

Additional file 1

Table 1 - Influence of batch effect on LLM prediction performance

Dataset modification^f	Accuracy^a	Recall^b	Precision^c	Specificity^d	NPV^e
None	80%	90%	82%	57%	72%
FSVA Batch adjusted^f	74%	78%	83%	65%	58%

^a Accuracy is the fraction of correctly classified patients and overall classified patients.

^b Recall is the fraction of correctly classified good outcome patients and the overall predicted good outcome patients

^c Precision is the fraction of correctly classified good outcome patients and the predicted good outcome patients.

^d Specificity is the fraction of correctly classified poor outcome patients and the overall poor outcome patients

^e NPV (Negative predictive value) is the fraction of correctly classified poor outcome patients and the overall predicted poor outcome patients

^f The dataset was modified according to the FSVA batch effect removal method described in the Materials and Methods.