High-dimensional biomarkers as treatment modifiers in randomized clinical trials

Stefan Michiels, PhD
Head of Methodology
Department of Biostatistics and Epidemiology
Gustave Roussy, Paris-Sud University, Villejuif, France
stefan.michiels@gustaveroussy.fr

Disclosure

I have no financial relationships to disclose.

I will not discuss off label use and/or investigational use in my presentation.

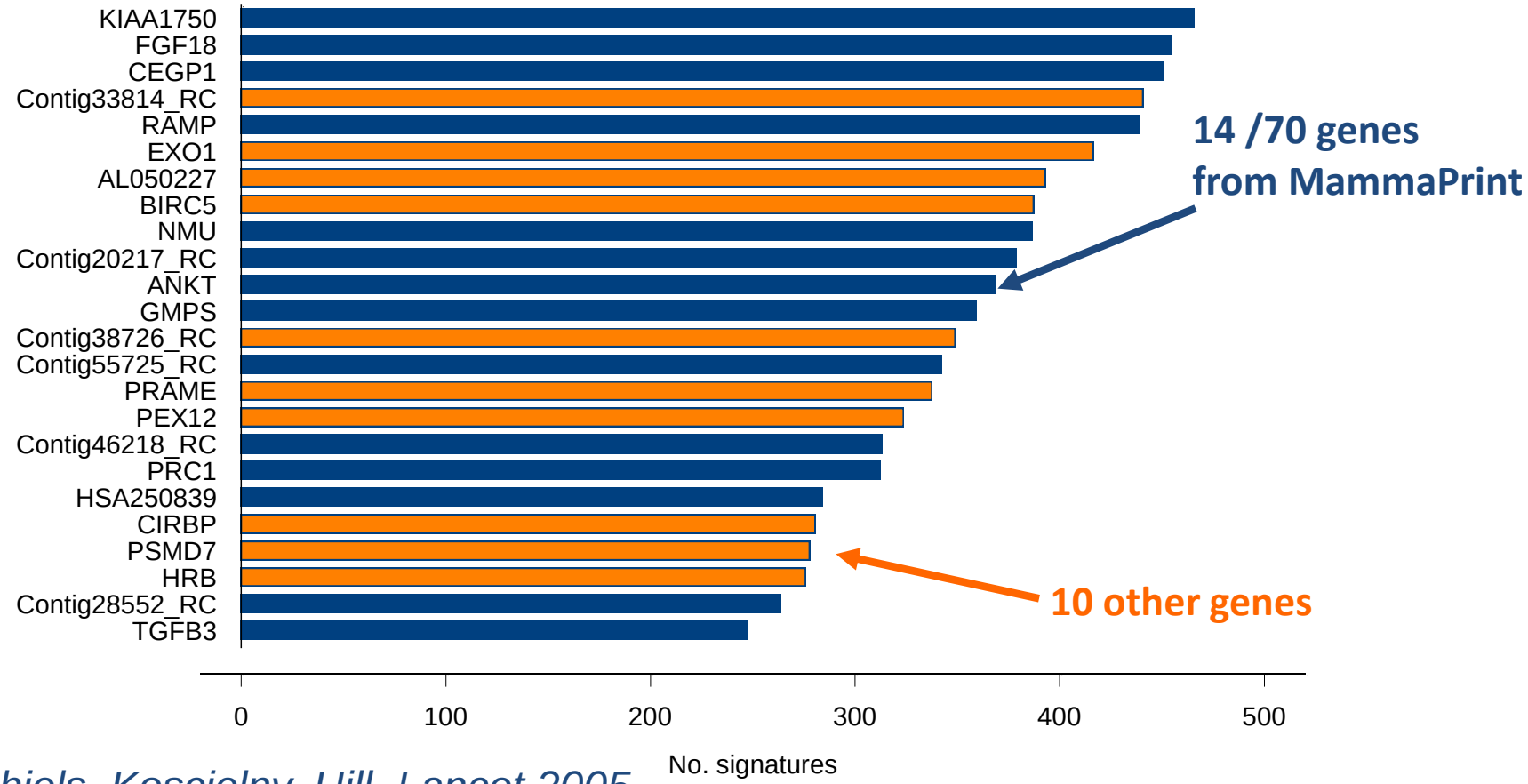
Lessons from early breast cancer: prognostic signatures

Available gene signatures for a price of 400-4175$…

- IHC4: 4 genes
- Oncotype Dx: 16 cancer genes
- PAM50 (ROR): 50 genes
- Mammaprint Dx: 70 genes
- Endopredict: 8 cancer genes
- Mapquant Dx (GGI) : 97 genes
- …

Instability of gene selection in original Mammaprint training data

Genes included in at least 250 out of 500 (50%) signatures for a training set size of 78 patients



Michiels, Koscielny, Hill. Lancet 2005

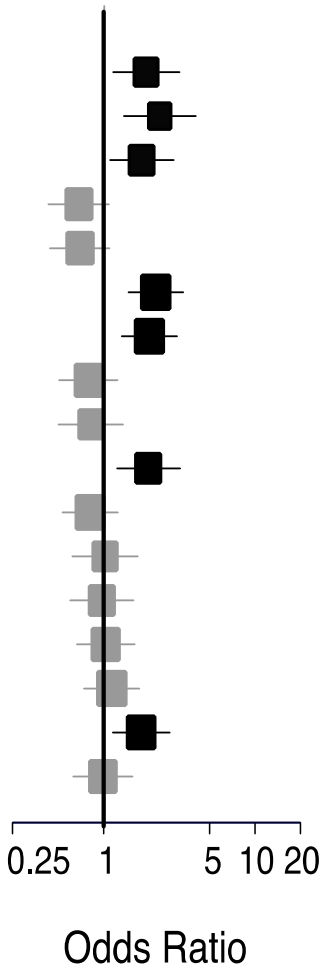
Association of gene modules with pathological complete response after anthracyclines, beyond clinicopathological factors

OR: odds ratio for a 1-unit change in gene module, adjusted for clinicopathological factors

FDR: false discovery rate

ALL
(845 pts, 189 pCR)

	OR	95% CI	P	FDR
GGI	1.7	(1.12,2.6)	1.3E−02	3.7E−02
Gene70	2.02	(1.29,3.2)	2.4E−03	1.3E−02
CIN70	1.61	(1.08,2.42)	2.1E−02	5.1E−02
Stroma1	0.73	(0.49,1.06)	1.0E−01	2.1E−01
Stroma2	0.74	(0.5,1.07)	1.1E−01	2.1E−01
Immune1	1.92	(1.36,2.73)	2.2E−04	3.7E−03
Immune2	1.78	(1.25,2.53)	1.3E−03	1.1E−02
RAS	0.82	(0.57,1.18)	3.0E−01	4.9E−01
MAPK	0.85	(0.56,1.27)	4.2E−01	6.0E−01
PTEN	1.75	(1.18,2.62)	5.8E−03	2.5E−02
AKTmTOR	0.84	(0.59,1.19)	3.2E−01	4.9E−01
PIK3CA	1.01	(0.67,1.53)	9.5E−01	9.5E−01
IGF1	0.97	(0.65,1.45)	8.9E−01	9.5E−01
SRC	1.02	(0.71,1.47)	9.1E−01	9.5E−01
MYC	1.1	(0.78,1.56)	5.8E−01	7.6E−01
E2F3	1.6	(1.12,2.3)	1.1E−02	3.7E−02
BetaCatenin	0.98	(0.68,1.43)	9.4E−01	9.5E−01



Ignatiadis et al JCO 2012

Move to treatment-effect modifiers

Past:

- Development of prognostic signatures by a model

$$Outcome \sim biomarker$$

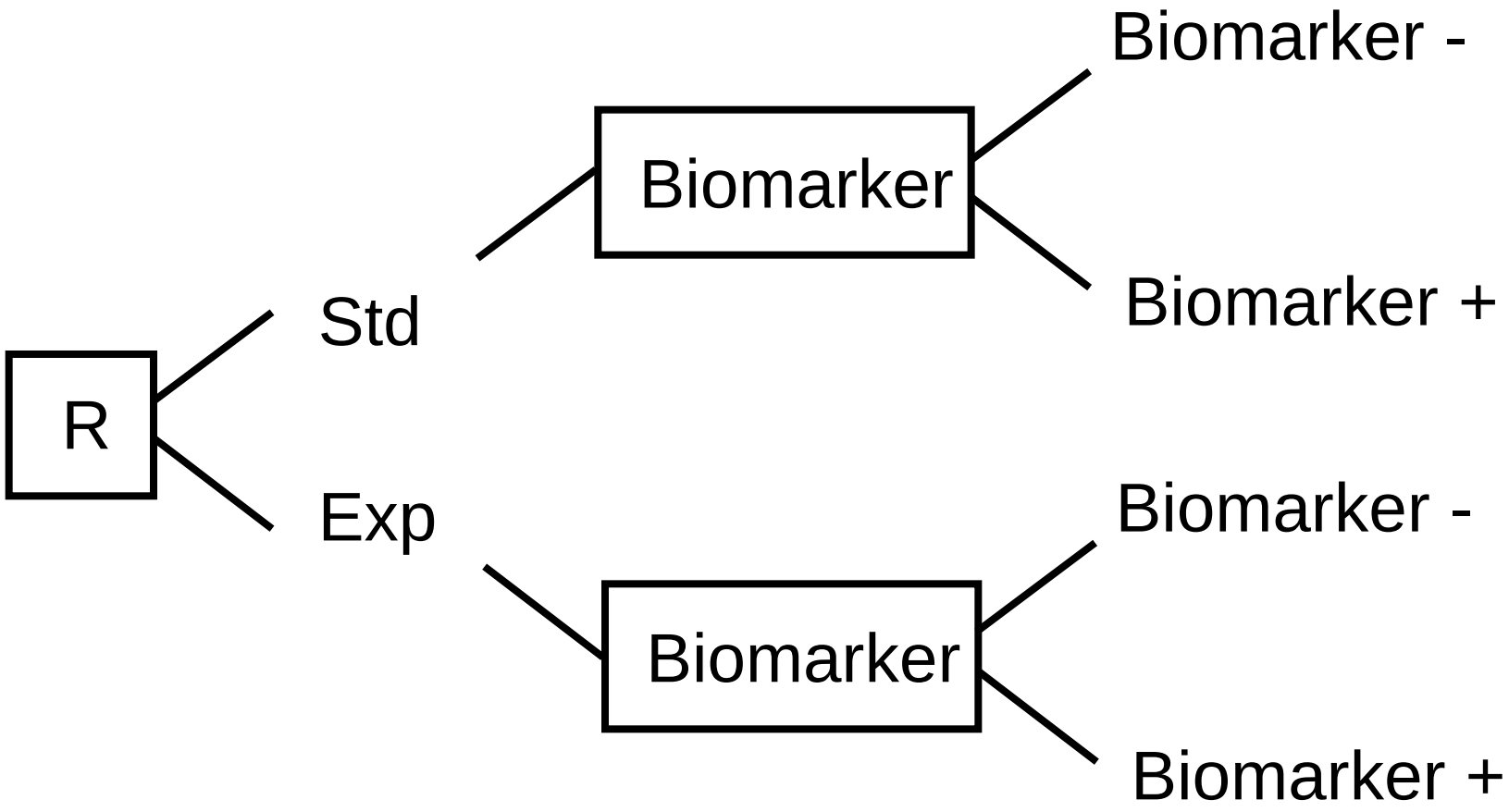
- e.g. 99% of published breast cancer signatures

Future:

- Development of “predictive” or treatment-effect modifying signatures

$$Outcome \sim biomarker + treatment + \mathbf{treatment \times biomarker}$$

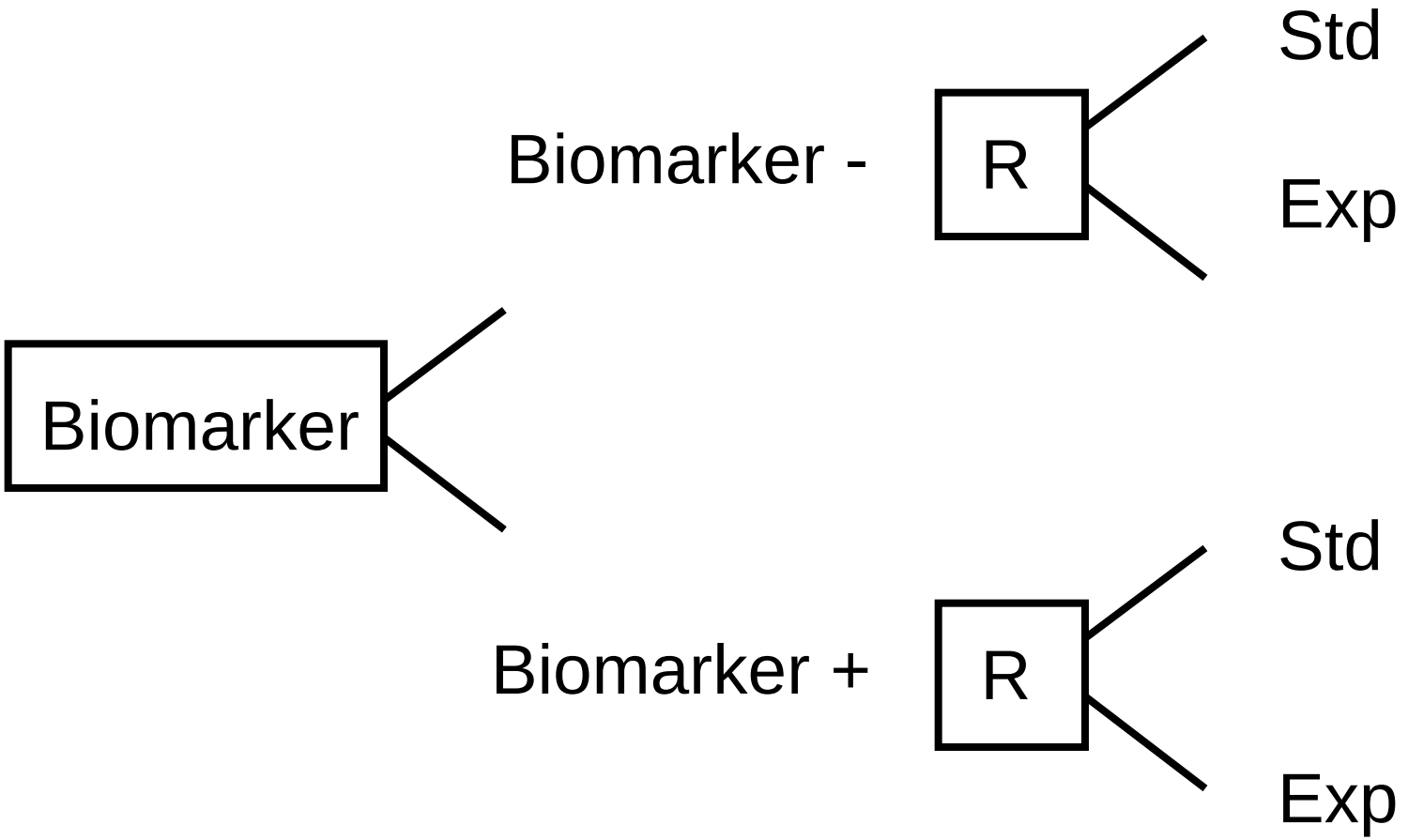
Randomize-all design



R: Randomization; Std: standard arm; Exp: experimental arm

Buyse, Michiels et al, Expert Rev Mol Diag 2011

Interaction (biomarker-stratified) randomized design

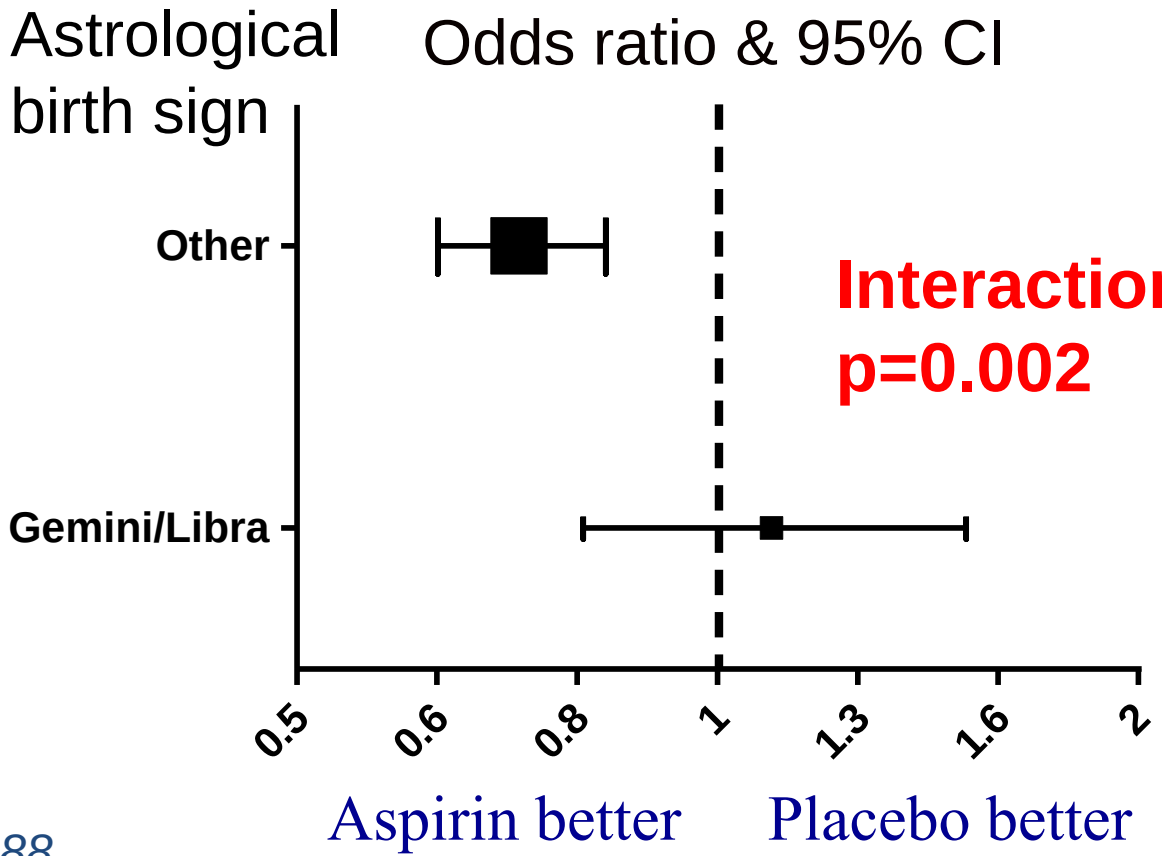


Typically large trials are needed!
AND/OR more sensitive endpoints (such as tumour measurements)

Buyse, Michiels et al, Expert Rev Mol Diag 2011

Most famous subgroup?

ISIS-2: aspirin vs control - effects on vascular death in 17,187 patients with acute myocardial infarction



Interaction p-value p=0.002

Peto et al Lancet 1988

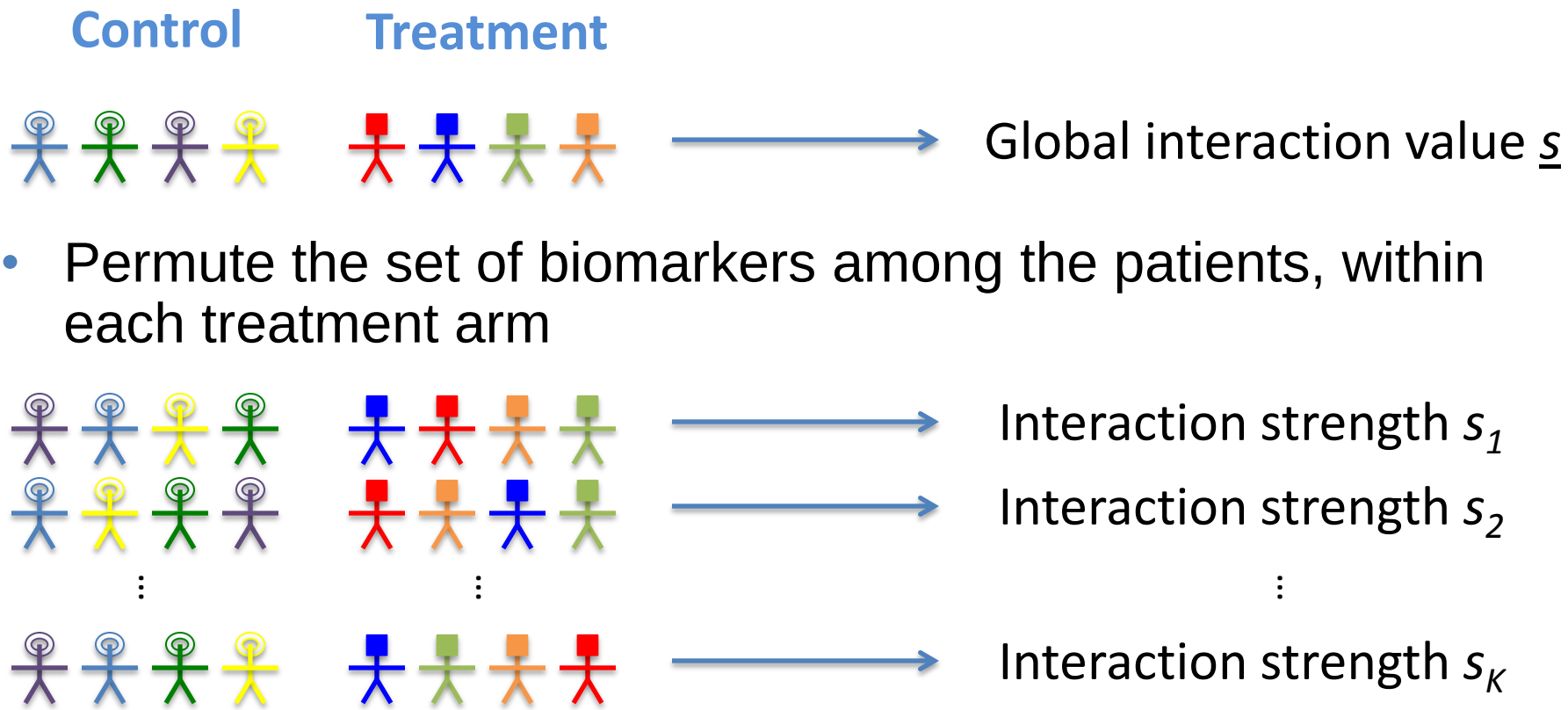
Stepwise strategy for high-dimensional data in a randomized clinical trial (RCT)

- **Step 1:** Perform a global interaction test for control of the global Type I error at a prespecified α level (e.g. 5%)

- **Step 2:** Only when global test is significant, develop classifier in a survival model with interaction effects
  - Use 10-fold crossvalidation to estimate treatment effects in composite biomarker score defined subsets (similar to Matsui CCR 2012)
  - Applying survival model on full RCT data: **indication classifier** for future patients (Simon Stat Med 2012)

Michiels, Rotolo in Matsui, Buyse, Simon 2014

Step I: Global interaction test by permutation

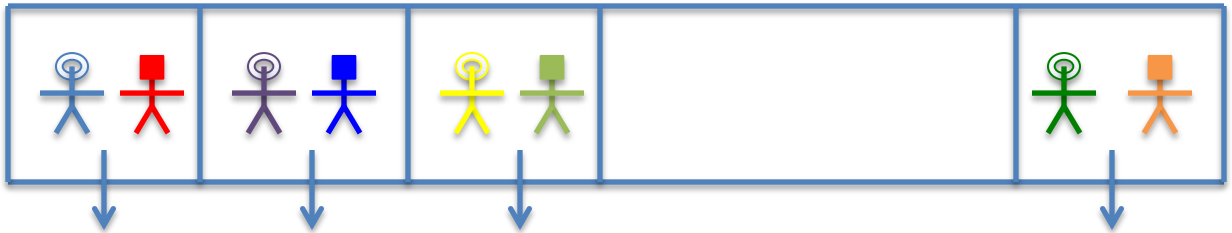


- Permute the set of biomarkers among the patients, within each treatment arm

- p-value = the proportion of *K* permutations in which the test statistic for global interaction exceeds the test statistic s for the original data

Michiels, Potthoff, George Stat Med 2011

Step II: Develop a composite biomarker classifier through 10-fold crossvalidation



Divide the RCT in 10 parts

Part 1 Part 2 Part 3 Part 10

Develop biomarker score (training set, 9/10 parts)

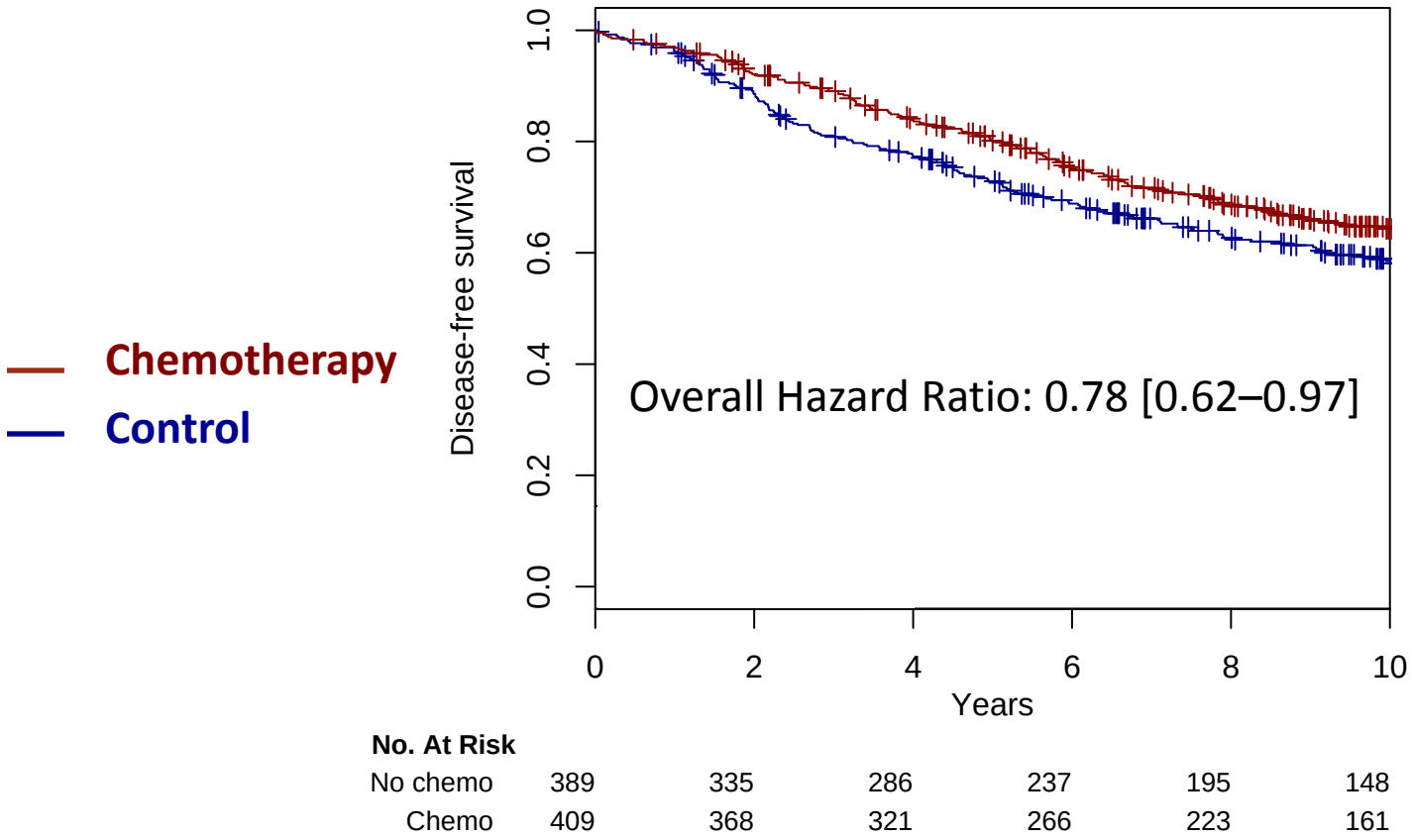
Evaluate biomarker score (validation set, 1/10 parts)

- Repeat this process 10 times
- Estimate treatment effects according to composite biomarkers scores in the 10 validation sets

Michiels, Rotolo in Matsui, Buyse, Simon 2014

French breast RCTs example

- Tissue-array from two French breast cancer RCTs of adjuvant chemotherapy with long term follow-up (798 pts)
- 11 biomarkers (ER,PR, HER2, EGFR, p53, p27…)
- Disease-free survival, 320 events

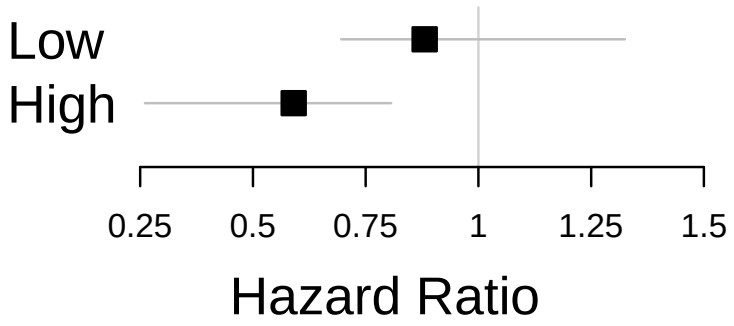


French breast RCTs example

Step I: Global test, 2000 permutations

- Three different proposed global statistics yield p-values = 0.045, 0.009 and 0.013

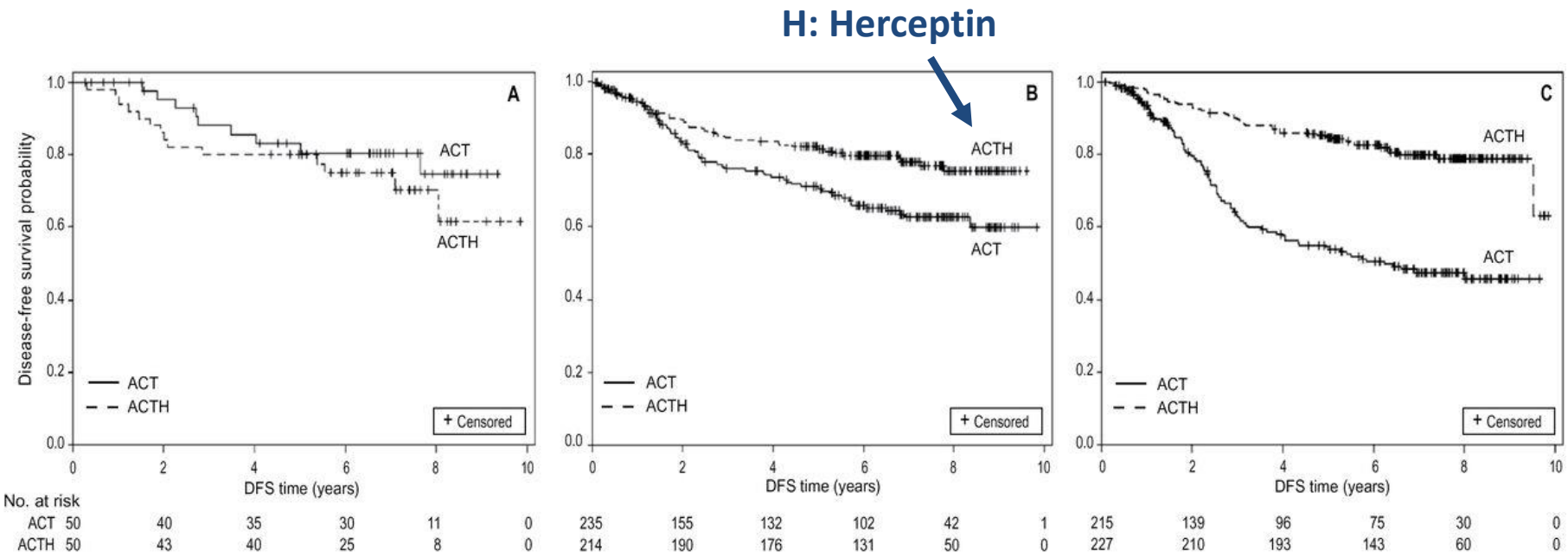
Step II: Crossvalidated treatment effects according to composite biomarker score (cut-off at 0.5)



Michiels, Rotolo in Matsui, Buyse, Simon 2014

Prediction of benefit of adjuvant trastuzumab

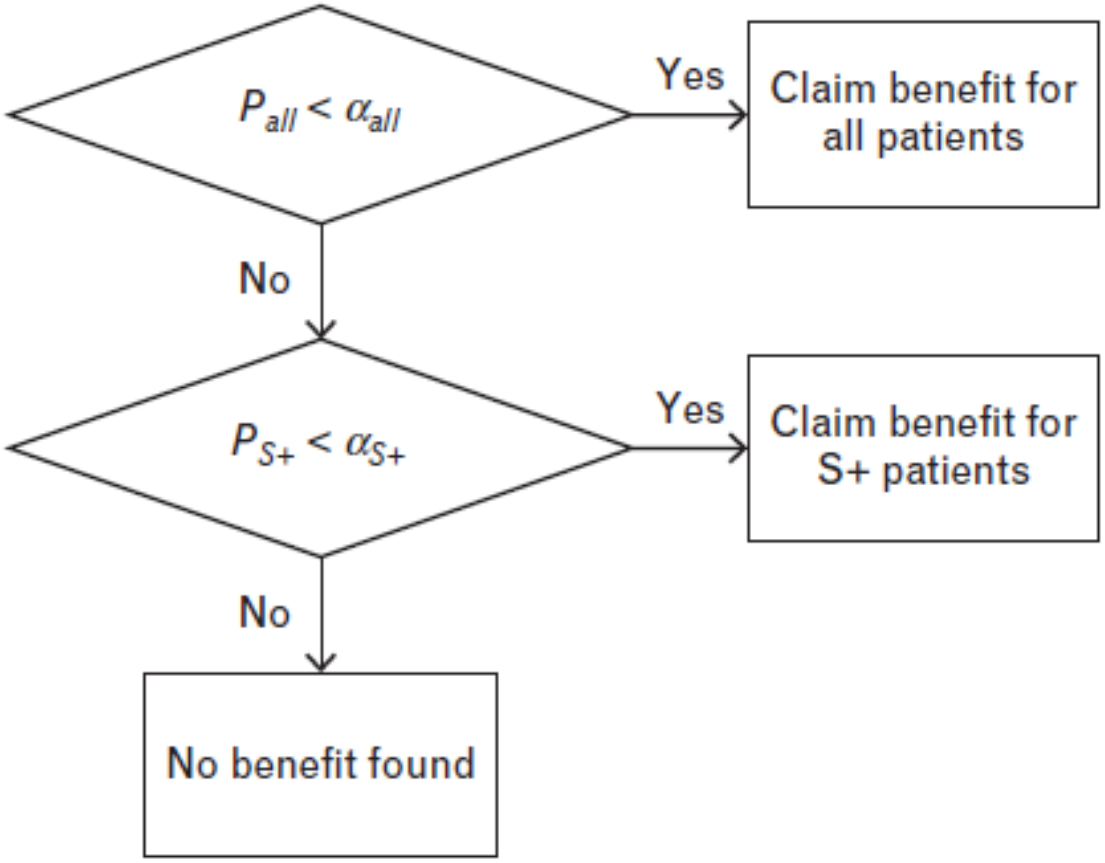
- Training-validation strategy in NSABP B31 trial, 8-gene signature
  Interaction between trastuzumab and signature: p<0.001



HR=1.58 (0.67 -3.69) **HR=0.60 (0.41-0.89)** **HR=0.28 (0.20-0.41)**

Pogue-Geile K L et al. J Natl Cancer Inst 2013

Alternative 1: Prospective subset testing



S+ : signature-positive patients; all: all patients

Buyse, Michiels, Curr Op Onc 2013

Alternative 1: Prospective subset testing

- Simplest approach: split significance level:

  $\alpha = \alpha_{all} + \alpha_{S+}$

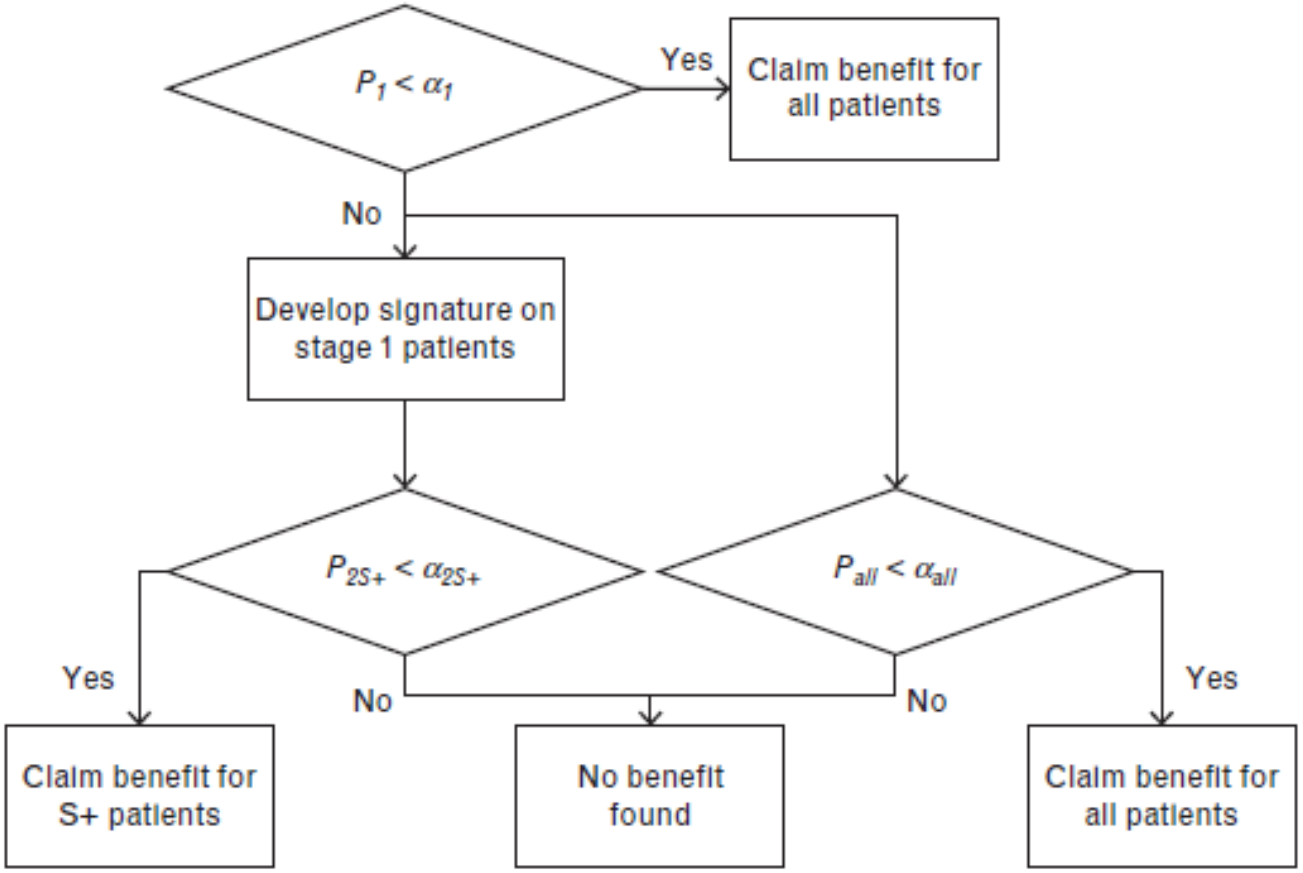
  - the new treatment is compared with the control in the overall population, ignoring the biomarker
  - if $p_{all} \leq \alpha_{all}$ claim effectiveness for all patients
  - if not, the new treatment is compared with the control in biomarker + patients only, and if $p_{S+} \leq \alpha_{S+}$, claim effectiveness for biomarker + patients only

- There are less conservative, yet properly controlled, ways of adjusting α for both (correlated) tests

Wang, Pharm Stat 2007; Jiang, JNCI 2007; Alosh, Stat Med 2009; Wang, Biom J 2009; Spiessens, Contr Clin Trials 2010

Alternative 2: Two-stage adaptive signature design

1: stage 1 patients
2S+: stage 2 signature-patients
all: all patients



- There exists a crossvalidation version (*Freidlin et al CCR 2010*)

Buyse, Michiels 2013; extension from Freidlin, Simon, CCR 2005

Pros and contras

- Prospective subset testing: needs prespecified signature
- Gene signature can be developed on a first stage of the trial but statistical power could be small
- Challenge for two-stage adaptive design: possible heterogeneity of treatment effects before and after the adaptation (changes in patient recruitment or other temporal trends)
- Crossvalidation scheme : need for independent validation trial when all data is repeatedly used to develop the signature?

Conclusion

- Move on from *prognostic* gene signatures (trying to be predictive of treatment benefit) to gene signatures **developed on RCT data as treatment modifier**
- Global interaction test at significance level a by permuting a statistic among the patients within the treatment groups
- **Continuous** gene signature development in RCT by crossvalidation approach
- Strong **challenges in controlling confounding**: handling of specimens, measurement error, tumour heterogeneity, biopsy vs primary vs metastatic specimen

Main references

- Michiels, S and Rotolo, F. (2014) Evaluation of clinical utility and validation of gene signatures in clinical trials, in Design and Analysis of Clinical Trials for Predictive Medicine, CRC Press (eds Matsui S., Buyse M., Simon R).
- Buyse, M. and Michiels, S. (2013) Omics-based clinical trial designs. *Curr Opin Oncol* 25:289-295.
- Buyse M., Michiels S et al. (2011) Integrating biomarkers in clinical trials. Expert Rev Mol Diagn 11:171-82.
- Michiels, S., Potthoff, R.F. and George, S.L. (2011) Multiple testing of treatment modifying biomarkers in a randomized clinical trial with a survival endpoint, *Stat Med* 30:1502-1518.