Y1: performance of single uncalibrated vs calibrated model on subgroup-by-sex: female uncalibrated-one-model-for-all calibrated-one-model-for-all 100 -89.7 90 86.4 80 72.7 60.8 Score(%) 60 55.9 52.2 41.3 40 20 precision recall accuracy