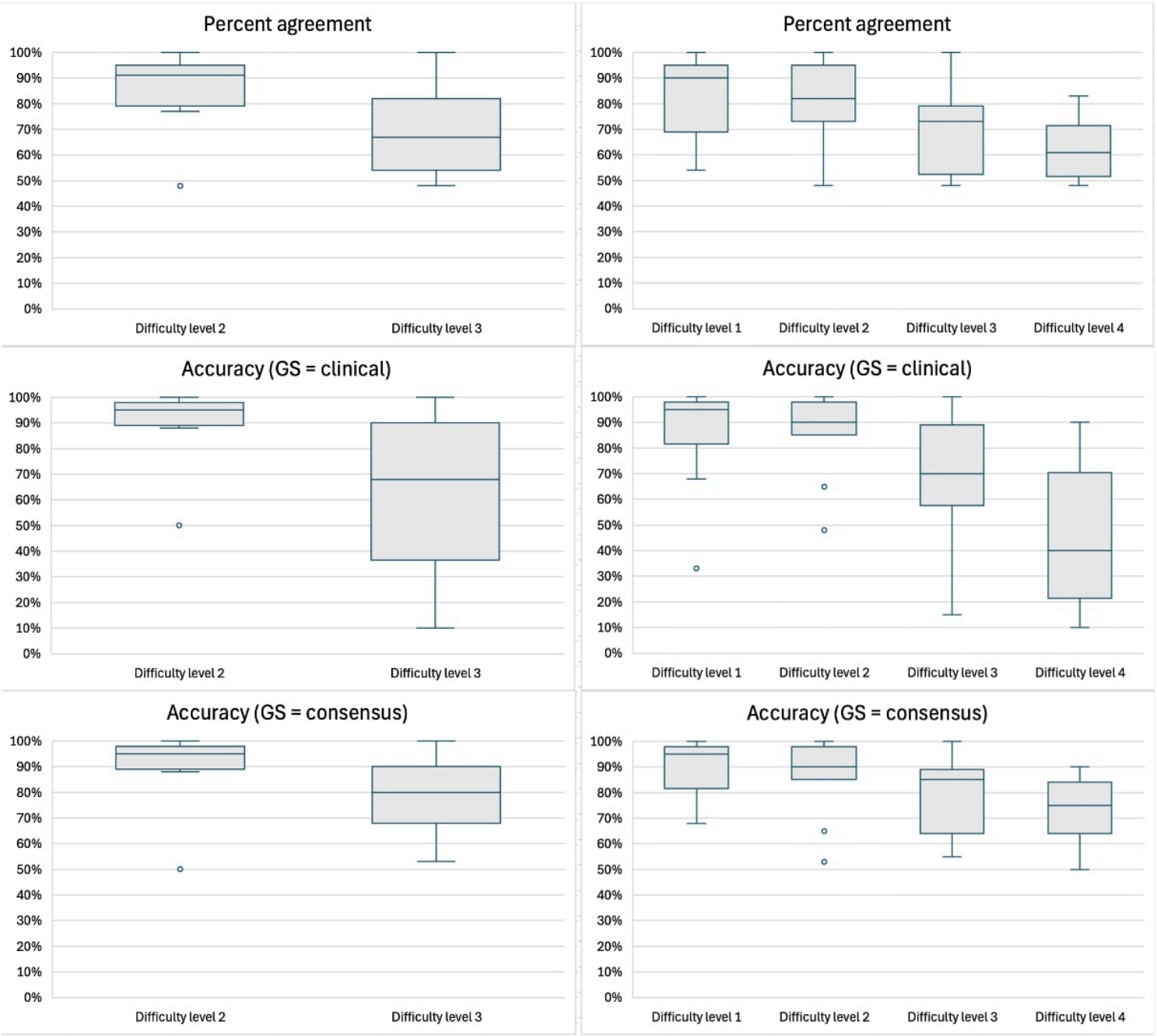

Percent agreement
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Difficulty level 2
Difficulty level 3
Percent agreement
Difficulty level 1
Difficulty level 2
Difficulty level 3
Difficulty level 4
Accuracy (GS = clinical)
Accuracy (GS = clinical)
Accuracy (GS = consensus)
Accuracy (GS = consensus)