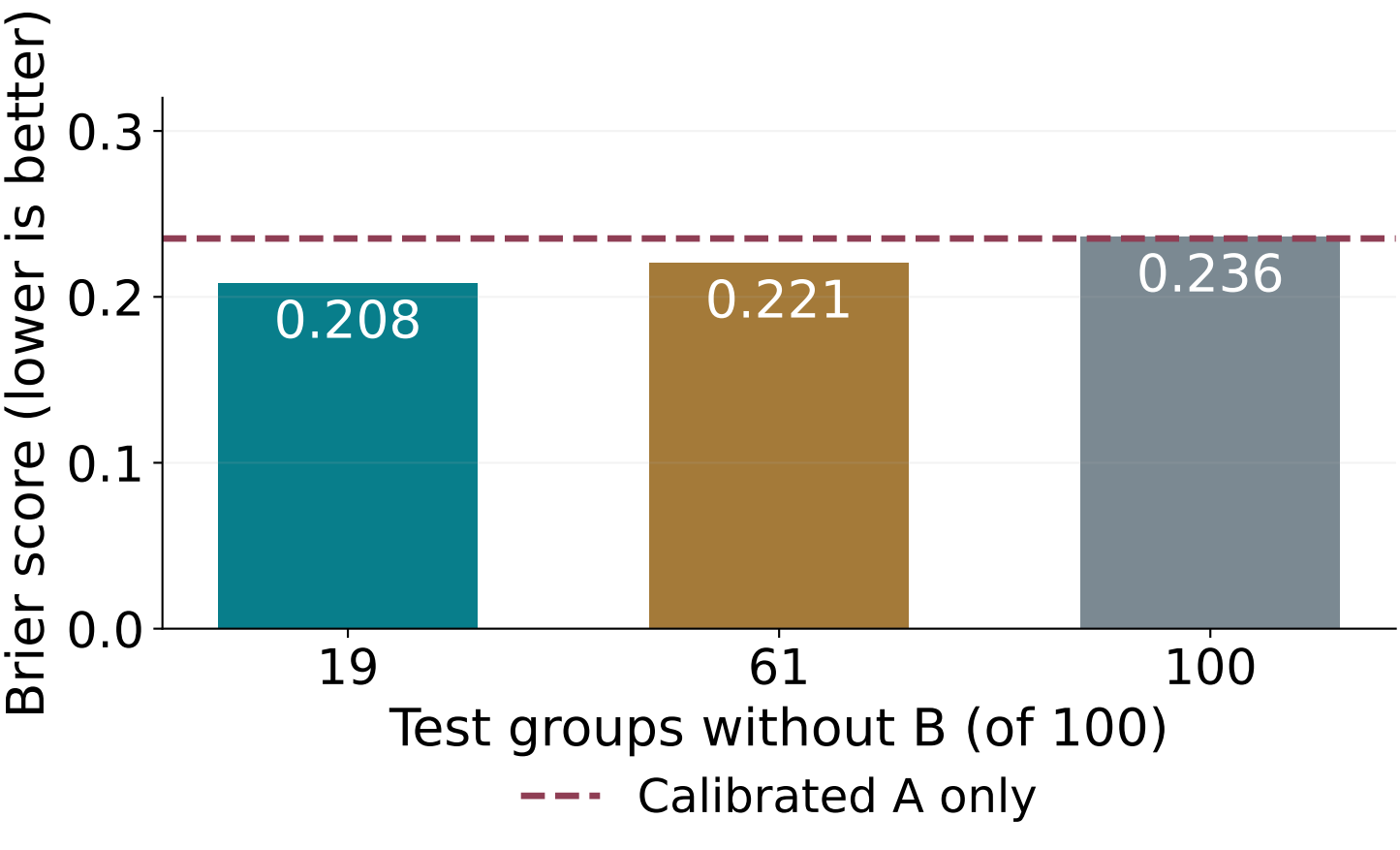

Brier score (lower is better)
0.3
0.2
0.1
0.0
0.208
0.221
0.236
19
61
100
Test groups without B (of 100)
Calibrated A only