

D-Calibration p-values by task and learner

Models tuned on Integrated Survival Brier Score (ISBS)

Learners ordered by average p-value. X denotes $p < 0.05$

