Y1: performance of single uncalibrated vs calibrated model on subgroup-by-race: non-white