ADAPTIVE SEAMLESS DESIGNS FOR SUBPOPULATION SELECTION BASED ON TIME TO EVENT ENDPOINTS

Emmanuel Zuber[1], Werner Brannath[2], Michael Branson[1], Frank Bretz[1], Paul Gallo[3], Martin Posch[2], Amy Racine-Poon[1]

[1] Novartis Pharma AG, Basel, Switzerland
[2] Section of Medical Statistics, Medical University of Vienna, Austria
[3] Novartis Pharmaceuticals Corp, East Hanover, NJ, USA

Outline

- Real life example:
  Confirmatory adaptive trial using bayesian decision tools
- Rationale and development strategy
- Study design and methodology
- Bayesian decision tools and processes
- Case study design characteristics and simulations
- Conclusions: advantages of the chosen strategy

Rationale for an adaptive design strategy

Targeted therapy might primarily benefit a subpopulation

- Pathway implicated in oncogenesis, progression, resistance to treatment

Evidence of activity

- Preclinically & Clinically
- But requires better definition of biological characteristics of benefiting patients

Traditional approach to identify & confirm a sensitive subpopulation:

- Exploratory trial(s) to identify subpopulation with greater benefit
- Phase II to confirm greater benefit in identified subpopulation
- Phase III trial in the chosen target population (full or subpopulation)

Ethical and strategic relevance of allowing

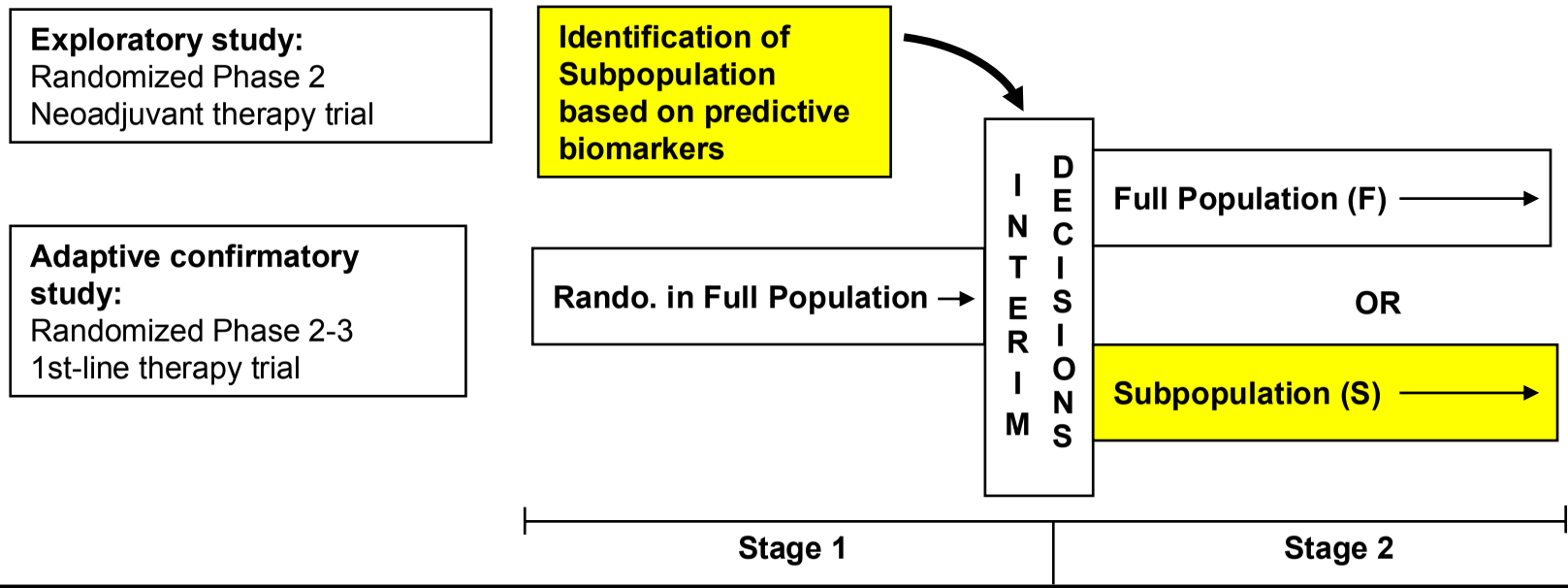
- Focus as early as possible on subpopulation, if it can be defined
- Efficient use of data from patients needed to confirm the subpopulation

=> **Integrate Phase II & III objectives in a single adaptive trial**

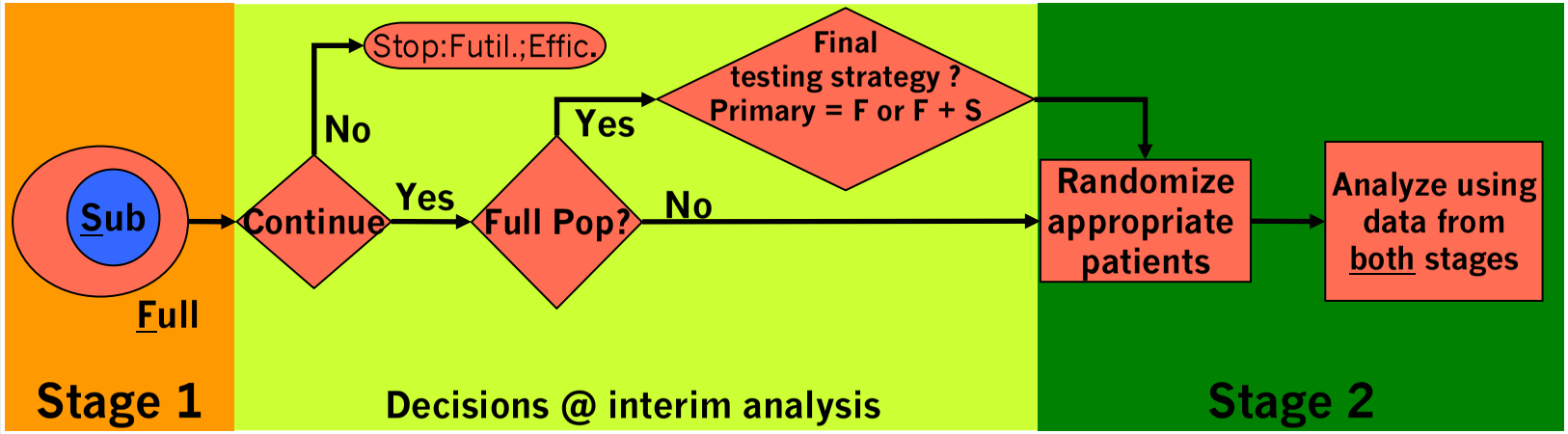
Development strategy

Adaptive trial: two stages, with an interim analysis, to simultaneously meet

- Phase II objectives
  - to confirm greater benefit in independently identified Subpopulation
  - to decide whether or not to adapt trial to focus on that Subpopulation
- Phase III objective
  - to demonstrate superiority on time to event (phase III) endpoint



Adaptive design



Stage 1: Futility stop or subpopulation selection (Bayesian tools)

- Subpopulation defined prior to interim analysis (external to trial)
- Probabilities of false positive and false negative decisions described a-priori via simulations

Stage 2: Confirmation of benefit while maintaining integrity

- Combining evidence from first and second stage
- False positive rate controlled by method, simulation used to explore power

Formal testing: type I error rate control (1)

Multiplicity issues

- Testing in 2 populations, group sequential testing (2 stages)
- Stage 2 adapted based on stage 1 data

Adaptive design methodology

- Independent p-values from 2 stages combined: inverse normal method (Lehmacher & Wassmer 1999, Cui, Hung & Wang 1999)

  $C(p_1, p_2) = w_1 \Phi^{-1}(1-p_1) + w_2 \Phi^{-1}(1-p_2)$ (at stage 2)

  - Pre-specified weights: $w_1 = (n_1/N)^{1/2}$, $w_2 = (n_2/N)^{1/2}$, $w_1^2 + w_2^2 = 1$
    n_i/N event fraction at stage i
  - Nominal α-levels: O'Brien-Fleming α-spending function
- Time to event: Independent p-values?
  - Logrank asymptotic independent increments property

Formal testing: type I error rate control (2)

Closed testing procedure

- Hochberg p-value q_i for intersection hypothesis $H_{0F} \cap H_{0S}$ at stage i
- Final testing strategies (analysis in F or S: primary/secondary):
  - Different p-values (q_i, p_{iS} or p_{iF}) in $C(p_1, p_2)$,
  - for intersection and elementary hypotheses testing

Example

- At stage i= 2, trial continued in the Full population
  tests in F and S of equal interest (H_{0F} & H_{0S} co-primary)
  - Intersection hypothesis $H_{0F} \cap H_{0S}$ tested using Hochberg p-value for i=1, 2:
    $q_i = \min\{2 \min(p_{iS}, p_{iF}), \max(p_{iS}, p_{iF})\}$
  - Testing strategy at stage 2:

reject $H_{0F} \cap H_{0S}$ if		$C(q_1, q_2) \geq c_2$	(c_2 from α-spending function)
then	reject H_{0F} if	$C(p_{1F}, p_{2F}) \geq c_2$	
	reject H_{0S} if	$C(p_{1S}, p_{2S}) \geq c_2$	

Adaptation decisions: Bayesian tools (1)

Bayesian tools:

- Predictive powers:
  - Probability of success in each of the possible stage 2 situations (F or S)
- Posterior probability:
  - Probability that the patients in S^c (outside the Subpop.) do not benefit

Decision rules:

- Predictive powers in F and in S < common threshold $\pi\{F, S\}$
  ⇨ stop for futility
- Only predictive power in S > threshold $\pi\{S\}$
  or
  Probability (treat. effect in S^c < target) > threshold $\pi\{S^c\}$
  ⇨ go with Subpopulation
- Otherwise
  ⇨ go with Full population

Adaptation decisions: Bayesian tools (2)

Calculated using Logrank statistics and ln(HR) assumptions

- Logrank = $U/\sqrt{I}$ ~ $\mathcal{N}(\theta\sqrt{I}, 1)$, U: efficient score, I : Fisher's information
- Ln(HR) $\approx U/I$ ~ $\mathcal{N}(\theta, I^{-1})$, with $I^{-1} \approx 4/n$, n: Number of events

Prior and likelihood distributions

- Some prior belief (or no prior belief) on expected treatment effect(s)

  $[\theta] \sim \mathcal{N}(\mu, 4/n_0)$, with θ=ln(HR), n_0=number of events in the prior
- Data at the interim analysis (after n_1 observed events), let $U_1/I_1 = \Delta_1$

  $[\Delta_1 \mid \theta] \sim \mathcal{N}(\theta, 4/n_1)$, likelihood of observed effect Δ_1 at stage 1

Adaptation decisions: Bayesian tools (3)

The posterior distribution of θ is given by:

- $[\theta \mid \Delta_1] \sim \mathcal{N}(\mu_1, \tau_1^2)$, with

  $\mu_1 = \tau_1^2[(4/n_1)^{-1} \Delta_1 + (4/n_0)^{-1} \mu]$, weighted average (weights=Information)

  $\tau_1^2 = [(4/n_1)^{-1} + (4/n_0)^{-1}]^{-1}$, Information=sum of prior & likelihood information

- For patients outside the Subpopulation, $\Delta_1 = \Delta_{NotS}$ and $n_1 = n_{NotS}$

Adaptation decisions: Bayesian tools (4)

Logrank asymptotic independent increments property

$p_{2F} = \Phi\{-(U_{2F}-U_{1F})/\sqrt{(I_{2F}-I_{1F})}\}$ (at stage 2, in the Full population)

The predictive powers are calculated (closed forms), from:

- Predictive distribution of Δ_2

  - Δ_2: « incremental » ln(HR) vector to be observed in stage 2 after n_{2F} additional events in Full population, n_{2S} in Subpopulation (information to come)

$$[\Delta_2|\Delta_1] = \mathcal{N}\left(\begin{pmatrix}\mu_{1S}\\ \mu_{1F}\end{pmatrix}, \begin{pmatrix}\tau_{1S}^2+4/n_{2S} & \tau_{1F}^2+4/n_{2F}\\ \tau_{1F}^2+4/n_{2F} & \tau_{1F}^2+4/n_{2F}\end{pmatrix}\right)$$

Adaptation decisions: Bayesian tools (5)

The predictive powers are calculated (closed forms), from:

- Hypothesis testing strategy
  - rejection regions for stage 2 p-values, depending on the interim results *via* the combination function
  - *i.e.* inverting $C(p_1, p_2) = w_1 \Phi^{-1}(1-p_1) + w_2 \Phi^{-1}(1-p_2) \geq c_2$
  - taking into account the closed testing procedure

=> Predictive power = probability of stage 2 incremental effect to give a stage 2 p-value inside the appropriate rejection region

Case study

- **Advanced/metastatic disease setting**
- **Phase III endpoint**: Progression Free Survival (PFS)
  - Target treatment effect: HR=0.77 (test/control)
  - Target number of events: 918 in Full population or 640 in Subpop.
  - Fixed total sample size of 1200 patients
- **Interim analyses**
  - First interim (Efficacy, Futility, Adaptation):

    - 170 Full pop. events (~19% target Full pop. events)
  - Second interim (Efficacy only):

    - Classical interim analysis, added to the initial 2-stage adaptive design

    - 551 Full pop. events, or 384 Subpop. events (if Subpop. selected)

    - i.e. 60% of target number of events, ½ way through stage 2

Optimization of trial design (simulation)

- **Sample size, power, trial timeframe** (event driven, time-to-event endpoint)
  - Different final sample sizes in F and S (fixed total N)
  - Different recruitment rates in F and S (bounded timeframe)
- **Timing of 1st interim analysis**
  - Enough information for decision making
  - Enough time/patients left to recruit in stage 2
- **Operating characteristics** vs. Subpop. prevalence and decision thresholds
  - Chances to stop (futility, efficacy) vs. to continue
  - Chances to go with relevant Subpopulation vs. not to go with irrelevant Subpopulation

Power simulations

Under $H_{1S} \cap H_{1S}^c$ (All patients benefit equally from treatment):

- **Conventional phase III** (no interim analysis): 98% power
- **Conventional phase III with interim** (Effic./Futility): 88% power
- **Adaptive design phase III:** 87% power
  (across a variety of values of Subpopulation prevalence)

Under $H_{1S} \cap H_{0S}^c$ (If only S benefits):

		Overall power	
S prevalence	**Adaptive ph. III**	**Conventional sequential ph. III**	**Conventional seq. ph. III, test in F+S**
30%	57%	16%	39%
40%	65%	28%	52%
50%	71%	41%	62%

(with $\pi\{F, S\} = 35\%$, $\pi\{S^c\} = 90\%$)

Decision making simulations

- with $\pi\{F, S\} = 35\%$, $\pi\{S^c\} = 90\%$

S preval.	Effect in subpop. S	Effect in outside subpop. S^c	Probability to Stop for futility	Probability to Go with F	Probability to Go with S
30%	No	No	**50%**	27%	23%
	Yes	No	20%	38%	**42%**
	Yes	Yes	8%	**85%**	7%
50%	No	No	**60%**	24%	16%
	Yes	No	15%	45%	**40%**
	Yes	Yes	8%	**84%**	8 %

Effect = Yes: HR=0.77 test/control (target effect),
No: HR=1

Decision making simulations

- with $\pi\{F, S\} = 35\%$, $\pi\{S^c\} = 90\%$ / **$\pi\{S^c\} = 75\%$**
  (less conserv. against continuing in S)

	Effect in		Probability to		
S preval.	subpop. **S**	outside subpop. **S^c**	**Stop for futility**	**Go with F**	**Go with S**
30%	No	No	**50%**	27%	23%
	Yes	No	20%	38%	**42%**
	Yes	Yes	8%	**85%**	7%
50%	No	No	**60%**	24%	16%
	Yes	No	15%	45%/28%	**40%/57%**
	Yes	Yes	8%	**84%**/71%	8 %**/21%**

Effect = Yes: HR=0.77 test/control (target effect),
No: HR=1

Conclusion: The development plan using an adaptive design

- **More reliable decision**: Subpopulation confirmed on
  - Phase III endpoint,
  - Large sample in advanced disease setting
- **Efficient use of data from stage 1 patients**
  - Test regimen superiority shown on stage 1 & 2 patients at final analysis
  - Stage 1 patients contribute to both Phase II and Phase III objectives
- **Higher or equal probability of success** than conventional phase III
  - ignoring Subpopulation, or
  - testing in F & S but always going with Full population
- **Time savings**: Start Phase III (~1-2 years)
  - No waiting for the Subpopulation to be identified and confirmed
  - Subpopulation identified concomitantly in a separate exploratory trial
- **Patients savings**: No extra Phase II to confirm Subpopulation (~ 200pts)
  - Replaced by stage 1 with Phase II objectives

Conclusion: Bayesian tools

- **Flexible decision making**
  - No inflation of the type I error rate
  - Possibility to optimize thresholds up until interim analysis
- **Sophisticated decision making**
  - Relevance of treatment effect outside subpop. ? (no specific test planned)
  - Robustness of decision assessed using various prior distributions
- **Communication tool**: Realistic answers to clinical questions
  - Chances to be successful ? Risk of no clinically relevant treatment effect ?
  - Reflects uncertainty on prior belief, on current data, on data to come
- **Efficient and flexible use of all data available**
  - Exploratory study results could be used in prior distribution
  - Use of accumulating knowledge to refine threshold setting