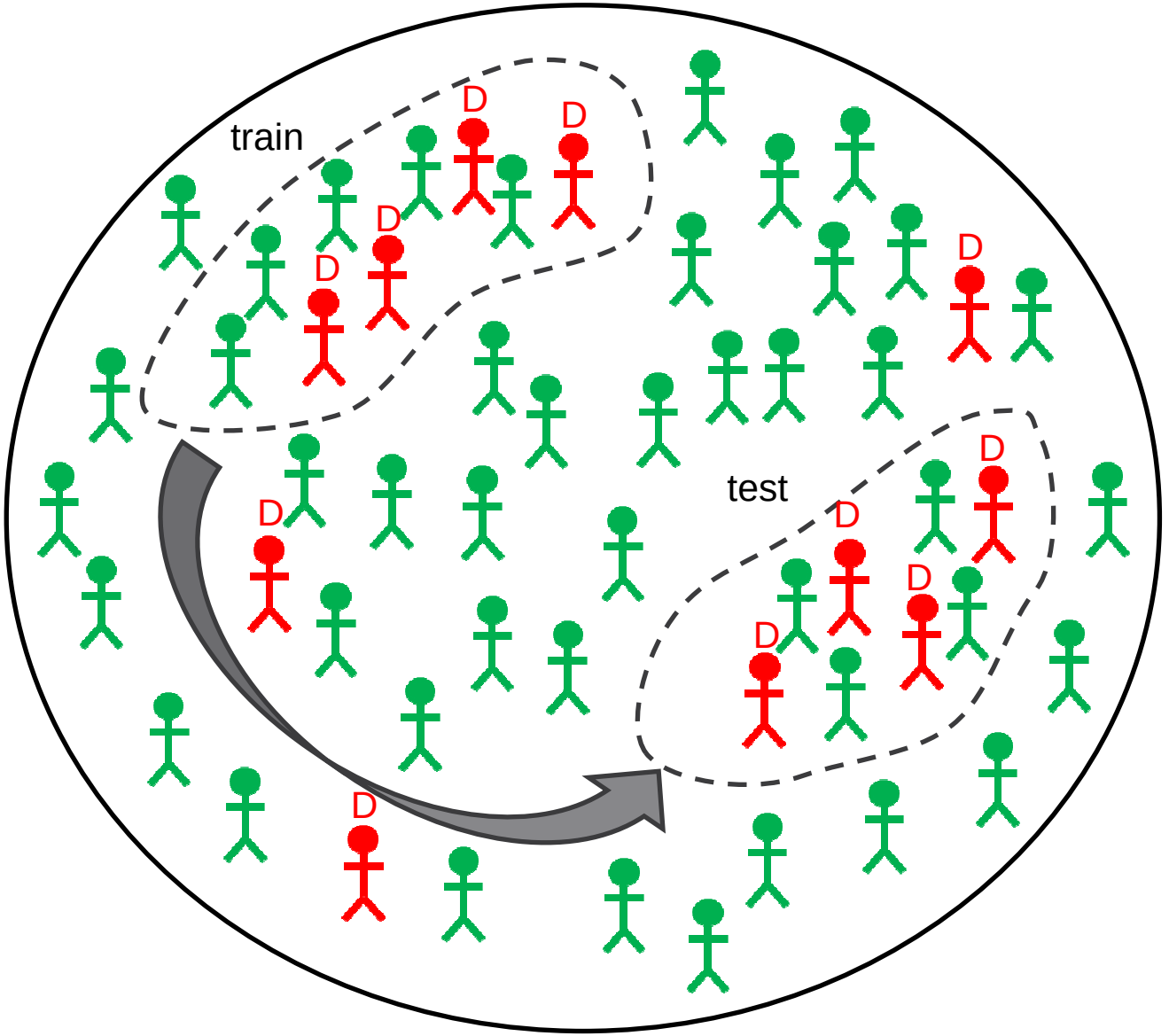
IMPROVER
systemsbiologyverification

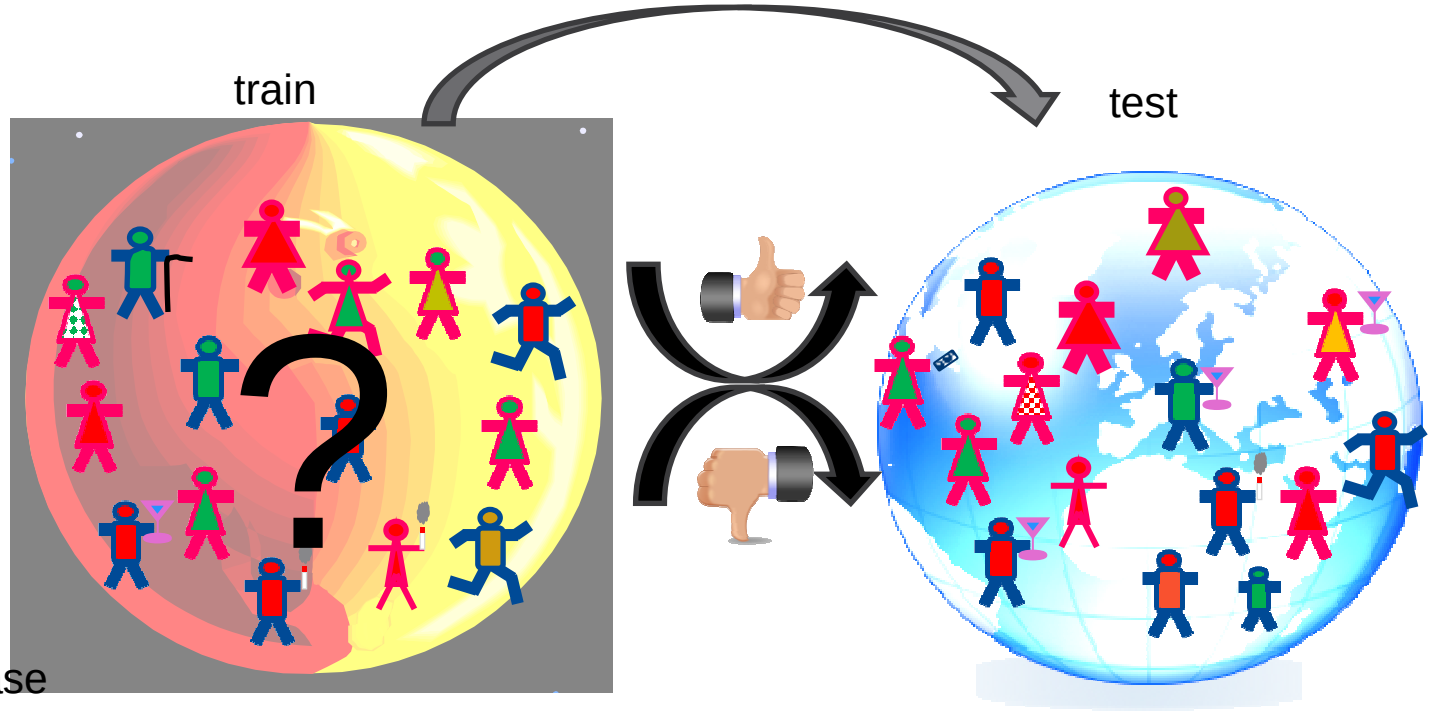
Lessons Learned from the Challenge
Messages from Biomedicine

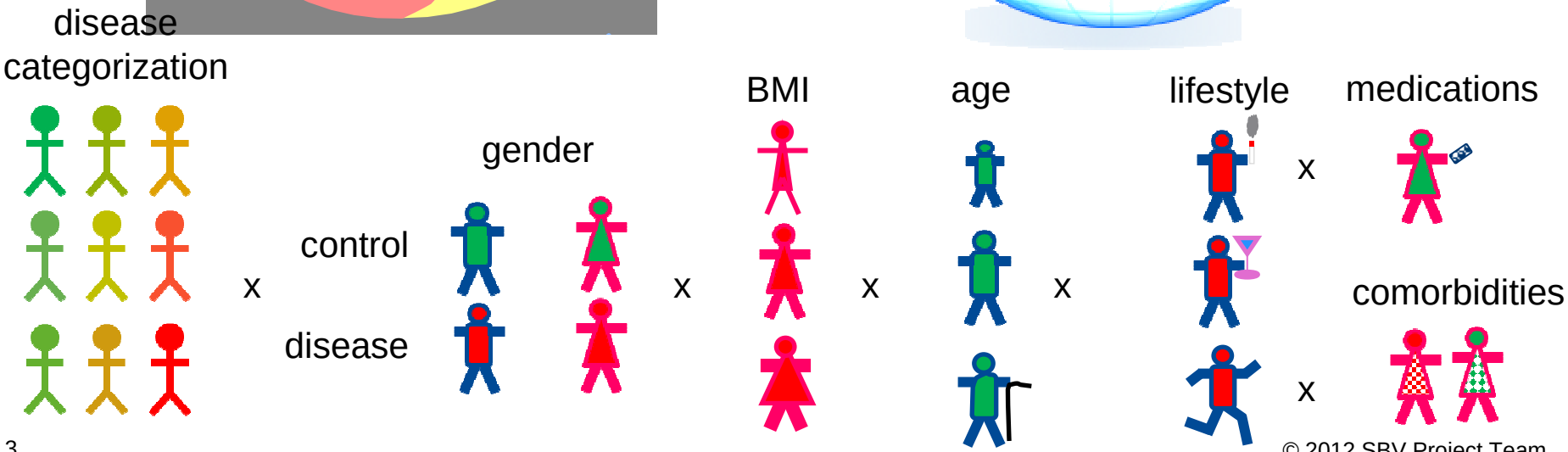
Stéphanie Boué, Ph.D.
PMI Research & Development
3rd October 2012

Diagnostic Signature Challenge - Theory

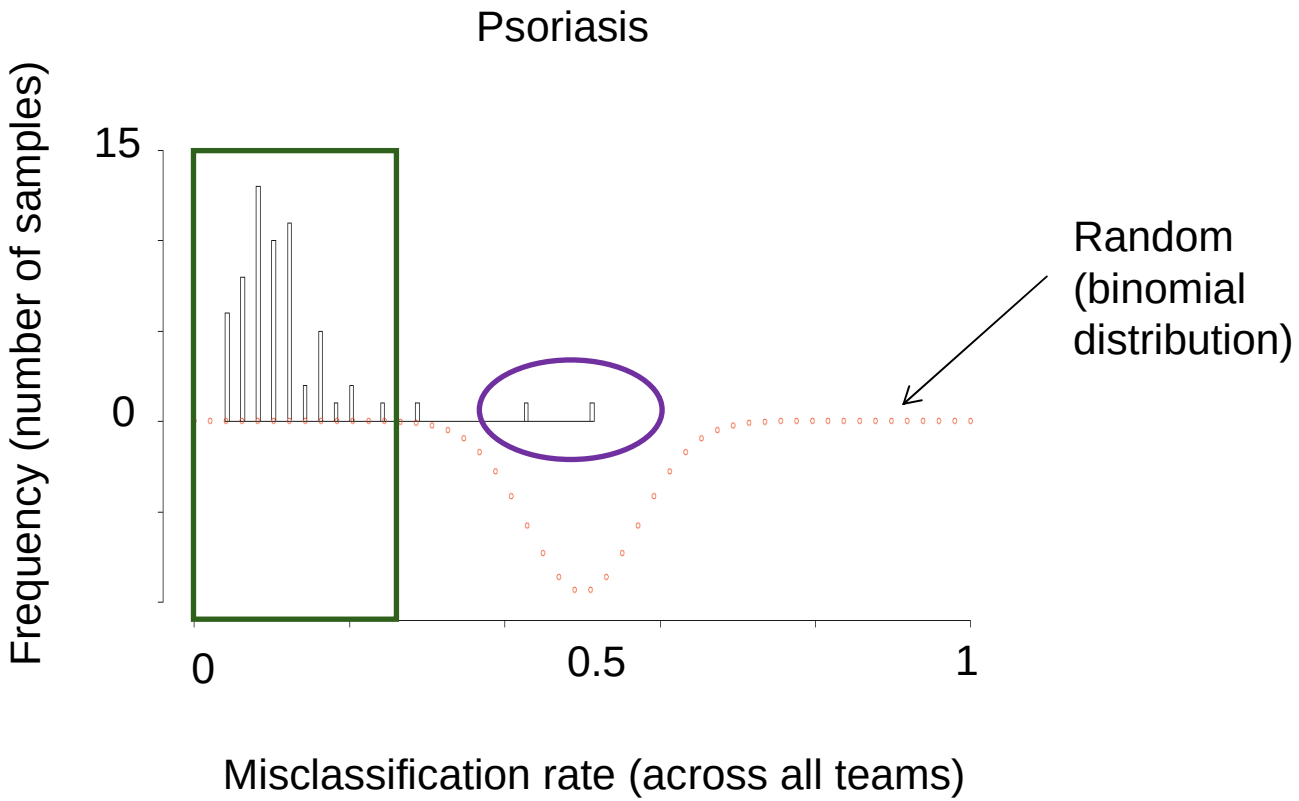


Diagnostic Signature Challenge - Practice





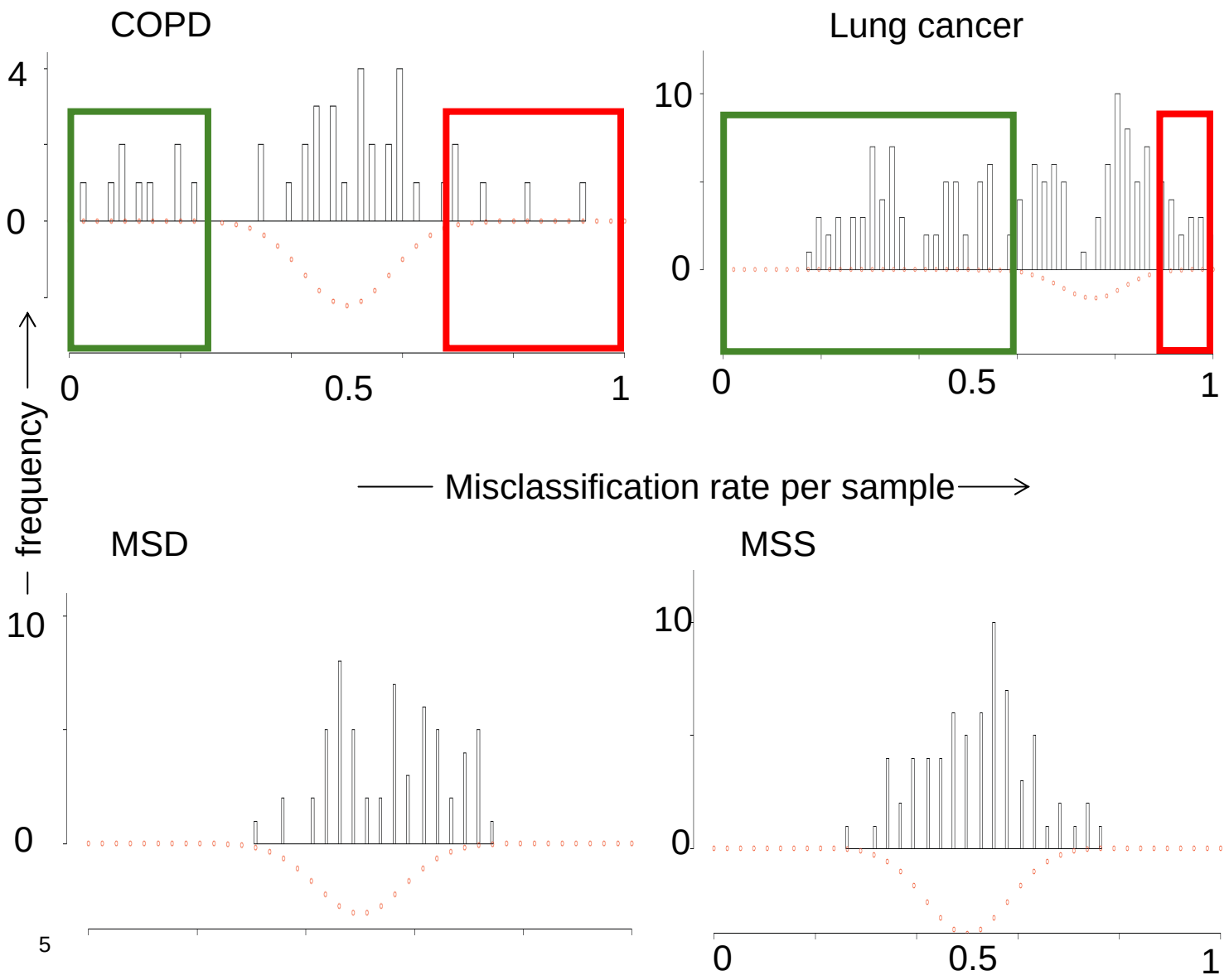
Misclassification Rate of Samples



While 58 samples are correctly classified by most teams, two samples are misclassified by more than 40% of the teams

Misclassification rate better than random

Misclassification Rate of Samples

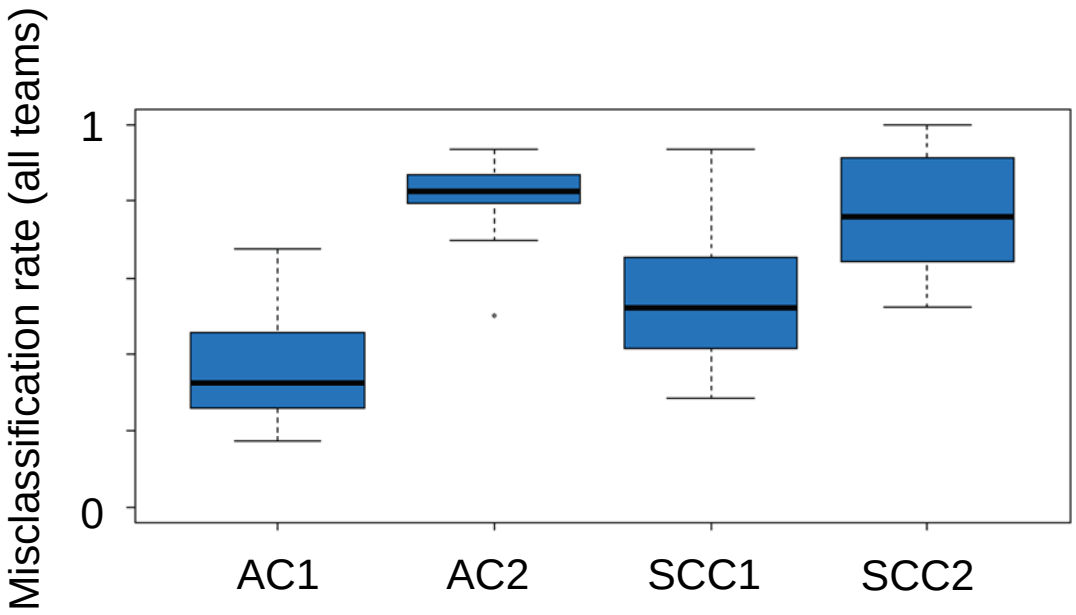


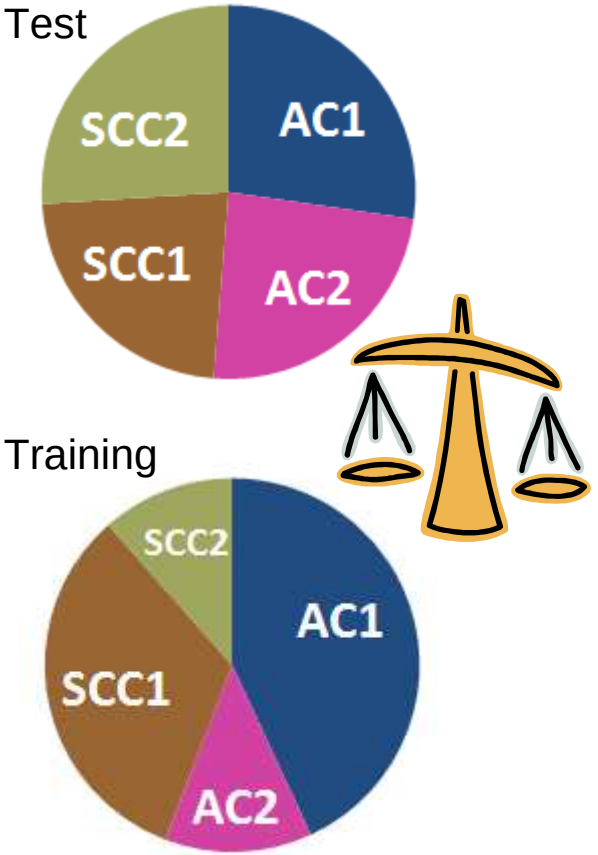
Misclassification rate better than random

Misclassification rate no different than random

Misclassification rate worse than random

Training and Test Datasets – Example of Lung Cancer





Test dataset

LC subtype	Stage	Nb Samples	
AC	I	41	27.3%
	II	36	24%
SCC	I	34	22.7%
	II	39	26%

Training dataset *

LC subtype	Stage	Nb Samples	
AC	I	112	43.1%
	II	33	12.7%
SCC	I	85	32.7%
	II	30	11.5%

* GSE2109, GSE10245, GSE18842, GSE37745

Possible Causes of Misclassification

In addition to possible technical challenges due to the acquisition of training and test data in different populations, causes of misclassification may be related to:

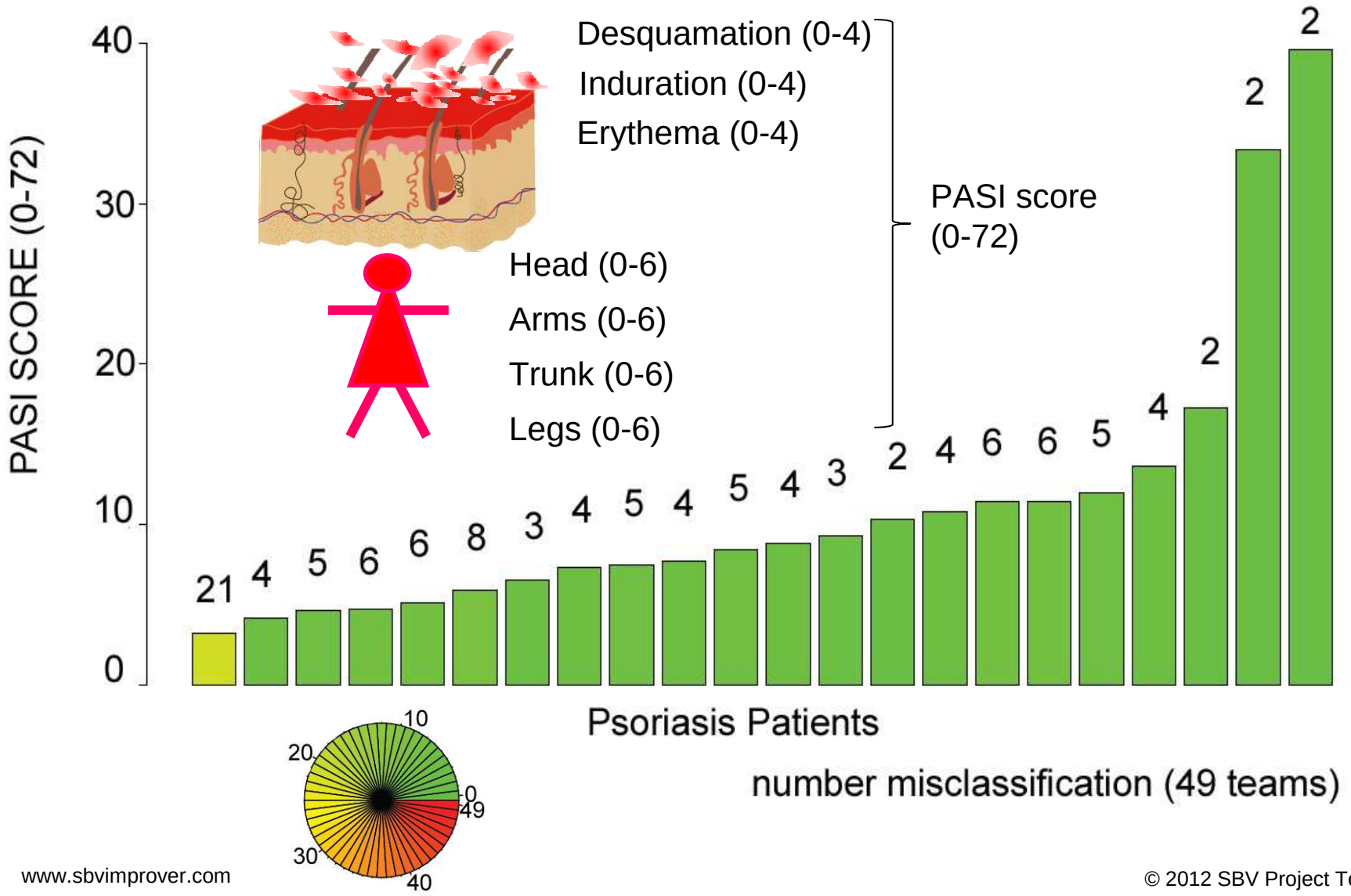
- The definition of classes, especially for disease stages, is based on discrete scales, whereas disease progression is usually continuous. Molecular staging may be more granular than the pathological/anatomical staging
- A disease phenotype could in fact encompass an array of related pathologies with different molecular origin/mechanisms
- The “Gold Standard”: How subjects are evaluated varies. How subjective is pathological evaluation?
- Confounding effects blurring disease signal: comorbidities, treatment, age, gender, lifestyle factors…

Possible Causes of Misclassification

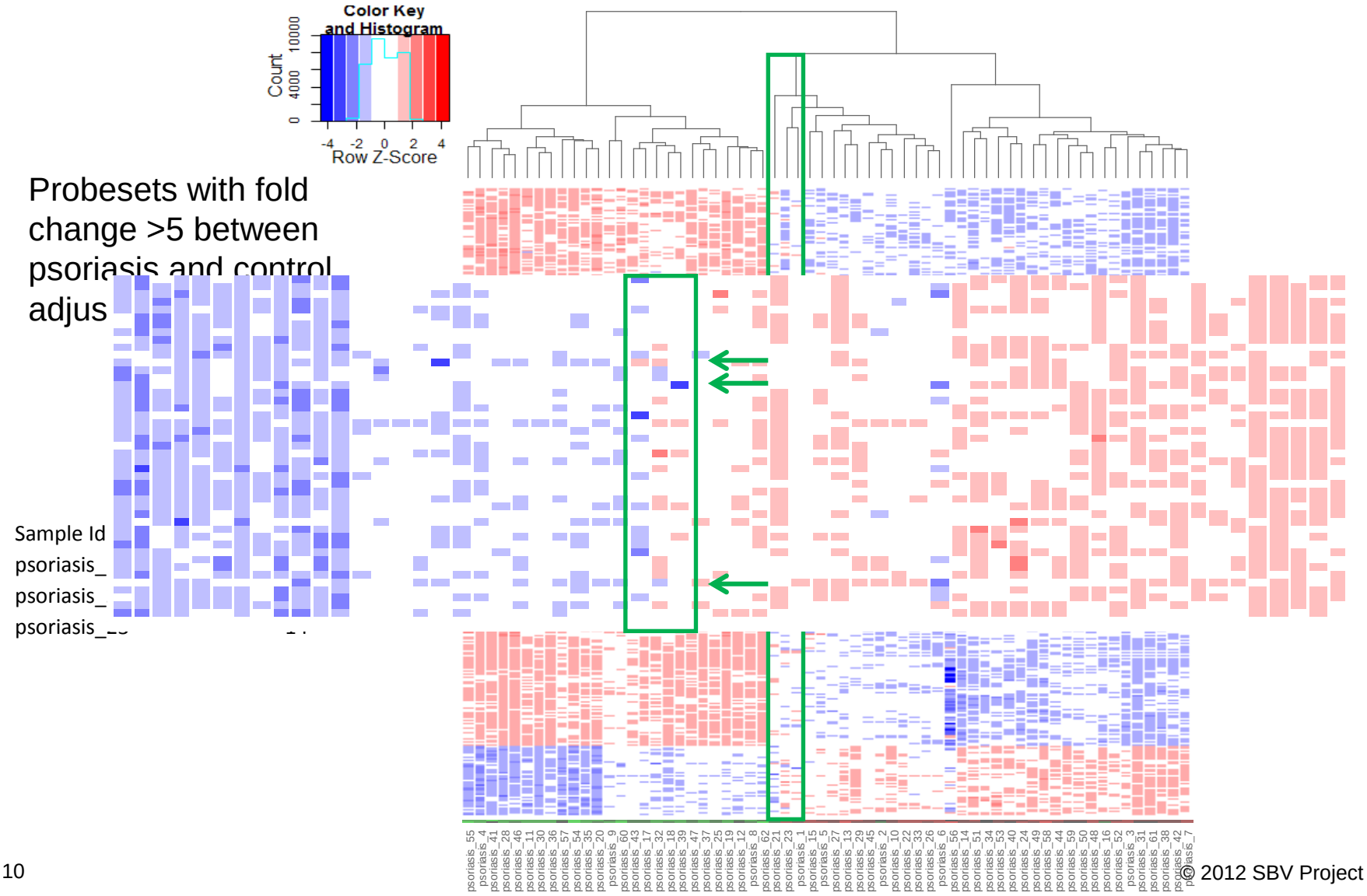
In addition to possible technical challenges due to the acquisition of training and test data in different populations, causes of misclassification may be related to:

- The definition of classes, especially for disease stages, is based on discrete scales, whereas disease progression is usually continuous. Molecular staging may be more granular than the pathological/anatomical staging
- A disease phenotype could in fact encompass an array of related pathologies with different molecular origin/mechanisms
- The “Gold Standard”: How subjects are evaluated varies. How subjective is pathological evaluation?
- Confounding effects blurring disease signal: comorbidities, treatment, age, gender, lifestyle factors…

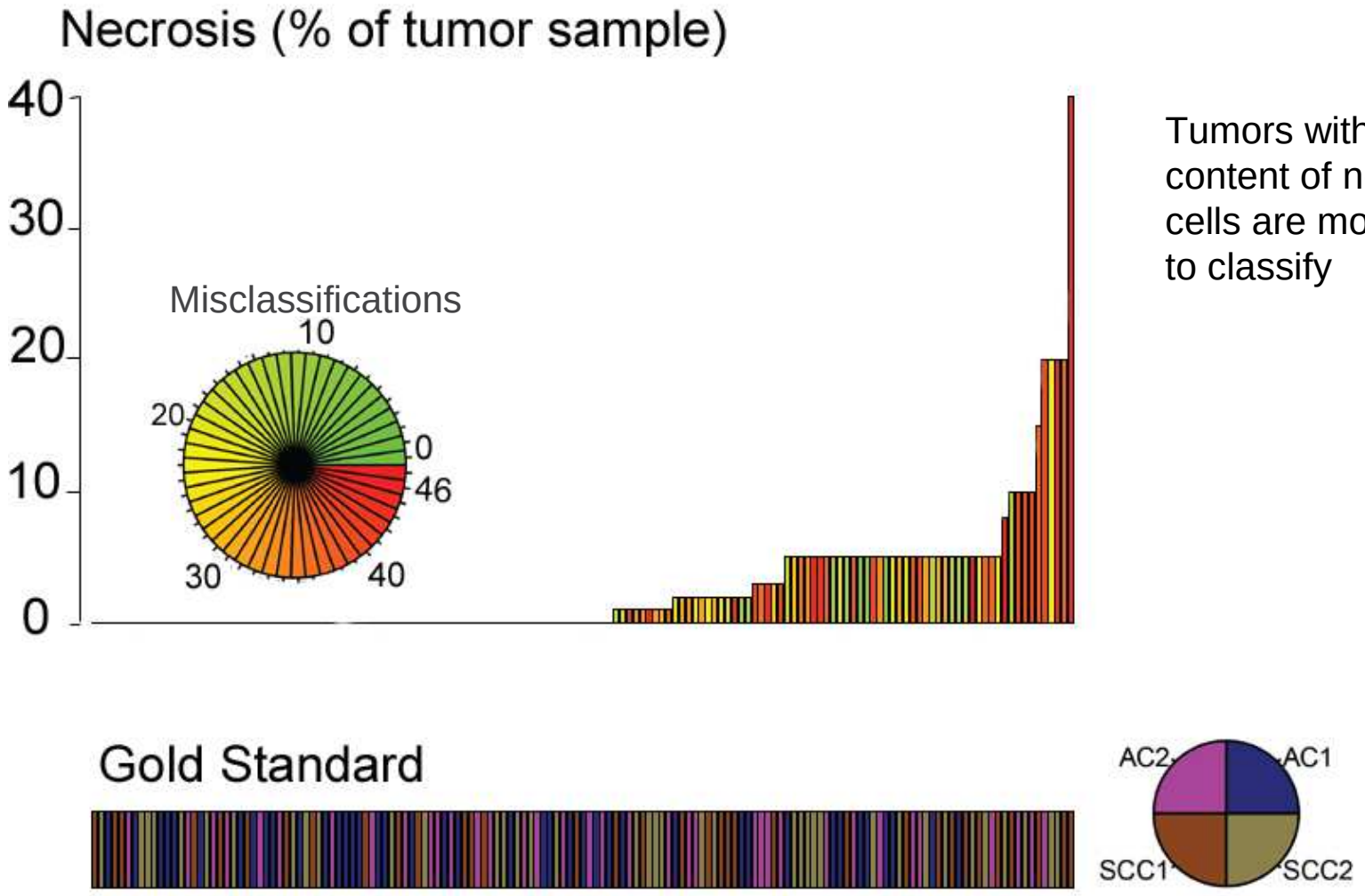
Misclassification Related to Disease Severity – Example of Psoriasis



Example Psoriasis – Gene Expression Genes Differentiating Controls-Psoriasis



Disease Specificities Could Cause Misclassification – Example of Necrosis Level in Lung Cancer Samples



Tumors with higher content of necrotic cells are more difficult to classify

Possible Causes of Misclassification

In addition to possible technical challenges due to the acquisition of training and test data in different populations, causes of misclassification may be related to:

- The definition of classes, especially for disease stages, is based on discrete scales, whereas disease progression is usually continuous. Molecular staging may be more granular than the pathological/anatomical staging
- A disease phenotype could in fact encompass an array of related pathologies with different molecular origin/mechanisms.

- The “Gold Standard”: How subjects are evaluated varies. How subjective is pathological evaluation?

- Confounding effects blurring disease signal: comorbidities, treatment, age, gender, lifestyle factors…

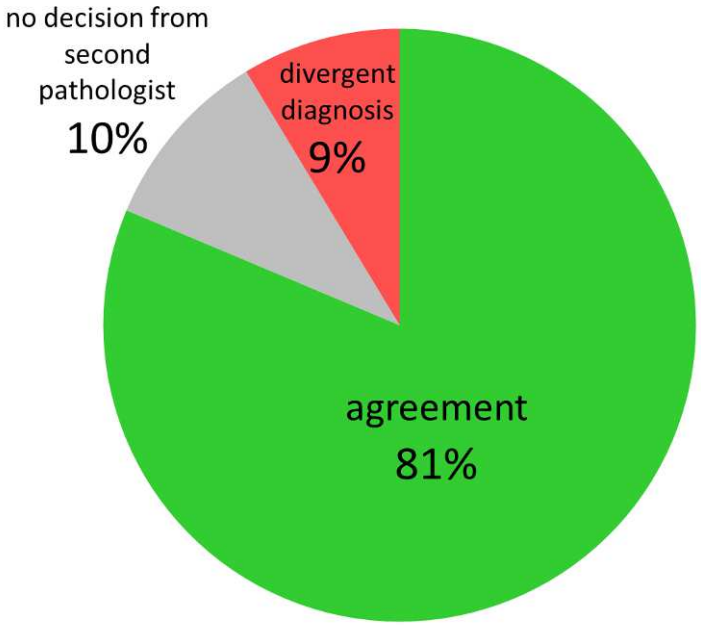
Gold Standard Variability – Example of Pathological Evaluation in Lung Cancer

Lung cancer tumor histopathology slides were reevaluated by an independent pathologist.

Based on the histology only, he could provide following details on the tumor:

- Type of tumor (except in very undifferentiated tumors where immunohistochemistry would be needed)
- Grade of differentiation of tumor
- Estimation of the degree of necrosis in the tumor
- Estimation of the number of lymphocytes in the tumor
- Estimation of the proliferation rate of the tumor (counting number of mitosis)
- Estimation of vascularization of the tumor

Variability in Pathological Evaluation

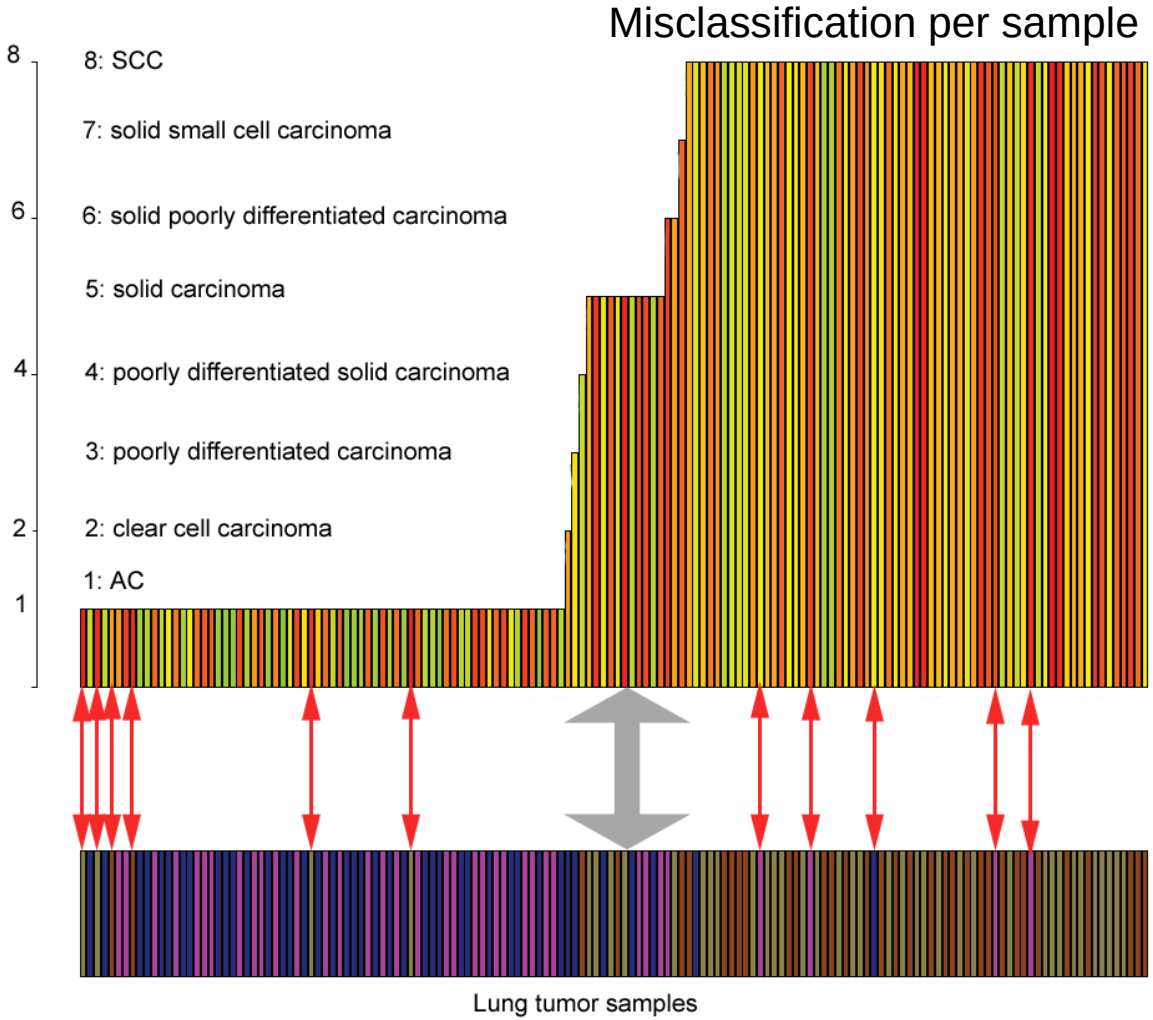


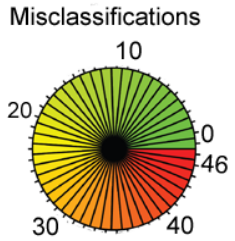
Gold Standard

Histological Reevaluation

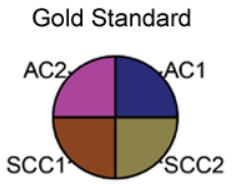
LC subtype	AC1	AC2	SCC1	SCC2
AC	34	28	2	4
Clear cell carcinoma	1			
Poorly differentiated carcinoma	1	1	1	1
Solid Carcinoma	3	3	1	4
Small cell carcinoma			1	
SCC	2	4	29	30

Uneasy Diagnostic May Be Reflected In Classification Performance





Samples on which pathologist evaluation do not agree (red arrows) or when the second evaluation cannot be decisive (grey arrow) are often misclassified



Possible Causes of Misclassification

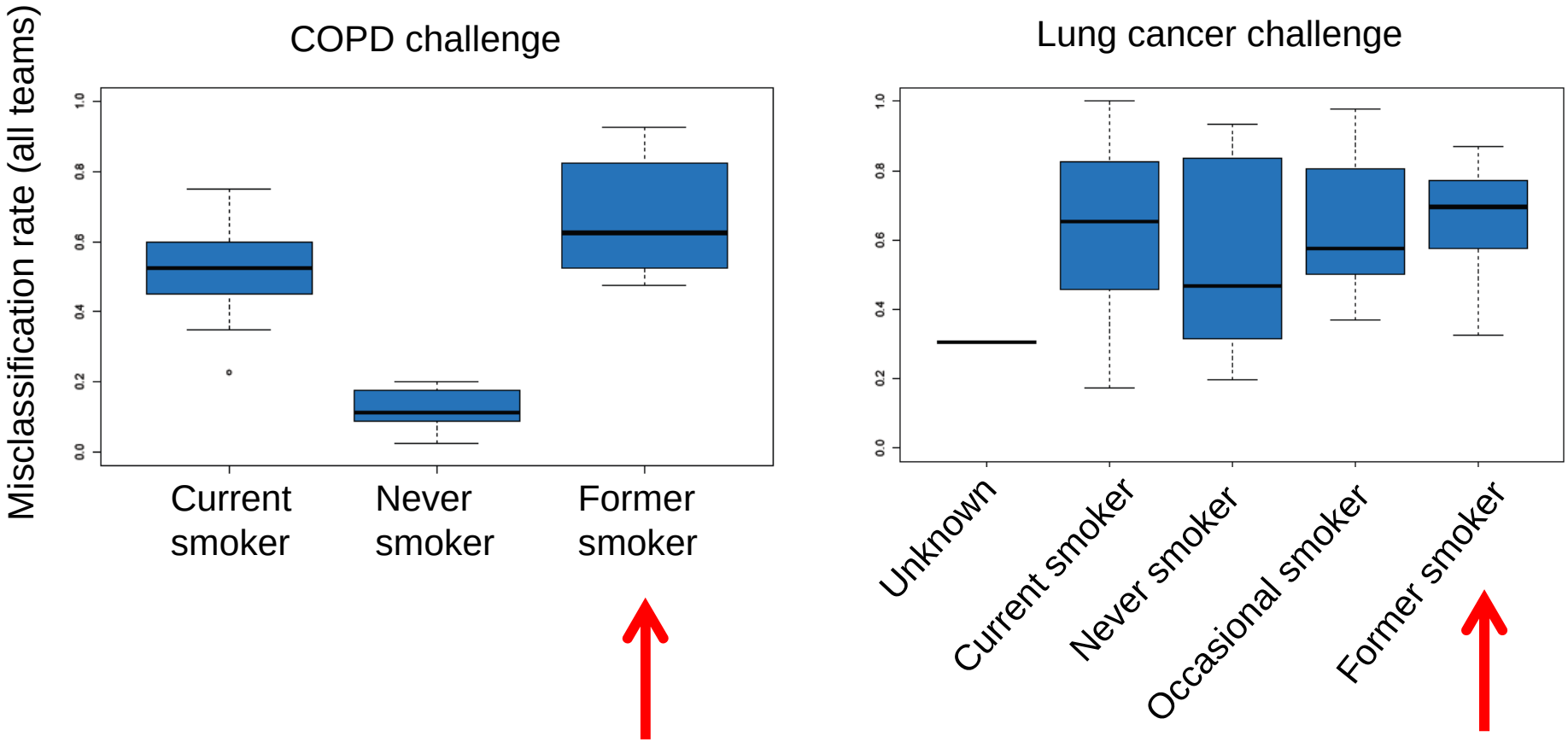
In addition to possible technical challenges due to the acquisition of training and test data in different populations, causes of misclassification may be related to:

- The definition of classes, especially for disease stages, is based on discrete scales, whereas disease progression is usually continuous. Molecular staging may be more granular than the pathological/anatomical staging
- A disease phenotype could in fact encompass an array of related pathologies with different molecular origin/mechanisms.

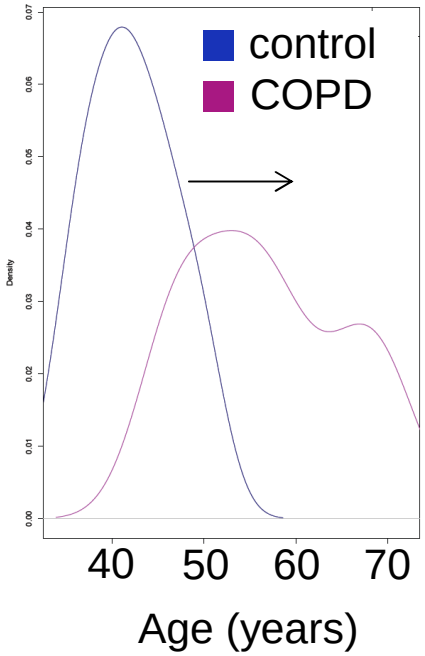
- The "Gold Standard": How subjects are evaluated varies. How subjective is pathological evaluation?

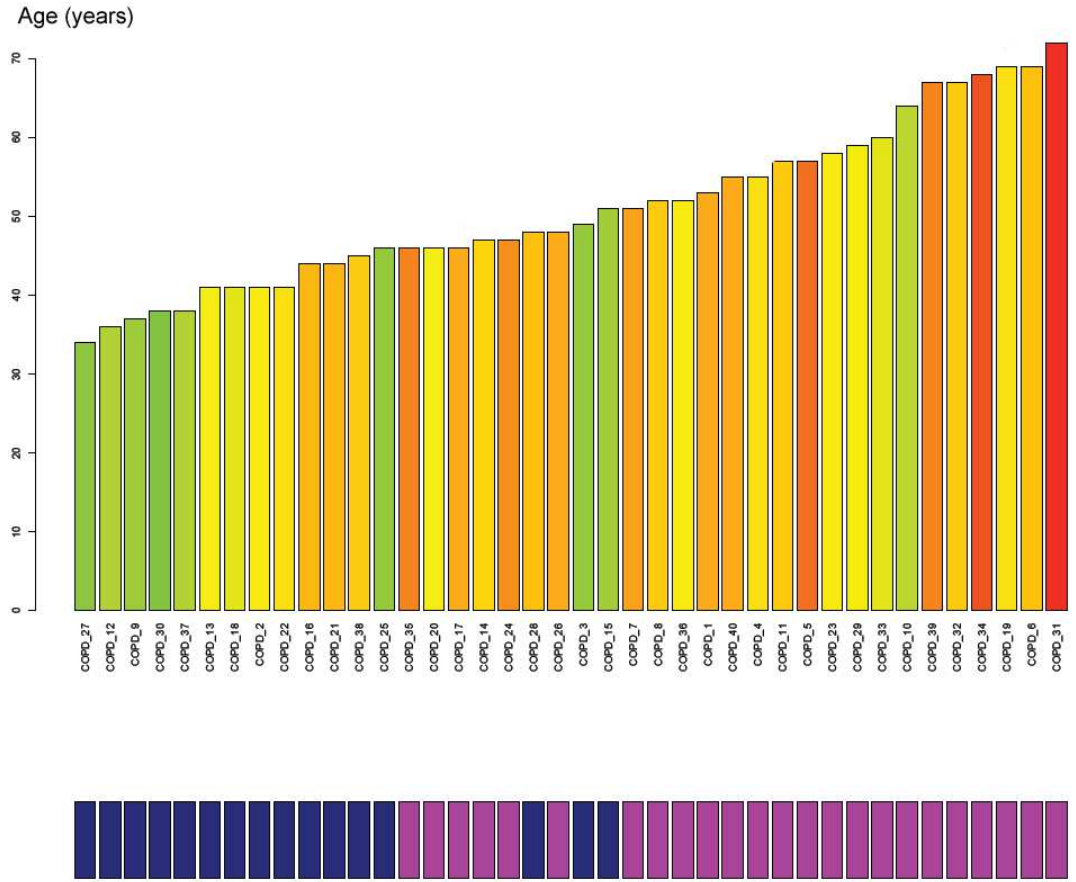
- Confounding effects blurring disease signal: comorbidities, treatment, age, gender, lifestyle factors…

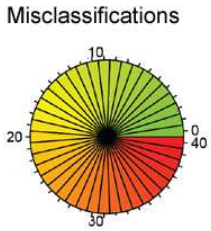
Smoking History And Classification Performance in Smoke-Related Diseases

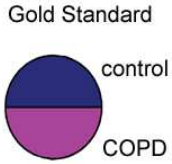


Confounding Factors – Age and COPD









Lessons Learned I – Messages from Biomedicine

- Even though there are many hurdle, it is possible, in some cases, to classify samples solely based on their gene expression profile.
- Classification is more straightforward when gene expression profiling is done on the primary tissue (e.g. skin in psoriasis) rather than on a surrogate tissue (e.g. blood in MS)
- While disease staging based on pathology is not easily retrievable from gene expression profiles, molecular staging may also be very helpful for diagnosis/prognosis and prediction of treatment benefit

- Classification of samples may be complicated by:
  - Poor availability of suitable training sets
  - The heterogeneity of the populations (training, test) and/or the disease
  - Confounding factors (gender, age, lifestyle factors…)

- Further investigations at the molecular, pathway and network levels is needed to understand if the biology of the samples actually caused misclassification, or if a different classification may emerge from gene expression profiles