Test recall
Classical approach on subset data
showing mean tall. dev over 5 runs

