Significance analysis of the the performance of the 0% and 30% pre-training dataset size setups / split = random Permeability Solubility HLM - NS - NS - NS MLM_0% | MLM_0% MLM_0% 0.008 0.032 0.016 - p < 0.05 -p < 0.05- p < 0.05p < 0.01 p < 0.01 MLM_30% MLM_30% A 0.008 0.032 0.016 p < 0.001 MLM_30% MLM_30% WFW _00/0 MLM_30% MFW _00/0 MLM_0% hPPB rPPB **RLM** - NS - NS - NS MLM_0% 1 MLM_0% MLM_0% 1 0.841 0.841 0.008 - p < 0.05 - p < 0.05 -p < 0.05p < 0.01 p < 0.01 - p < 0.01 MLM_30% MLM_30% 0.841 0.841 0.008 - p < 0.001 MLM_30% MLM_30% MLM_30% MLM_0% MFW _00/0 Mrw _00/0