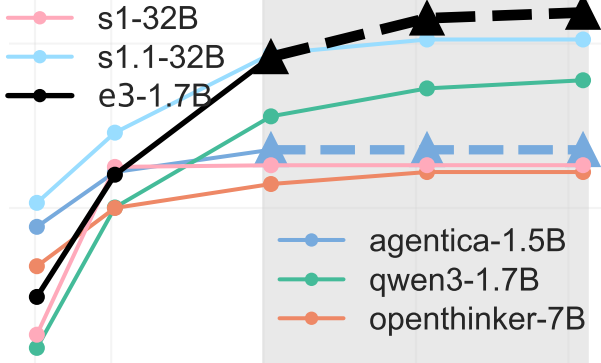


e3 vs. open-source models

Acc. on AIME '25

40

20



4k

8k

16k

24k

32k

Test-time compute (tokens)

agentica-1.5B

qwen3-1.7B

openthinker-7B