MMLU Virology Brier Score Correlation: spearman Correlation: pearson Correlation: kendall — OLMo 7B Param. (Data Scaling) 1.0 Density 8.0 0.40.750.751.00 0.751.00 1.00 -1.00-1.00Correlation Between FLOPs and Scores (Per Sample) Correlation Between FLOPs and Scores (Per Sample) Correlation Between FLOPs and Scores (Per Sample)

Model Family — Cerebras (Param. and Data Scaling)

— INCITE 7B Param. (Data Scaling)

LLM360 Amber 7B Tokens (Param Scaling)

— Pythia 12B Param. (Data Scaling)

— Pythia 300B Tokens (Param. Scaling)