Y1: performance of single uncalibrated vs calibrated model on subgroup-by-age-group: sixties 120 uncalibrated-one-model-for-all calibrated-one-model-for-all 100 100 -94.7_93.1 90.9 80 66.7 Score(%) 60 -52.6 48 40 31.6 20 f1 precision recall accuracy