

# Ensembles

## Main ensembles

Main ensembles considered are: unweighted mean/median; mean/median weighted by inverse scaled relative WIS.

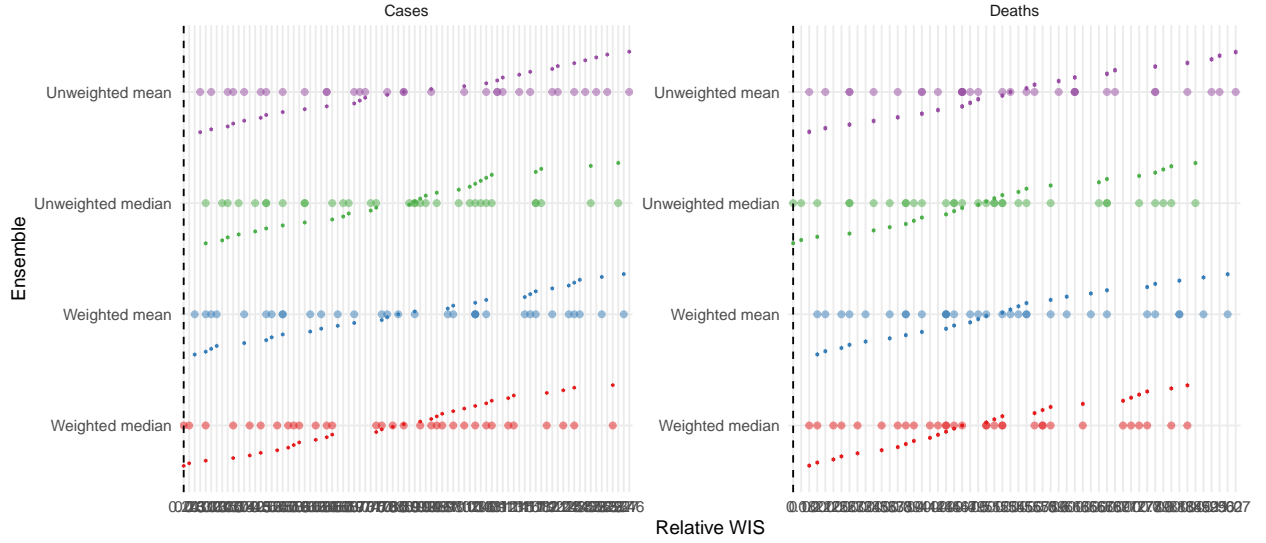


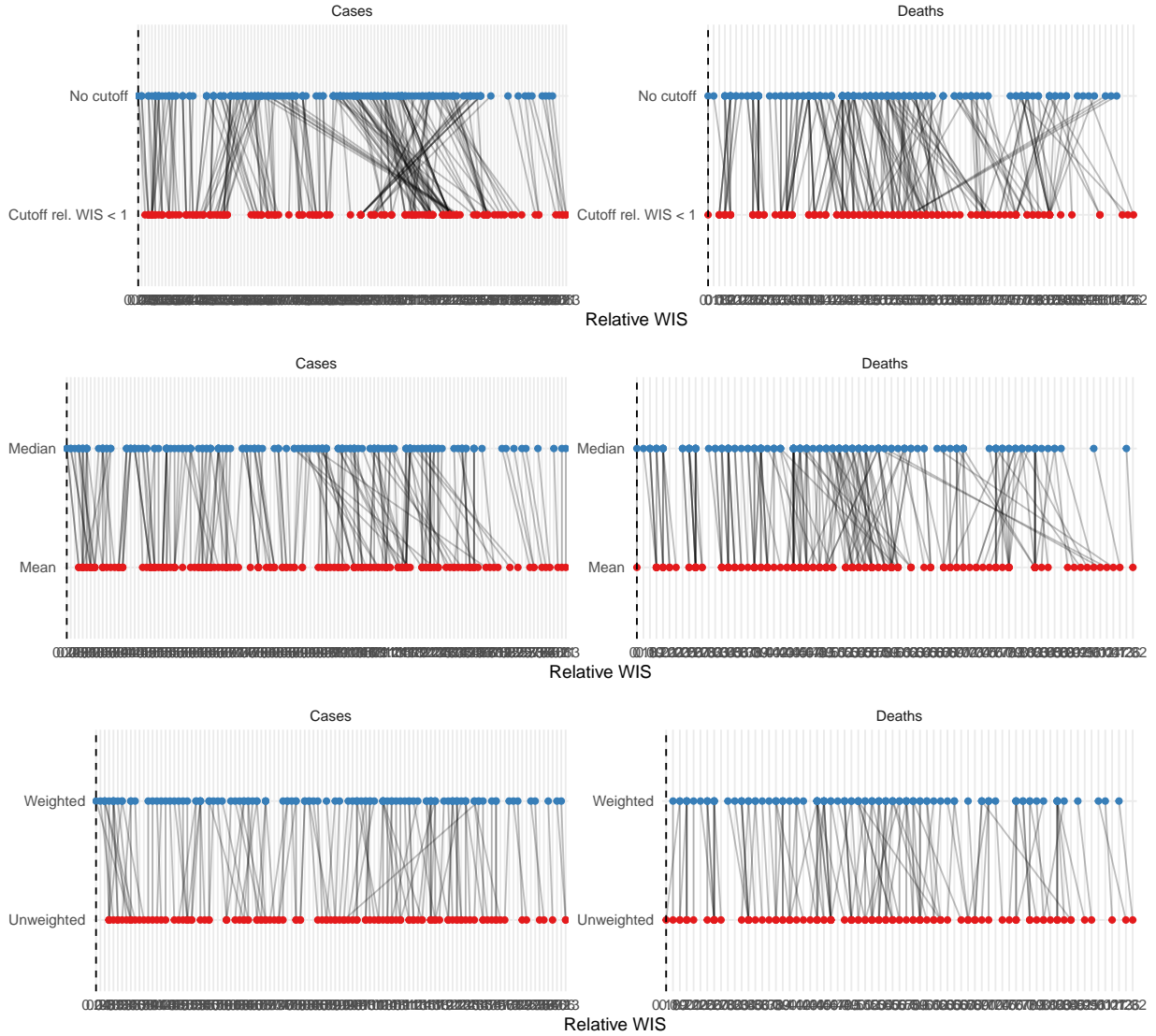
Figure 1: Main ensembles, using all component models and considering all of the history. Each point is one country.

Table 1: Predictive performance of main ensembles, as measured by the scaled relative WIS.

Horizon	Weighted mean	Weighted median	Unweighted mean	Unweighted median
<b>Cases</b>				
1 week	0.67	0.66	0.69	0.66
2 weeks	0.87	0.8	0.92	0.81
3 weeks	1.1	0.95	1.18	0.96
4 weeks	1.43	1.03	1.49	1.04
<b>Deaths</b>				
1 week	0.51	0.5	0.52	0.5
2 weeks	0.4	0.4	0.41	0.4
3 weeks	0.39	0.37	0.4	0.37
4 weeks	0.45	0.38	0.47	0.38

## Variations on main ensembles

We further considered: a *cutoff* (i.e., creating only ensembles from models that had relative scaled WIS  $< 1$  across all forecast horizons for a given target/country); and *limited history* (i.e., only considering the last 5/10 weeks of scores for the weighted ensemble, as opposed to the whole history). Each point in the countries below represent one country/method/target triplet at the 2-week horizon, with only one aspect of the ensemble varying between the top and bottom rows, connected by lines.



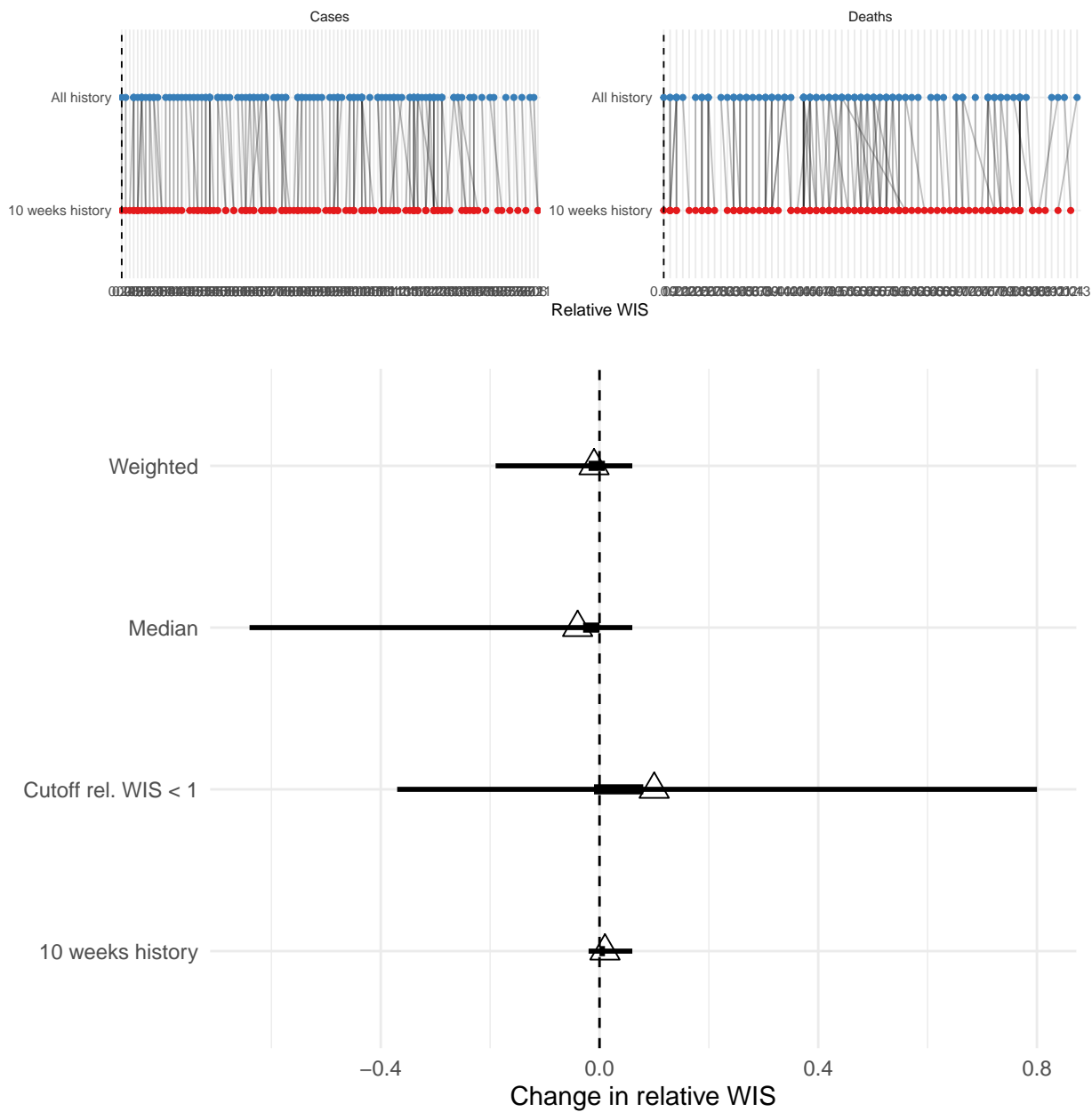


Figure 2: Aggregate differences between ensemble variants; shown are the central 48% and 96% intervals (thick/thin lines) and means of the distribution of difference in rel. WIS across all methods/countries.