

2279 biopsies – 615 slides  
graded by  $\geq 4$  pathologists

Exclude 65 biopsies  
( $< 75\%$  concordance)

2214 biopsies  
with valid labels

Train/test split

Remaining to  
train set

Train set  
( $AUMC_{\text{train}}$ ,  $n=1940$ )

Add all DC  
+ matching OE/NM  
+  $\sim 2\times$  WT

Test set  
( $AUMC_{\text{test}}$ ,  $n=274$ )

