

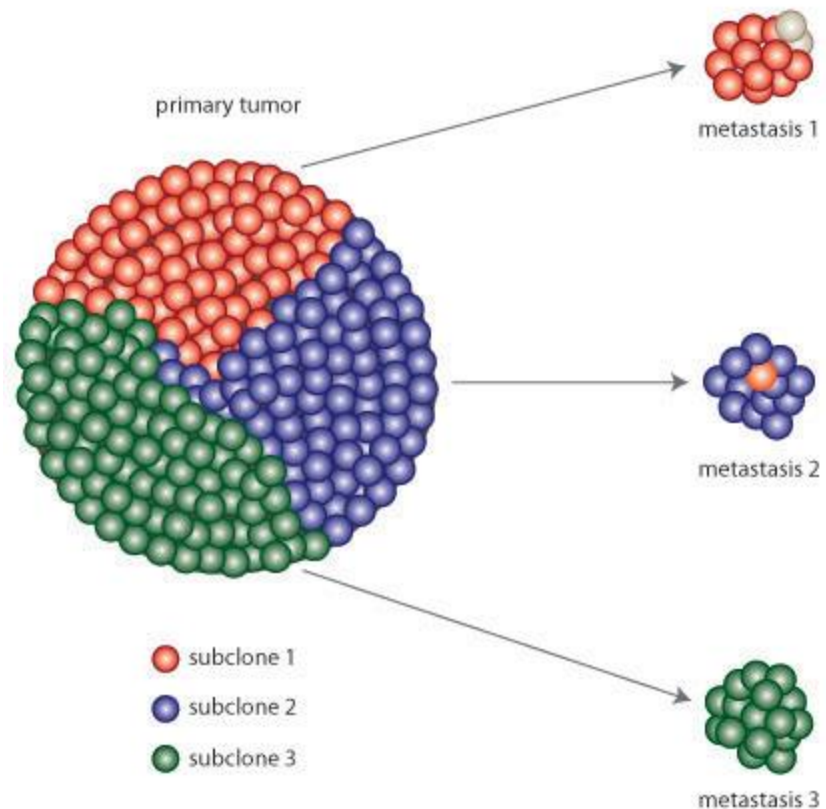
A Measurement Error Model Framework for Biomarkers based on Emerging Technologies

*Dong Wang
National Center for Toxicological Research
US Food and Drug Administration*

The opinion expressed in this presentation is of the author. It does not necessarily reflect the position of the US Food and Drug Administration.

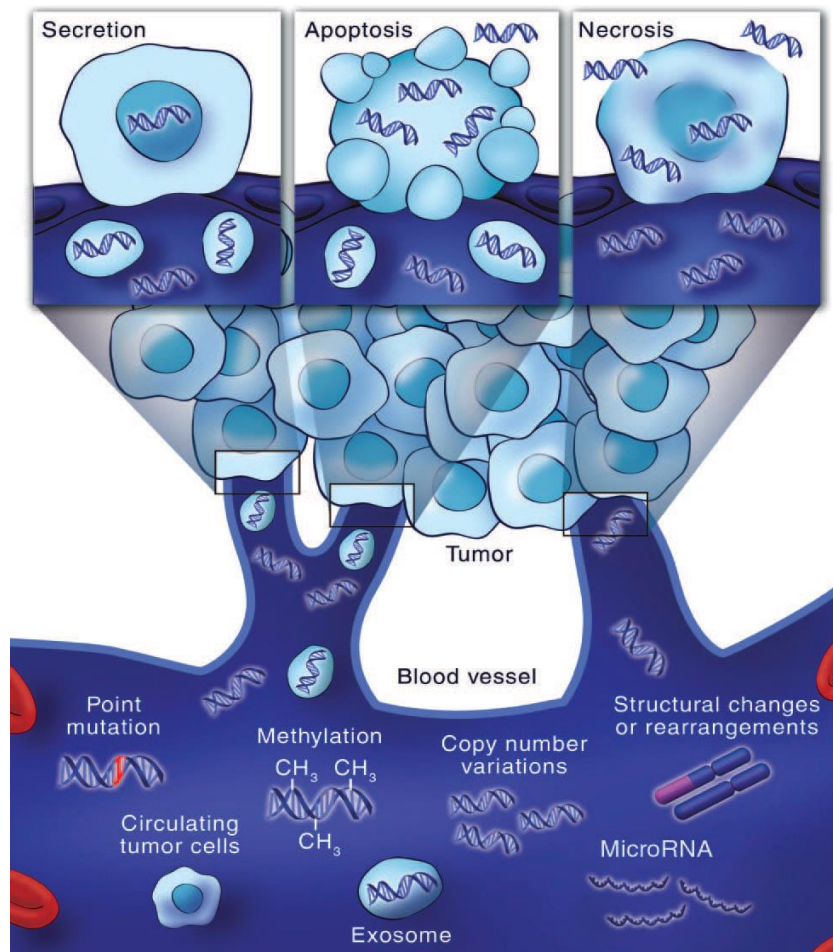
The drawbacks of surgical tissue biopsy

- Difficulty in acquiring adequate amounts of tissue
- Limited repeatability
- Biases that result from sampling a small area of a single tumor, including those due to tissue heterogeneity
- Being invasive, costly and time-consuming
- possibility for clinical complications



Roche

Circulating Tumor DNA (ctDNA) and Liquid Biopsy



- ☐ ctDNA is tumor DNA that has been shed into the bloodstream
- ☐ Variants of Next Generation Sequencing technology are able to detect tumor specific mutations from the plasma sample
- ☐ Liquid biopsy is a minimally invasive test
- ☐ Less biased assessment of the tumor genetic profiles

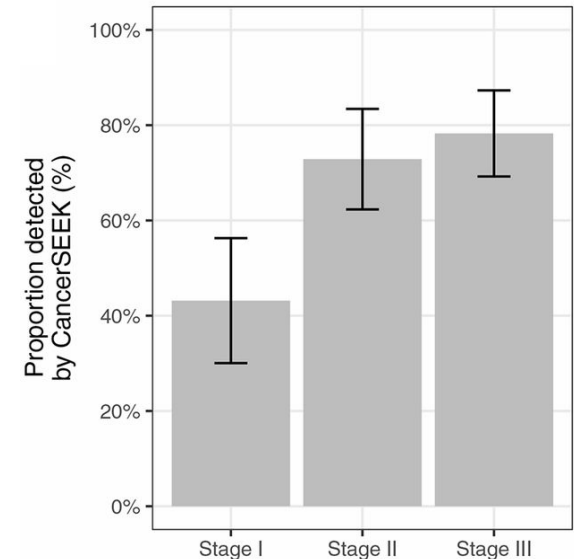
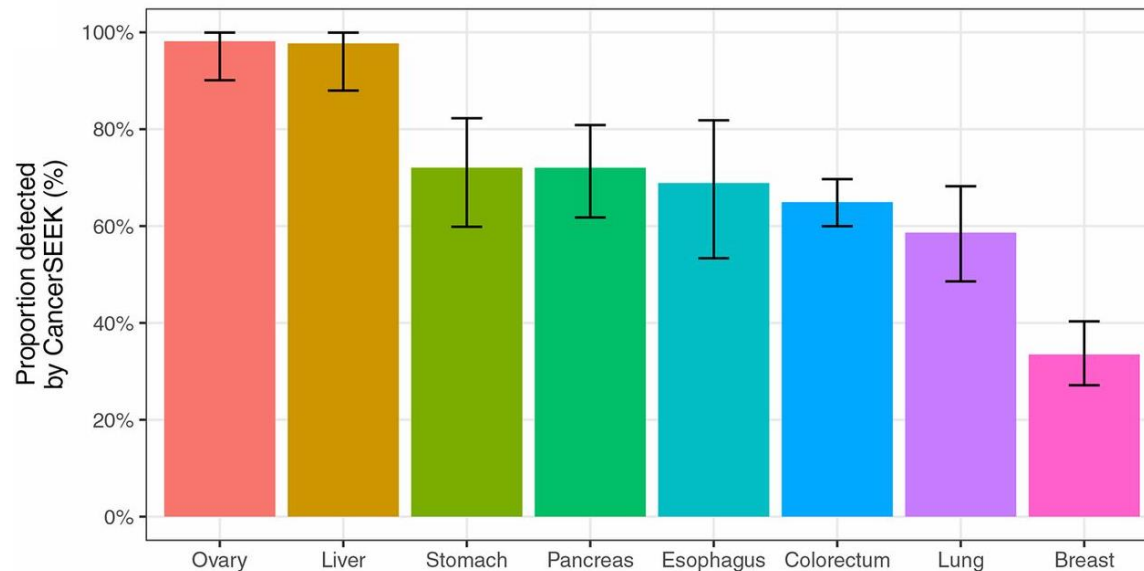
Diaz and Bardelli, 2014

Challenges of Liquid Biopsy



- The amount of tumor-derived ctDNA varies in different patients.
- Early stage patients probably generate less ctDNA.
- Normal tissue also sheds DNA into the blood, which constitutes the overwhelming majority of cell free DNA.
- Even with highly sensitive NGS technology (with barcoding), the actual sensitivity varies widely (<50%-90%).
- It is harder for early stage detection.
- Current statistical methods usually assumes that biomarker status can be determined without error.
- How do we incorporate misclassification in liquid biopsy in statistical analysis?

An Example Using CancerSeek (Cohen et al., 2018)



- Sensitivity varies widely between cancer types and stages
- Specificity is high (with caveat)
- How can we accommodate the misclassification in biomarker validation, subgroup selection, and clinical utility assessment?

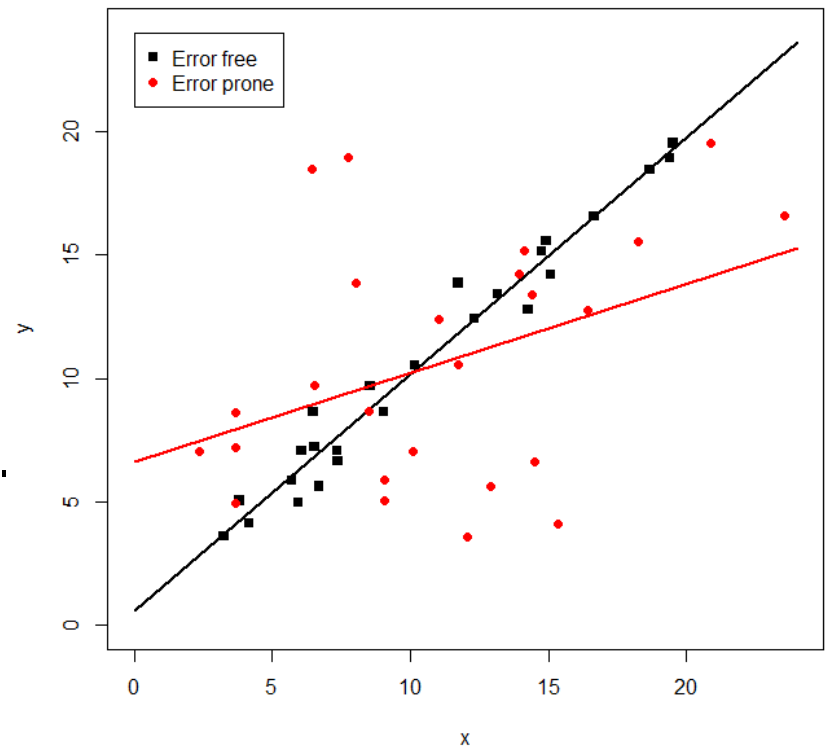
Measurement error and misclassification



- Often impossible to measure some covariates accurately:
e.g. Dietary intake, blood pressure, weight
- Instead, we have error-prone observations
- How these relate to the underlying true values is the focus of 'measurement error model'
- Common model: classical additive error:
$$\times \text{Observed} = \text{True} + \text{Measurement Error}$$
- ...but other models are available
- Misclassification model for categorical variables

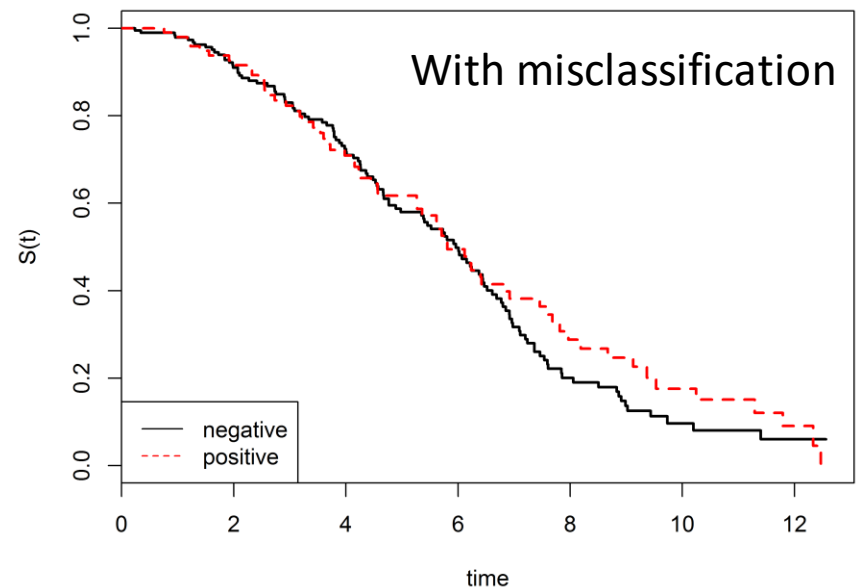
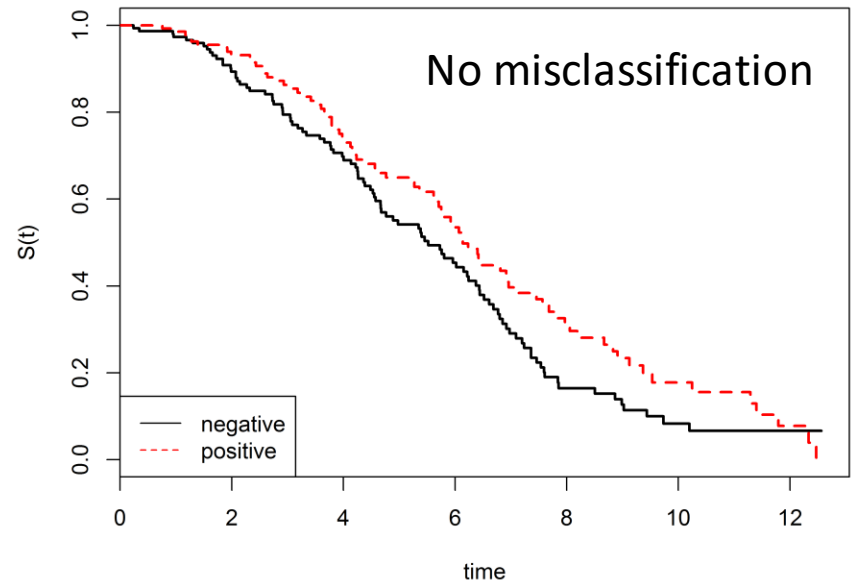
The effect of measurement errors in covariates

- Simple linear regression: $Y = \beta_0 + \beta_x X + \epsilon$
- Classical measurement error (continuous covariate):
 $W = X + U$
- Regress Y on W to obtain an estimate of β_x^* where
- $\beta_x^* < \beta_x$
- Simple linear regression: attenuation, the regression coefficient is smaller in magnitude and less significant.
- But things are more complicated for other models



An example of misclassified covariate

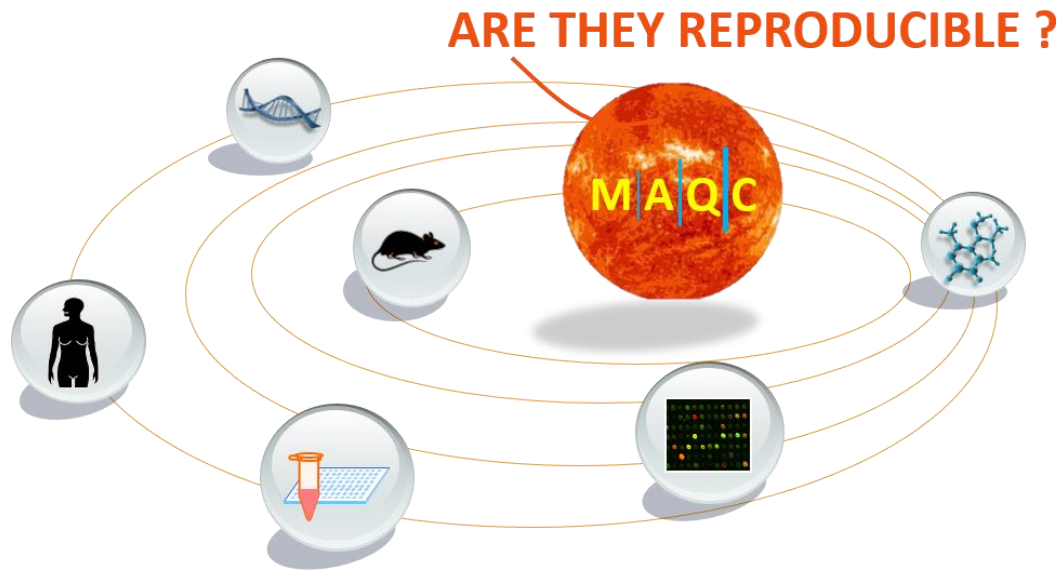
- Proportional hazards model for a single arm trial with a binary biomarker
- The biomarker X is predictive of survival outcomes under this treatment
- However X cannot be measured exactly, instead a potentially misclassified version X^* is available
- In this simulation, X^* has sensitivity of 0.7 and specificity of 0.98
- The coefficient is significant for the error free biomarker, but is not significant for the misclassified version



Impact of measurement errors and misclassification

- In the simple linear regression setting, measurement errors in the covariate leads to smaller coefficient values and less significant result (attenuation)
- For more than one covariate or nonlinear models, the situation is more complex (both the magnitude and sign of the coefficients could change)
- For binary covariate, misclassification generally leads to attenuation (with probability tending to one)
- ❖ Needs for extensive characterization of misclassification of deep sequencing based biomarkers in precision medicine applications
- ❖ Can we make better decisions with information on misclassification

The MAQC project at FDA and other institutions



MAQC-IV/SEQC2

2015 - Present

QC and reliable use of Whole Genome Sequencing (WGS) and Targeted Gene Sequencing (TGS) in clinical application and regulatory science research

- Current study will generate detailed performance characteristics for deep sequencing in identifying somatic mutations

SEQC2 Study for Performance Metrics of Targeted Sequencing



Sample	Dilution Ratio (A by B)	Mutant Allele Fraction Range
A	-	2.5% - 15%
B	-	0%
C	1:2	1.25% - 7.5%
D	1:5	0.5% - 3%
E	1:25	0.1% - 0.6%
F	1:125	0.02% - 0.12%

Sample B is a normal sample for an individual to minimize mutation background and maximize homogeneity.

- ❑ Carefully designed benchmark samples with cell line DNA to characterize performance in realistic settings
- ❑ Contrived cfDNA samples with synthetic plasma
- ❑ Cross-site comprehensive QC assessment for both gDNA and ctDNA deep sequencing
- ❑ Clinical trials to compare tissue biopsy and liquid biopsy results

Can we make better decisions with a detailed understanding of misclassification regarding NGS based biomarkers?

- There is a variety of methods to correct the bias in the setting of measurement error or misclassification
- We will focus on a class of methods called simulation extrapolation (SIMEX)
- Consider the linear regression setting: $Y = \beta_0 + \beta_x X + \epsilon$, where the true predictor X cannot be observed. Instead, we observe $W = X + U$, where $U \sim N(0, \sigma_u^2)$ is independent of (X, Y) .
- Using W in place of X will result in biased estimate for the regression coefficient.
- We cannot obtain the error-free X , but we can make it worse by adding more random error

The SIMEX algorithm

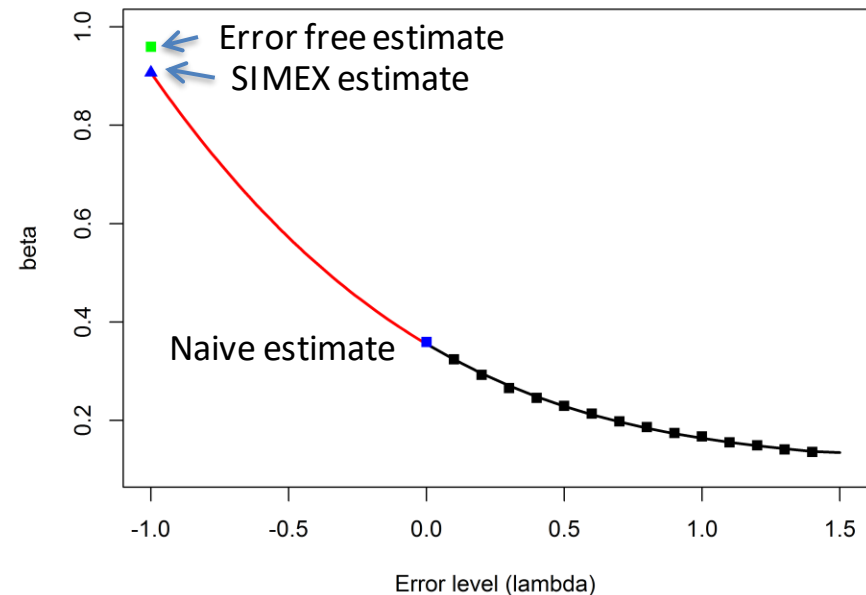
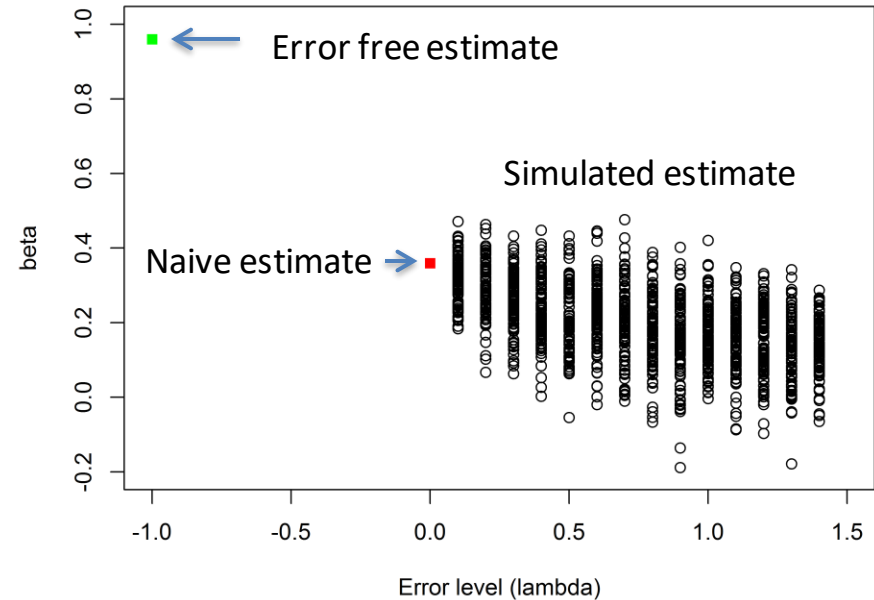
- Use simulation to generate additional data sets with increasingly larger errors $(1 + \lambda)\sigma_u^2$

$$W_{b,i}(\lambda) = W_i + \sqrt{\lambda}U_{b,i}, \quad i = 1, \dots, n, \quad b = 1, \dots, B$$

- $E\{W_{b,i}(\lambda)|X_i\} = X_i, \quad \text{Var}\{W_{b,i}(\lambda)|X_i\} = (1 + \lambda)\text{Var}(W_i|X_i)$
- $E\{[W_{b,i}(\lambda) - X_i]^2|X_i\}$ converges to zero as $\lambda \rightarrow -1$
- Calculate the estimator $\hat{\beta}_b(\lambda)$ with each pseudo data set
- Average $\hat{\beta}(\lambda) = B^{-1} \sum_{b=1}^B \hat{\beta}_b(\lambda)$
- Extrapolate $\hat{\beta}(\lambda)$ to $\lambda = -1$ to obtain the SIMEX estimator

An Example for Linear Regression with Additive errors

- If we know the form of the extrapolation function exactly, the SIMEX estimator is consistent
- In practice, we have to guess what is the function form
- The quadratic function is often used



Misclassification Simulation Extrapolation: MC-SIMEX

- Misclassified predictors (biomarkers based on deep sequencing) for nonlinear models (logistic, proportional hazards ...)
- *A General Method for Dealing with Misclassification in Regression: the Misclassification SIMEX* (Kuchenhoff et al. 2006)
- Consider the model $h(y_i) = \beta_0 + b_{1j}T + \beta_2(Tx_i^*) + \epsilon_i$ where T is the treatment indicator, x_i^* is an error prone biomarker and $h(y)$ is a link function. For biomarker validation, we want to test whether β_2 is significant.
- With MC-SIMEX, we will generate X^* with even larger misclassification rate, calculate a series of estimates, then extrapolate to the case of no misclassification

MC-SIMEX algorithm

- Let $\pi_{ij} = P(X^*|X = j)$, and the misclassification matrix is $\Pi = [\pi_{ij}]$.
- For the binary case

$$\Pi = \begin{bmatrix} \pi_{00} & 1 - \pi_{11} \\ 1 - \pi_{00} & \pi_{11} \end{bmatrix}$$

- $\Pi^\lambda = E\Lambda^\lambda E^{-1}$, where Λ is the diagonal matrix of eigenvalues and E is the matrix of corresponding eigenvectors.
- If we further misclassify X^* by Π^λ to get X^{**} , then X^{**} is a misclassified version of X by $\Pi^{1+\lambda}$
- Generate a series of new data sets with increased misclassification indexed by λ
- Extrapolate $\hat{\beta}(\lambda)$ back to $\lambda = -1$ to give the MC-SIMEX estimator as $\Pi^0 = I$

Variance estimation

- Asymptotic variance can be derived for some cases like logistic regression
- Assuming the maximum likelihood estimator can be obtained by solving the estimating equations $\sum_{i=1}^n \psi(Y_i, X_i, \beta) = 0$
- Let $\hat{\beta}(\lambda) = g(\lambda, \gamma)$, where γ is the parameter for the extrapolation function
- Asymptotically, $\hat{\beta}_{SIMEX} = \frac{\partial}{\partial \gamma} g(0, \gamma) \Sigma(\gamma) \frac{\partial}{\partial \gamma} g(0, \gamma)^T$
- $\Sigma(\gamma)$ has a pretty complex form based on the asymptotic expansion of estimating functions
- For more complex models, the asymptotic variance based estimator is not practical, bootstrap or jackknife will be more useful.

Biomarker validation

- Consider a proportional hazards model with data from the treatment arm only regarding a single biomarker, the conditional hazard function for survival time Y_i given $\{X_i, Z_i\}$ is

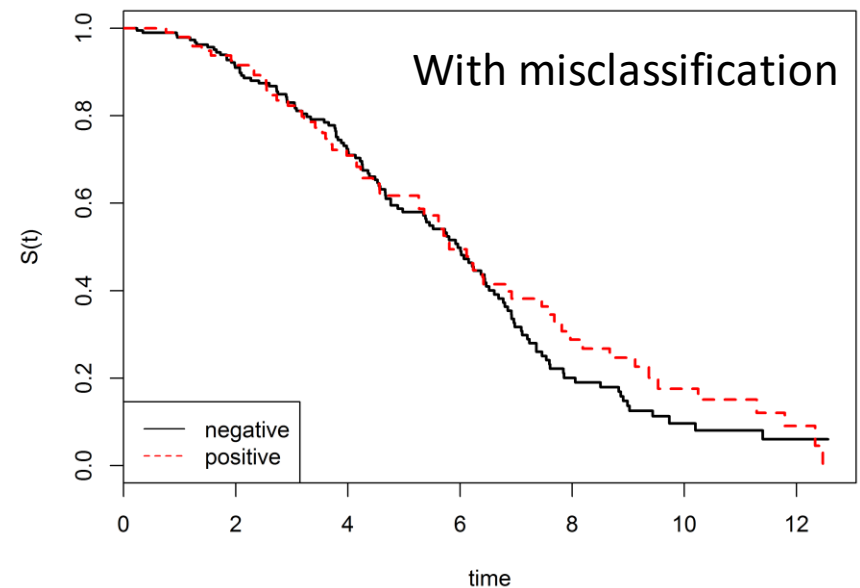
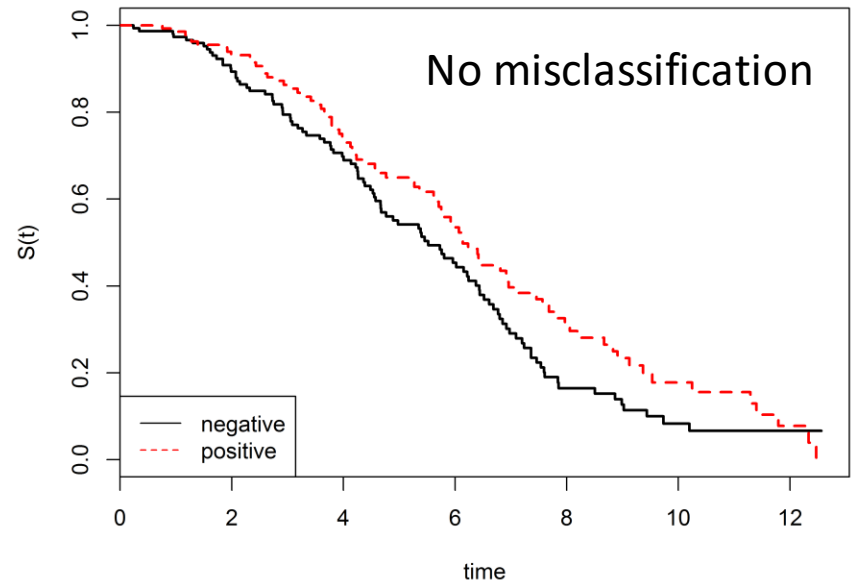
$$\lambda(y|X_i, Z_i) = \lambda_0(y) \exp(\beta_x^T X_i + \beta_z^T Z_i),$$

where X_i is the biomarker state and Z_i represents other variables for the i th patient.

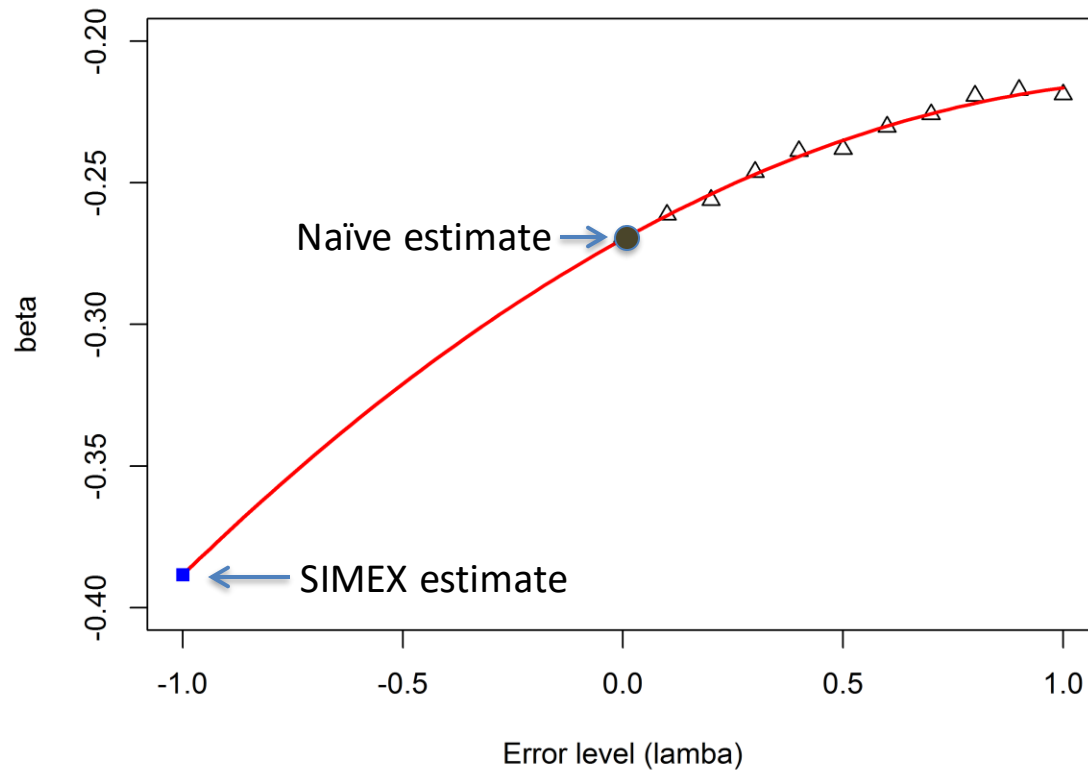
- Simulation: data were generated for 150 patients in the biomarker positive and biomarker negative group respectively. Censoring rate is around 40%. A Weibull baseline hazards function is used and $\beta_x = -0.4$.
- The true biomarker status X is not available, instead a misclassified version X^* is used with the sensitivity being 0.7 and the specificity being 0.98.

An example of misclassified covariate

- Proportional hazards model for a single arm trial with a binary biomarker
- The biomarker X is predictive of survival under this treatment
- However X cannot be measured exactly, instead a potentially misclassified version X^* is available
- In this simulation, X^* has sensitivity of 0.7 and specificity of 0.98
- The coefficient is significant for the error free biomarker, but is not significant for the misclassified version

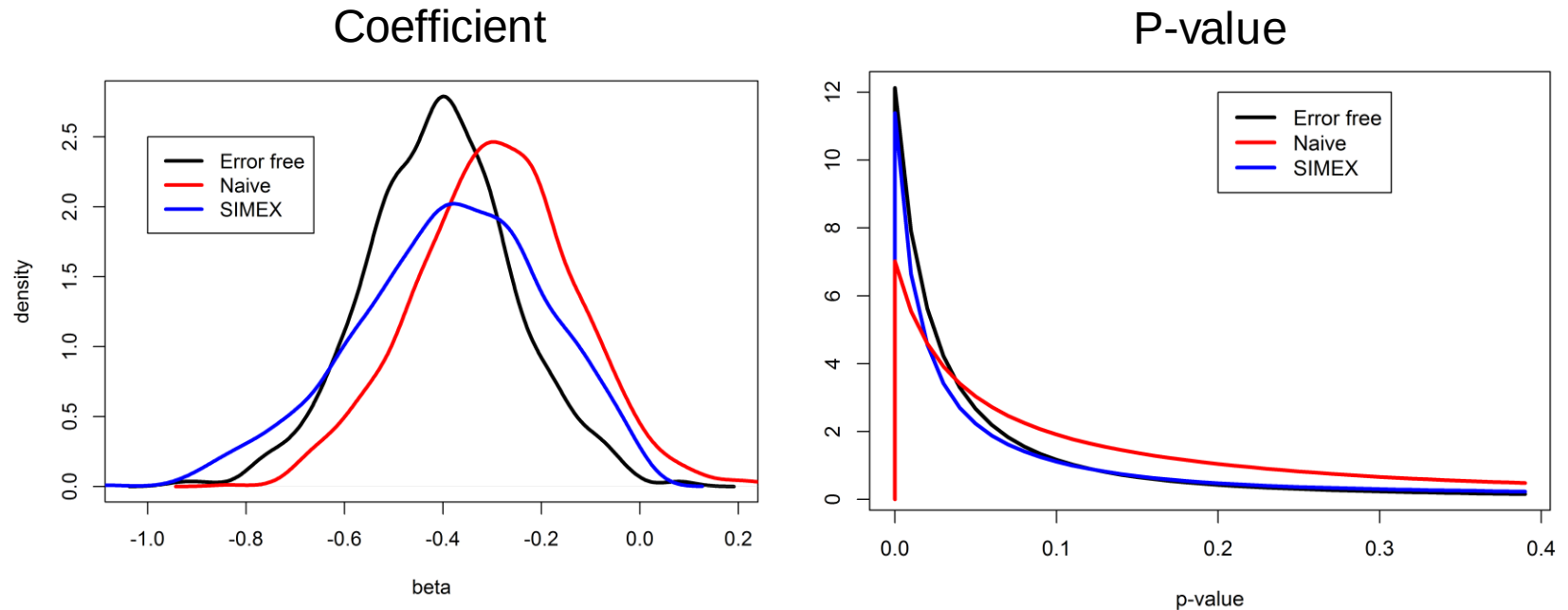


An example of misclassification in a biomarker for survival outcomes



- SIMEX estimate for β_x using a simulated data set. The true coefficient used in the simulation is -0.4

Simulation with 1000 data sets



- SIMEX performs as expected for the validation of a single biomarker with misclassification
- Bias was reduced with SIMEX
- The variance is larger, so no need to use it if misclassification is minimal

Summary

- Learnings from MAQC and other projects provide great opportunity to improve decision making in precision medicine applications
- Measurement error model based approaches such as SIMEX can be used to sharpen the results for biomarker validation
- It has the potential to improve the efficiency of identifying promising biomarkers for subsequent trials and provide guidance for further development
- This approach can facilitate decisions regarding which technology to use and what application scenario to focus
- The implication for subgroup selection and clinical utility assessment will be explored further

Acknowledgement

NCTR

Joshua Xu

Weida Tong

CDRH

Pamela Gallagher

Jincao Wu

Jai Pandey

CDER

Sue-Jane Wang

Gene Pennello

FDA/OCS

Samir Lababidi

Thank you!

