

Bias Category

Estimation

Dataset  
MedQA  
NEJM

Hypothesis Assessment

Decision

Implicit

Accuracy Impact (%)

-17.5 -15.0 -12.5 -10.0 -7.5 -5.0 -2.5 0.0

