Why we do not have an omics–based clinical risk score?

▸ Most ML methods apply standardisation on both training and validation data

▸ But in clinical prediction, you might only have single samples to predict on
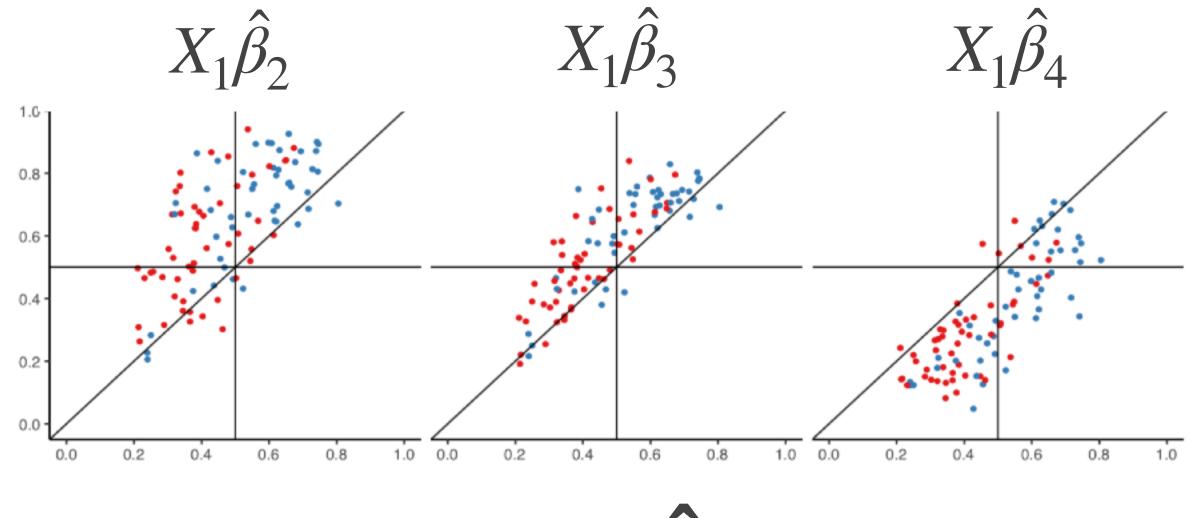
New X	Sample 4
Gene 1	4.2
Gene 2	3.8
Gene 3	8.4
Gene 4	3.1

Statistical challenges

1. Single-patient prediction

2. Different scaling on genes between datasets

3. Concordance in feature selection/coefficient estimates

Transferability
The prediction on one gene expression platform should be equivalent to another platform

$X_1\hat{\beta}_2$ $X_1\hat{\beta}_3$ $X_1\hat{\beta}_4$

$$X_1\hat{\beta}_1$$