

Expert Alignment vs Accuracy

Maps	0.058	0.09	0.072	0.078
ernova	0.0026	0.00053	0.027	0.03
teness	-0.0059	0.038	-0.048	-0.0038
motion	0.13	0.04	0.046	0.095
ctomy	0.32	0.22	0.27	0.37
ardiac	-0.12	-0.11	-0.11	-0.076
Sepsis	-0.00074	0.0015	-0.017	-0.021

Chain-of-
Thought

Socratic
Prompting

SubQ
Decomp

Vanilla



Pearson Correlation