MedGemma-4B: Prompt Strategy Comparison on MedMCQA (n=4,183)

- - - Random baseline (25%)

47.6%

-5.7%
(CoT hurts)

43.0%

41.9%

-11.9%
(Few-shot hurts)

40.8%

35.7%

Accuracy (%)

Zero-shot Direct | Answer Only | Zero-shot CoT | Few-shot CoT | Few-shot Direct