Y2: performance of single vs ensemble model on subgroup-by-sex: female 100 one-model-for-all one-model-per-subgroup ensemble-on-all ensemble-on-subgroup 80.5 79.9 79.1 76.4 77.4 77.7 78.1 80 77.7 75.3 76.6 76.9 74.8 73.9 73 73.5 68.3 60 Score(%) 40 20 0 precision f1 recall accuracy