# How to report translational research results

**Stefan Michiels, PhD**
Director Team 2 CESP, INSERM U1018
Department of Biostatistics and Epidemiology
Gustave Roussy, Paris-Sud University, Villejuif, France
stefan.michiels@gustaveroussy.fr

**To consult the statistician after an experiment is finished is often merely to ask him to conduct a post mortem examination. He can perhaps say what the experiment died of.**

Sir R Fisher, Presidential Address to the First Indian Statistical Congress, 1938. Sankhya 4, 14-17

GUSTAVE ROUSSY
CANCER CAMPUS
GRAND PARIS

design

# How to report translational research results

**Stefan Michiels, PhD**
Director Team 2 CESP, INSERM U1018
Department of Biostatistics and Epidemiology
Gustave Roussy, Paris-Sud University, Villejuif, France
stefan.michiels@gustaveroussy.fr

Instituts thématiques Inserm
Institut national de la santé et de la recherche médicale

UNIVERSITÉ PARIS SUD
Comprendre le monde, construire l'avenir®

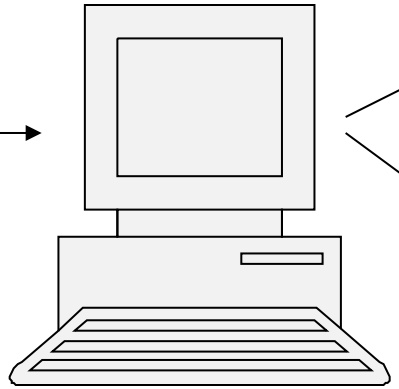
UNIVERSITE DE VERSAILLES SAINT-QUENTIN-EN-YVELINES

CESP Centre de recherche en Épidémiologie et Santé des Populations

# PROGNOSTIC GENE SIGNATURES

Measure ≈ 25,000 genes in RNA from breast tumors

Apply algorithm to identify classifier



Class of good prognosis

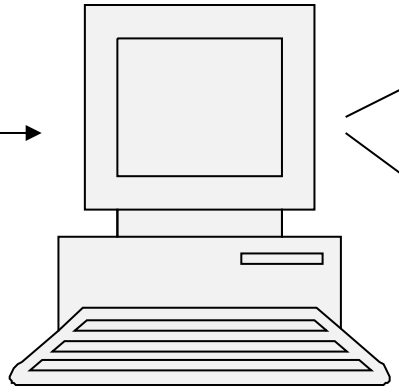
Class of poor prognosis

# MAMMAPRINT®

Measure ≈ 25,000 genes in RNA from breast tumors

Agendia
24,479 probe sets

Apply algorithm to identify classifier



Good prognosis
(no metastases at 5 years)

Poor prognosis
(metastases within 5 years)

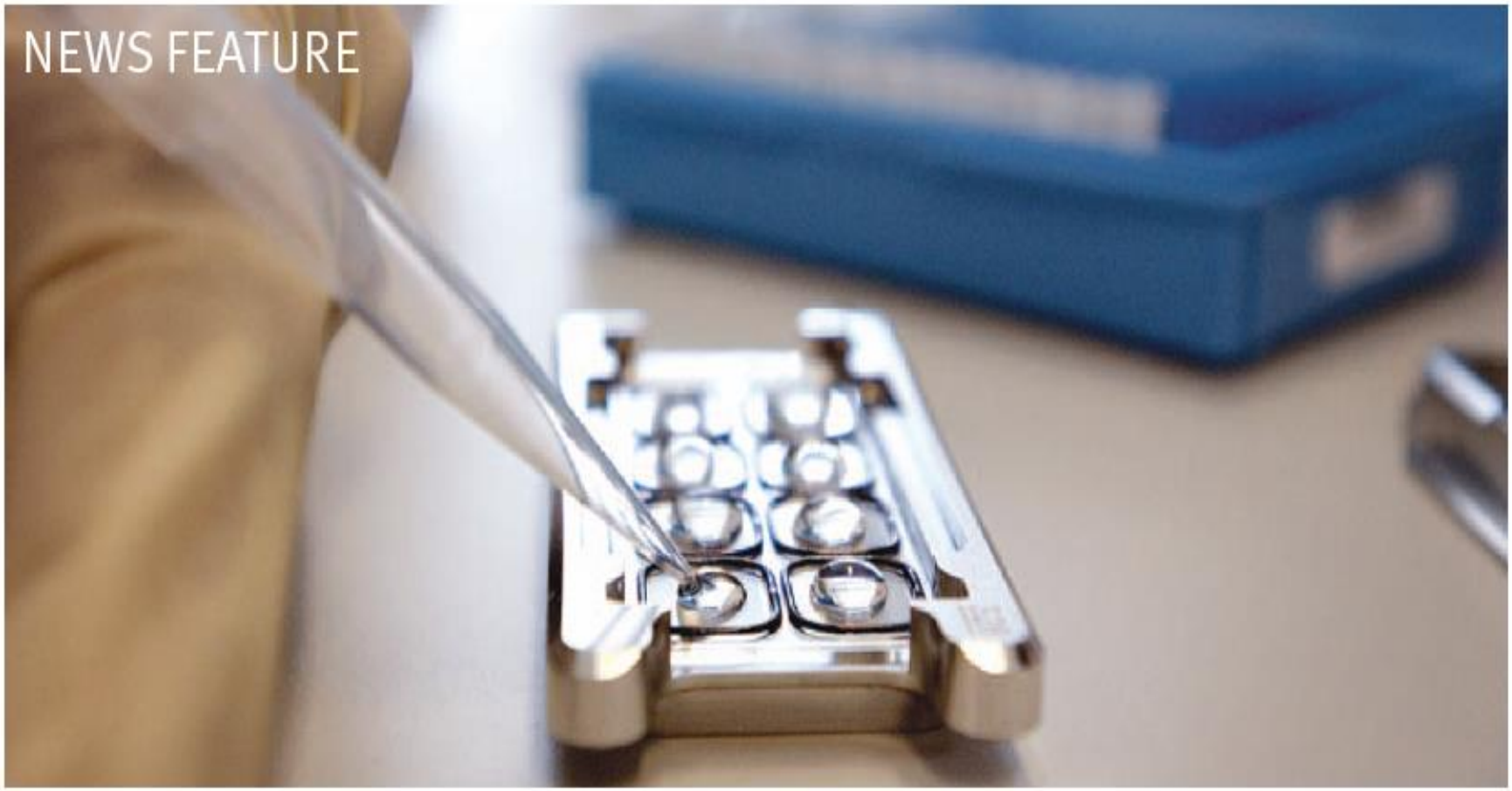
*Ref: van't Veer et al, Nature 2002; 415: 539.*

# GENE SIGNATURES IN BREAST CANCER

- 70-gene « Amsterdam » signature (*MammaPrint™, Agendia*)
- 76-gene « Rotterdam » signature (*Veridex*)
- 21-gene assay (*Oncotype DX™, Genomic Health*)
- 97-gene « genomic grade » (*MapQuant Dx™, Ipsogen*)
- and others…

These signatures were identified using different criteria and include different sets of genes.

They are « broadly » similar in their ability to classify patients to good or poor prognosis, but they may provide different predictions for a given patient!




# An array of problems

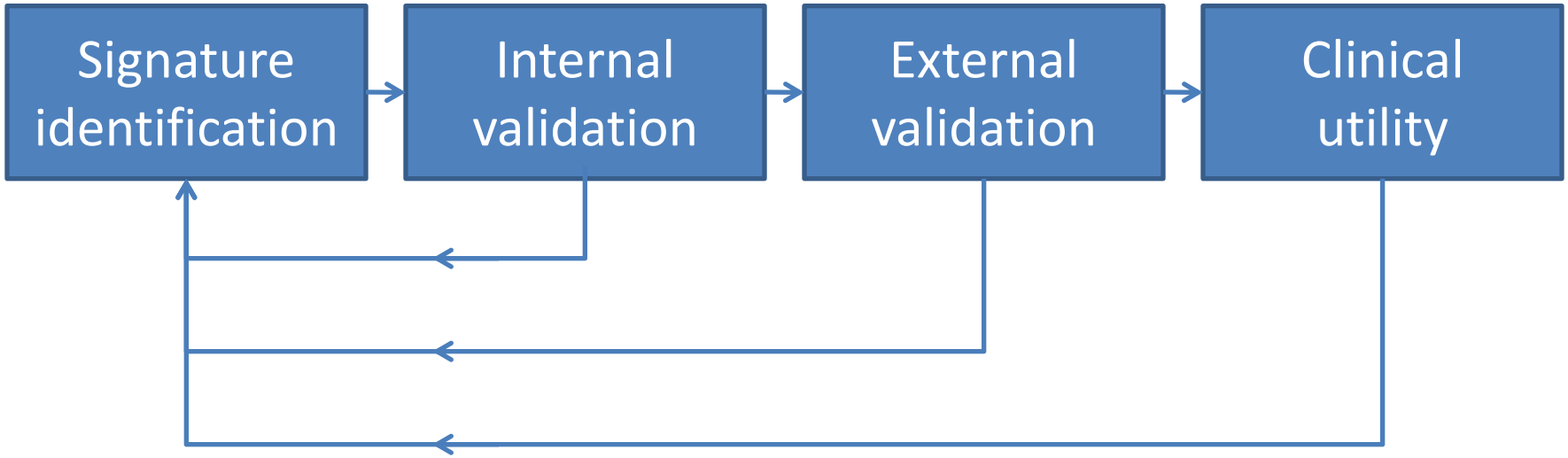
Despite the huge amount of published microarray data in cancer, little is being converted into clinical practice. Validating initial data is proving to be a key challenge, reports SIMON FRANTZ.

*Ref: Nature Reviews Drug Discovery 4, 362-363 (2005).*

# PROBLEMS…

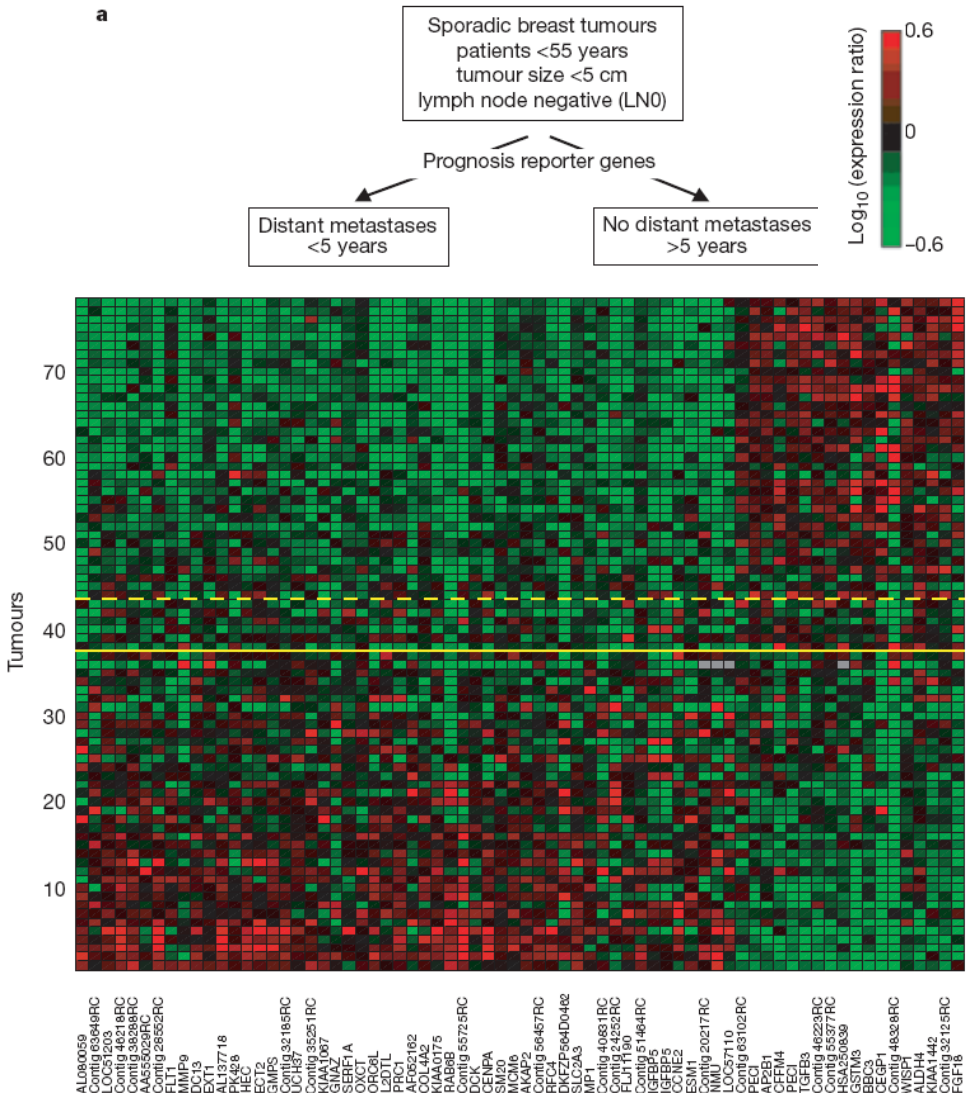
- Microarrays can identify the expression of 10,000 to 40,000 genes (sequences) at a cost of 500-1000 $
- They are typically applied to a small number of retrospective samples that happen to be available
- Frequent problems include
  - selection bias (unrepresentative samples)
  - multiplicity (many more genes than samples)
  - overfitting (impressive but unconfirmed results)

# SOLUTION:
# A RIGOROUS DEVELOPMENT PROCESS



# CLASS COMPARISON (supervised method)

- Training set: 78 pts, 34 with distant metastasis at 5 years
- Gene expression levels ranked by correlation coefficient with binary metastatic status at 5 years
- Gene selection (70 genes with highest correlations) = « molecular signature »



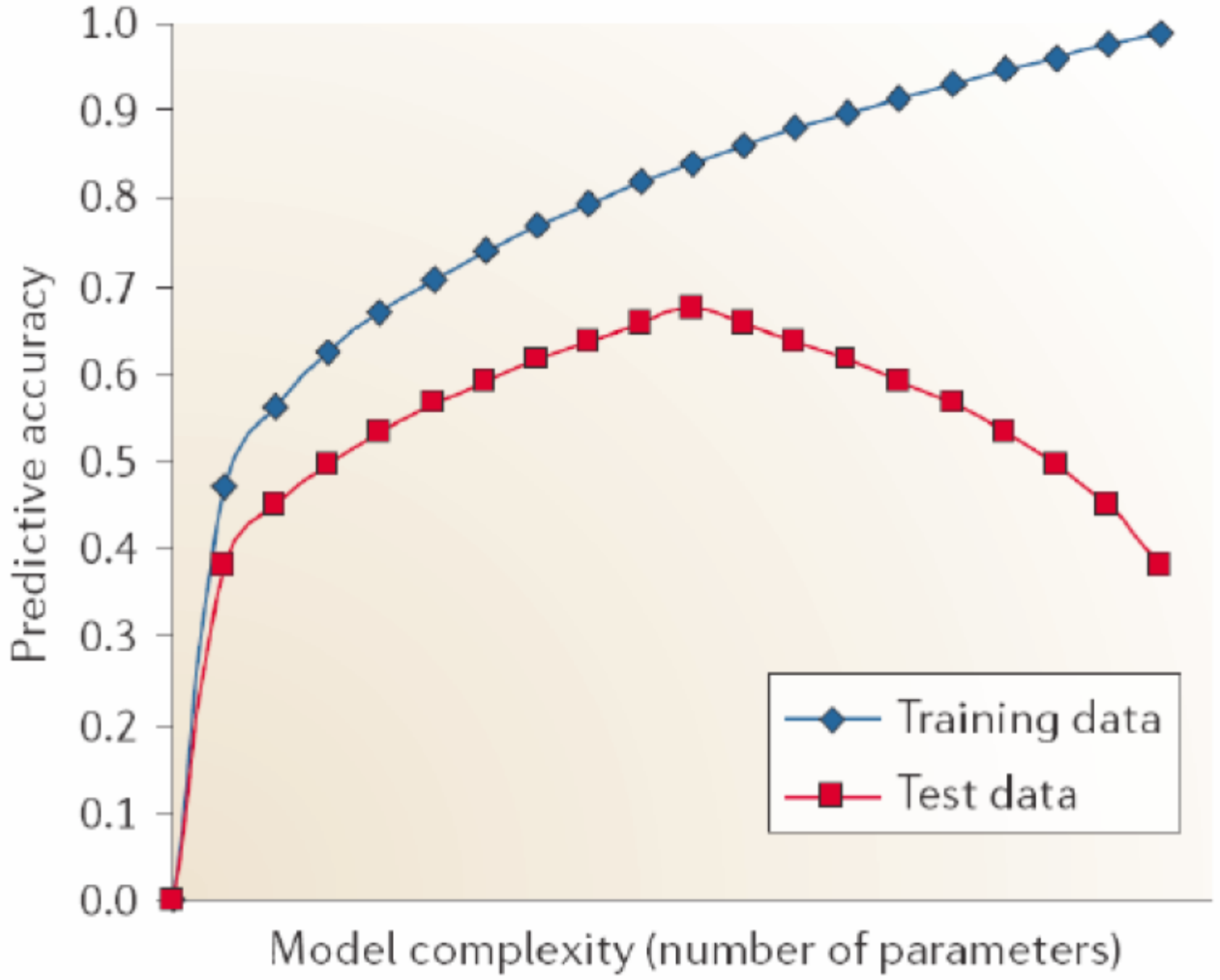

*Ref: van't Veer et al, Nature 2002; 415: 539.*

# DEVELOPMENT OF PREDICTION RULE: TRAINING SET

Most statistical methods were developed for settings where variables did not largely outnumber patients (in traditional analyses n>>p, here p>>n)

Components of Class Prediction

1. Gene selection:
   which genes to select for the model ?
2. Choice of a prediction rule:
   Linear Discriminant Analysis
   Nearest Neighbour, …
3. Determine cut-offs

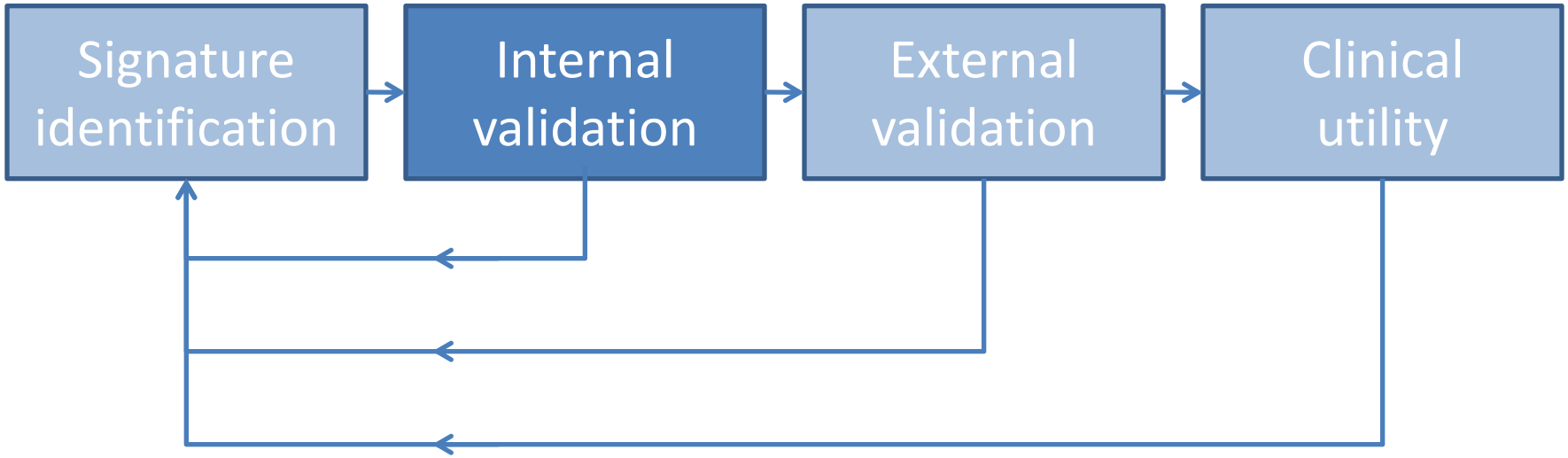



*Ref: Allison, Nature Rev Genet 2006.*

# CHOICE OF A PREDICTION RULE

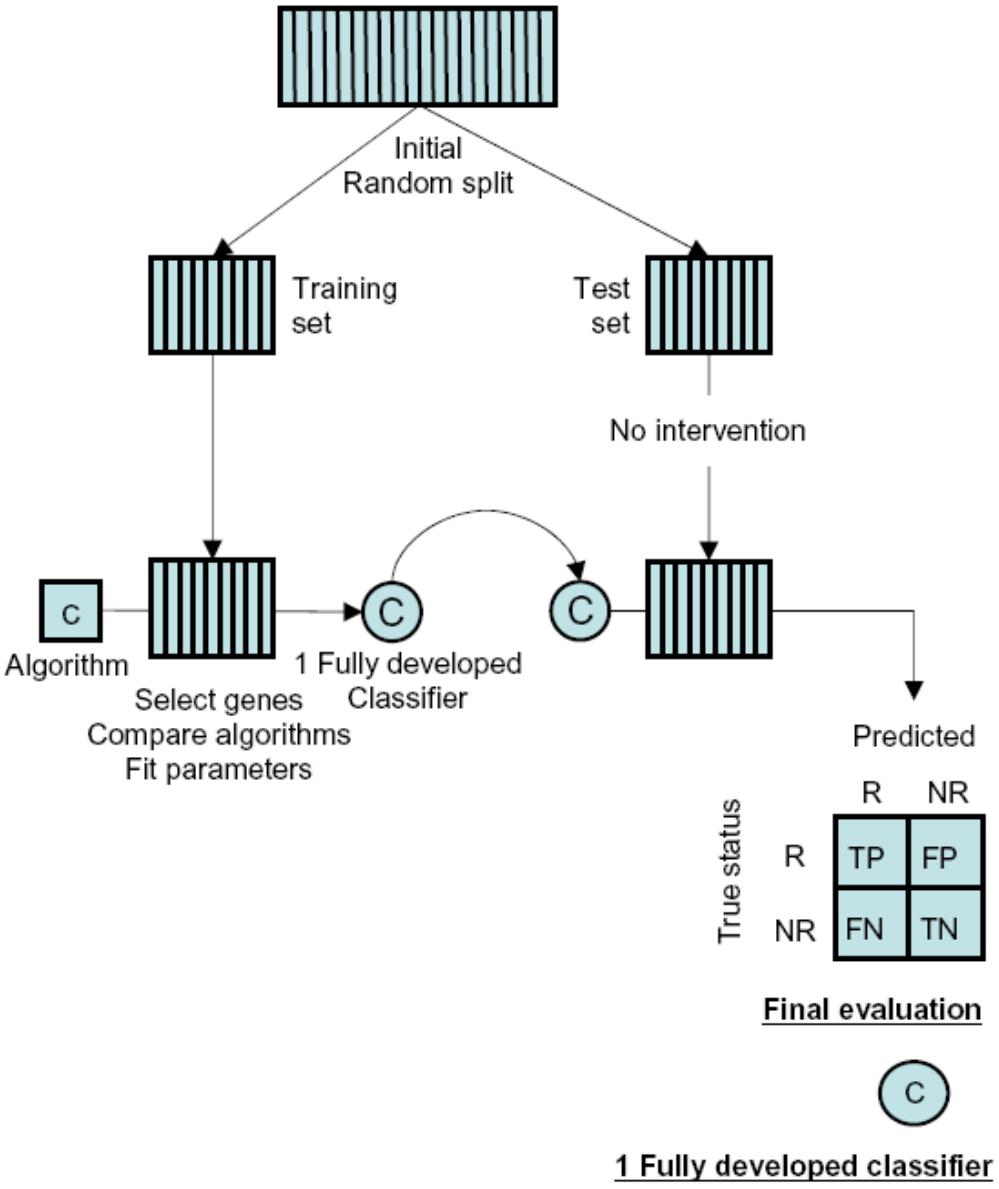
- MammaPrint: nearest centroid prediction rule (simple)
- Not much theoretical or empirical evidence that more complex models perform better (p>>n)
- Variations on univariate gene selection methods and prediction rules have only a modest impact on performance (MAQCII)
- “Human factor” very important: have your data analyzed by someone who understands statistics!

*Ref: Michiels et al, Lancet 2005; Hand, Stat Sci 2006;*
*Popovici et al, Breast Cancer Re 2010.*


Signature identification
Internal validation
External validation
Clinical utility

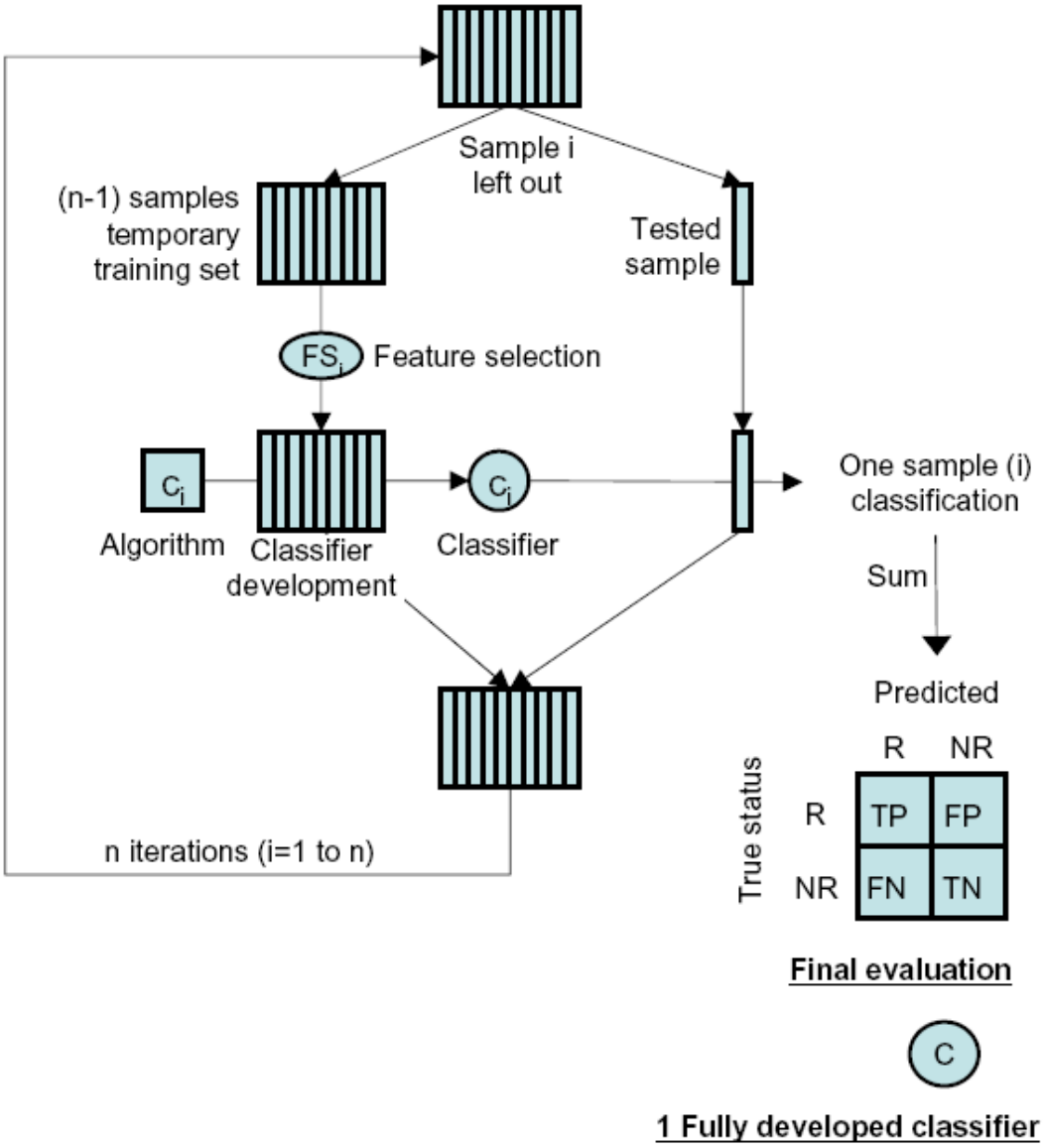
# INTERNAL VALIDATION: CROSS-VALIDATION

- Cross-validation simulates the process of separately developing a model on one set of data and predicting for a test set of data not used in developing the model
- The cross-validated estimate of misclassification error is an estimate of the prediction error for model fit using specified algorithm to full dataset
- Leave-one-out cross validation:
  - Omit sample 1
    - Develop multivariate classifier from scratch on training set with sample 1 omitted
    - Predict class for sample 1 and record whether prediction is correct
  - Repeat from 2 to n and take average missclassification rate

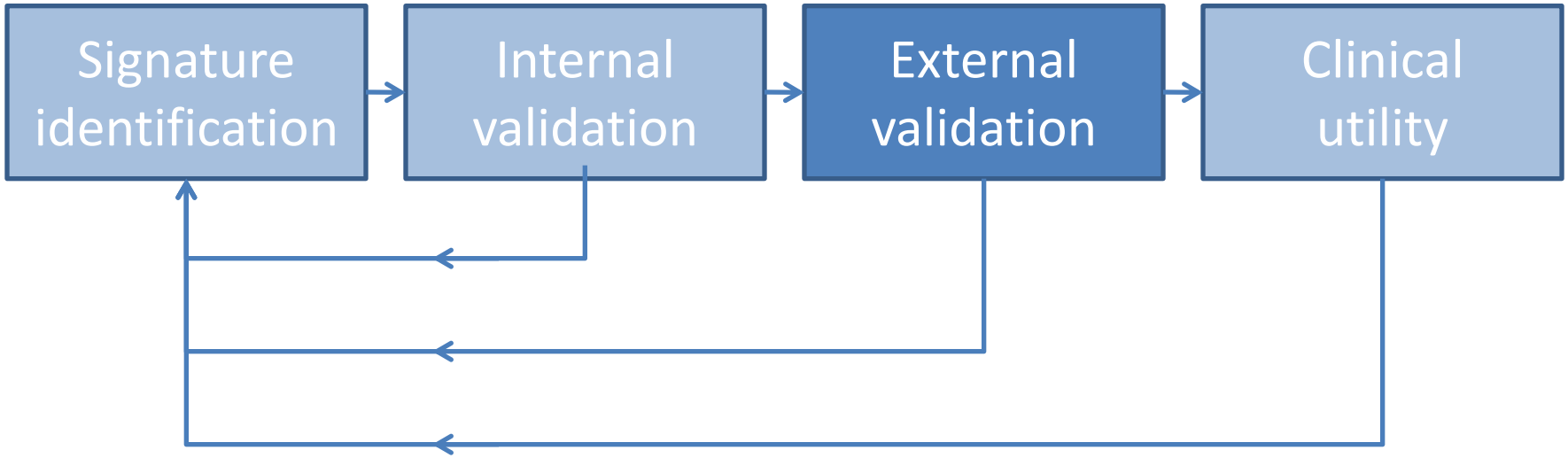



**A. Split-sample procedure**

*Ref: Dupuy and Simon JNCI 2007.*




**B. Leave-one-out cross-validation procedure**


Signature identification
Internal validation
External validation
Clinical utility

# EXTERNAL VALIDATION

External validation requires an independent study to be prospectively designed to confirm the results of a previous study, in order to reduce the play of chance and the potential for biases.

The same methodological guidelines apply as for the validation of other tumor markers (REMARK NCI-EORTC Guidelines)

*Refs: Ransohoff, Nat Rev Cancer 2004, 2005*
*McShane et al, JNCI 2005; Altman et al PLOS Med 2012*

# COMMON MISTAKES IN VALIDATION STUDIES

- To include part of the initial sample of patients in the validation study
- To include other types of patients in the validation study than in the initial sample
- To use another measurement technique (rt-PCR *vs.* microarray)
- To change the prediction rule by adapting it to the new sample of patients through changing the list of genes, the prediction rule, or the cutoff

*Ref: Koscielny et al JCO 2005; Michiels, Hill, NEJM 2007; Michiels et al. BJC 2007.*

# INDEPENDENT VALIDATION OF MAMMAPRINT?

The New England
Journal of Medicine



VOLUME 347 DECEMBER 19, 2002 NUMBER 25

## A GENE-EXPRESSION SIGNATURE AS A PREDICTOR OF SURVIVAL IN BREAST CANCER

MARC J. VAN DE VIJVER, M.D., PH.D., YUDONG D. HE, PH.D., LAURA J. VAN 'T VEER, PH.D., HONGYUE DAI, PH.D., AUGUSTINUS A.M. HART, M.SC., DORIEN W. VOSKUIL, PH.D., GEORGE J. SCHREIBER, M.SC., JOHANNES L. PETERSE, M.D., CHRIS ROBERTS, PH.D., MATTHEW J. MARTON, PH.D., MARK PARRISH, DOUWE ATSMA, ANKE WITTEVEEN, ANNUSKA GLAS, PH.D., LEONIE DELAHAYE, TONY VAN DER VELDE, HARRY BARTELINK, M.D., PH.D., SJOERD RODENHUIS, M.D., PH.D., EMIEL T. RUTGERS, M.D., PH.D., STEPHEN H. FRIEND, M.D., PH.D., AND RENÉ BERNARDS, PH.D.

31/54 events in the N- group in the *validation* set (*NEJM* 2002) came from the *training set* (*Nature* 2002).

”Independent prediction was not demonstrated; therefore these results may not be strongly reproducible and should not be interpreted as ‘definitive’ ”

*Ref: Ransohoff, Nat Rev Cancer 2004.*

# EXTERNAL VALIDATION OF MAMMAPRINT®

## Validation study 1 (*Van de Vijver et al NEJM 2002*)

- N=295 breast cancers from one single center
- Potential bias: inclusion of 61 pts of the training set
- Predictive accuracy of MammaPrint :

  Se = 93%  ($CI_{95\%}$ 81% to 99%)

  Sp = 53%  ($CI_{95\%}$ 44% to 61%)

*Ref: van de Vijver et al, NEJM 2002.*

# EXTERNAL VALIDATION OF MAMMAPRINT®

Validation study 1 (*Van de Vijver et al NEJM 2002*)

- N=295 breast cancers from one single center
- Potential bias: inclusion of 61 pts of the training set
- Predictive accuracy of MammaPrint :

  Se = 93% ($CI_{95\%}$ 81% to 99%)

  Sp = 53% ($CI_{95\%}$ 44% to 61%)

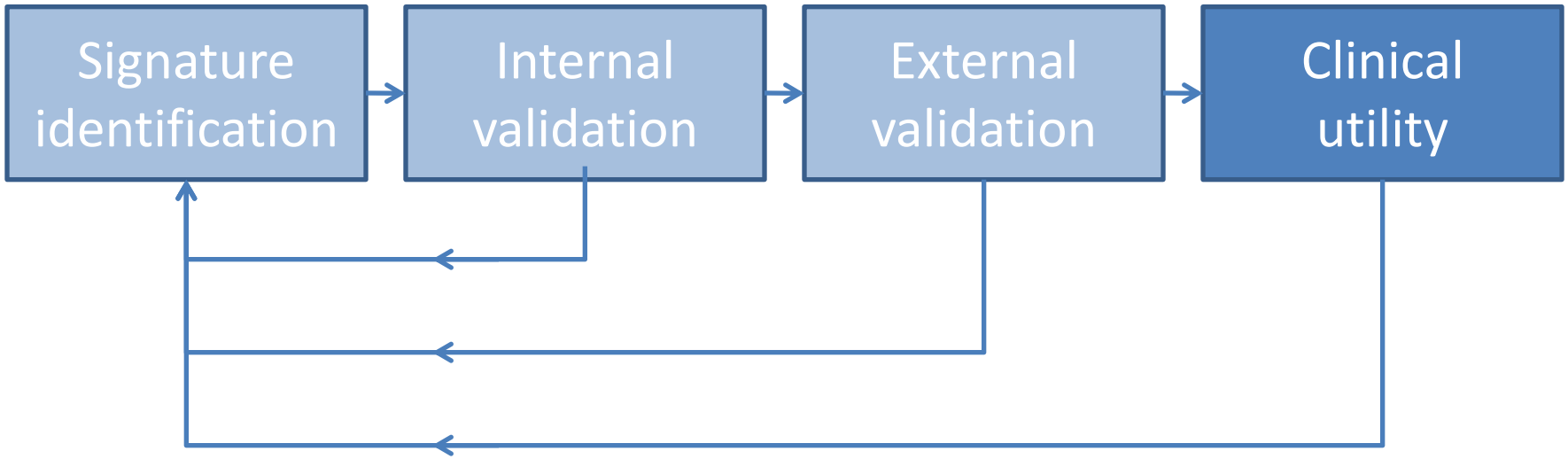
Validation study 2 (*TRANSBIG, Buyse et al JNCI 2006*)

- N=307 breast cancers from 5 European centers
- No bias but done on frozen samples available from >10 years ago
- Predictive accuracy of MammaPrint :

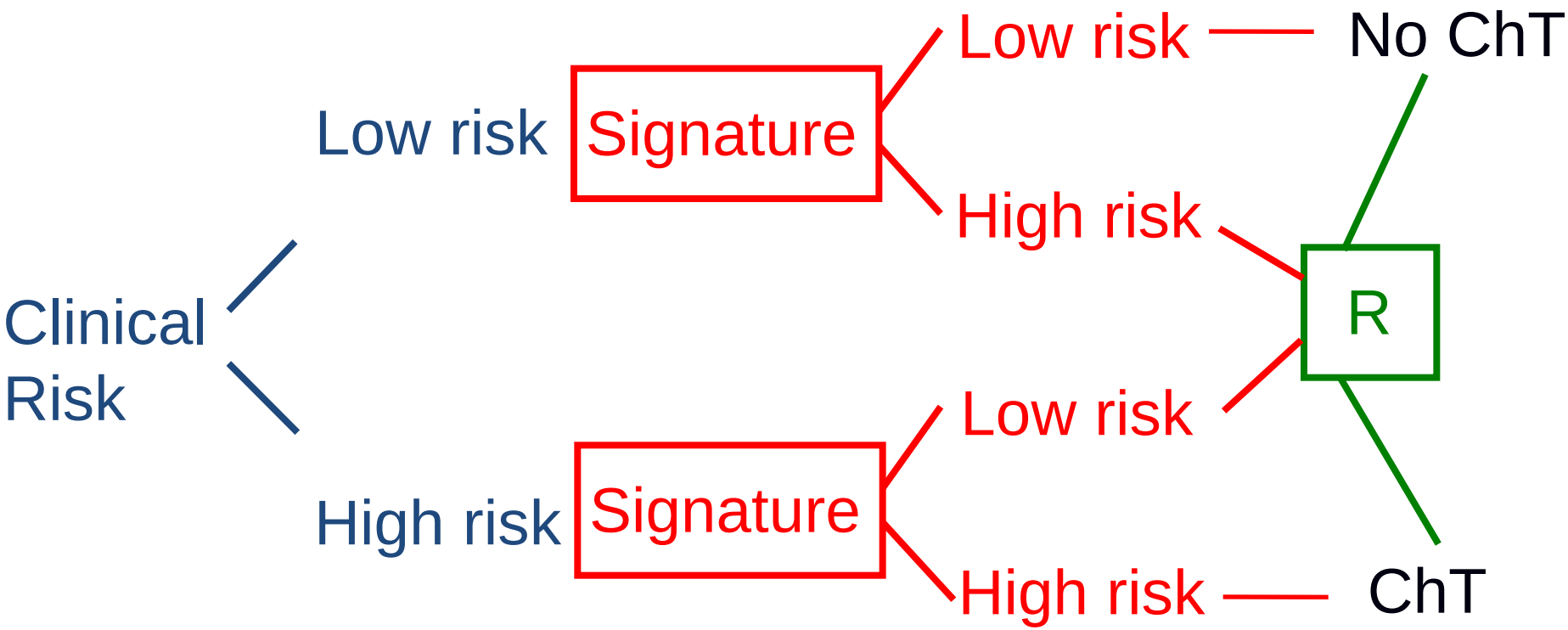
  Se = 90% ($CI_{95\%}$ 78% to 95%)

  Sp = 42% ($CI_{95\%}$ 36% to 48%)

*Ref: Buyse et al, JNCI 2006; 98: 1183.*

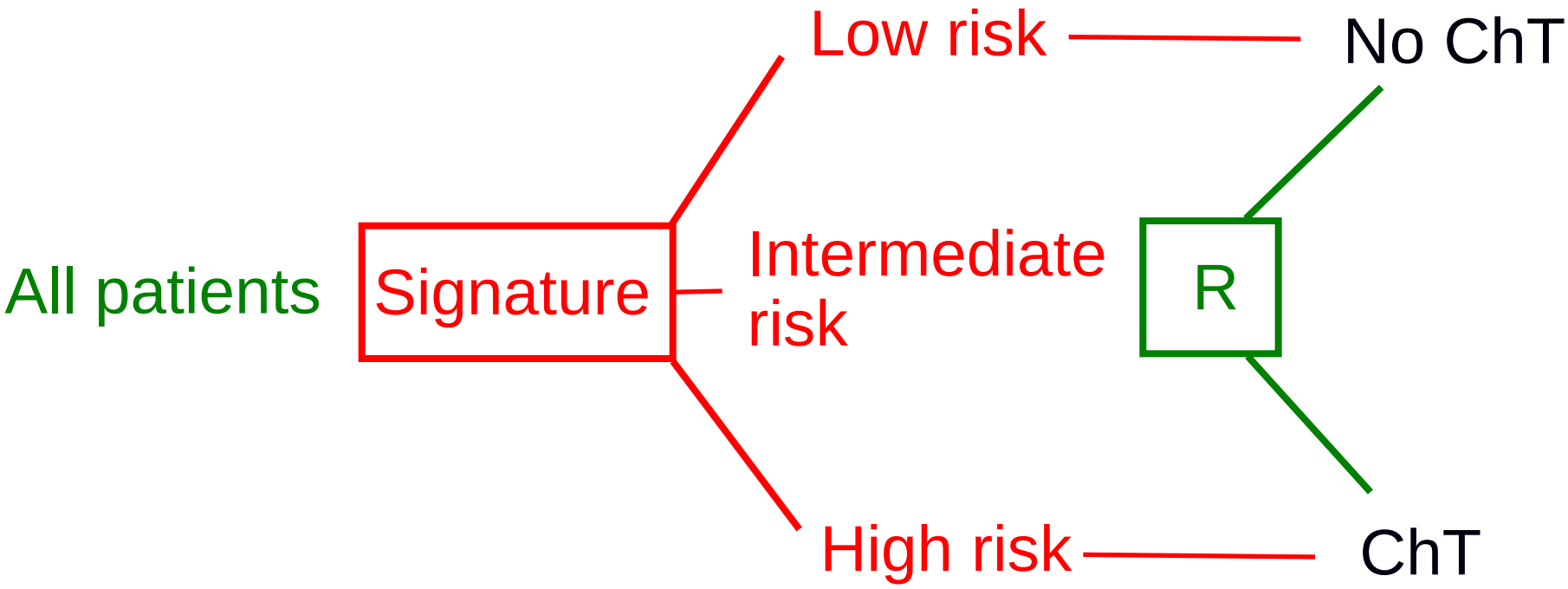

Signature identification
Internal validation
External validation
Clinical utility

# “DISCORDANT RISK” DESIGN




*Buyse, Michiels et al, Exp Rev Mol Diag 2011; ChT: chemotherapy*

# “INTERMEDIATE RISK” DESIGN



# BENEFITS OF LARGE RANDOMIZED TRIALS FOR THE VALIDATION OF GENE SIGNATURES

- Tumor samples will be collected prospectively
- Patients will be selected identically
- Patients will receive controlled treatments
- Randomization will make it possible to investigate the predictive value of signatures
- Micro-arrays and pathological reviews will be done centrally
- Large numbers will make the results reliable

PERSPECTIVE

OPEN

doi:10.1038/nature1256

# Criteria for the use of omics-based predictors in clinical trials

Lisa M. McShane[1], Margaret M. Cavenagh[1], Tracy G. Lively[1], David A. Eberhard[2], William L. Bigbee[3], P. Mickey Williams[4], Jill P. Mesirov[5], Mei-Yin C. Polley[1], Kelly Y. Kim[1], James V. Tricoli[1], Jeremy M. G. Taylor[6], Deborah J. Shuman[1], Richard M. Simon[1], James H. Doroshow[1] & Barbara A. Conley[1]

**The US National Cancer Institute (NCI), in collaboration with scientists representing multiple areas of expertise relevant to 'omics'-based test development, has developed a checklist of criteria that can be used to determine the readiness of omics-based tests for guiding patient care in clinical trials. The checklist criteria cover issues relating to specimens, assays, mathematical modelling, clinical trial design, and ethical, legal and regulatory aspects. Funding bodies and journals are encouraged to consider the checklist, which they may find useful for assessing study quality and evidence strength. The checklist will be used to evaluate proposals for NCI-sponsored clinical trials in which omics tests will be used to guide therapy.**

# Some Further Reading


- Buyse M., Michiels S, Sargent DJ., Grothey A., Matheson A., de Gramont A. Integrating biomarkers in clinical trials. Expert Rev Mol Diagn 2011;11:171-82.
- Freidlin,B., McShane,L.M. & Korn,E.L. (2010). Randomized Clinical Trials With Biomarkers: Design Issues. *Journal of the National Cancer Institute*, 102, 152-160.
- Michiels S, Koscielny S, Hill C. Interpretation of microarray data in cancer. Br J Cancer 2007; 96:1155-1158
- Michiels S, Kramar A, Koscielny S. Multidimensionality of microarrays: Statistical challenges and (im)possible solutions. Mol Onc 2011; 5: 190-196.

- Altman DG, McShane LM, Sauerbrei W, Taube SE. Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK): explanation and elaboration. PLoS Med. 2012;9(5):e1001216
- McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM; Statistics Subcommittee of the NCI-EORTC Working Group on Cancer Diagnostics. Reporting recommendations for tumor marker prognostic studies (REMARK). J Natl Cancer Inst. 2005 Aug 17;97(16):1180-4.
- McShane LM, Cavenagh MM, Lively TG, Eberhard DA, Bigbee WL, Williams PM, Mesirov JP, Polley MY, Kim KY, Tricoli JV, Taylor JM, Shuman DJ, Simon RM, Doroshow JH, Conley BA. Criteria for the use of omics-based predictors in clinical trials. Nature. 2013 Oct 17;502(7471):317-20. doi: 10.1038/nature12564.