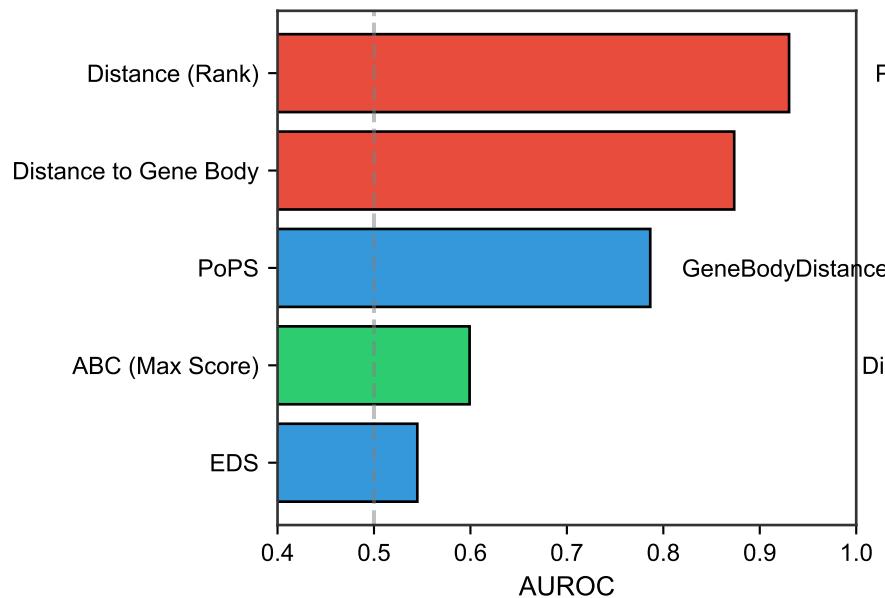
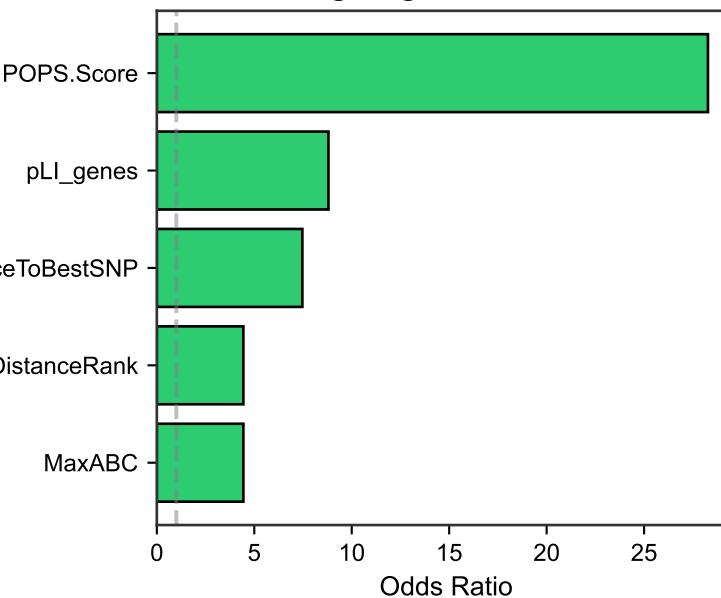


# RegulatoryBench: Task-Stratified Evaluation Reveals Hidden Performance Patterns

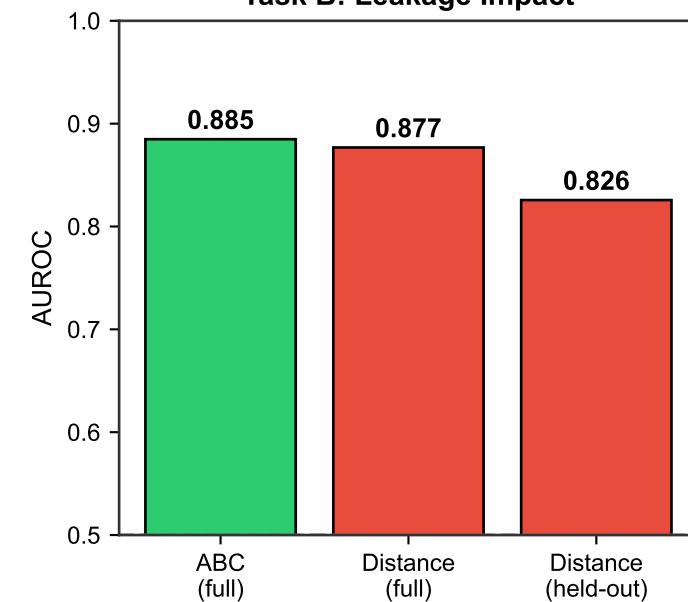
**Task A: Distance Dominates**



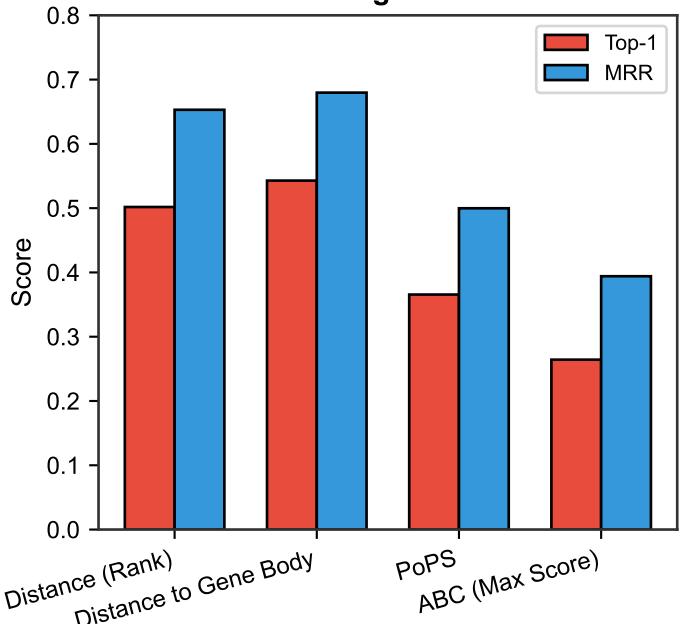
**Drug Target Enrichment**



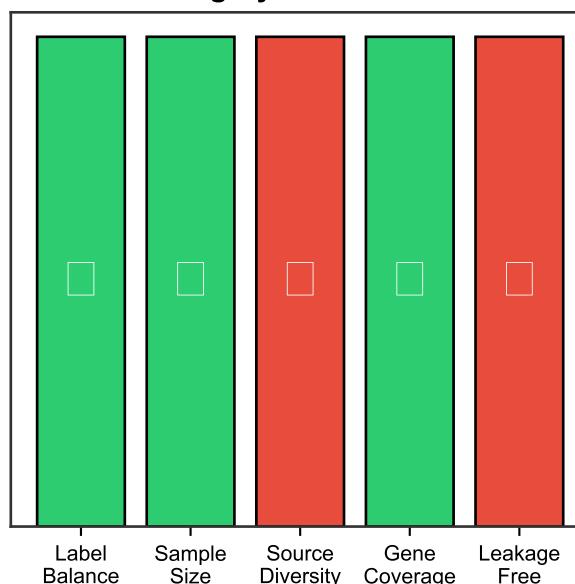
**Task B: Leakage Impact**



**Ranking Metrics**



**Integrity: 3/5 Passed**



**Summary**

## KEY FINDINGS

1. Distance (AUROC=0.930) dominates GWAS→Gene mapping
2. PoPS shows highest drug-target enrichment ( $OR=28.3$ ,  $p<0.001$ )
3. ABC Prediction shows NO drug-target enrichment ( $OR=0.64$ ,  $p=0.768$ )
4. 78% leakage in Task B inflates ABC performance