Study designs

Valter Torri, Luca Porcu

Laboratory of Methodology for Biomedical Research, Department of Oncology

IRCCS – Istituto di Ricerche Farmacologiche ‘Mario Negri’, Milan (Italy)

Declaration of disclosure

- No conflicts of interest to declare

Study designs in the last 15 years (1999-2013)

Phase II trials

N (%)		GIST		Non-GIST Sarcoma	
Endpoint	RR	19	(90)	108	(79)*
	PFS	2	(10)	28	(21)
Sample size	Median	30		41	
	Range	13-147		7-270	
Design	Single-arm	19	(90)	121	(89)
	Randomized	2	(10)	15	(11)
With stopping rules (only futility/efficacy stopping rules were applied)	Yes	12	(57)	114	(84)
	No	9	(43)	22	(16)
Statistical techniques	Frequentist	21	(100)	133	(97)^
	Bayesian	0	(0)	3	(3)

** Four studies with combined endpoint: PFS+RR*

^ One study with bayesian decision rules for randomization, frequentist techniques for other statistical issues

Study designs in the last 15 years (1999-2013)

Phase III trials, adjuvant setting

N (%)		GIST		Non-GIST Sarcoma	
Endpoint	DFS	1	(100)	3	(75)
	OS	0	(25)	1	(25)
Sample size	Range	713		81-504	
Design	Two-arm parallel	1	(100)	4	(100)
With stopping rules (only futility/efficacy stopping rules were applied)	Yes	1	(100)	3	(75)
	No	0		1	(25)
Statistical techniques	Frequentist	1	(100)	4	(100)
	Bayesian	0	(0)	0	(0)

Study designs in the last 15 years (1999-2013)

Phase III trials, advanced setting

N (%)		GIST		Non-GIST Sarcoma	
Endpoint	PFS	3	(75)	2	(33)
	RR	0		3	(50)
	OS	1	(25)	1	(17)
Sample size	Range	81-946		162-711	
Design	Two-arm parallel	4	(100)	5	(83)
				1: three-arm parallel	
With stopping rules (only futility/efficacy stopping rules were applied)	Yes	1	(25)	6	(100)
	No	3	(75)	0	
Statistical techniques	Frequentist	4	(100)	6	(100)
	Bayesian	0		0	

Interim monitoring: the futility design

Single arm phase II trial

Outcome: proportion of tumor responses (CR/PR)

Hypotheses: H0: p=0.20, H1: p=0.30

Errors: $\alpha_{\text{One-sided}} = 0.10$ $\beta = 0.20$

Without an interim analysis

Sample size: 88 pts

Expected sample size: 88 pts

Real errors: $\alpha_{\text{One-sided}} = 0.098$, $\beta = 0.183$

Performing the analysis ad interim (Stopping rule for futility*)

** Optimal Two-Stage design*

Sample size: 89 pts

Expected sample size (H0): 63.1 pts

Real errors: $\alpha_{\text{One-sided}} = 0.098$, $\beta = 0.199$

Providing evidences with small clinical trials

Statistical strategies:

- Improving design efficiency
- All or nothing bet
- Considering different the levels of evidence

Adaptive designs: definition

- "A clinical study design that uses accumulating data to decide how to modify aspects of the study as it continues,

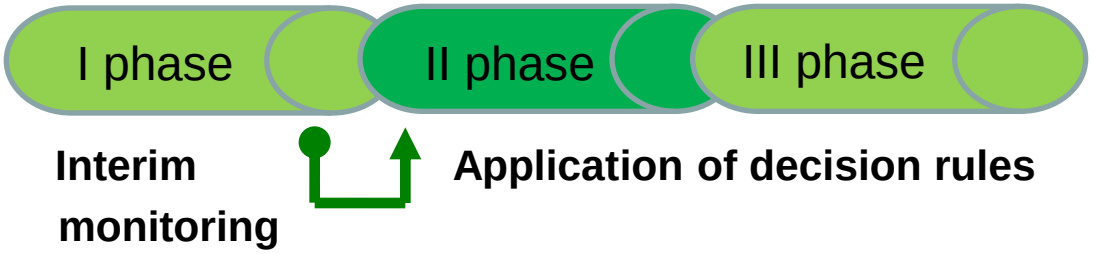
- without undermining the validity and integrity of the trial"

Gallo P. et al. Journal of Biopharmaceutical Statistics, vol.16: 275-283, 2006

Adaptive designs: the traditional model

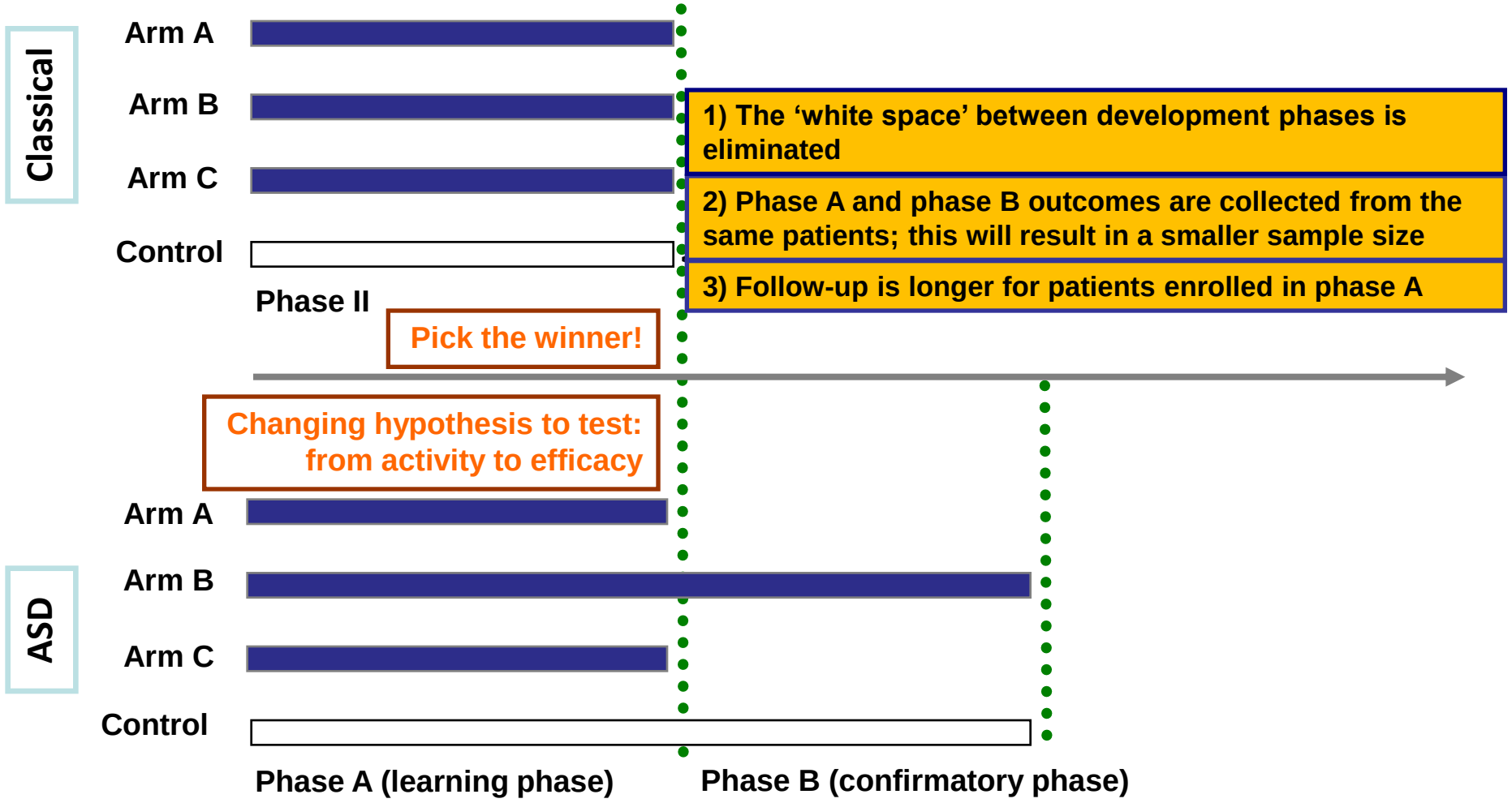
- **Clinical trial is performed in sequential phases**
- **At the end of each phase**

  1. data analysis is performed
  2. one or more planned decision rules are applied



Decision rules	
Name	**Decision rule for…**
Allocation rule	allocating patients to treatments
Sampling rule	determining the sample size for the subsequent phase
Stopping rule	stopping the trial for efficacy, safety or futility
Other decision rules	Statistical hypothesis to test (e.g. from superiority to non-inferiority); target population (e.g. changing eligibility criteria), etc.

Adaptive designs: ASD, an example



Adaptive designs: the bad model

"In such trials, changes are made 'by design' and ***not on an ad hoc basis***; therefore, adaptation is a design feature aimed to enhance the trial, ***not a remedy for inadequate planning***"

Gallo P. et al. Journal of Biopharmaceutical Statistics, vol.16: 275-283, 2006

Providing evidences with small clinical trials

Statistical strategies:

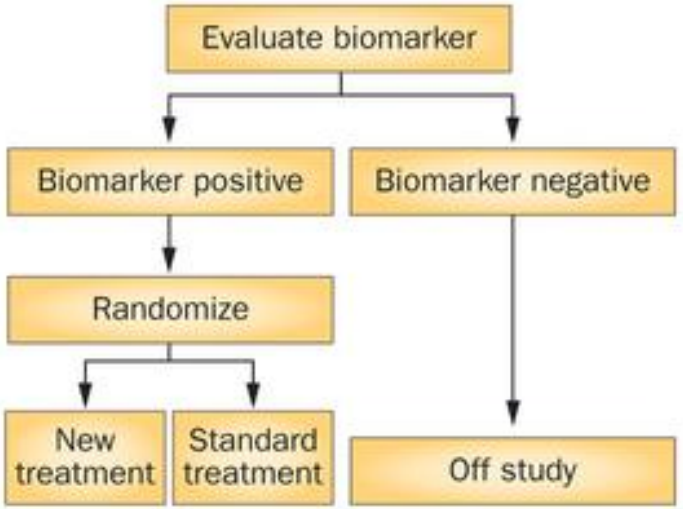
- Improving design efficiency
- All or nothing betting
- Considering different the levels of evidence

Enrichment design (ED)

Convincing evidence indicates that the benefits of the treatment are limited to the biomarker-positive subgroup

From strong proof of concept to ED

An ED evaluates the new treatment only in the biomarker positive subpopulation



Freidlin B., Korn E., Biomarker enrichment strategies: matching trial design to biomarker credentials, Nature Reviews, Clinical Oncology, 1-10, 2013

Providing evidences with small clinical trials

Statistical strategies:

- Improving design efficiency
- All or nothing betting
- Considering different the levels of evidence

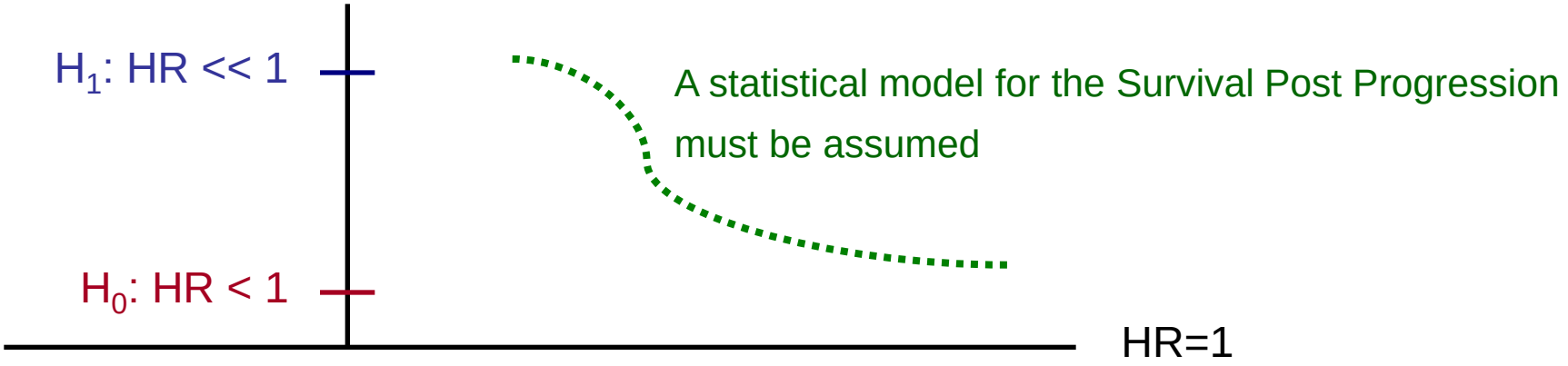
Overall Survival in small clinical trials

HR	Deaths
0.9	2829
0.8	631
0.7	247
0.6	121
0.5	66
0.4	38

Allocation ratio of 1:1;
$alpha_{Two\text{-}sided}$=5%; power=80%

In small clinical trials it's extremely difficult to demonstrate a risk reduction <50%

PFS: a screening endpoint for clinical benefit



Raising the PFS bar…

1. A more reasonable sample size is requested
2. Specifity, i.e.: prob. of rejecting uneffective treatments on OS is improved
3. Sensitivity, i.e .: prob. of accepting effective treatments on OS maybe not

Surrogate endpoint: is a feasible concept?

Prentice's criteria could be demonstrated?

1. Correlation between the surrogate endpoint and the true endpoint
2. An effect of treatment on the true endpoint should be detected
3. An effect of treatment on the surroga detected
4. The effect of treatment on the true e ost all explained by the surrogate endpoint

equivocal estimates

almost impossible

Increasing alpha and beta errors

Single arm phase II trial

Outcome: proportion of tumor responses (CR/PR)

Hypotheses: H0: p=0.20 H1: p=0.30 (One-sided test)

Alpha\Beta	0.10	0.15	0.20	0.25	0.30
0.05	**160**	137	116	99	86
0.10	127	101	88	74	62
0.15	105	83	70	60	48
0.20	91	72	55	46	**37**

Number of tumor responses

Min-max: 37-160; **Sample size reduction**: 23%

Increasing alpha and beta errors

Parallel group phase III trial (assignment ratio of 1:1)

Outcome: Overall Survival

Hypotheses: H0: HR=1 H1: p=0.70 (Two-sided test)

Alpha\Beta	0.10	0.15	0.20	0.25	0.30
0.05	**331**	283	247	219	195
0.10	270	227	195	170	148
0.15	233	193	164	141	122
0.20	207	169	142	121	**103**

Number of events according to the D.A.Schoenfeld formula

Min-max: 103-331; **Sample size reduction**: 31%

Indirect evidence: imatinib in GIST and non-GIST population

Imatinib mesylate (STI-571 Glivec®, Gleevec™) is an active agent for gastrointestinal stromal tumours, but does not yield responses in other soft-tissue sarcomas that are unselected for a molecular target: Results from an EORTC Soft Tissue and Bone Sarcoma Group phase II study

J. Verweij[a,*], A. van Oosterom[b], J.-Y. Blay[c], I. Judson[d], S. Rodenhuis[e], W. van der Graaf[f], J. Radford[g], A. Le Cesne[h], P.C.W. Hogendoorn[i], E.D. di Paola[j], M. Brown[j], O.S. Nielsen[k]

Primary endpoint: Respose Rate (RR)

Design: One-stage design (P_0=10%, P_1=30%, α=10%, β=10%),

2 strata (GIST e non-GIST sarcoma)

Results	Activity
GIST RR:	19/27
Non-GIST sarcomas RR:	0/24

European Journal of Cancer 39 (2003) 2006–2011

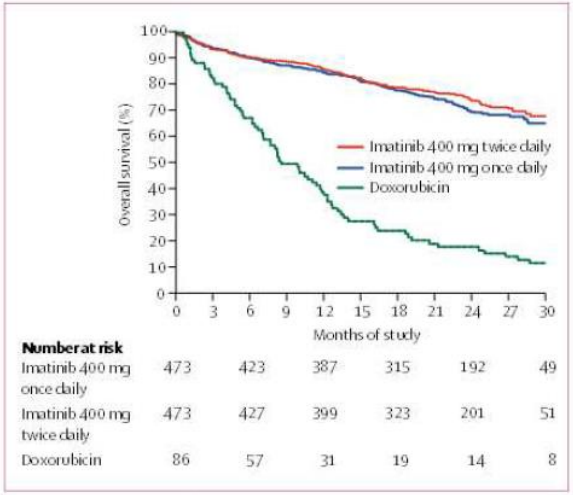
Non Randomized controls: imatinib in GIST

Progression-free survival in gastrointestinal stromal tumours with high-dose imatinib: randomised trial

*Jaap Verweij, Paolo G Casali, John Zalcberg, Axel LeCesne, Peter Reichardt, Jean-Yves Blay, Rolf Issels, Allan van Oosterom, Pancras C W Hogendoorn, Martine Van Glabbeke, Rossella Bertulli, Ian Judson, for the EORTC Soft Tissue and Bone Sarcoma Group, the Italian Sarcoma Group, and the Australasian Gastrointestinal Trials Group**

Summary

Background Imatinib is approved worldwide for use in gastrointestinal stromal tumours (GIST). We aimed to assess dose dependency of response and progression-free survival with imatinib for metastatic GIST.



Number at risk						
Imatinib 400 mg once daily	473	423	387	315	192	49
Imatinib 400 mg twice daily	473	427	399	323	201	51
Doxorubicin	86	57	31	19	14	8

Figure 6: **Overall survival for total study population**
Data are compared with historical (GIST) controls from the EORTC database. Dox=doxorubicin-based regimen

"We compared survival data (OS) of our study with those from the EORTC database on pts who received doxorubicin based CT for GIST as first line treatment"

"Even in view of the limitations to this approach the difference in OS is so striking that to attribute this finding to chance is difficult"

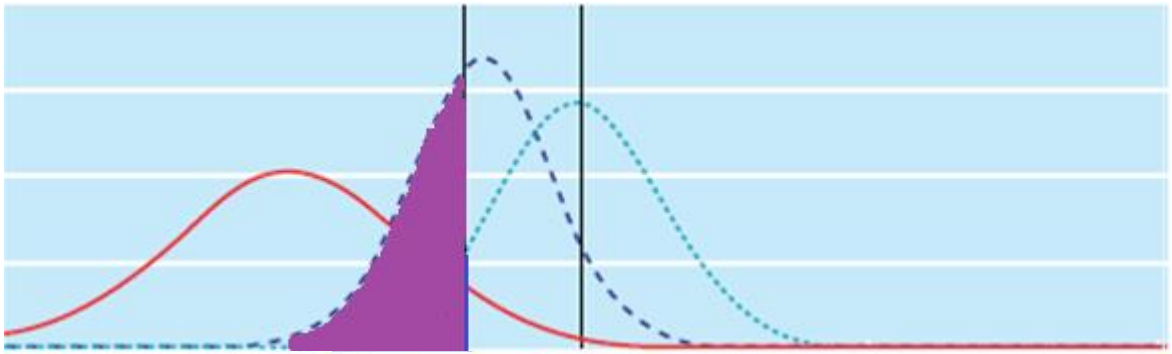
Nonrandomized controls

The experimental drug must demonstrate a remarkable benefit.

Selection bias

- Overt bias is controlled by matching, stratification, regression models
- Hidden bias can be controlled by sensitivity analysis

Indirect evidence: Bayes machine



Δ Null effect

Posterior ~ likelihood data x Prior

$Weight_{prior}$ and $weight_{likelihood}$: inverse variance

$Weight_{prior} = 0$ → Posterior = likelihood

From bayesian statistics to classical statistics

Prior probability H_1: 5%

Posterior probability H_1: 45%

Indirect evidence: similar diseases

VOLUME 27 · NUMBER 19 · JULY 1 2009

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Phase II Multicenter Trial of Imatinib in 10 Histologic Subtypes of Sarcoma Using a Bayesian Hierarchical Statistical Model

Rashmi Chugh, J. Kyle Wathen, Robert G. Maki, Robert S. Benjamin, Shreyaskumar R. Patel, Paul A. Myers, Dennis A. Priebat, Denise K. Reinke, Dafydd G. Thomas, Mary L. Keohan, Brian L. Samuels, and Laurence H. Baker

A B S T R A C T

Purpose

The purpose of this trial was to assess the efficacy of imatinib in patients with one of 10 different subtypes of advanced sarcoma.

Primary endpoint: clinical benefit response (CR/PR within 16 weeks or SD lasting at least 16 weeks)

One-sided interval under investigation (H_1): clinical benefit response (CBR) > 0.30