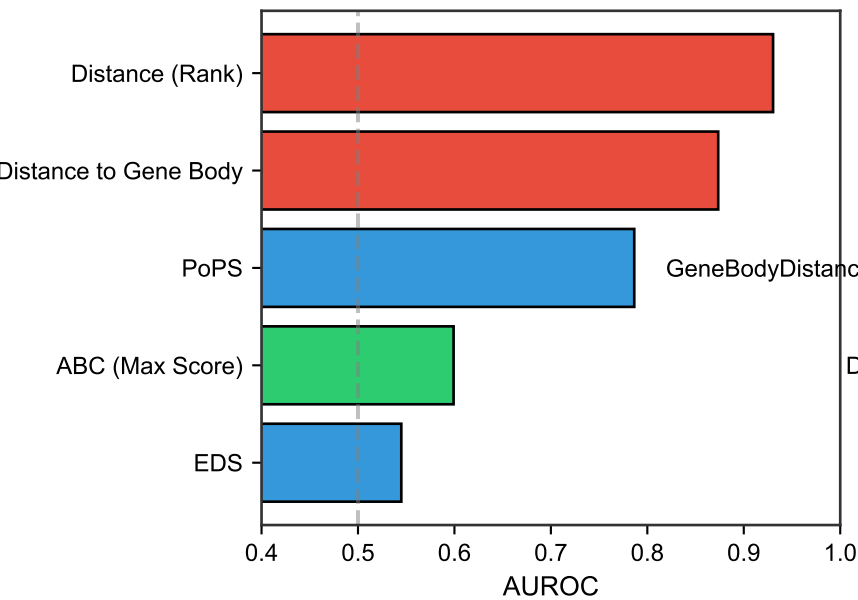
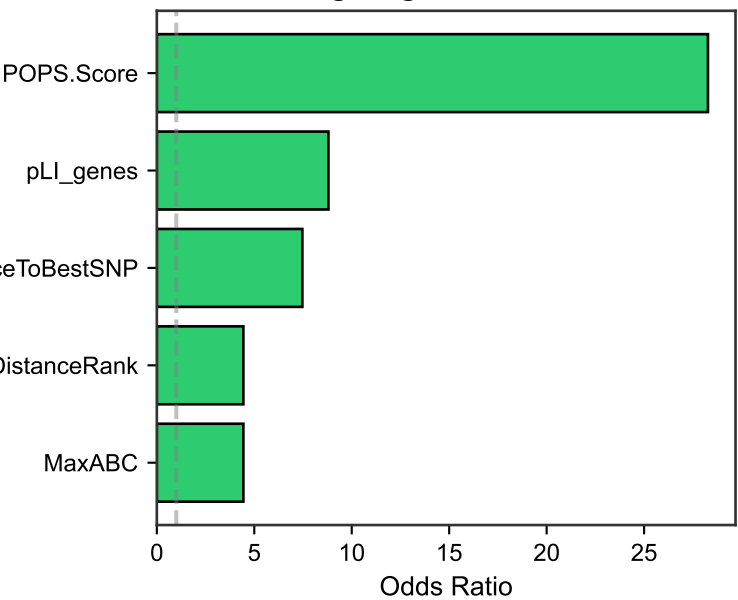


RegulatoryBench: Task-Stratified Evaluation Reveals Hidden Performance Patterns

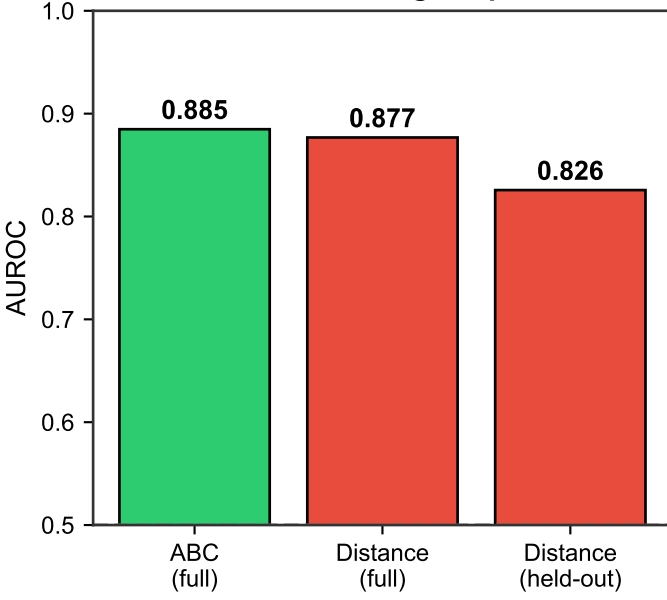
Task A: Distance Dominates



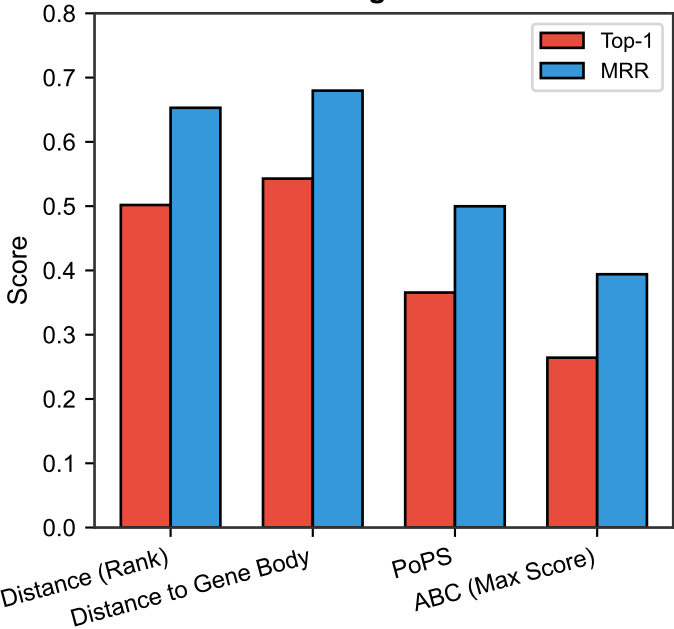
Drug Target Enrichment



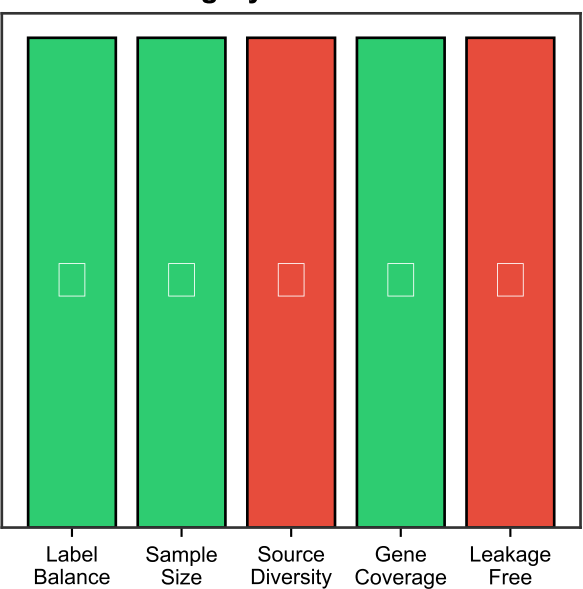
Task B: Leakage Impact



Ranking Metrics



Integrity: 3/5 Passed



Summary

- KEY FINDINGS
- Distance (AUROC=0.930) dominates GWAS→Gene mapping
  - PoPS shows highest drug-target enrichment (OR=28.3, p<0.001)
  - ABC Prediction shows NO drug-target enrichment (OR=0.64, p=0.768)
  - 78% leakage in Task B inflates ABC performance