Evaluation & Implementation Challenges with Genomic Signatures in Clinical Drug Development

V. Devanarayan, Ph.D.
Exploratory Statistics, Pharmaceutical R&D
Abbott

EMA Workshop on Pharmacogenomics: From Science to Clinical Care
European Medicines Agency, London, UK
October 8-9, 2012

Abbott
A Promise for Life

Disclosure Information

I have the following financial relationships to disclose:

- I am a minor stockholder in Abbott Laboratories
- I am an Employee of Abbott Laboratories, and
- I will not discuss off label use in my presentation.

Abbott funded all work related to preparation of this presentation.

V. Devanarayan

October 8, 2012

Typical uses of biomarkers in drug development

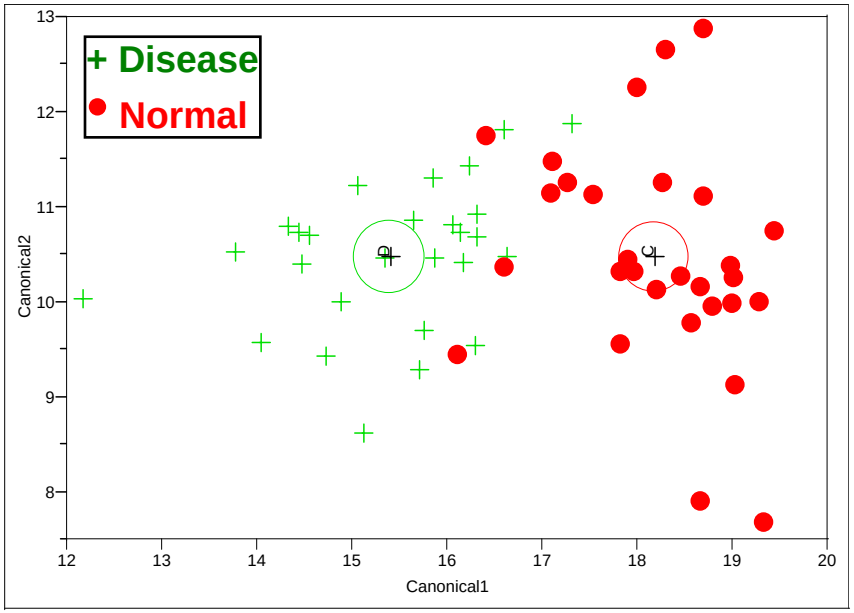
- Predict responders & non-responders to a drug.
- Predict safety events such as liver and kidney injury.
- Patient-selection for clinical trial.
  - – Better specificity in disease diagnosis (e.g., AD vs. FTD vs. VD)
  - – Identify which patients are likely to progress in disease
    - Reduce variability, placebo response, etc.
- Dose selection (PK-PD modeling)
- Proof of Mechanism & Concept in early drug development
  - – Pharmacodynamic, Target engagement (receptor occupancy), etc.

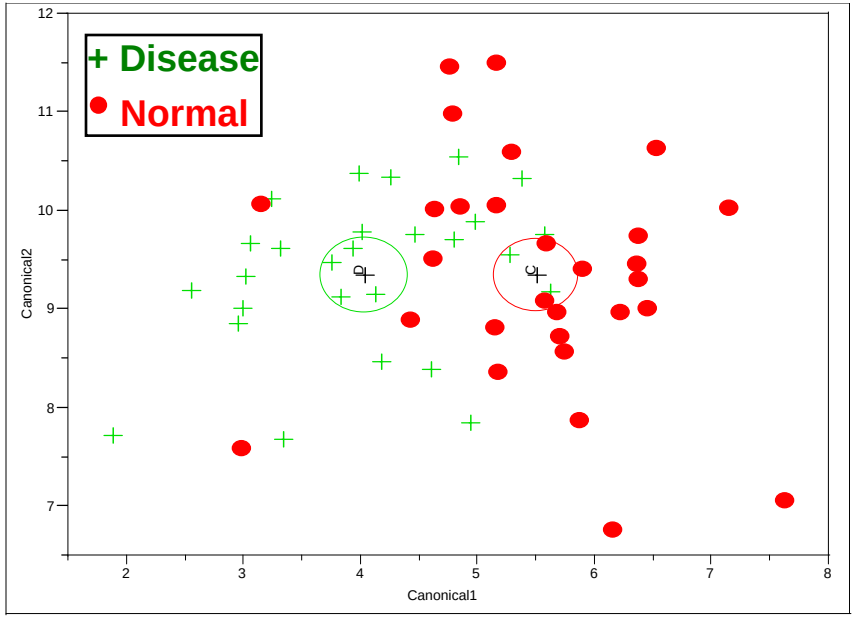
Some Practical Challenges

1. Variability (Analytical + Biological)
2. Biological Relevance
3. Biomarker performance evaluation
   - Internal & External Verification
     - Predictive Accuracy (disease progression, adverse events, …)
     - P-values (patient response/non-response), treatment differentiation, …)
4. Robustness
5. Translation
   - Animals to Humans, between human subpopulations (gender, race/region, age, disease severity and subtypes, etc.)

I will now briefly review some of these topics via illustrations.

Analytical + Biological Variability ➔ Biomarker Performance: Example 2





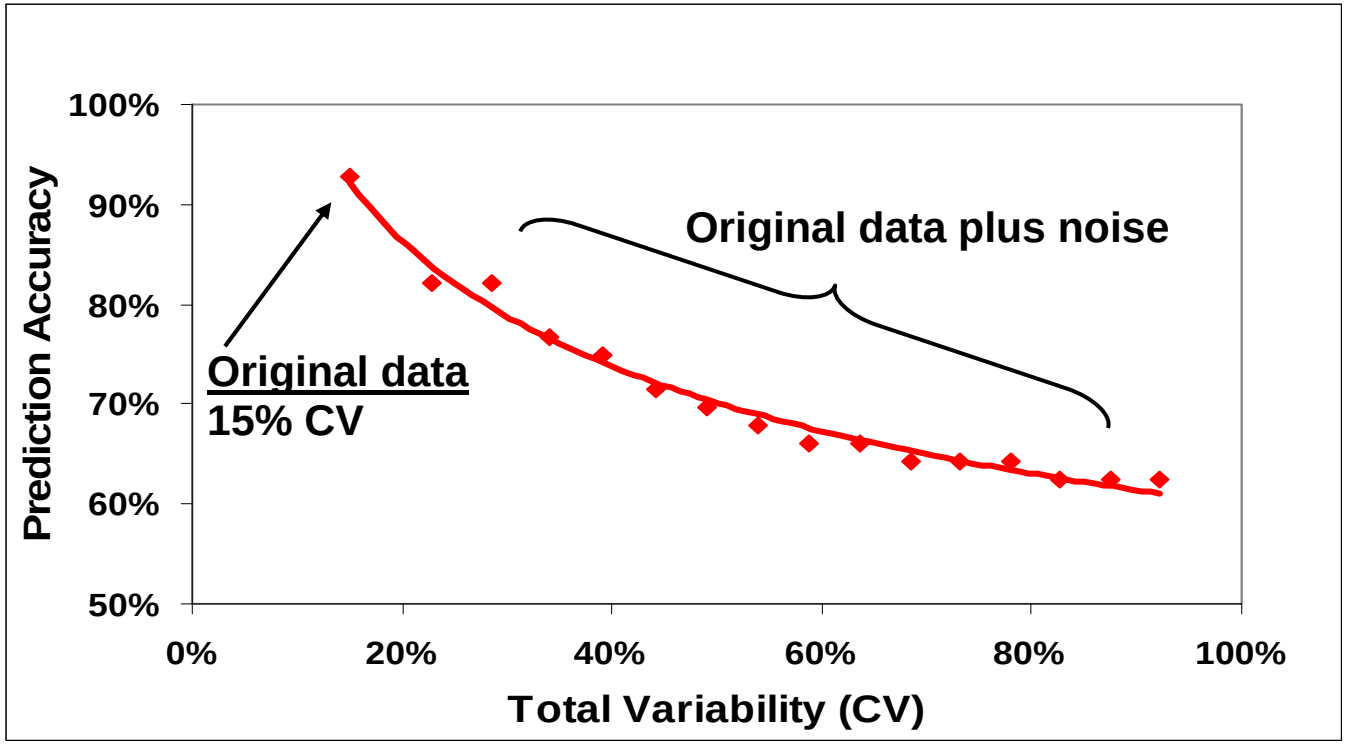
Discriminant Analysis

- **Marker X with 15% CV is a key predictor from the multi-analyte panel.**
- **Prediction Accuracy ~ 85%**

- **Same Marker X in the panel from another lab has 35% CV**
- **Prediction Accuracy ~ 65%**

Biomarker performance drops greatly when a different assay is used!

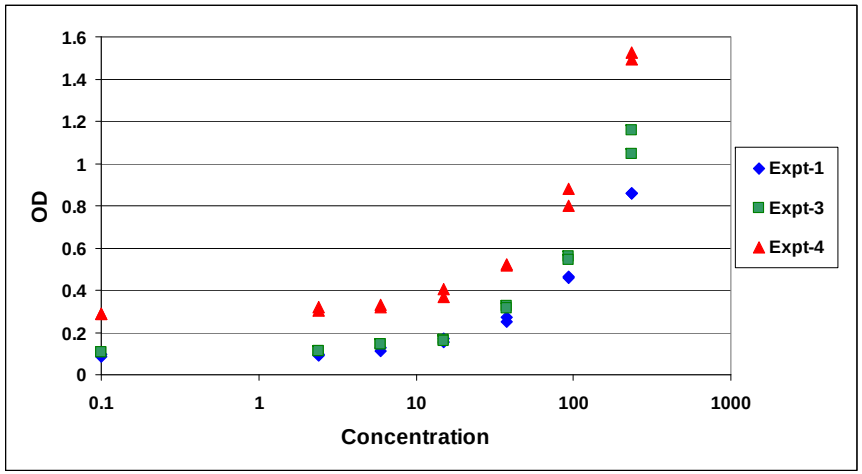
Analytical + Biological Variability ➔ Biomarker Performance: Generalization

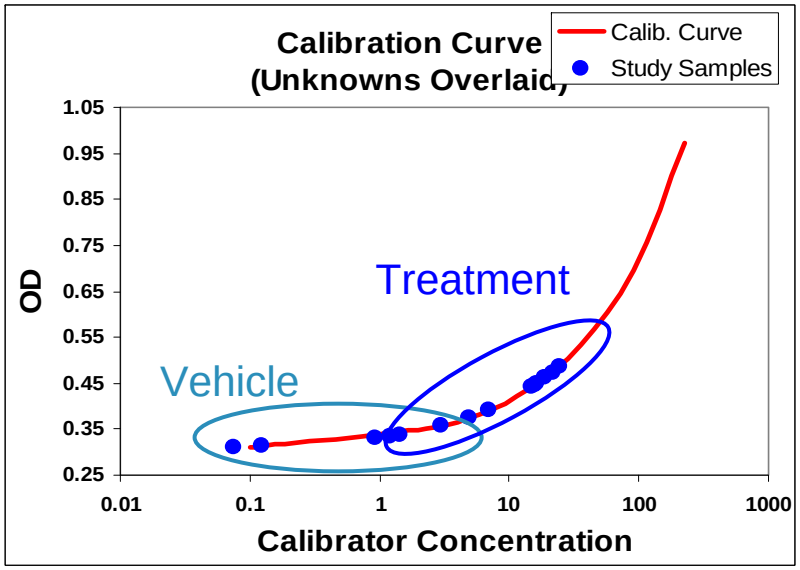


Variability artificially added to the original data in increasing increments. (via simulation).

Biomarker performance decreases with increasing variability.

Assay quality impacts biomarker utility in Clinical Proof-of-Concept study



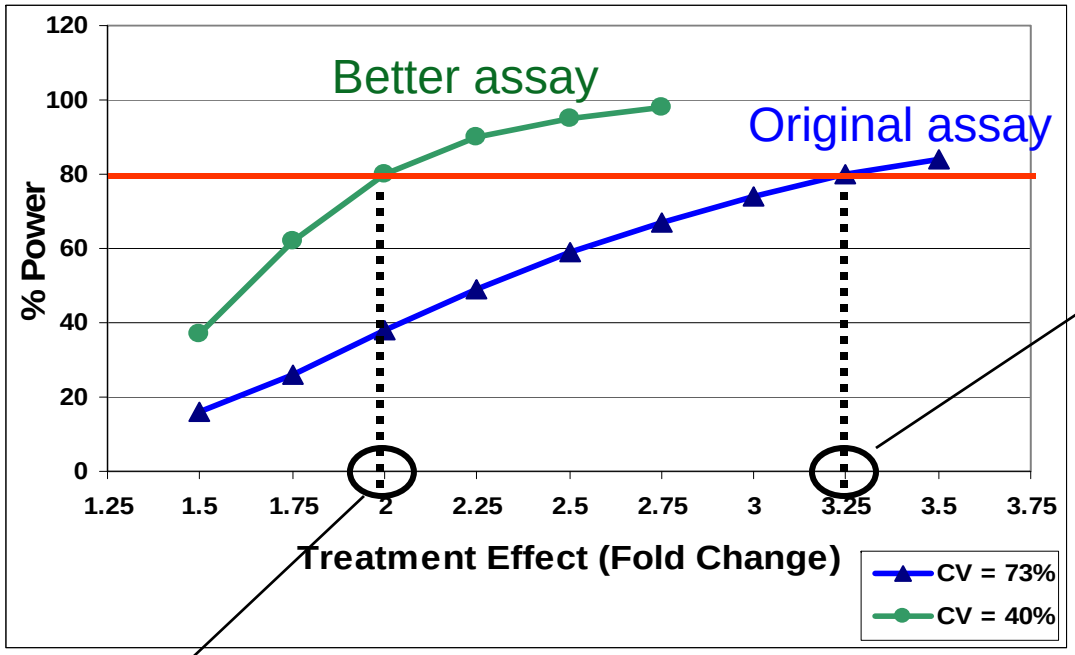


- ELISA calibration curve data from some experiments for measuring a critical PD marker.
- Significant lower plateau in most calibration curves.
- *Need to evaluate where the study samples fall on the curve.*

- Most samples fall on the lower plateau of the curve.
- High variability!
- ***Need to re-optimize this assay to improve sensitivity.***

Assay quality impacts biomarker use in Clinical Proof-of-Concept study (contd.)

Power Analysis



Poor assay sensitivity results in 73% CV.

→ fold-change > 3.25 can be detected with 80% power.

But expected fold-change is 2-fold.

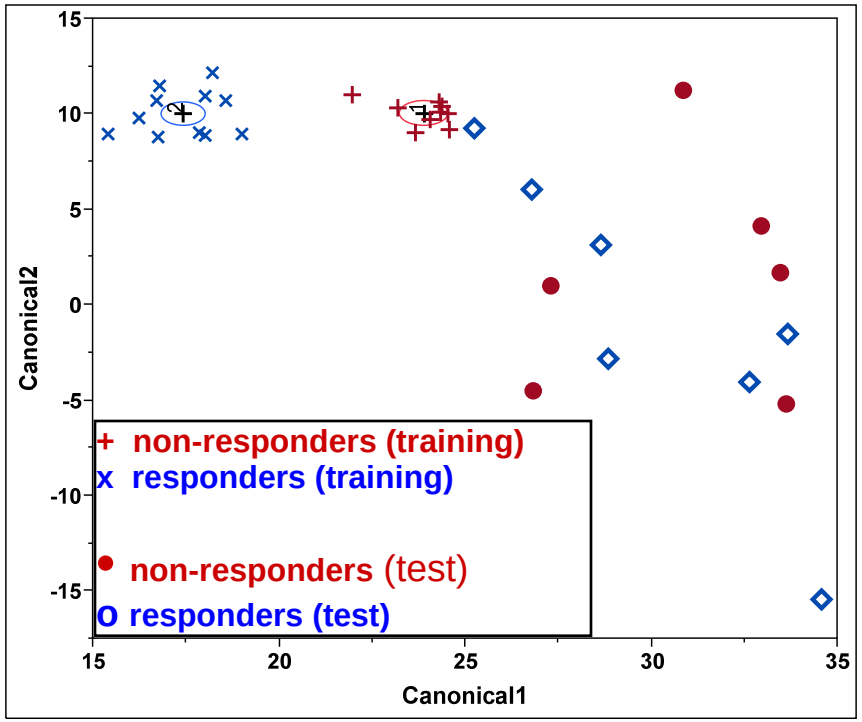
So ***this biomarker is not suitable for PoC study***.

Improving assay sensitivity & reducing CV to 40% enables 2-fold change to be detected with 80% power.

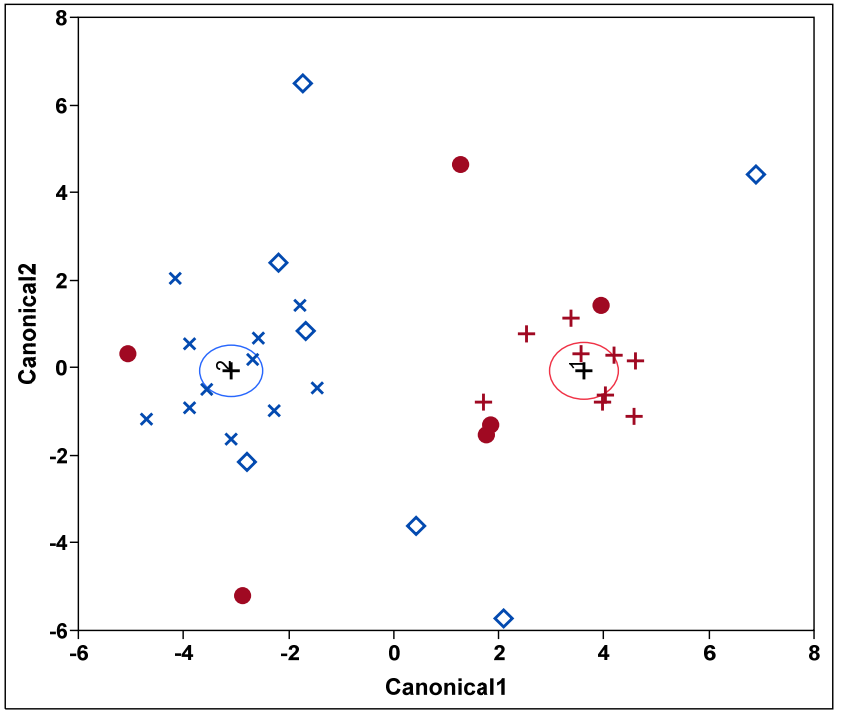
Biomarker is now ready for use in the Clinical PoC study.

Analytical batch-effect impacts biomarker confirmation: Example

Before Normalization



After Normalization



Before normalization, all "responders" are incorrectly predicted.

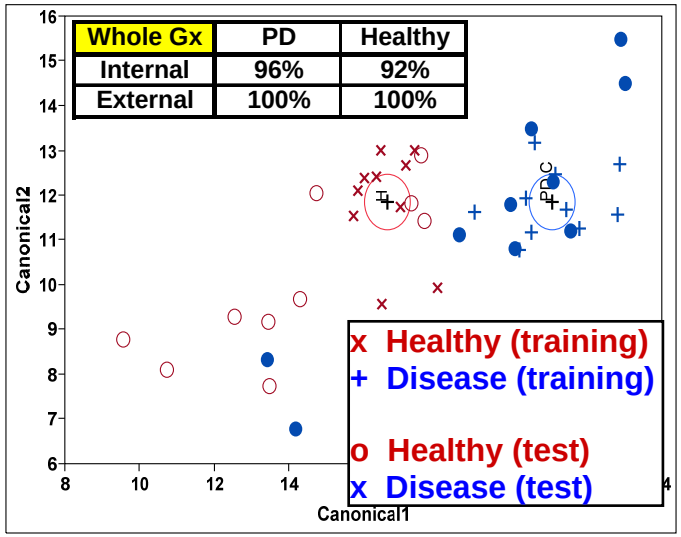
Normalization results in significant improvement, although far from perfect.

- ***Due to other issues (more heterogeneity in external set).***

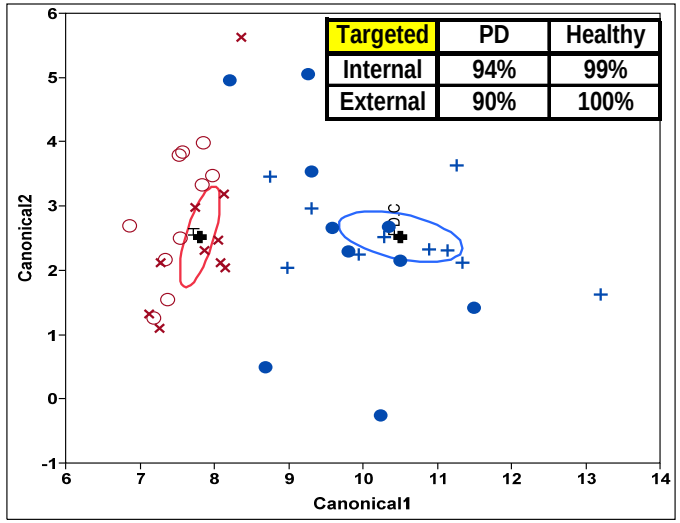
Biological relevance, assay availability, etc. Example

Biomarker signatures from the whole genome may include genes that are not in the biological pathway, or sensitive assays may not be available.

Optimal signature derived from the entire genomic array.



Signature derived from only a subset of genes in the biological pathway and for which sensitive assays were available



Targeted signature performs almost as well (in this example), and is more likely to be accepted for routine implementation.

Biomarker Performance Evaluation
Internal Validation

- Using same data to identify and evaluate a biomarker signature will inflate the performance metrics (e.g., ROC AUC).
- Cross-Validation/Resampling methods help reduce the bias.
- k-fold cross-validation (CV):
  - Original data divided randomly into k equal parts
    - If N=100, k=5, obtain 5 random subsets of 20 each.
  - Leave first part out, "train" on the remaining, "test" on the left-out.
  - Repeat this for each of the other parts;
  - Aggregate predictions from all left-out parts.
  - Calculate performance (e.g., sensitivity/specificity, p-value, …)
  - Repeat this procedure 25 times. Report Mean & SD of the metrics.

Biomarker Performance Evaluation
Internal Validation (contd.)

- Example of Questionable results:
  - Dave et al. *"Prediction of survival in follicular lymphoma based on molecular features of tumor infiltrating cells".* NEJM, Nov. 18, 2004 vol. 35set 2:2159-2169
  - Reasons are explained and illustrated at:
    - http://www-stat.stanford.edu/~tibs/FL/report/index.html

→Unfortunately, poor cross-validation is quite common in biomarker publications.

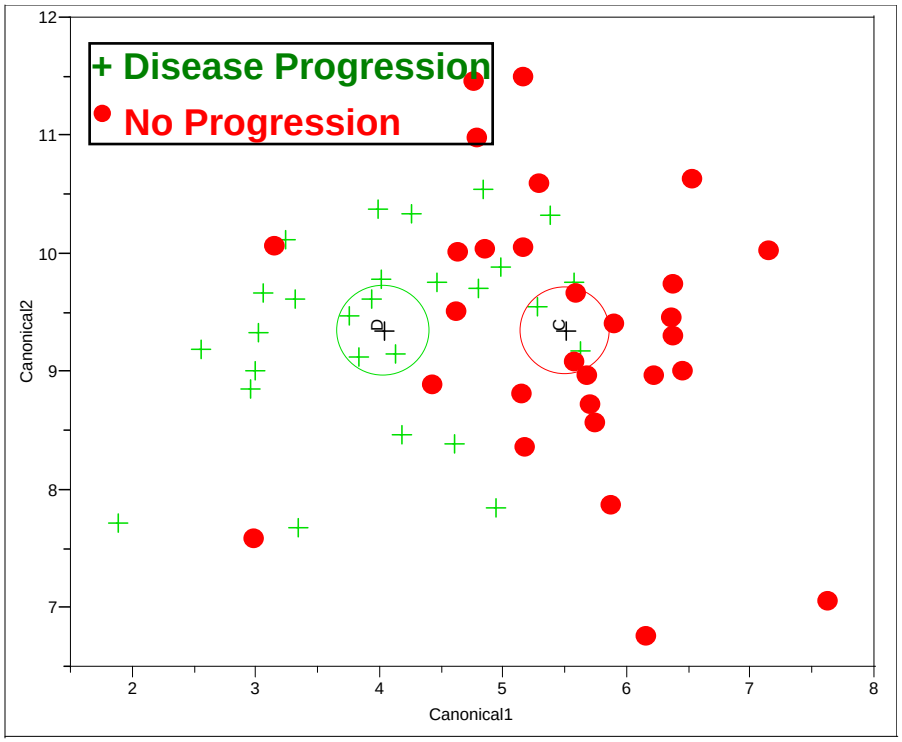
→***Can't take publication/literature claims for granted.***

Biomarker Performance Evaluation
External Validation

- After rigorous internal cross-validation, *test the signatures in independent external cohorts*.
  - Should adequately represent the target population with respect to several features (gender, race, age, disease severity, …)

- Samples in training & external sets are seldom run together.
- So *batch-effect normalization* may be necessary.
  1. Normalize the training & external data.
     - A method that works well in my experience: *Eigen-Strat.*
  2. Apply previously derived signature on the normalized training set.
  3. Use this model on normalized external data to predict the response.

Example 1: Evaluation of Biomarker Performance



6-marker proteomic multiplex signature for possible use in selecting patients for a Clinical Trial

Predictive Accuracy:

- **Internal Cross-Validation:**
  - **No CV: 84%**
  - **Partial CV: 72%**
  - **Full CV: 65%**
- *External Validation (new study): 63%*

Biomarker performance biased by improper Cross-Validation

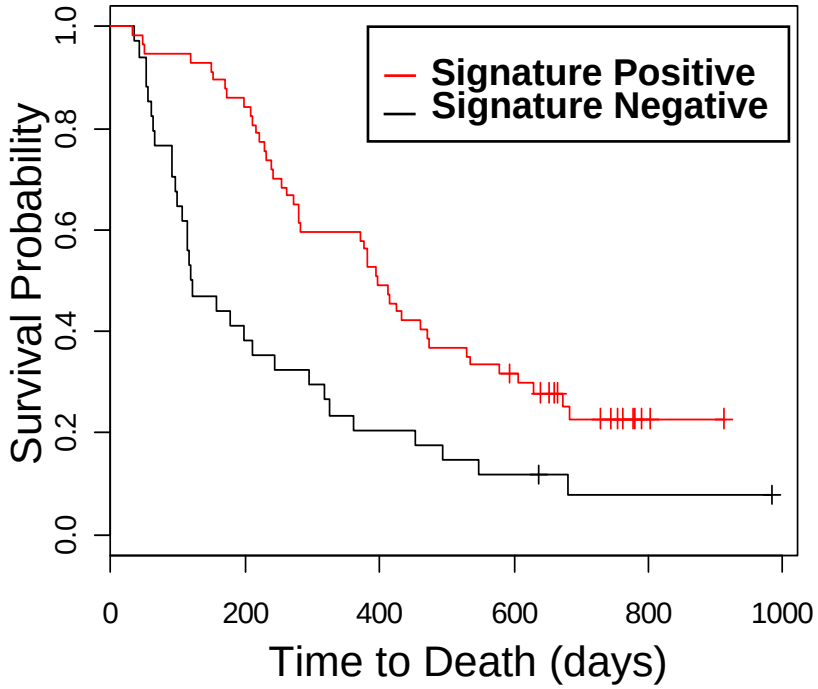
Example 2: Evaluation of Biomarker Performance

4-SNP Genotype Signature for Predicting Patient Response to Treatment

- **Derived from a large genotype array (100s of SNPs) via a Statistical Algorithm**

Signature Positive: SNP-1 ≠ WT,SNP-2 ≠ WT, SNP-3 = WT, SNP-4 ≠ WT

- **Patients in this Signature Positive group are expected to respond better.**



p-value of Treatment Effect in Signature Positive vs. Negative:

- **Internal Validation:**
  - **No Cross-Validation: p < 0.0001**
  - **10-fold Cross-Val: p = 0.06**
- **External Validation (independent clinical study): p =0.1**

Improper Cross-Validation exaggerates biomarker performance.

Robustness

During a study, additional variability can be introduced (unavoidable factors)

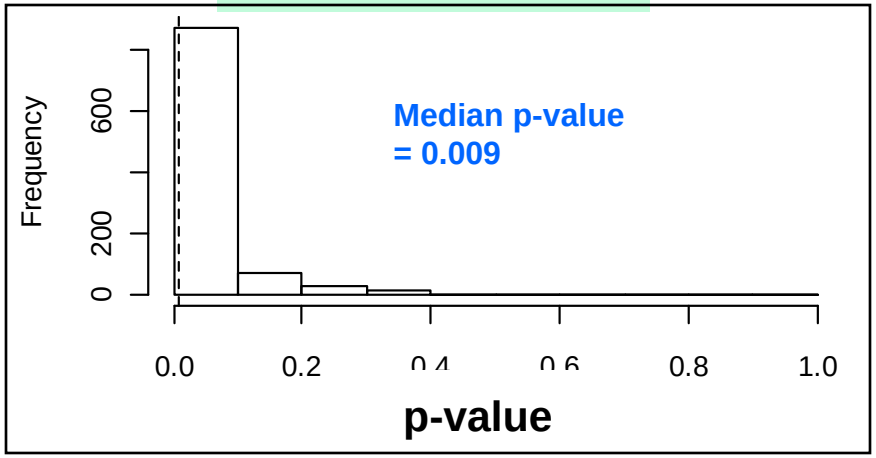
- changes in reagents, instruments, operators, sample collection/storage, …
- *This is typically not accounted for during biomarker validation/evaluation.*

Example: 5-marker Signature for identifying patients more likely to respond to treatment. Robustness of this signature is evaluated via Simulations.

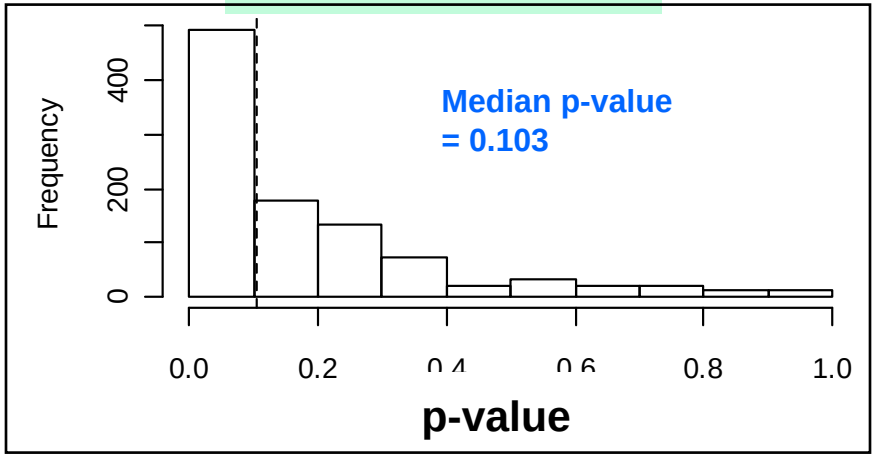
15% CV & 30% random noise are artificially added to the original data.

Distribution of p-values for Treatment Effect evaluated via 1000 iterations.

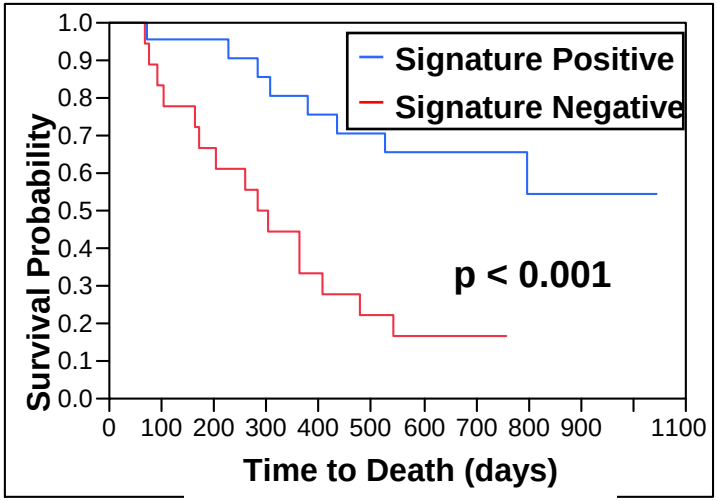
Additional 15% CV



Additional 30% CV

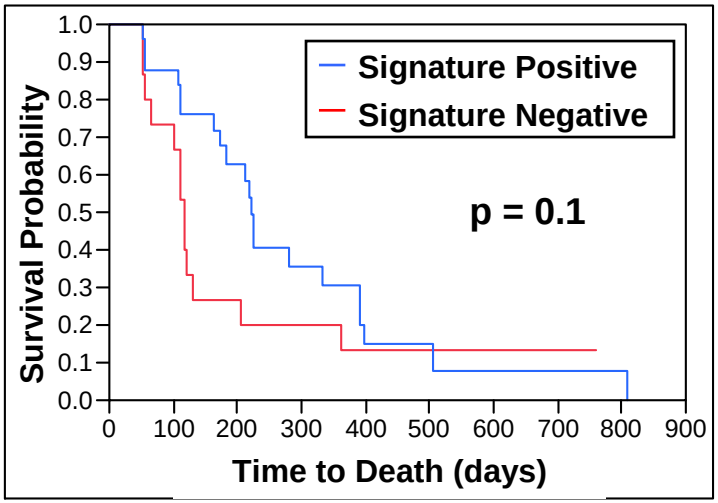


Translation



Biomarker Signature derived & evaluated in male cancer patients

Confirmed via external validation on same population



Same Biomarker Signature does not perform well when tested in a different study (females, older age group, more severe cancer)

This gets more challenging between animals & humans!

Summary

- For most diseases & treatments, biomarkers are critical for clinical drug development.
  - Non-responders, disease progression, safety monitoring, …
- Some practical challenges:
  1. Variability (Analytical & Biological)
  2. Biological Relevance, Assay availability, etc.
  3. Predictive performance evaluation
     - Internal Validation (cross-validation methods)
     - External Verification
  4. Robustness, Reproducibility, etc.
  5. Translation (species, demographics, disease subtypes, etc.)
- ***Consideration of these & other challenges is critical for successful biomarker strategy.***