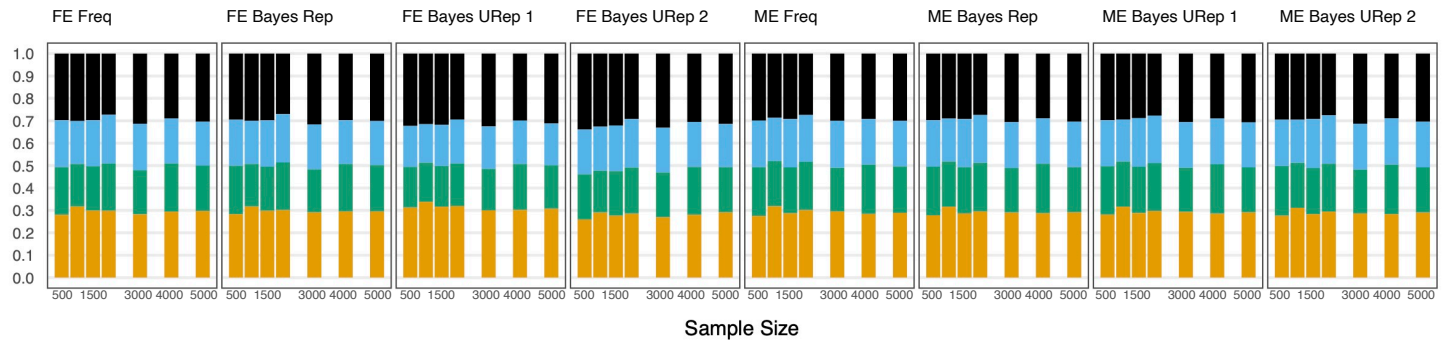
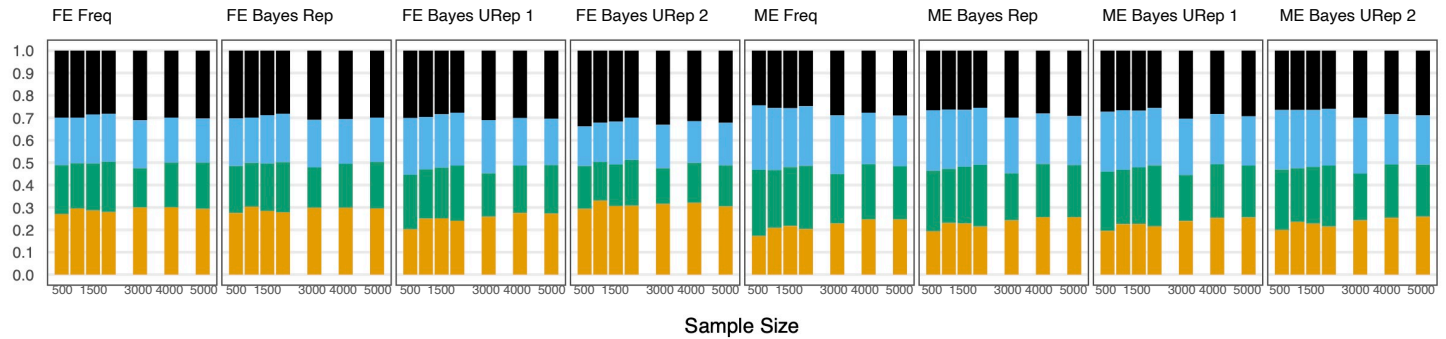


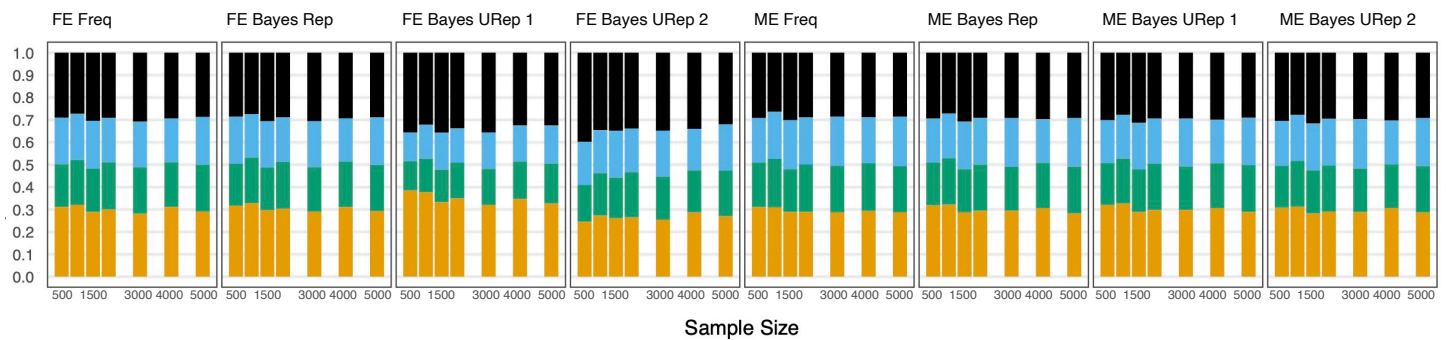
a) Sc 1.1



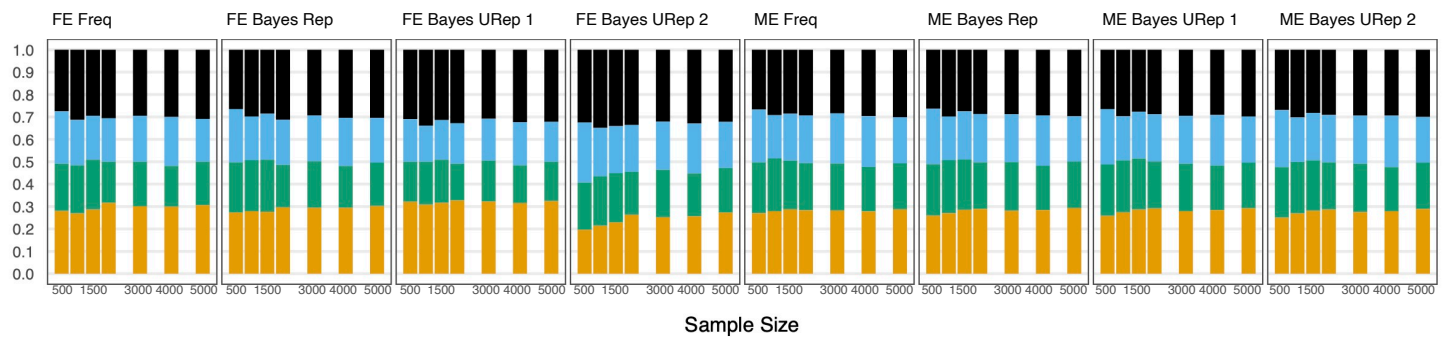
b) Sc 2.1



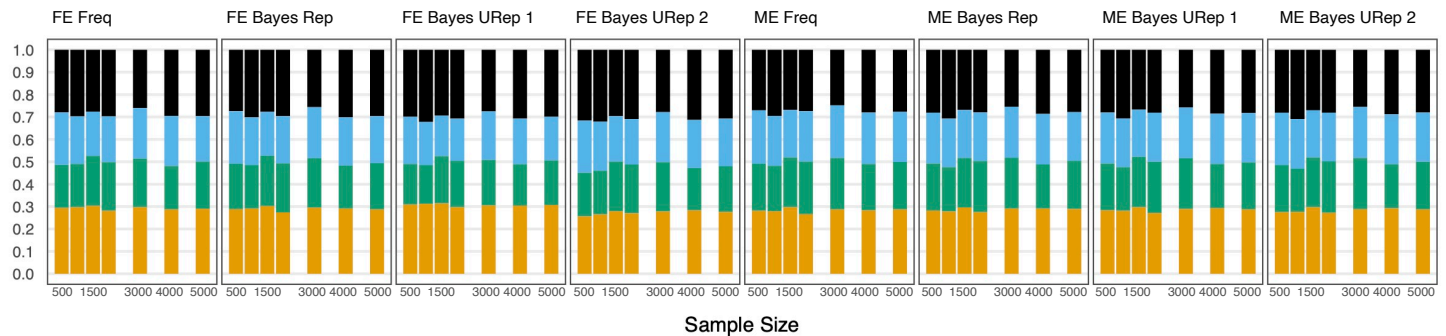
c) Sc 3.1



d) Sc 3.3



e) Sc 4.1



■ treatment A ■ treatment B ■ treatment C ■ treatment D

Fig. SM 2 Proportion of scenarios where each treatment (A–D) were incorrectly predicted as best for five null scenarios, when using a fixed and mixed effects using frequentist and Bayesian approaches for sample sizes ranging between $N = 500$ – $5,000$.