Q&A Accuracy vs AUROC: no_abst_all (r = 0.85)