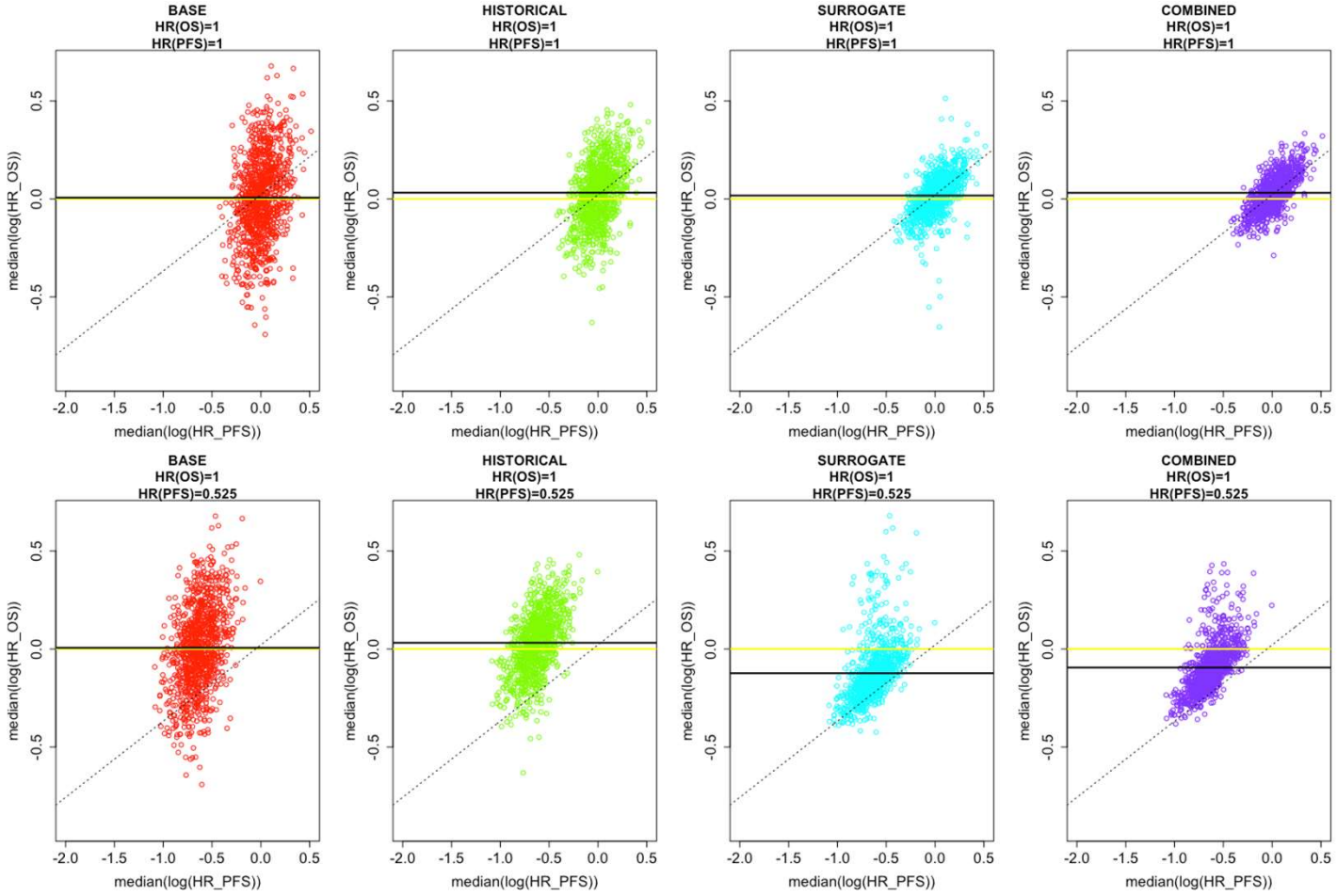
- When current HR(PFS) and HR(OS) present no drift with respect to PFS/OS historical relationship there is an increase in the chance to get AA with respect to base approach
- Under the design hypothesis (scenario 2B) minor impact on AA is observed with respect to base approach, while a sensible reduction in AA is observed in presence of large or extreme prior-data conflict (scenarios 3B-4B)
- Conversion Rate is close to 100% in all scenarios, meaning there is a high correlation between AA and FA
- Power is maintained around 90% (not impacted by PFS)

Null Hypothesis – Surrogate Approach

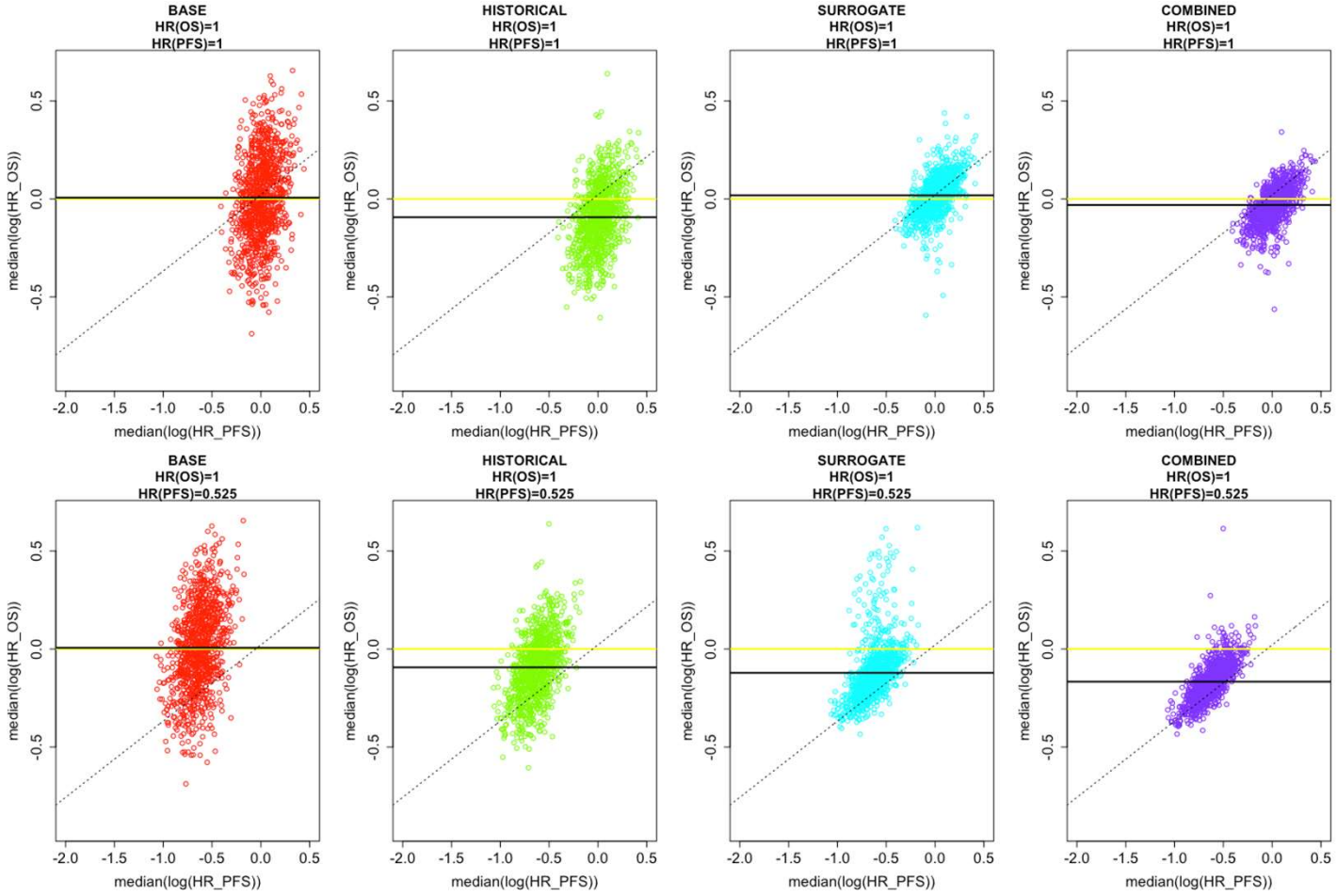
	HR(PFS)	HR(OS)	Control Median(OS)	%AA	%FA	α^G
Scenario 5B	1	1	8.5 months	<0.1	1.8	1.8
Scenario 6B	0.525	1	8.5 months	1.4	1.8	2.9

- When current HR(PFS) and HR(OS) present no drift with respect to PFS/OS historical relationship (scenario 5B) the global type I error is maintained below 2.5%
- When there is a large prior-data conflict with respect to PFS/OS historical relationship we have an increase in the global type I error rate, which is however maintained fairly low
- The contribution of the %FA on the global Type I error does not depend on HR(PFS), in fact the FA analysis is uniquely driven by OS concurrent data

IA Posterior Inference (comparable controls)



IA Posterior Inference (superior controls)



IA Posterior Inference (inferior controls)