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.