Y1: performance of single uncalibrated vs calibrated model on subgroup-by-age-group: seventies uncalibrated-one-model-for-all calibrated-one-model-for-all 100 -91.7 91.7 89.5 80 63.2 Score(%) 60 53.7 50 41.4 37.9 40 20 f1 precision recall accuracy