

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

- ▶ Gene expression platforms are measured on a **relative scale**
 - ▶ Sequencing depth in RNA-Seq:

$$\hat{y} = X\hat{\beta} \qquad \hat{y} = (X - \mathbf{1})\hat{\beta}$$

- ▶ Use of different reagents between experiments:

$$\hat{\beta} = (X^{\top}X)^{-1}Xy$$

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

X	Sample 1	Sample 2	Sample 3
Gene 1	1.2	2.1	1.5
Gene 2	5.6	4.6	7.1
Gene 3	9.2	10.1	6.9
Gene 4	4.1	3.6	2.7