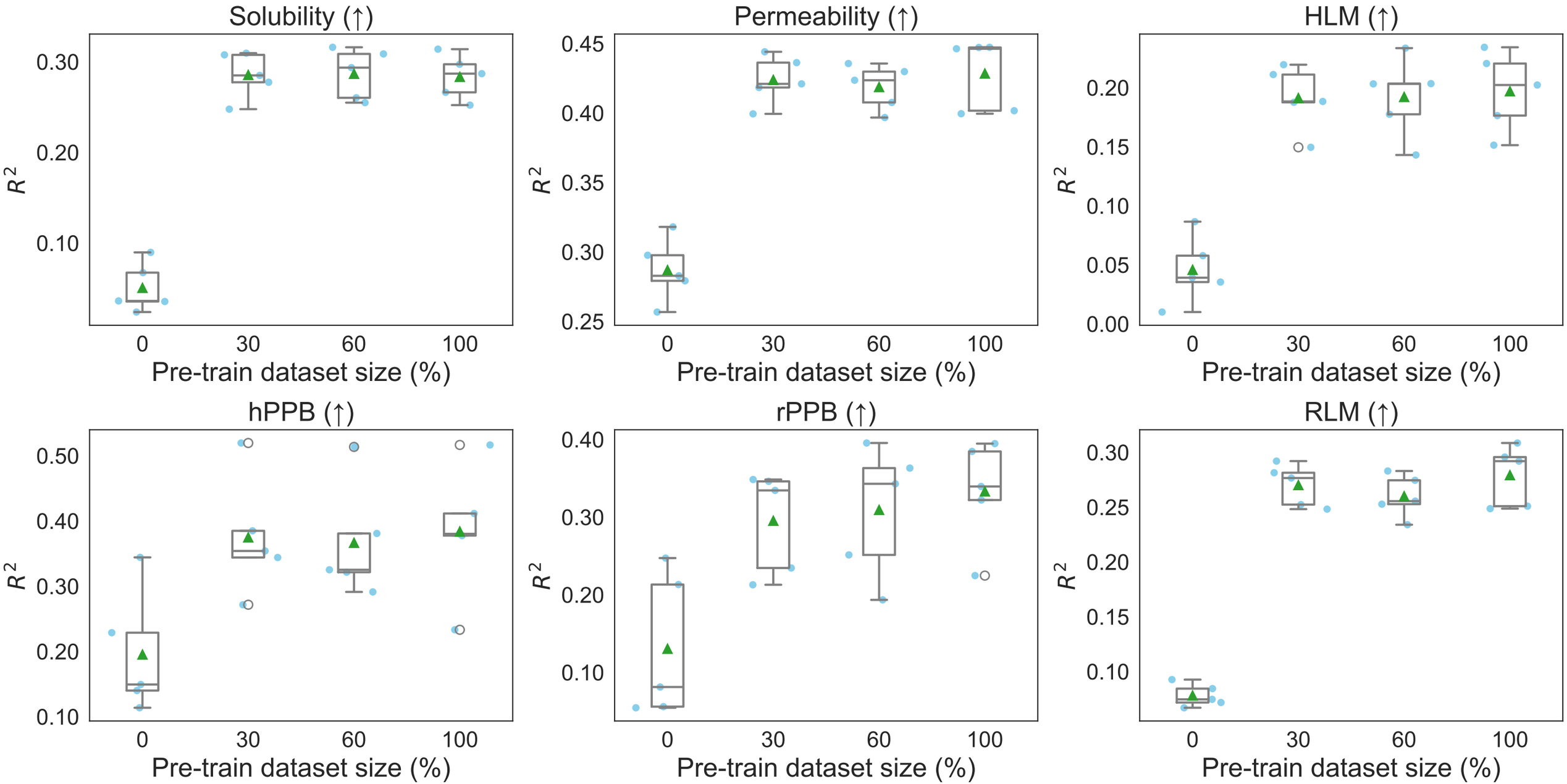
Comparing pre-training dataset sizes / pre-training objective = MTR / split = random
Solubility (↑)
Permeability (↑)
HLM (↑)
hPPB (↑)
rPPB (↑)
RLM (↑)
R^2
Pre-train dataset size (%)
0
30
60
100
0.10
0.20
0.30
0.25
0.35
0.40
0.45
0.00
0.05
0.15
0.50