&A Accuracy vs AUROC: yes_abst_norm_logits_first_prompt, truthfulqa (r = 0