Appendix A

Appendix A

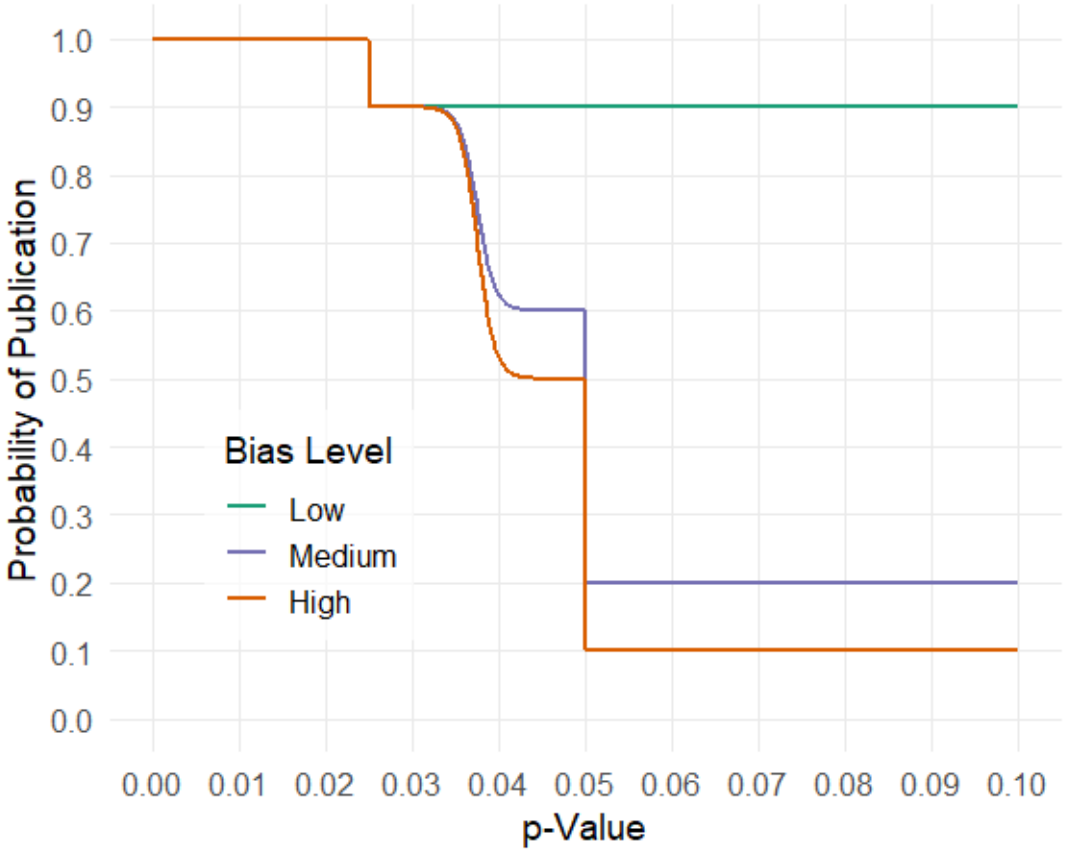

Figure S1 Publication bias mechanism used in the simulation. The curves show how the probability of publication varies with *p*-values under three bias conditions. Green, purple, and orange lines represent low, medium, and high levels of publication bias, respectively.

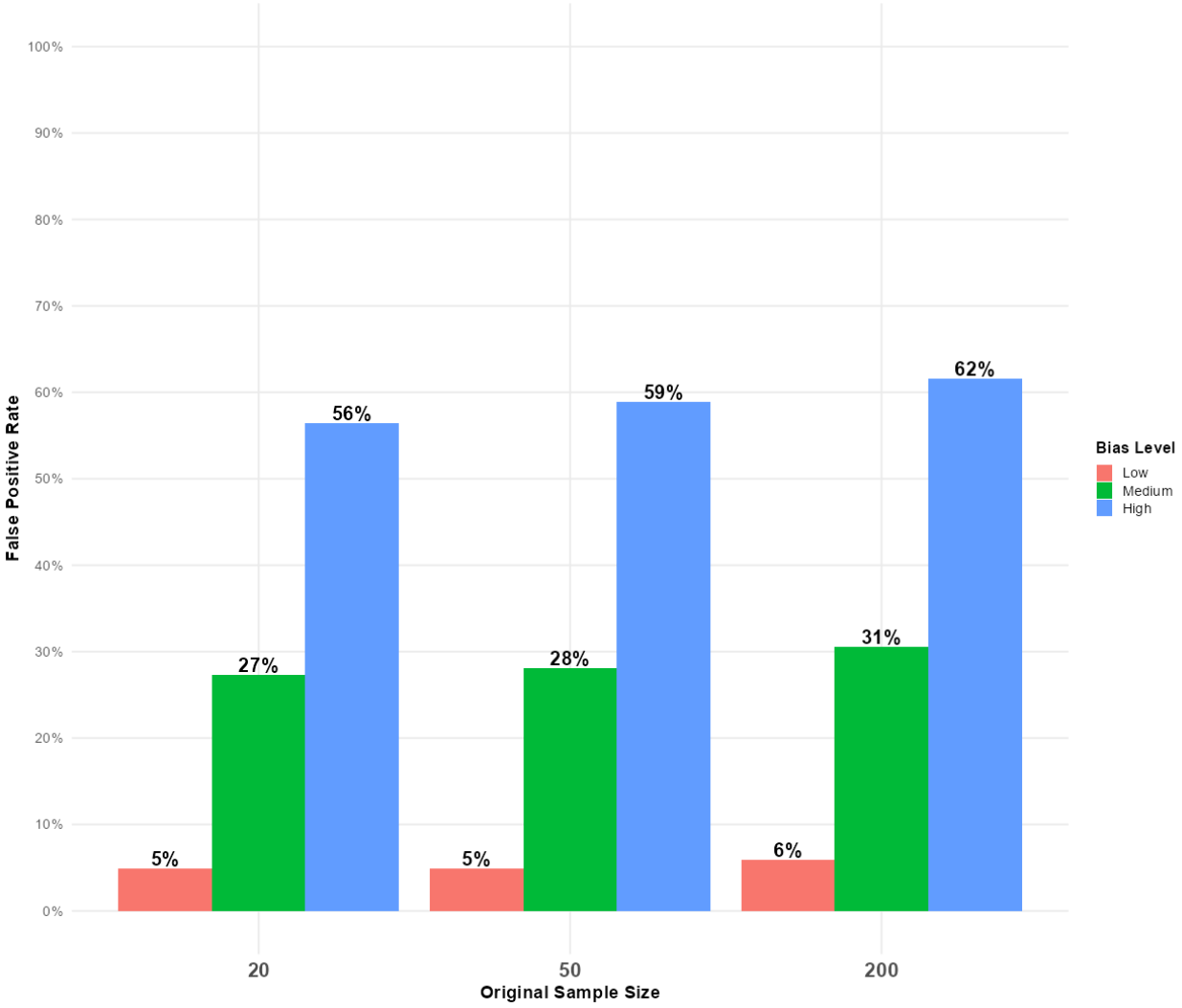

Figure S2 False positive rates of original study outcomes across different original sample sizes and bias mechanism levels, given a nominal significance level of $\alpha = .05$.

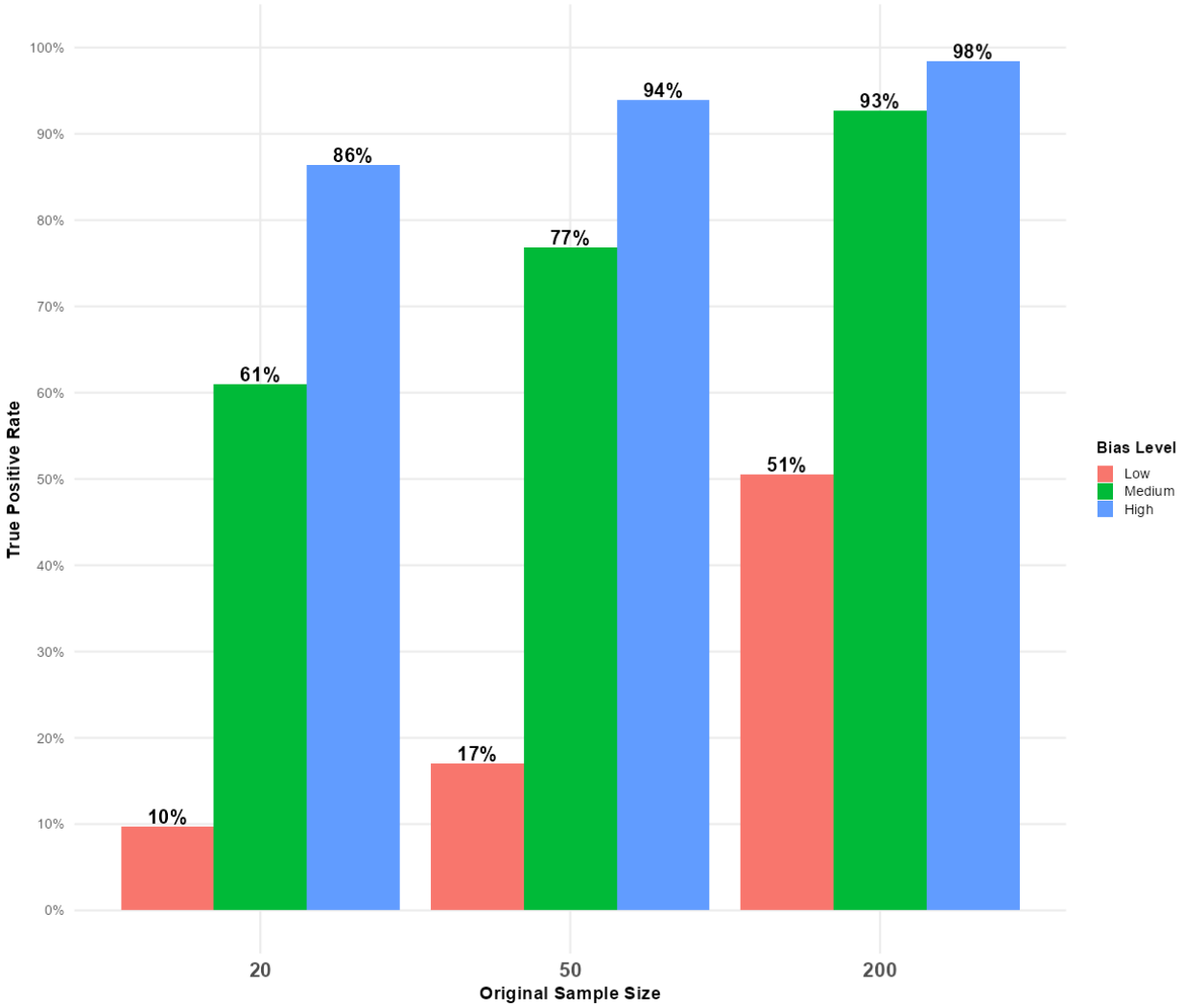

Figure S3 True positive rates of original study outcomes across different original sample sizes and bias mechanism levels when the underlying effect is $\theta = 0.2$, given a nominal significance level of $\alpha = .05$.

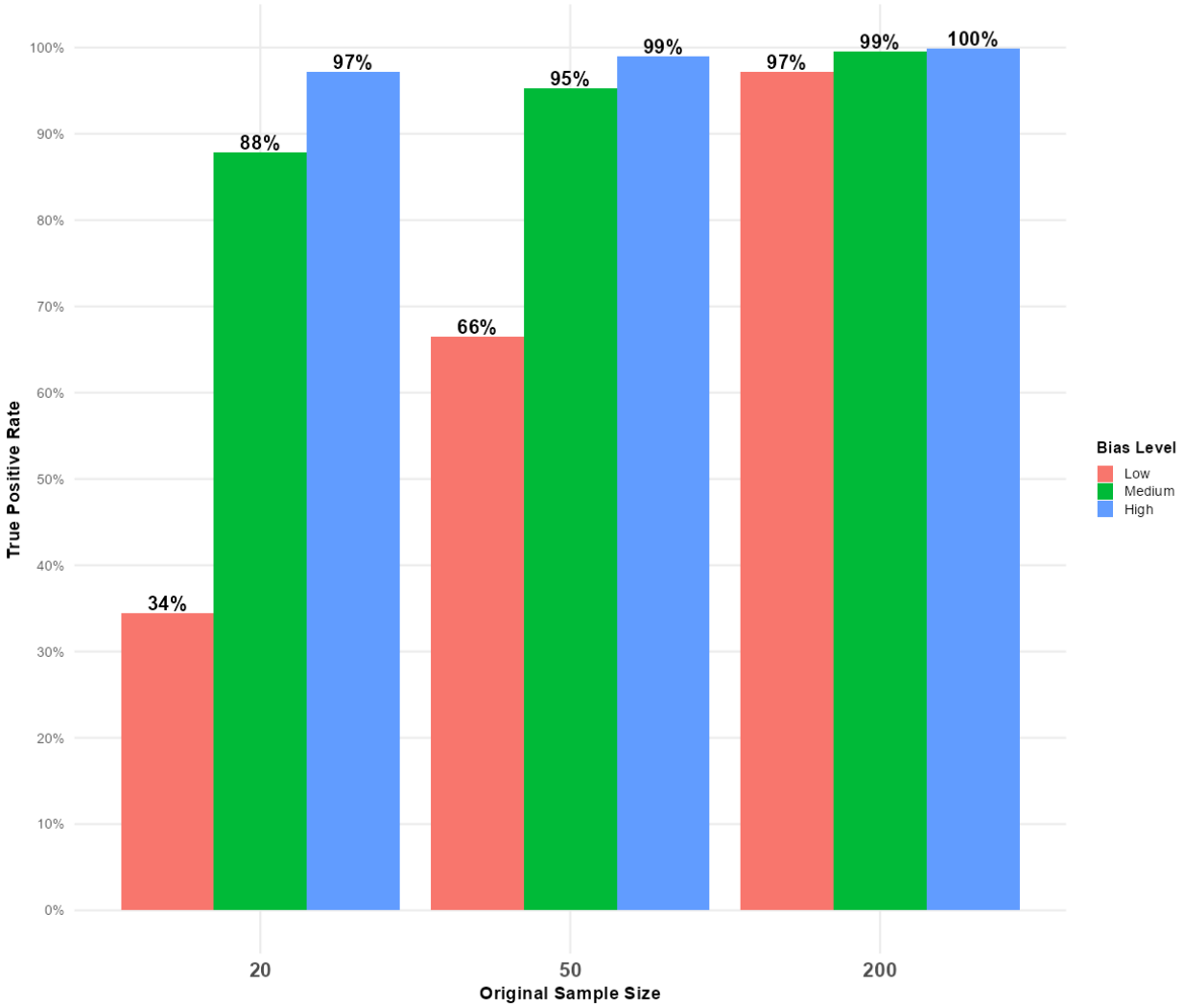

Figure S4 True positive rates of original study outcomes across different original sample sizes and bias mechanism levels when the underlying effect is $\theta = 0.5$, given a nominal significance level of $\alpha = .05$.

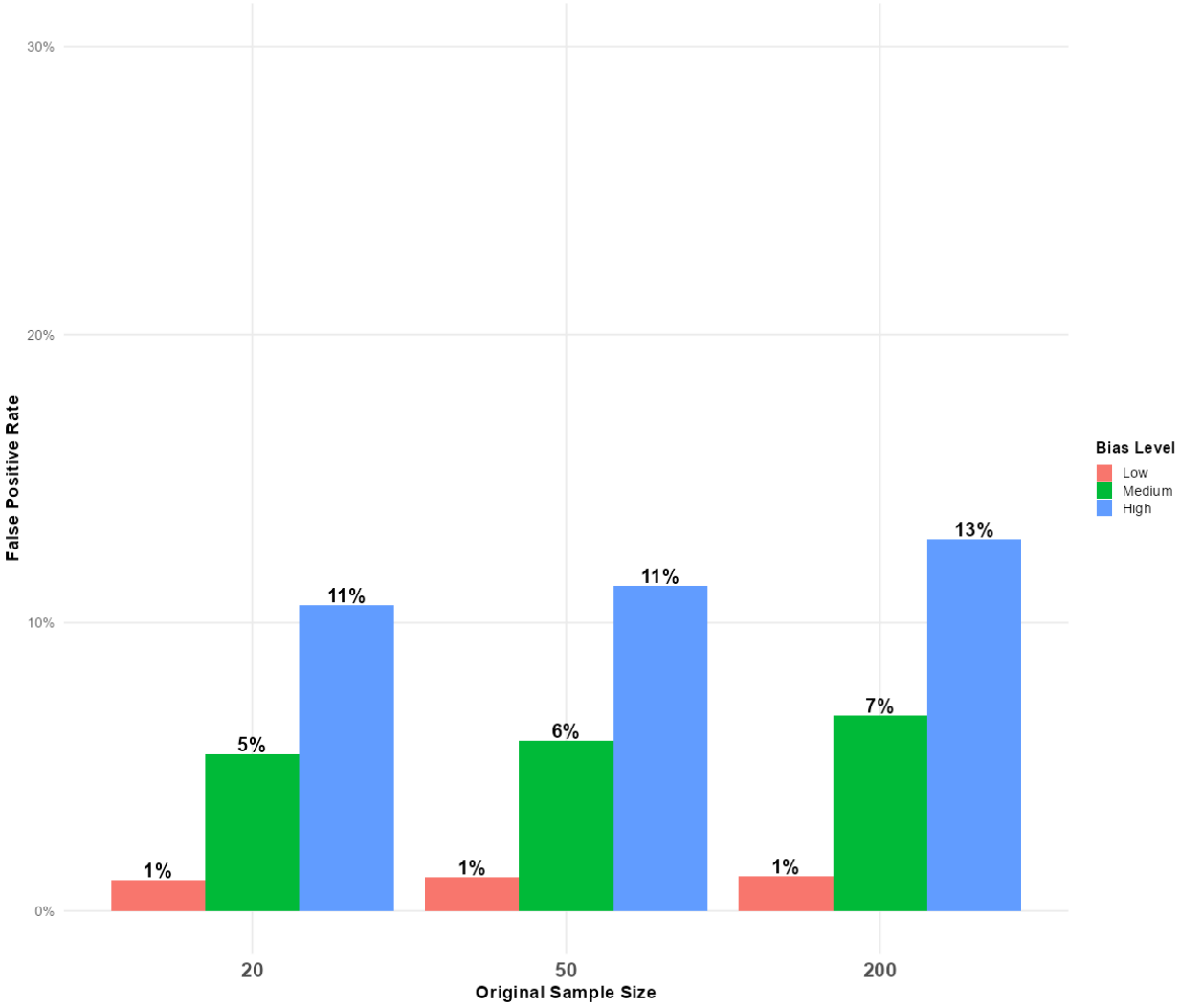

Figure S5 False positive rates of original study outcomes across different original sample sizes and bias mechanism levels, given a nominal significance level of $\alpha = .01$.

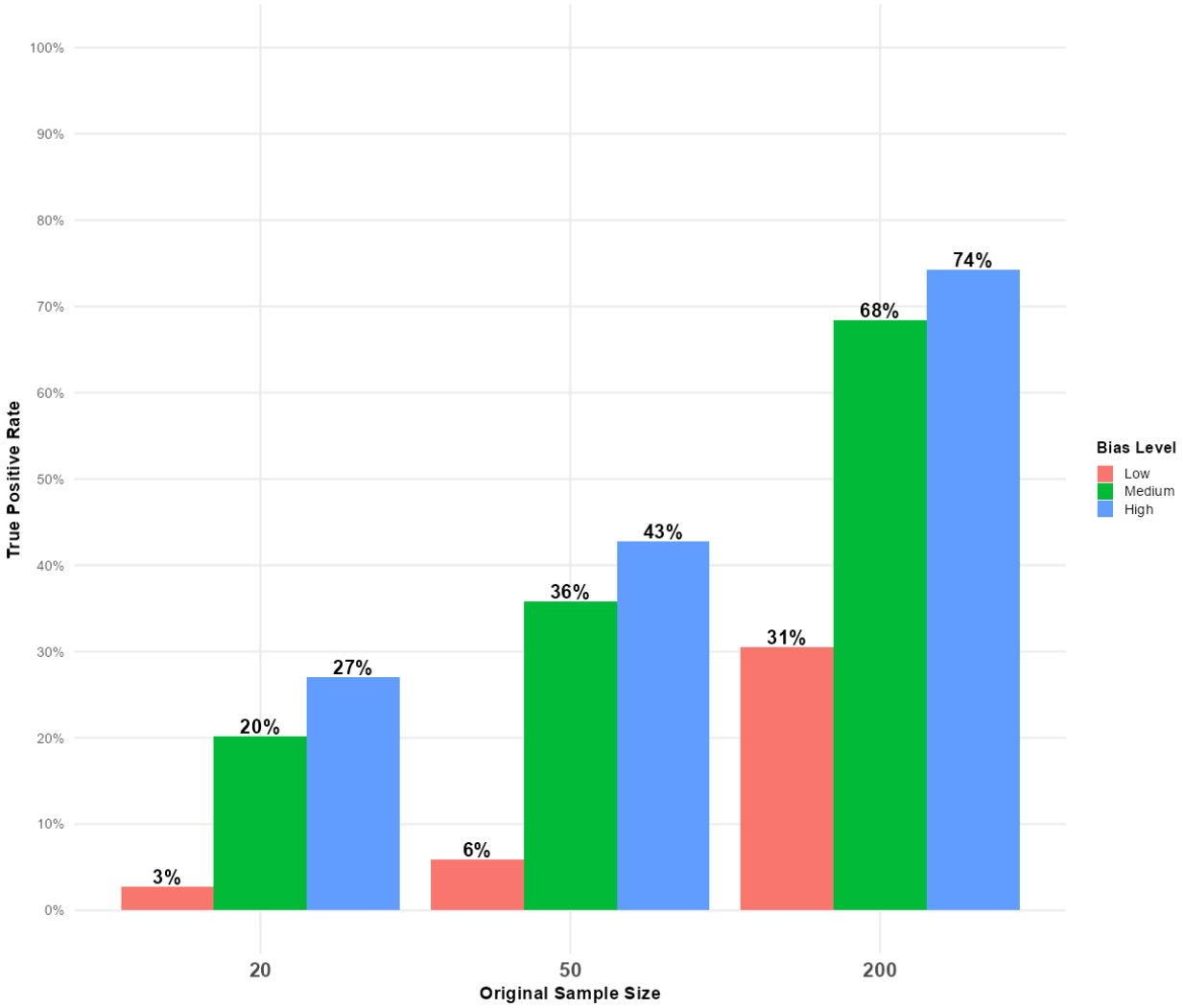

Figure S6 True positive rates of original study outcomes across different original sample sizes and bias mechanism levels when the underlying effect is $\theta = 0.2$, given a nominal significance level of $\alpha = .01$.

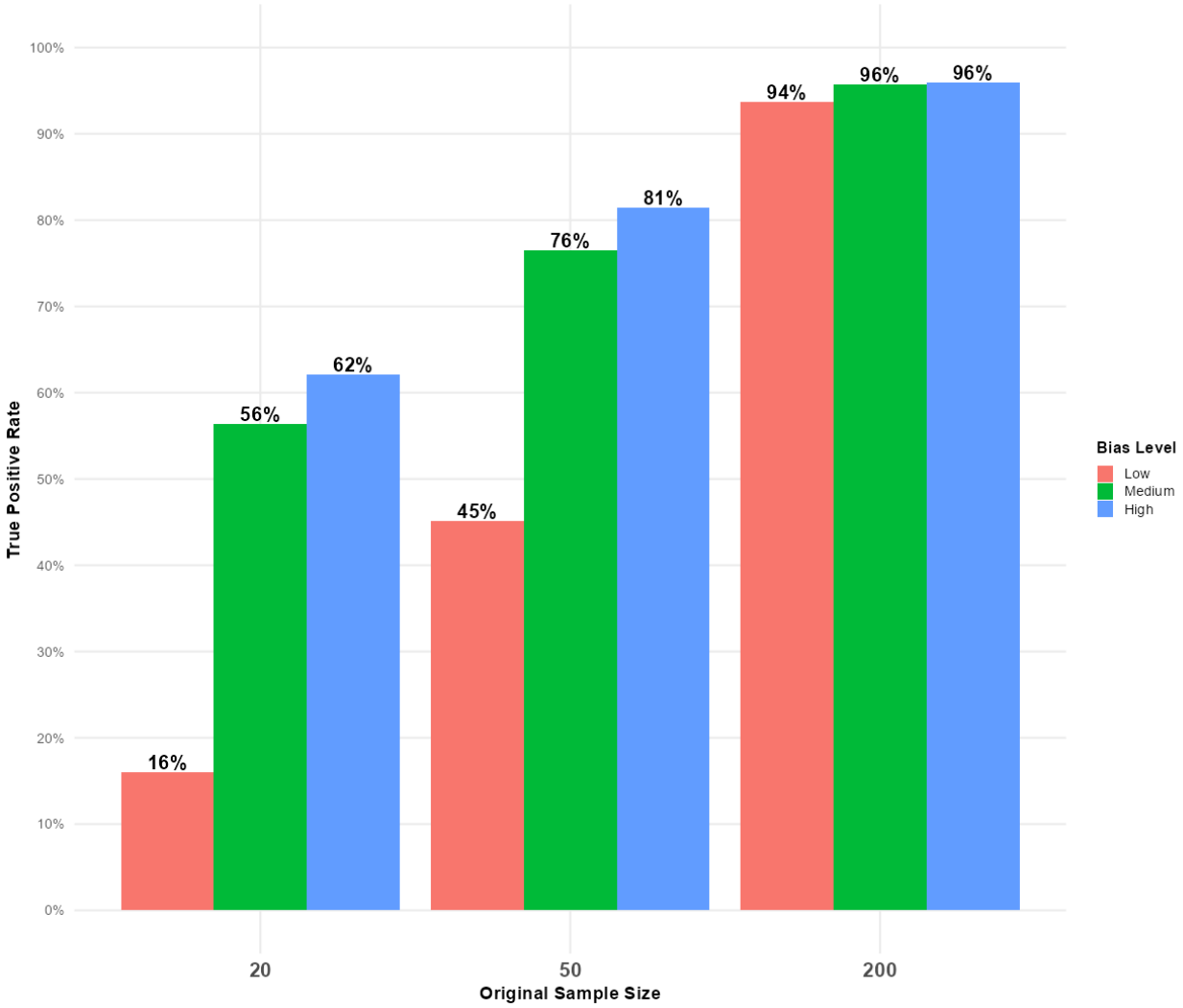

Figure S7 True positive rates of original study outcomes across different original sample sizes and bias mechanism levels when the underlying effect is $\theta = 0.5$, given a nominal significance level of $\alpha = .01$.

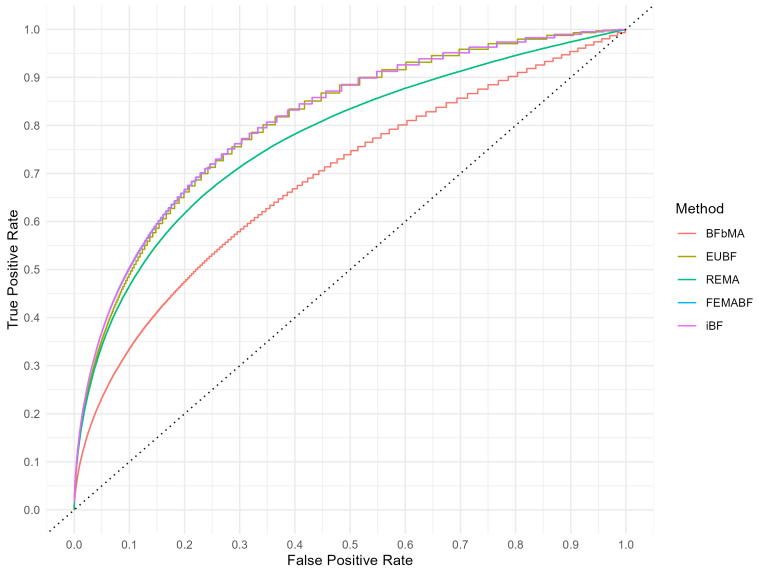

Figure S8 ROC curves of the MABF methods (one-sided test) and the random-effects meta-analysis when the number of replication is 2 and per-group sample size is 40 (N_{rep} = 2, n_{rep} = 40), given a true effect size of $\theta = 0.2$.

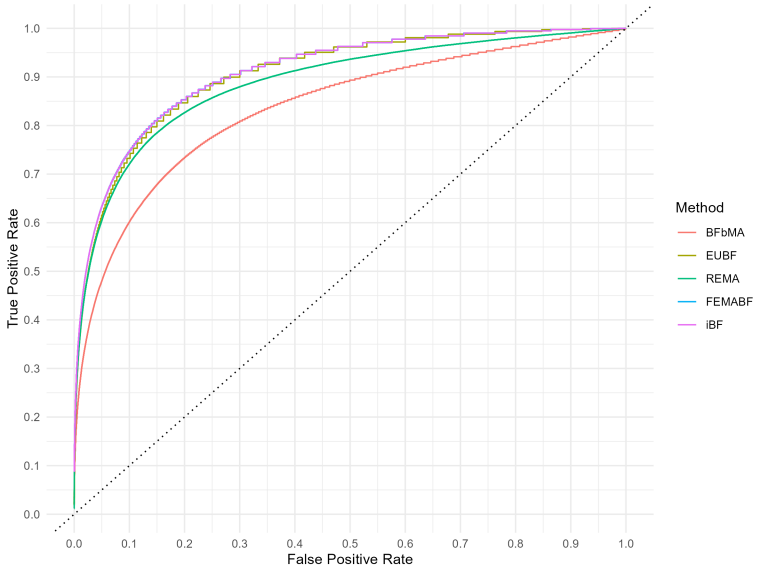

Figure S9 ROC curves of the MABF methods (one-sided test) and the random-effects meta-analysis when the number of replication is 5 and per-group sample size is 40 (N_{rep} = 5, n_{rep} = 40), given a true effect size of $\theta = 0.2$.

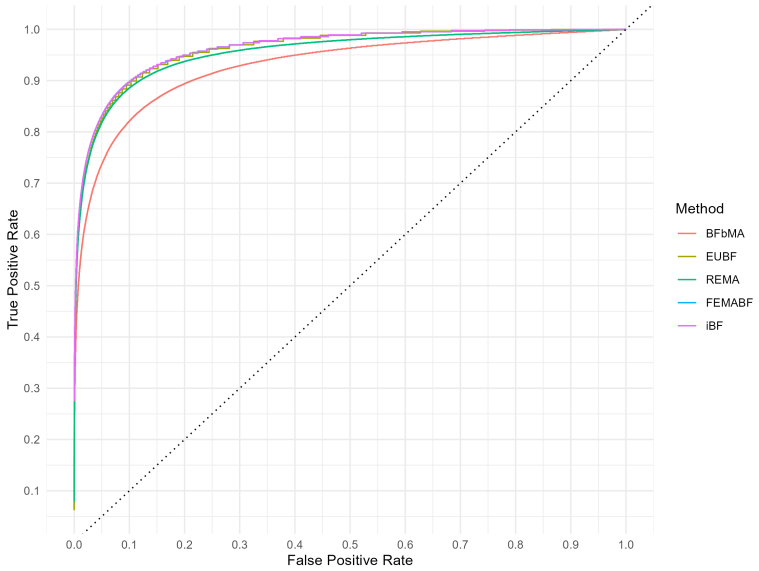

Figure S10 ROC curves of the MABF methods (one-sided test) and the random-effects meta-analysis when the number of replication is 10 and per-group sample size is 40 (N_{rep} = 10, n_{rep} = 40), given a true effect size of $\theta = 0.2$.

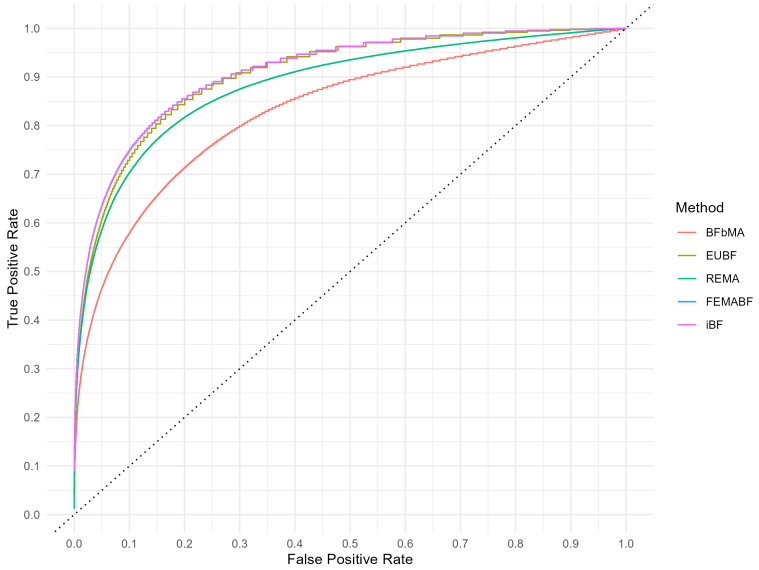

Figure S11 ROC curves of MABF methods (one-sided test) and the random-effects meta-analysis when the number of replication is 2 and per-group sample size is 100 (N_{rep} = 2, n_{rep} = 100), given a true effect size of $\theta = 0.2$.

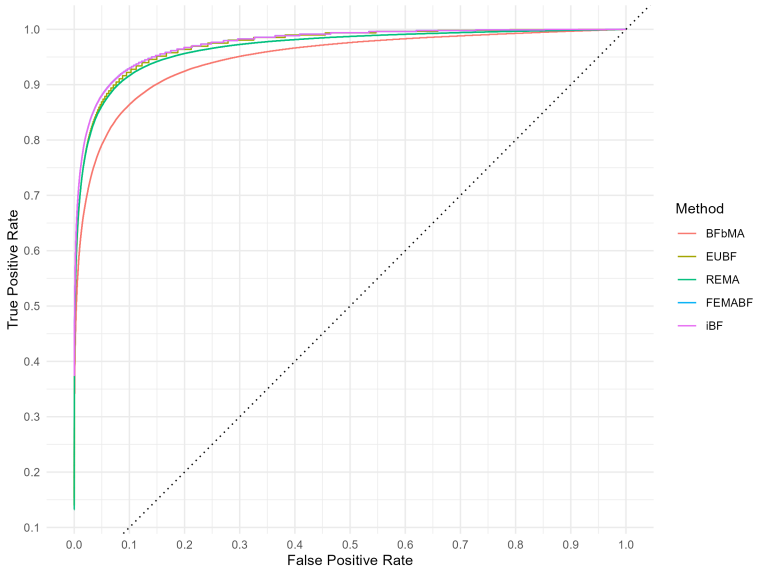

Figure S12 ROC curves of the MABF methods (one-sided test) and the random-effects meta-analysis when the number of replication is 5 and per-group sample size is 100 (N_{rep} = 5, n_{rep} = 100), given a true effect size of $\theta = 0.2$.

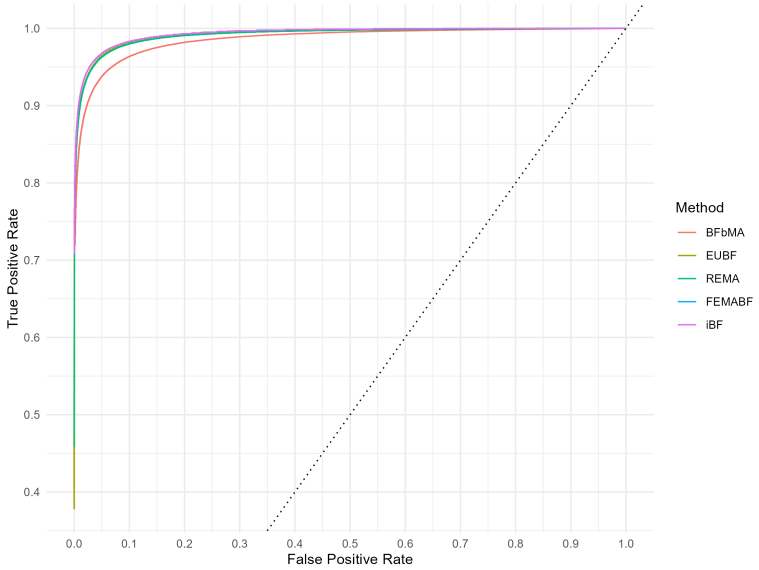

Figure S13 ROC curves of the MABF methods (one-sided test) and the random-effects meta-analysis when the number of replication is 10 and per-group sample size is 100 (N_{rep} = 10, n_{rep} = 100), given a true effect size of $\theta = 0.2$.

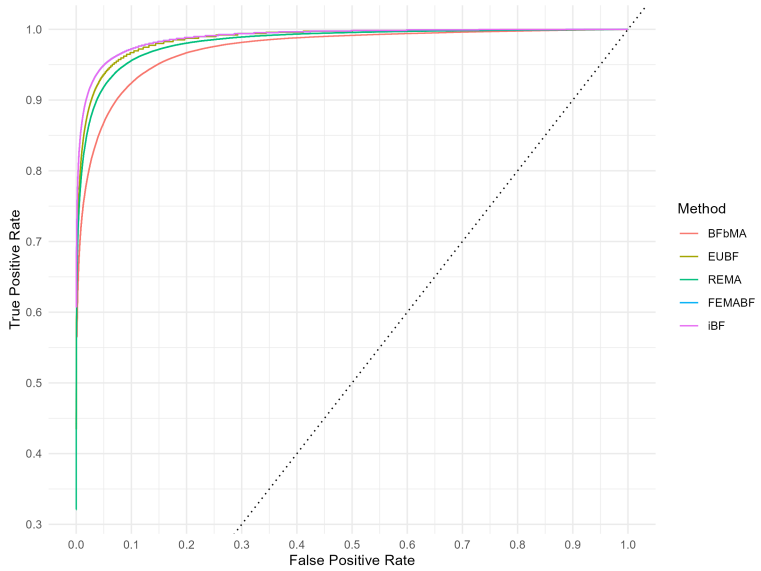

Figure S14 ROC curves of the MABF methods (one-sided test) and the random-effects meta-analysis when the number of replication is 2 and per-group sample size is 400 (N_{rep} = 2, n_{rep} = 400), given a true effect size of $\theta = 0.2$.

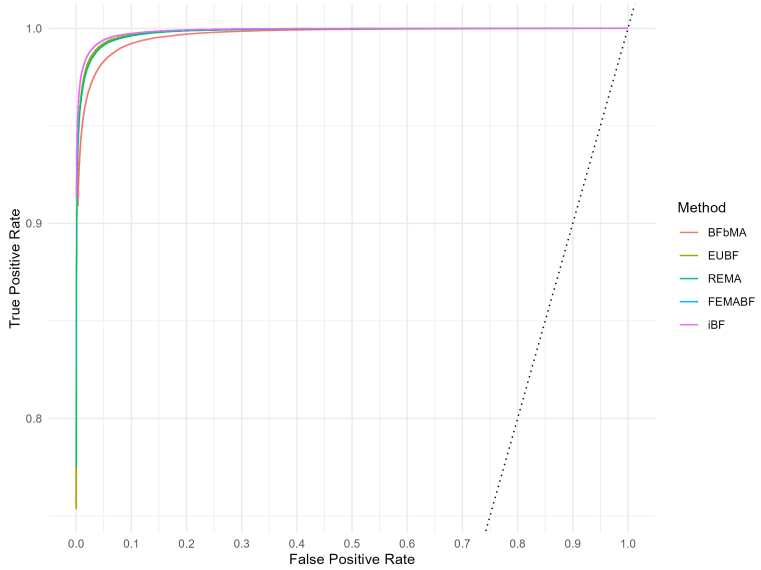

Figure S15 ROC curves of the MABF methods (one-sided test) and the random-effects meta-analysis when the number of replication is 5 and per-group sample size is 400 (N_{rep} = 5, n_{rep} = 400), given a true effect size of $\theta = 0.2$.

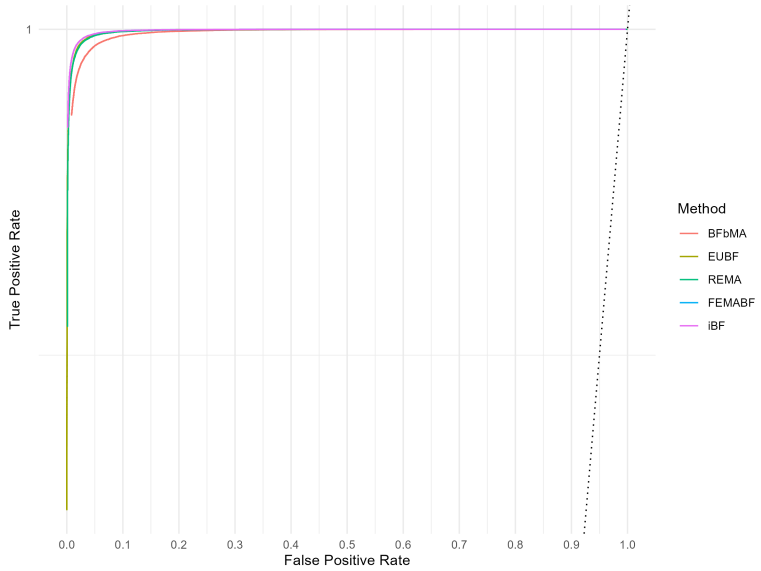

Figure S16 ROC curves of the MABF methods (one-sided test) and the random-effects meta-analysis when the number of replication is 10 and per-group sample size is 400 (N_{rep} = 10, n_{rep} = 400), given a true effect size of $\theta = 0.2$.

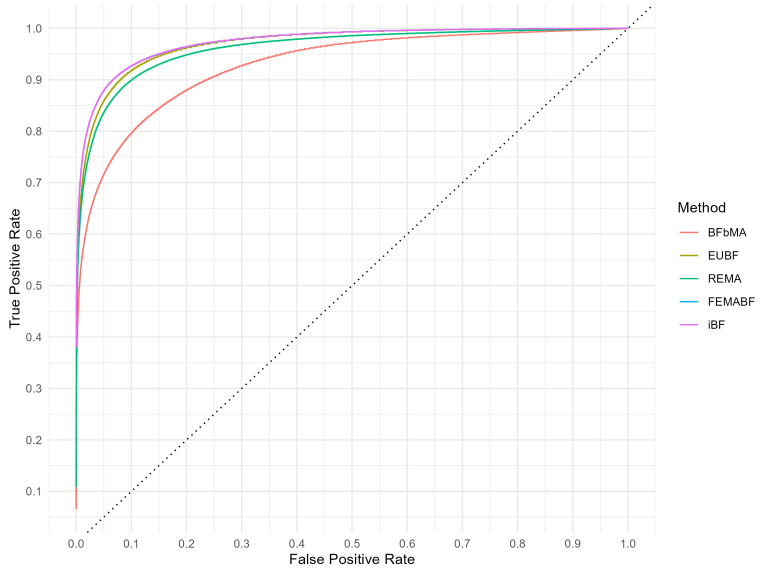

Figure S17 ROC curves of the MABF methods (one-sided test) and the random-effects meta-analysis when the number of replication is 2 and per-group sample size is 40 (N_{rep} = 2, n_{rep} = 40), given a true effect size of $\theta = 0.5$.

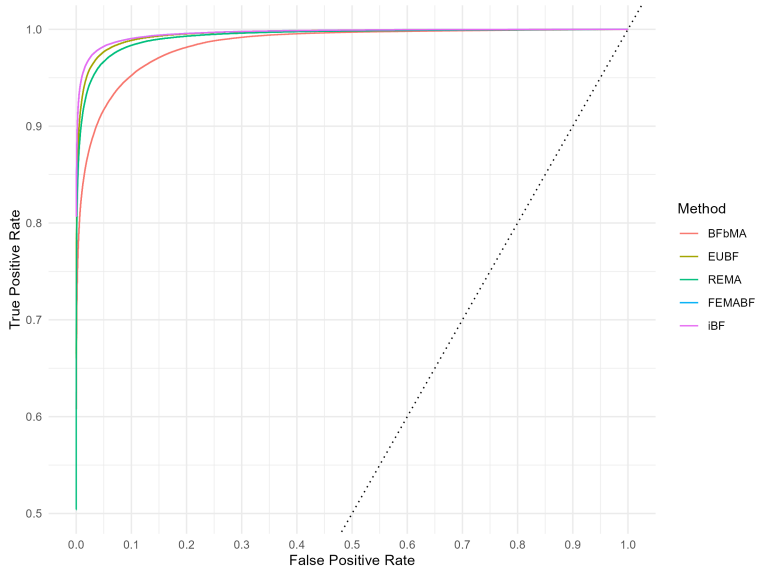

Figure S18 ROC curves of the MABF methods (one-sided test) and the random-effects meta-analysis when the number of replication is 2 and per-group sample size is 100 (N_{rep} = 2, n_{rep} = 100), given a true effect size of $\theta = 0.5$.

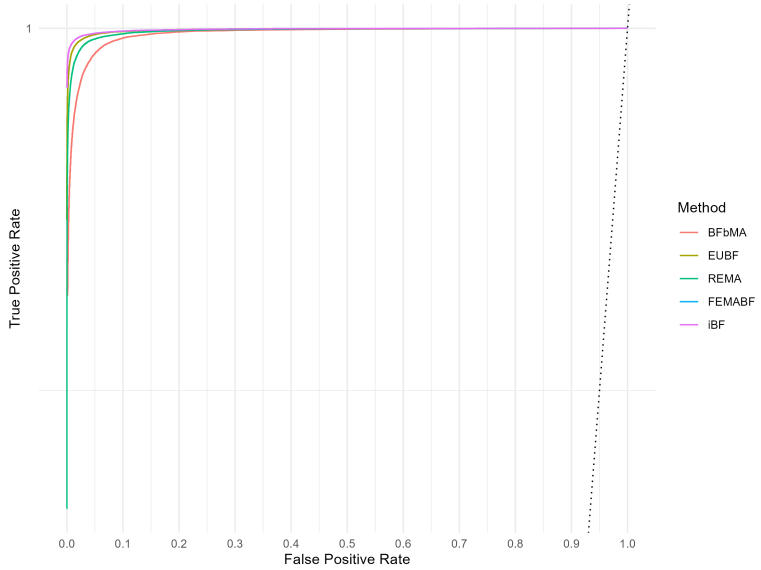

Figure S19 ROC curves of the MABF methods (one-sided test) and the random-effects meta-analysis when the number of replication is 2 and per-group sample size is 400 (N_{rep} = 2, n_{rep} = 400), given a true effect size of $\theta = 0.5$.

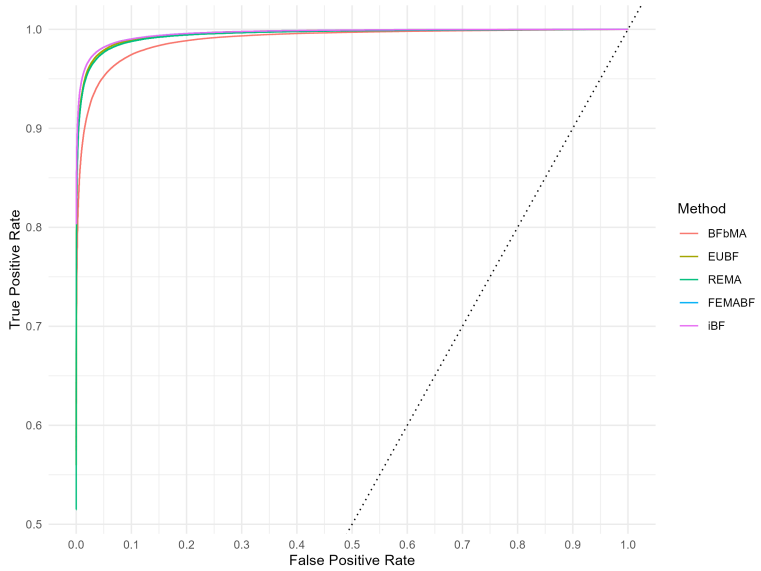

Figure S20 ROC curves of the MABF methods (one-sided test) and the random-effects meta-analysis when the number of replication is 5 and per-group sample size is 40 (N_{rep} = 5, n_{rep} = 40), given a true effect size of $\theta = 0.5$.

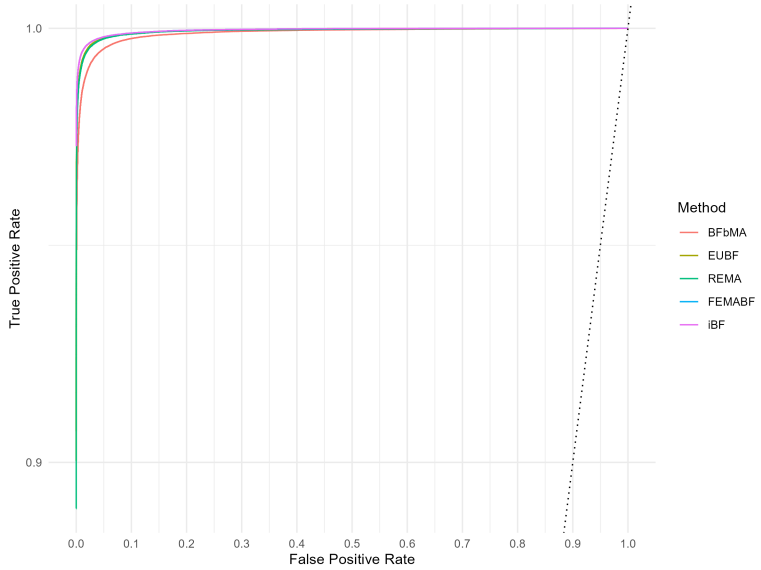

Figure S21 ROC curves of the MABF methods (one-sided test) and the random-effects meta-analysis when the number of replication is 5 and per-group sample size is 100 (N_{rep} = 5, n_{rep} = 100), given a true effect size of $\theta = 0.5$.

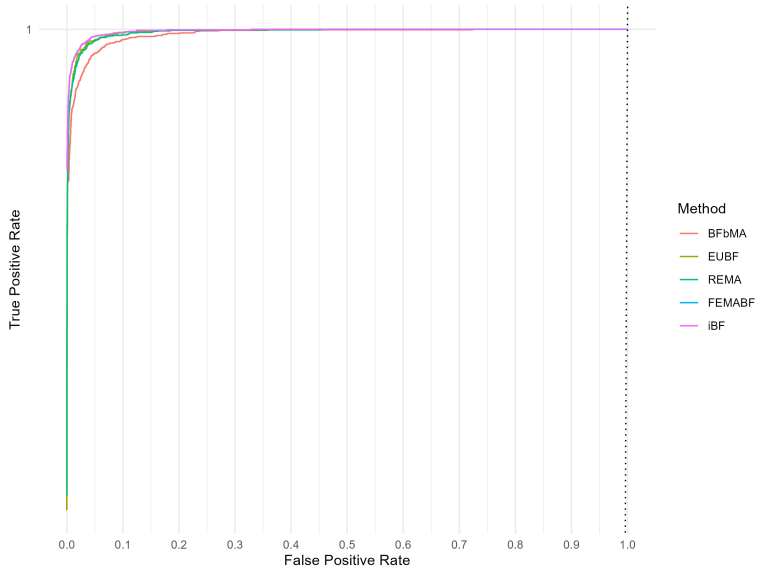

Figure S22 ROC curves of the MABF methods (one-sided test) and the random-effects meta-analysis when the number of replication is 5 and per-group sample size is 400 (N_{rep} = 5, n_{rep} = 400), given a true effect size of $\theta = 0.5$.

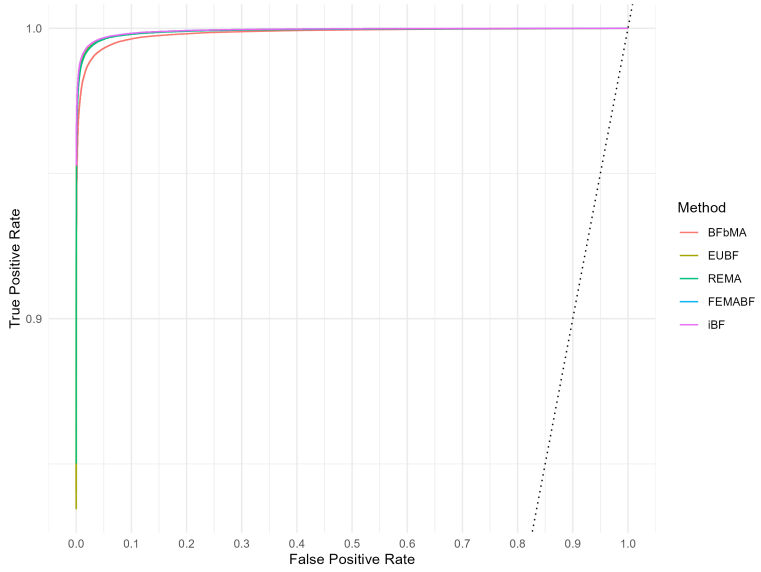

Figure S23 ROC curves of the MABF methods (one-sided test) and the random-effects meta-analysis when the number of replication is 10 and per-group sample size is 40 (N_{rep} = 10, n_{rep} = 40), given a true effect size of $\theta = 0.5$.

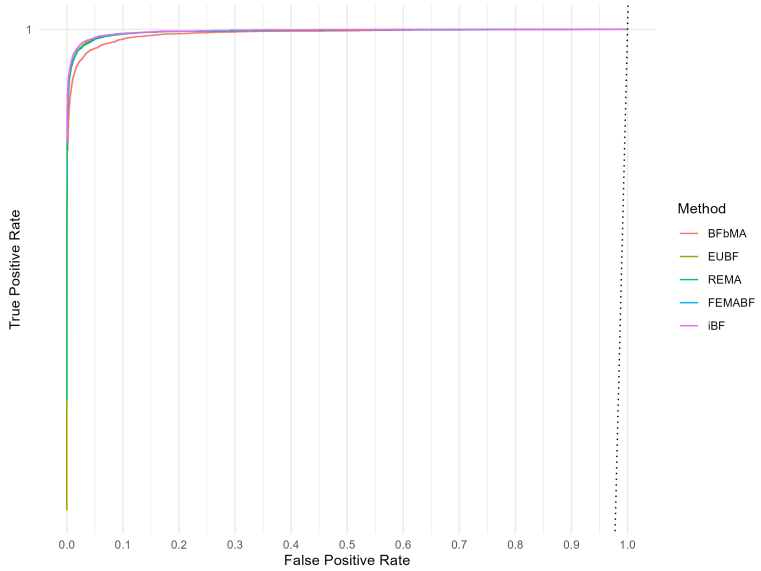

Figure S24 ROC curves of the MABF methods (one-sided test) and the random-effects meta-analysis when the number of replication is 10 and per-group sample size is 100 (N_{rep} = 10, n_{rep} = 100), given a true effect size of $\theta = 0.5$.

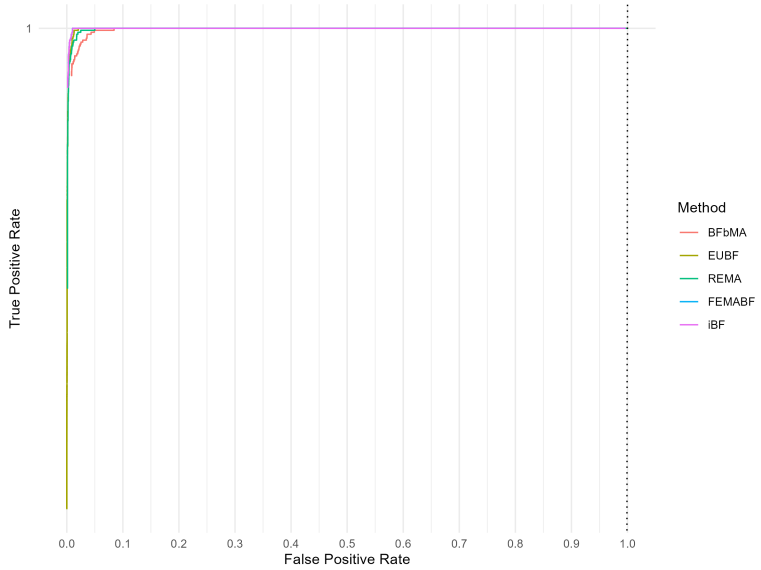

Figure S25 ROC curves of the MABF methods (one-sided test) and the random-effects meta-analysis when the number of replication is 10 and per-group sample size is 400 (N_{rep} = 10, n_{rep} = 400), given a true effect size of $\theta = 0.5$.

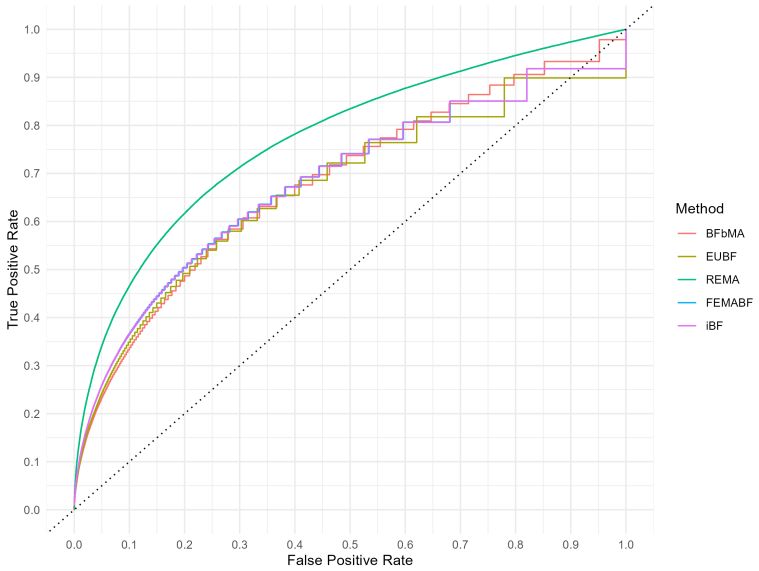

Figure S26 ROC curves of the MABF methods (two-sided test) and the random-effects meta-analysis when the number of replication is 2 and per-group sample size is 40 (N_{rep} = 2, n_{rep} = 40), given a true effect size of $\theta = 0.2$.

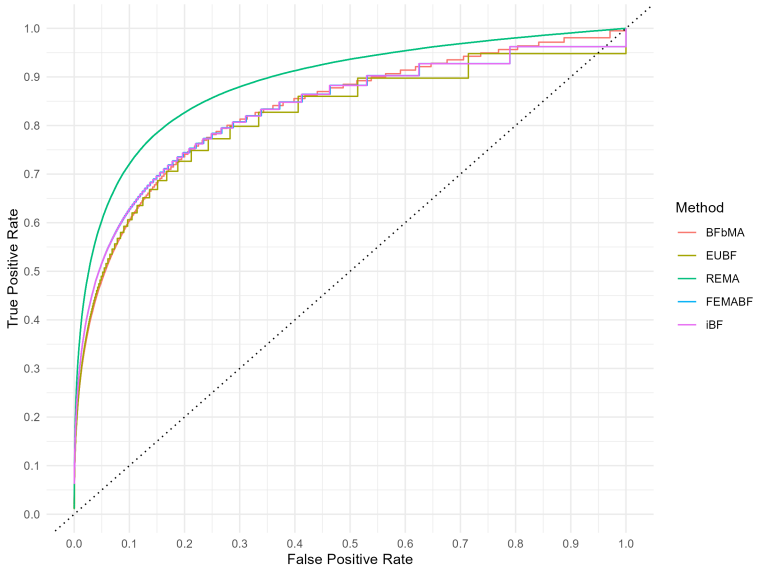

Figure S27 ROC curves of the MABF methods (two-sided test) and the random-effects meta-analysis when the number of replication is 5 and per-group sample size is 40 (N_{rep} = 5, n_{rep} = 40), given a true effect size of $\theta = 0.2$.

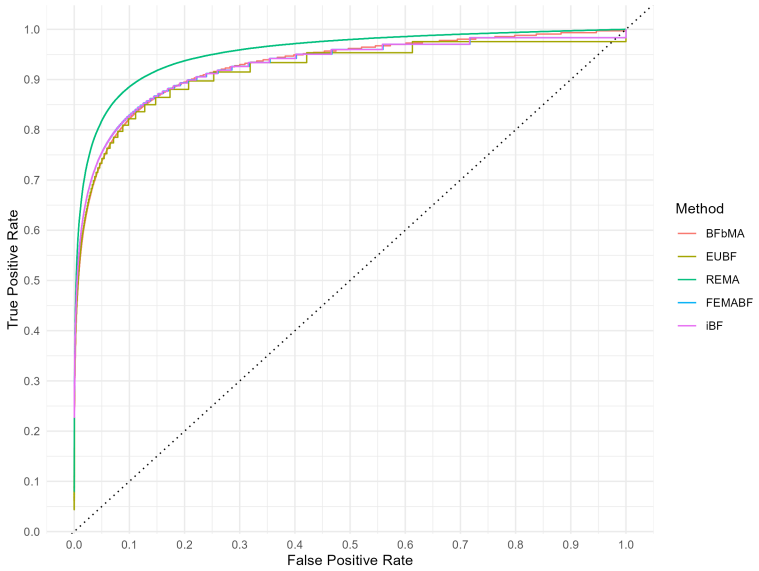

Figure S28 ROC curves of the MABF methods (two-sided test) and the random-effects meta-analysis when the number of replication is 10 and per-group sample size is 40 (N_{rep} = 10, n_{rep} = 40), given a true effect size of $\theta = 0.2$.

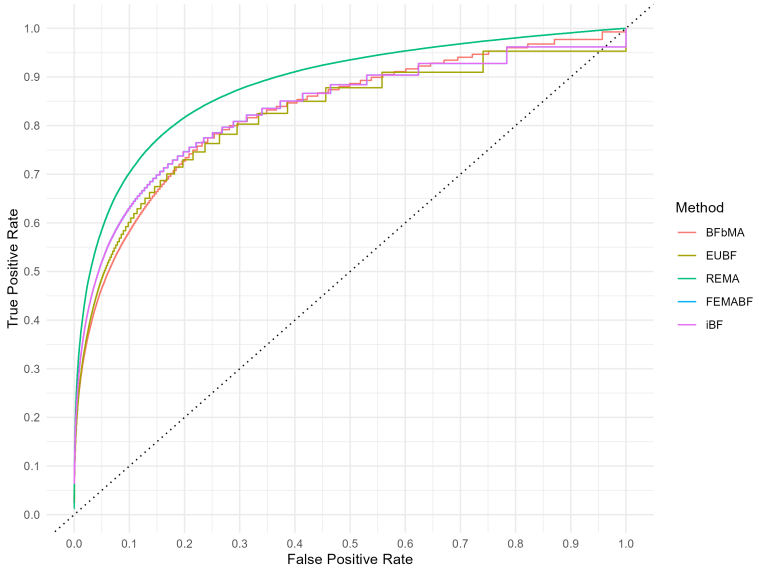

Figure S29 ROC curves of the MABF methods (two-sided test) and the random-effects meta-analysis when the number of replication is 2 and per-group sample size is 100 (N_{rep} = 2, n_{rep} = 100), given a true effect size of $\theta = 0.2$.

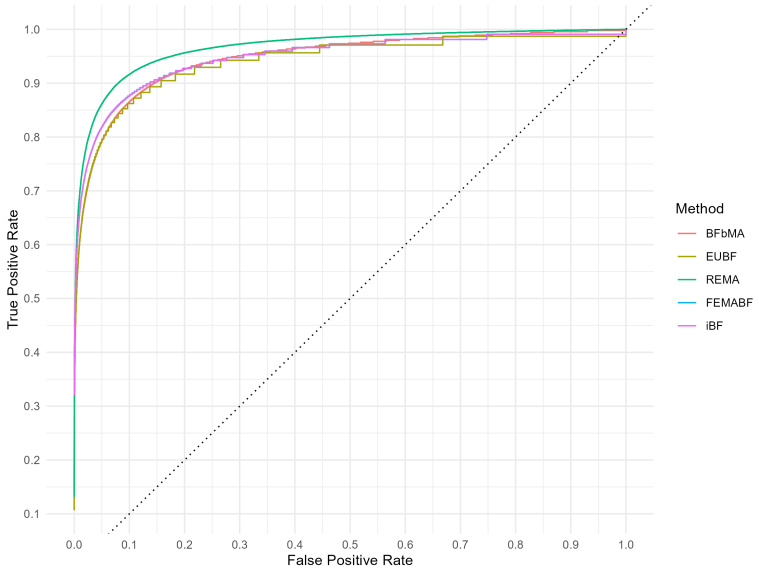

Figure S30 ROC curves of the MABF methods (two-sided test) and the random-effects meta-analysis when the number of replication is 5 and per-group sample size is 100 (N_{rep} = 5, n_{rep} = 100), given a true effect size of $\theta = 0.2$.

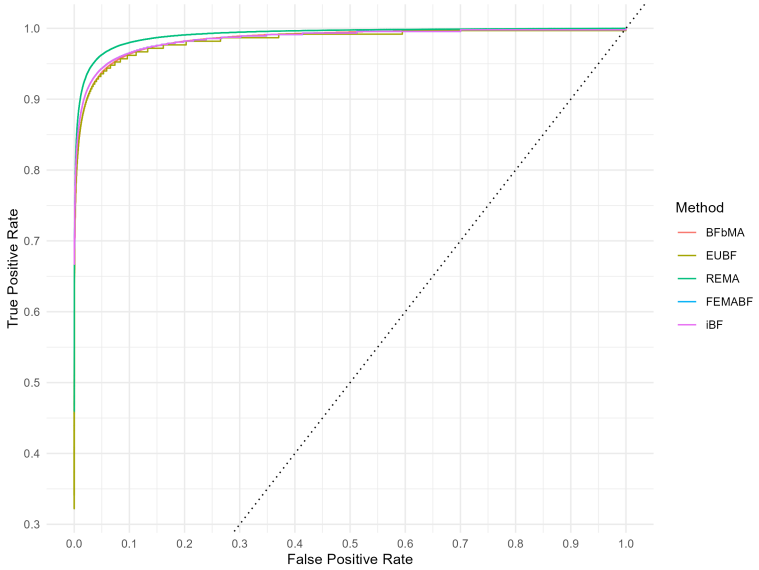

Figure S31 ROC curves of the MABF methods (two-sided test) and the random-effects meta-analysis when the number of replication is 10 and per-group sample size is 100 (N_{rep} = 10, n_{rep} = 100), given a true effect size of $\theta = 0.2$.

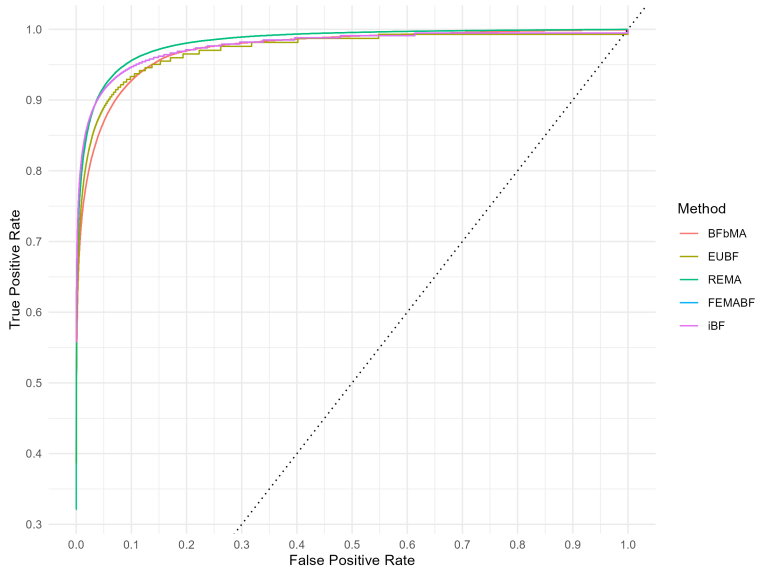

Figure S32 ROC curves of the MABF methods (two-sided test) and the random-effects meta-analysis when the number of replication is 2 and per-group sample size is 400 (N_{rep} = 2, n_{rep} = 400), given a true effect size of $\theta = 0.2$.

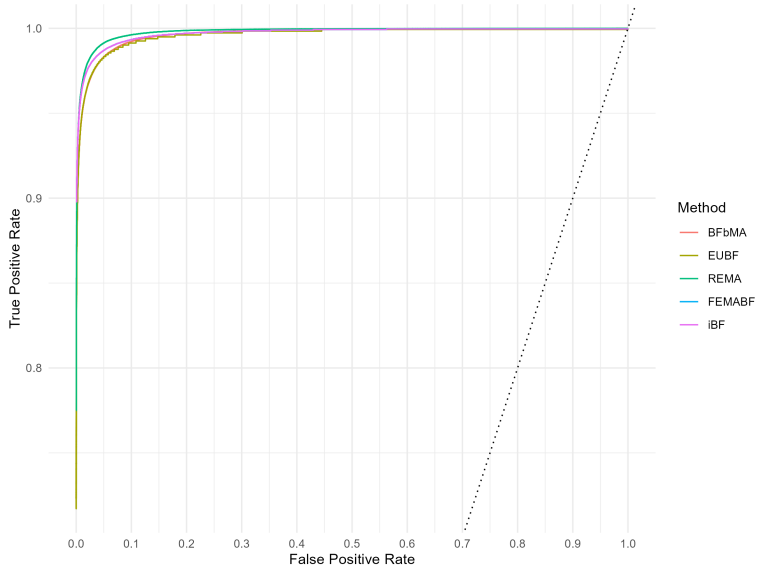

Figure S33 ROC curves of the MABF methods (two-sided test) and the random-effects meta-analysis when the number of replication is 5 and per-group sample size is 400 (N_{rep} = 5, n_{rep} = 400), given a true effect size of $\theta = 0.2$.

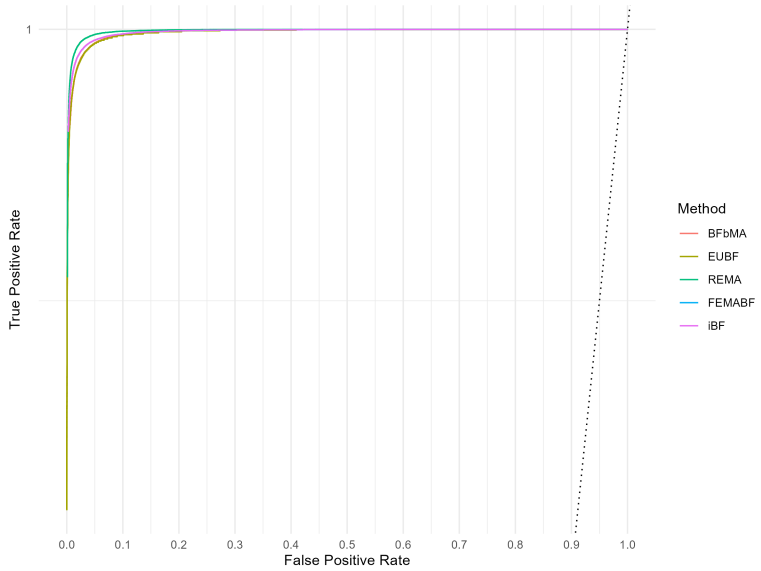

Figure S34 ROC curves of the MABF methods (two-sided test) and the random-effects meta-analysis when the number of replication is 10 and per-group sample size is 400 (N_{rep} = 10, n_{rep} = 400), given a true effect size of $\theta = 0.2$.

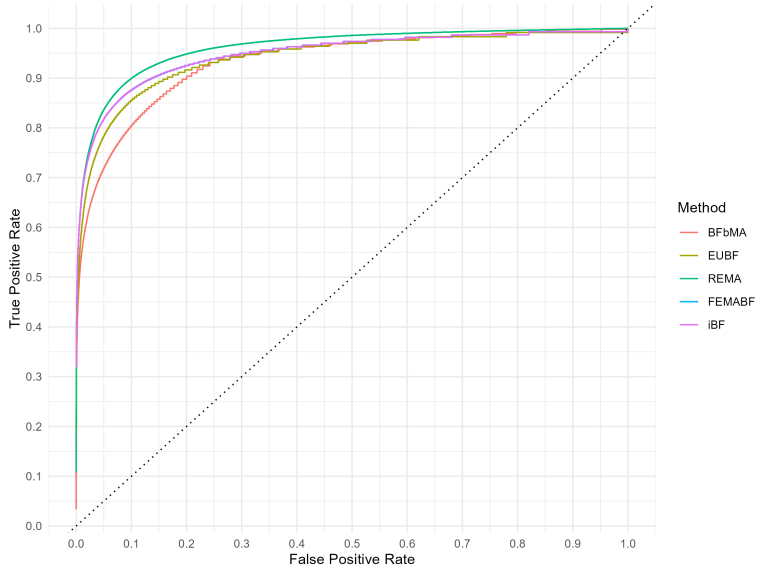

Figure S35 ROC curves of the MABF methods (two-sided test) and the random-effects meta-analysis when the number of replication is 2 and per-group sample size is 40 (N_{rep} = 2, n_{rep} = 40), given a true effect size of $\theta = 0.5$.

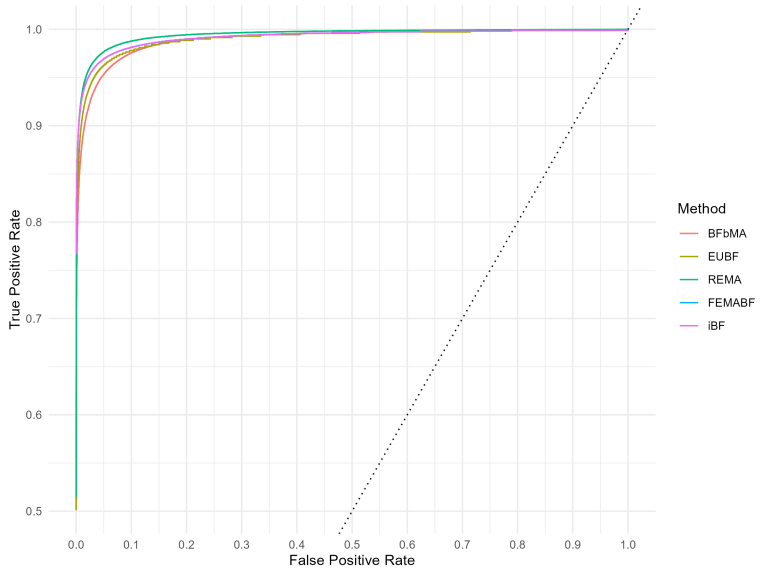

Figure S36 ROC curves of the MABF methods (two-sided test) and the random-effects meta-analysis when the number of replication is 5 and per-group sample size is 40 (N_{rep} = 5, n_{rep} = 40), given a true effect size of $\theta = 0.5$.

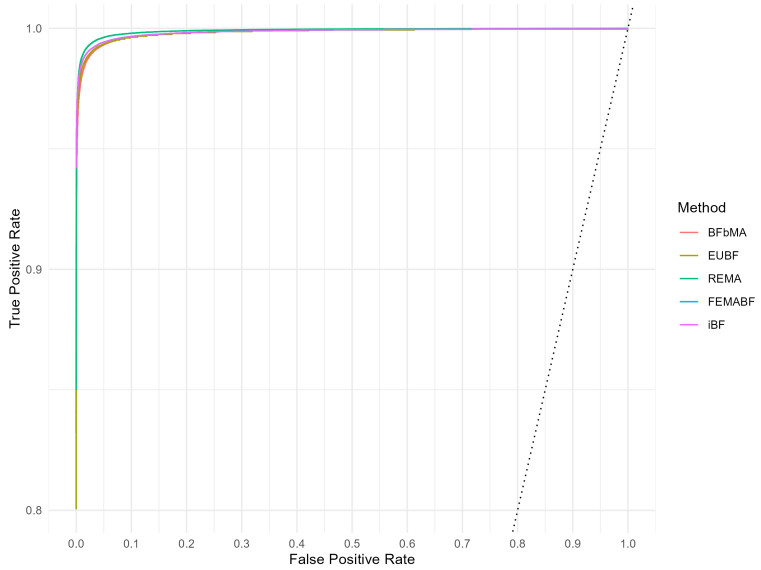

Figure S37 ROC curves of the MABF methods (two-sided test) and the random-effects meta-analysis when the number of replication is 10 and per-group sample size is 40 (N_{rep} = 10, n_{rep} = 40), given a true effect size of $\theta = 0.5$.

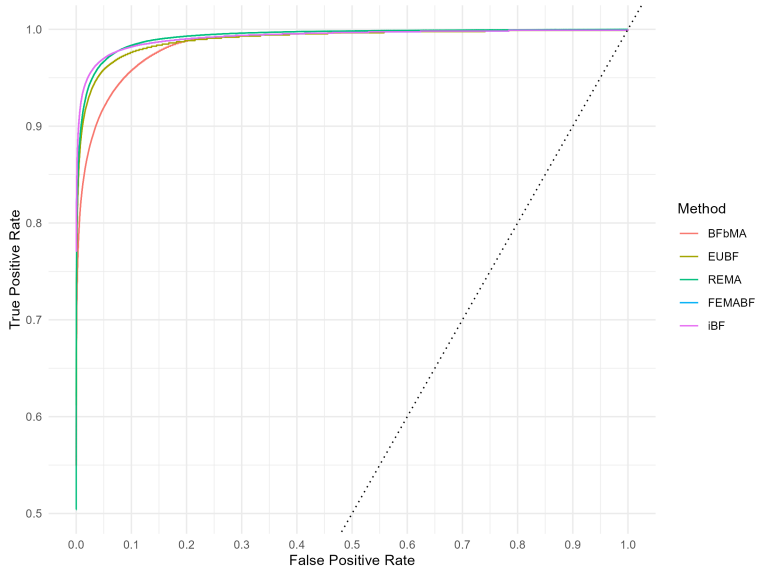

Figure S38 ROC curves of the MABF methods (two-sided test) and the random-effects meta-analysis when the number of replication is 2 and per-group sample size is 100 (N_{rep} = 2, n_{rep} = 100), given a true effect size of $\theta = 0.5$.

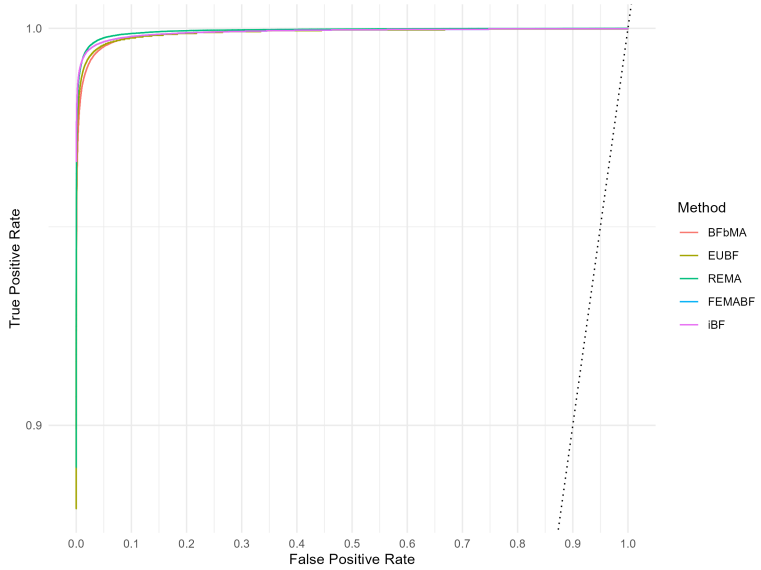

Figure S39 ROC curves of the MABF methods (two-sided test) and the random-effects meta-analysis when the number of replication is 5 and per-group sample size is 100 (N_{rep} = 5, n_{rep} = 100), given a true effect size of $\theta = 0.5$.

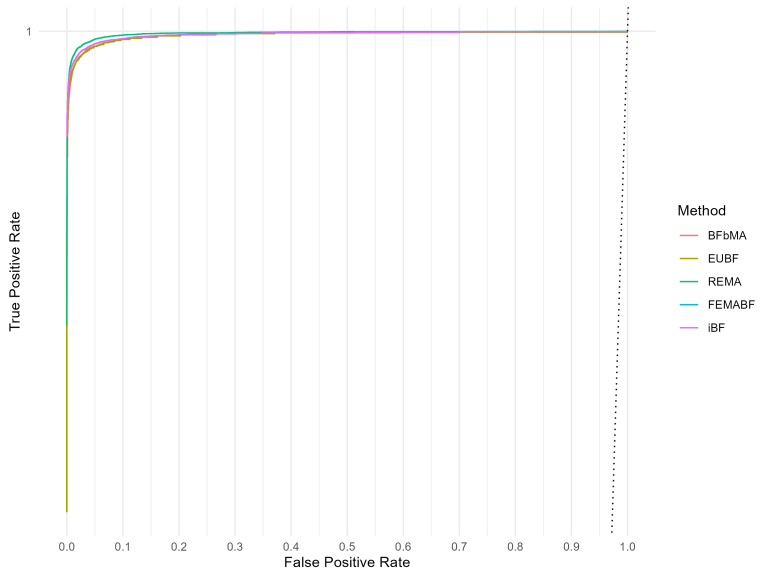

Figure S40 ROC curves of the MABF methods (two-sided test) and the random-effects meta-analysis when the number of replication is 10 and per-group sample size is 100 (N_{rep} = 10, n_{rep} = 100), given a true effect size of $\theta = 0.5$.

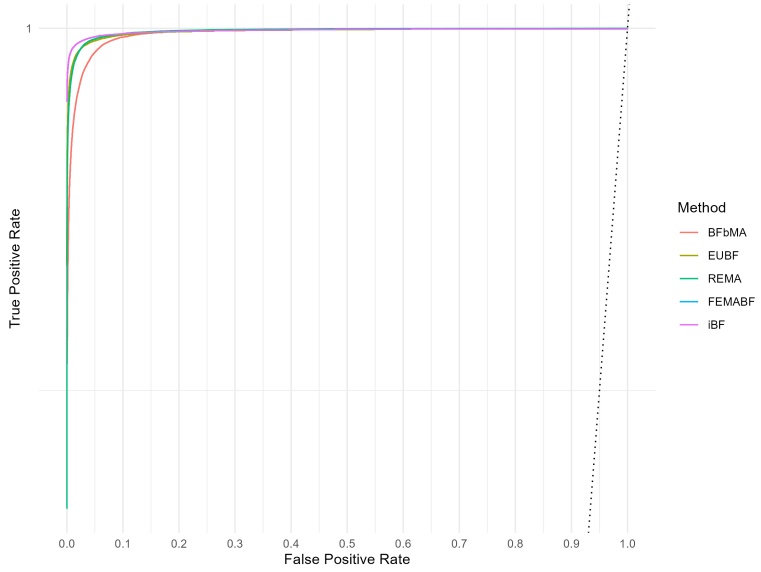

Figure S41 ROC curves of the MABF methods (two-sided test) and the random-effects meta-analysis when the number of replication is 2 and per-group sample size is 400 (N_{rep} = 2, n_{rep} = 400), given a true effect size of $\theta = 0.5$.

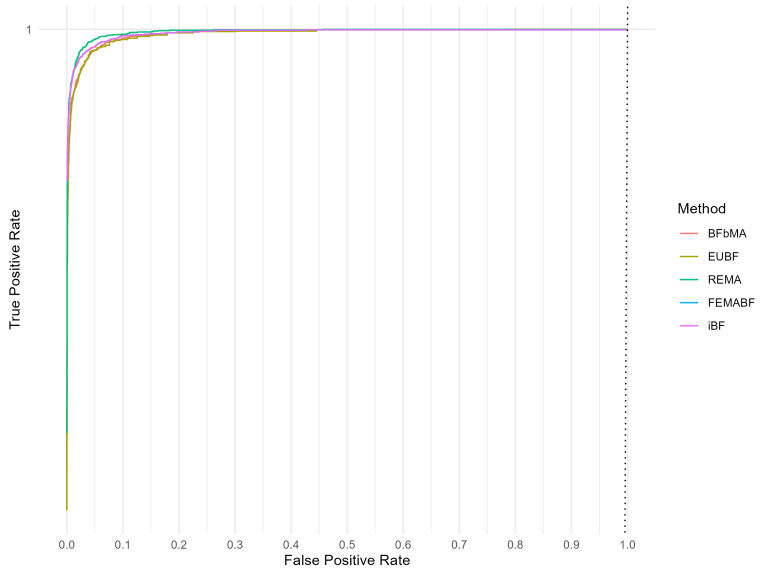

Figure S42 ROC curves of the MABF methods (two-sided test) and the random-effects meta-analysis when the number of replication is 5 and per-group sample size is 400 (N_{rep} = 5, n_{rep} = 400), given a true effect size of $\theta = 0.5$.

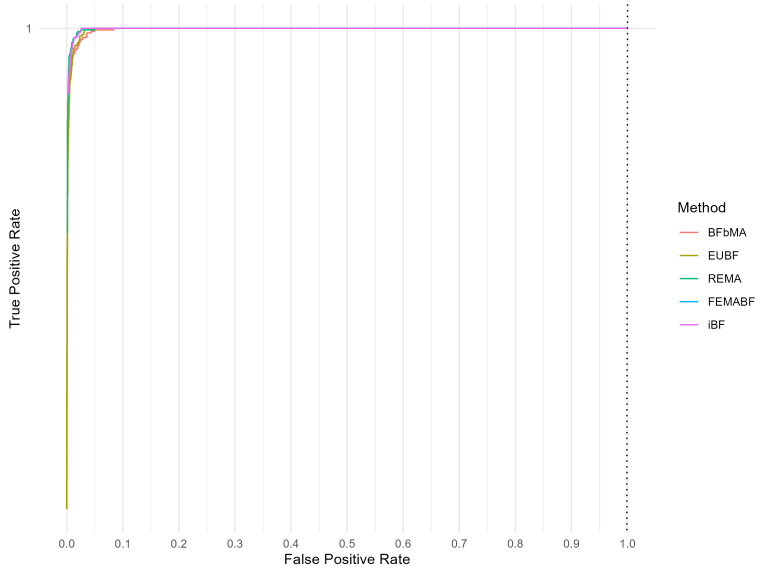

Figure S43 ROC curves of the MABF methods (two-sided test) and the random-effects meta-analysis when the number of replication is 10 and per-group sample size is 400 (N_{rep} = 10, n_{rep} = 400), given a true effect size of $\theta = 0.5$.

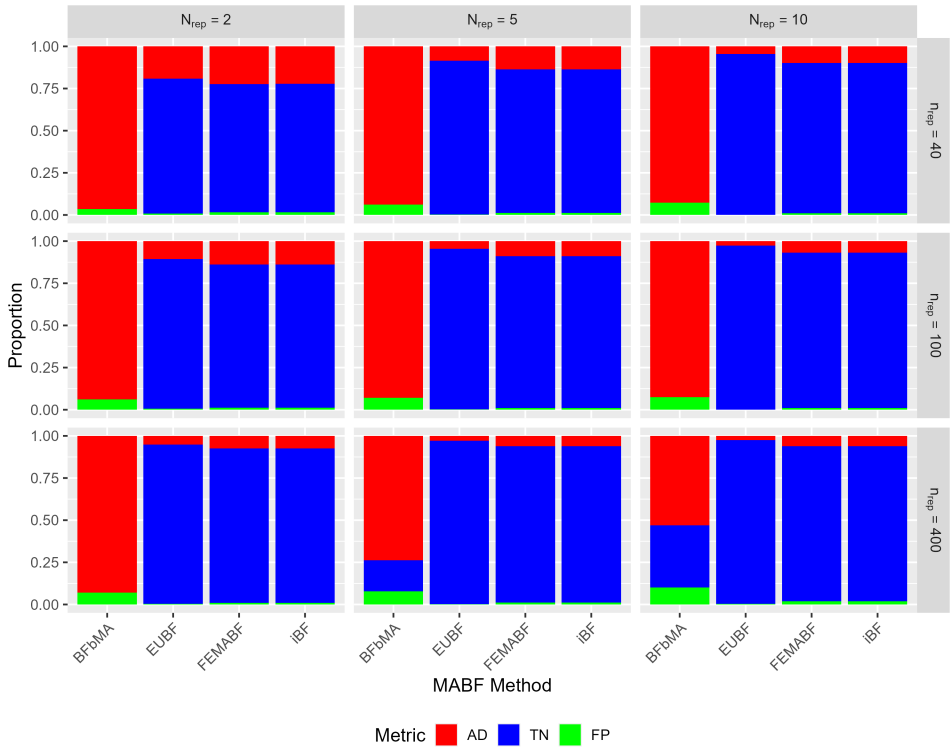

Figure S44 Proportions of anecdotal evidence, false positives, and true negatives produced by the MABF methods (one-sided test) across different replication designs when $\theta = 0$.

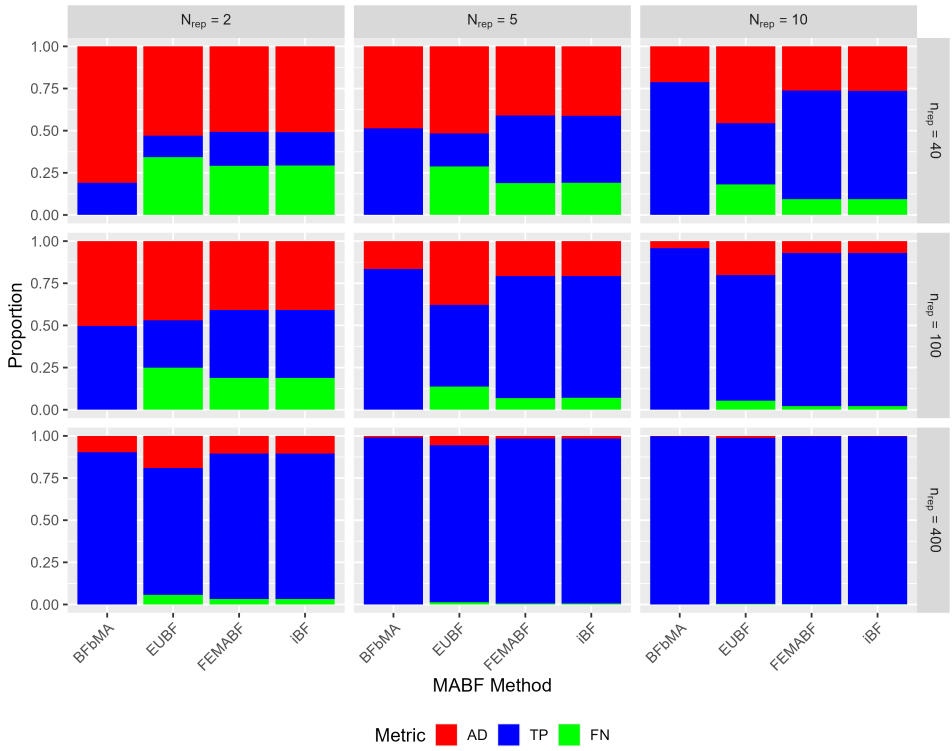

Figure S45 Proportions of anecdotal evidence, true positives, and false negatives produced by the MABF methods (one-sided test) across different replication designs when $\theta = 0.2$.

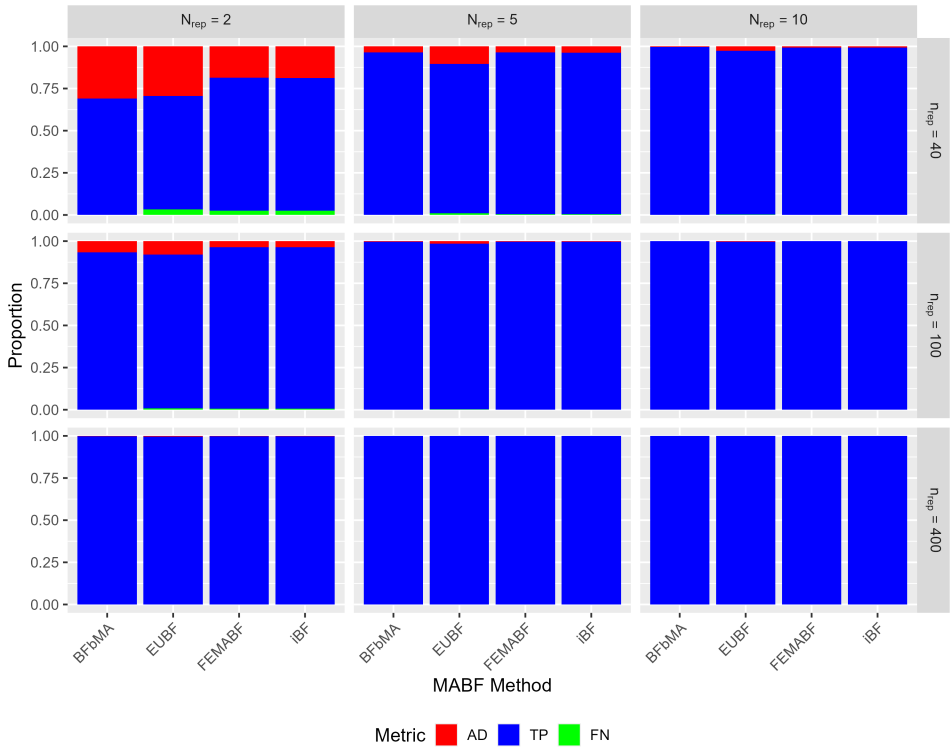

Figure S46 Proportions of anecdotal evidence, true positives, and false negatives produced by the MABF methods (one-sided test) across different replication designs when $\theta = 0.5$.

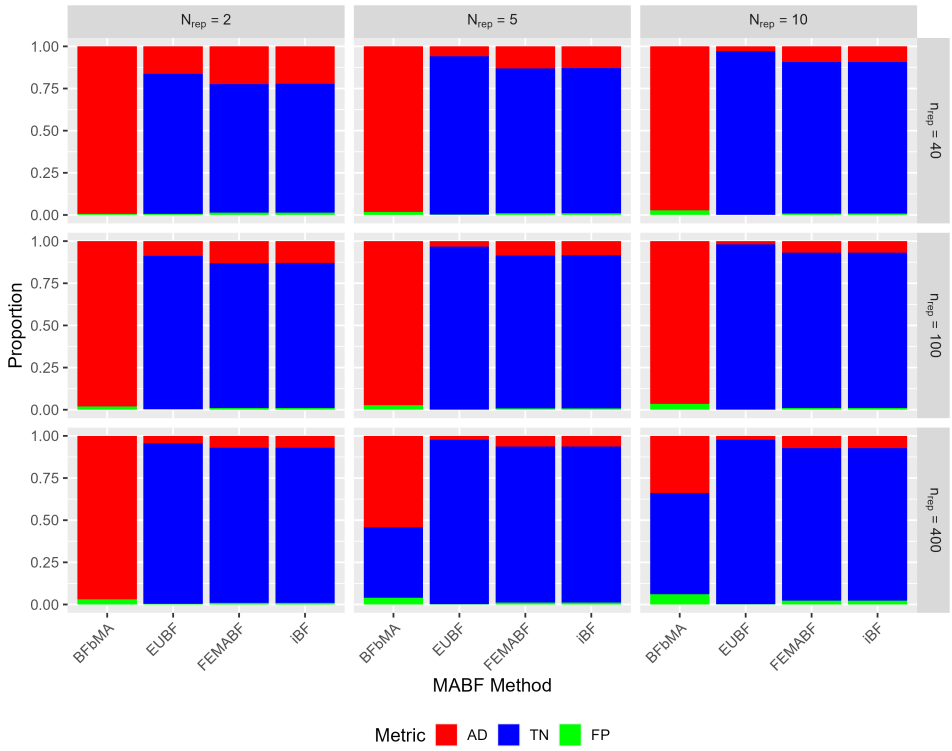

Figure S47 Proportions of anecdotal evidence, false positives, and true negatives produced by the MABF methods (two-sided test) across different replication designs when $\theta = 0$.

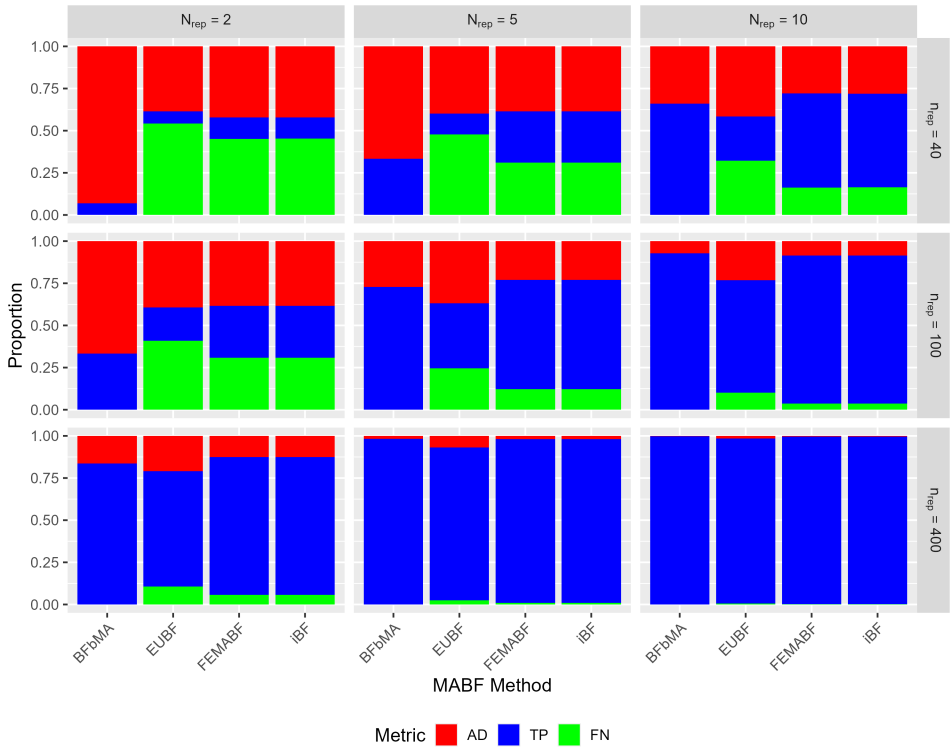

Figure S48 Proportions of anecdotal evidence, true positives, and false negatives produced by the MABF methods (two-sided test) across different replication designs when $\theta = 0.2$.

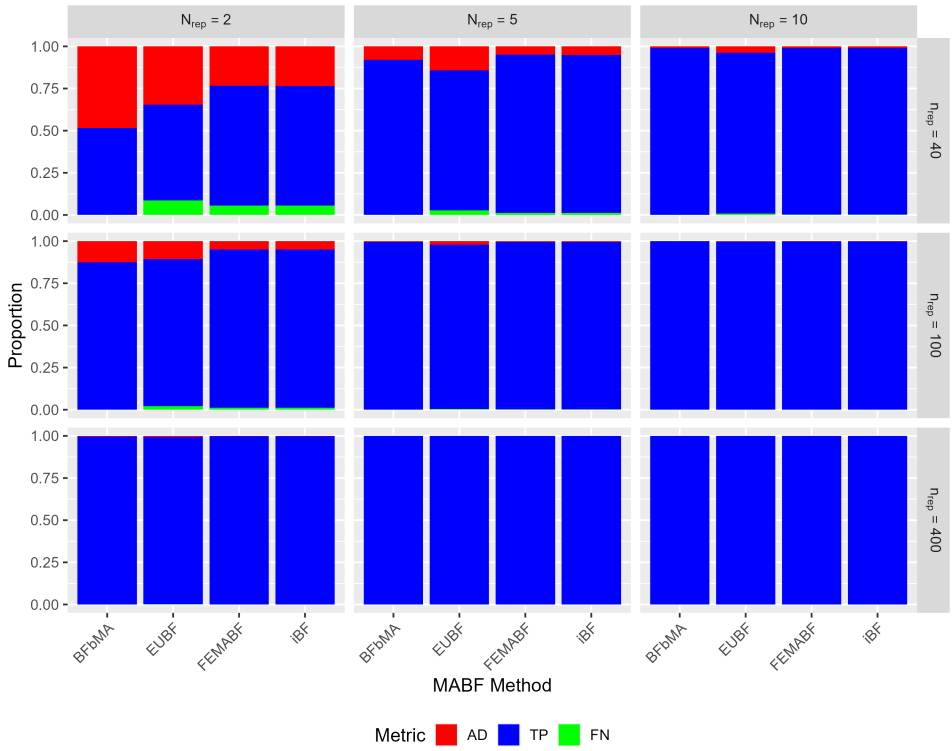

Figure S49 Proportions of anecdotal evidence, true positives, and false negatives produced by the MABF methods (two-sided test) across different replication designs when $\theta = 0.5$.

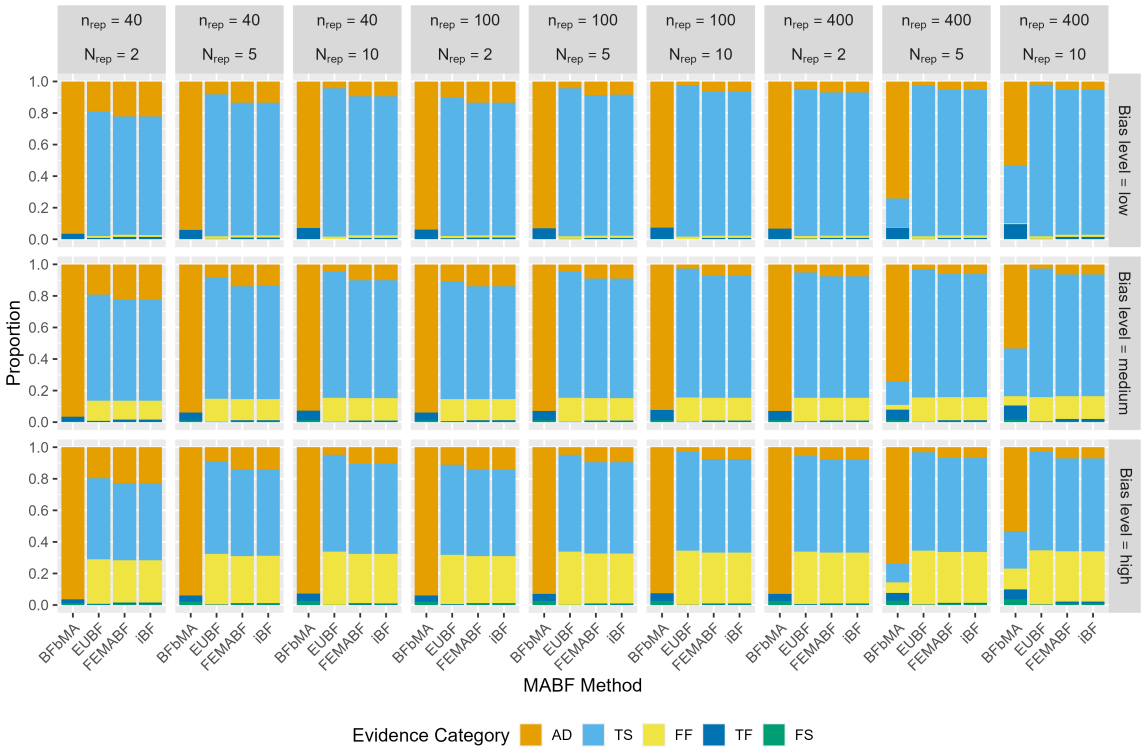

Figure S50 Proportions of anecdotal evidence, true success, false failure, true failure, and false success produced by the MABF methods (one-sided test) across different replication designs separated by levels of biasing mechanisms when $\theta = 0$.

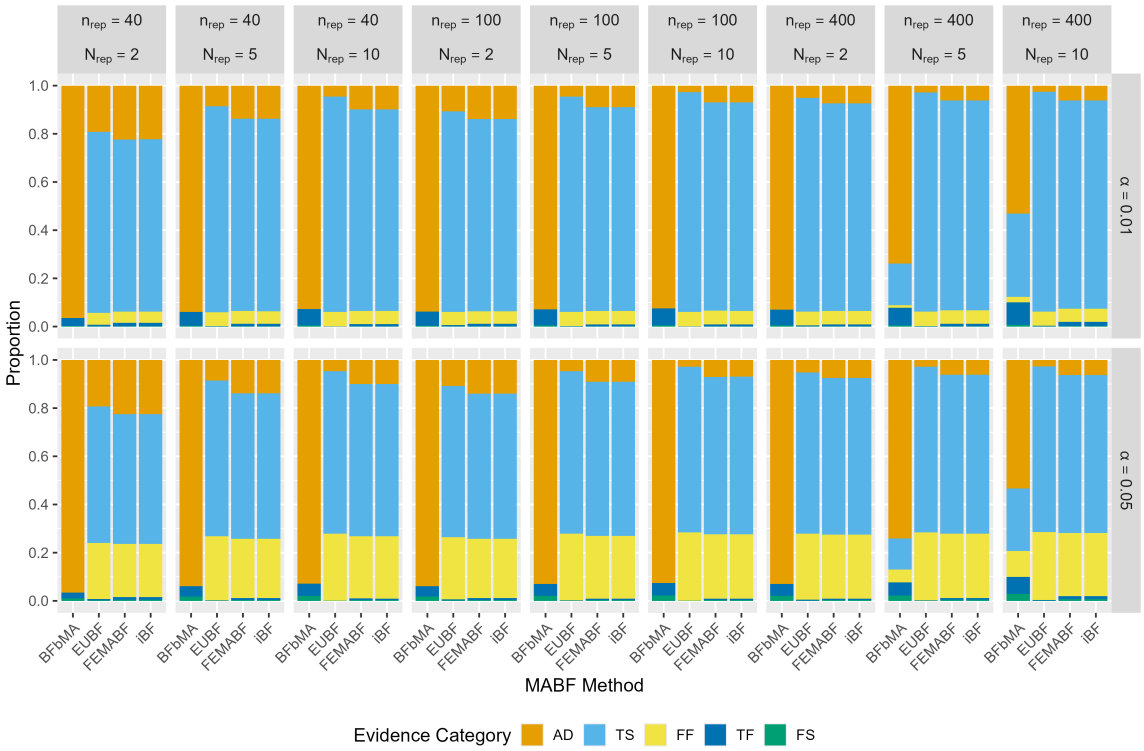

Figure S51 Proportions of anecdotal evidence, true success, false failure, true failure, and false success produced by the MABF methods (one-sided test) across different replication designs separated by original study nominal significance levels when $\theta = 0$.

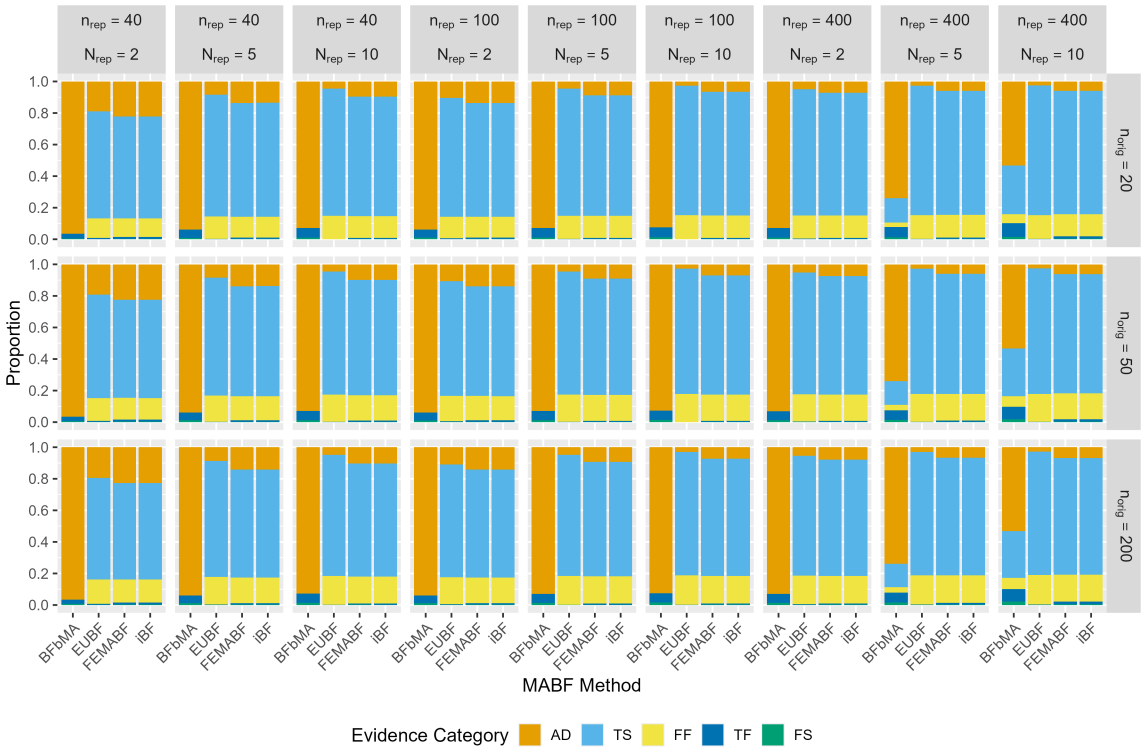

Figure S52 Proportions of anecdotal evidence, true success, false failure, true failure, and false success produced by the MABF methods (one-sided test) across different replication designs separated by original study sample sizes when $\theta = 0$.

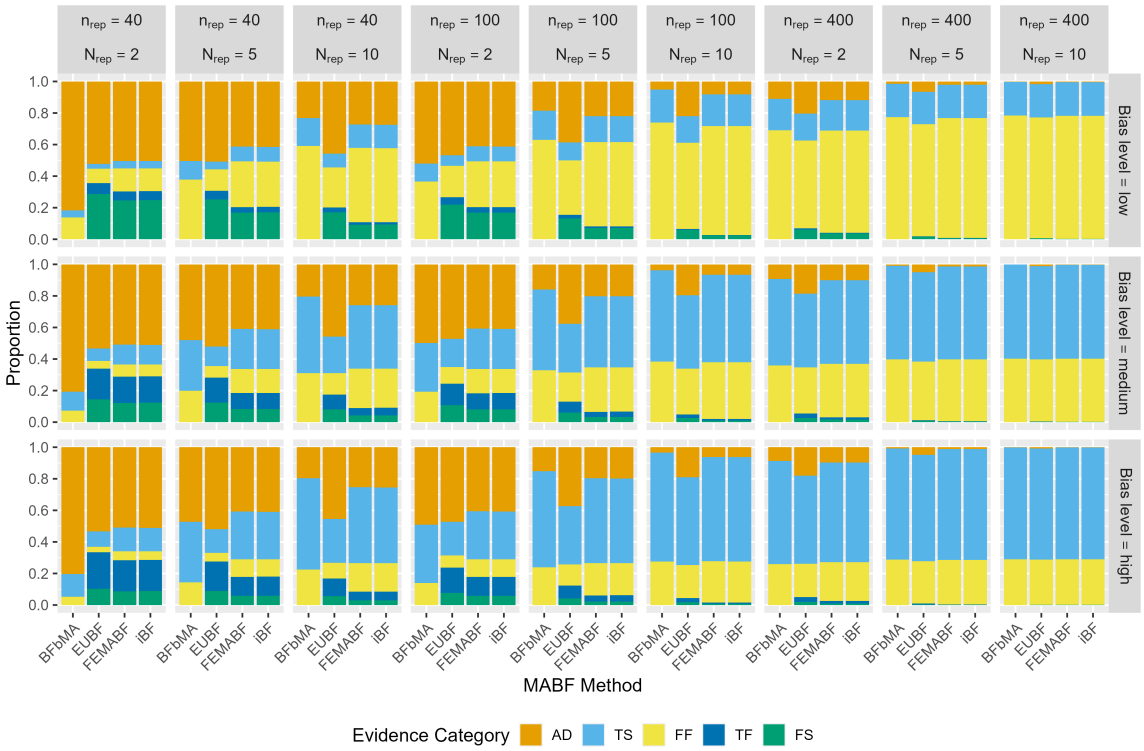

Figure S53 Proportions of anecdotal evidence, true success, false failure, true failure, and false success produced by the MABF methods (one-sided test) across different replication designs separated by levels of biasing mechanisms when $\theta = 0.2$.

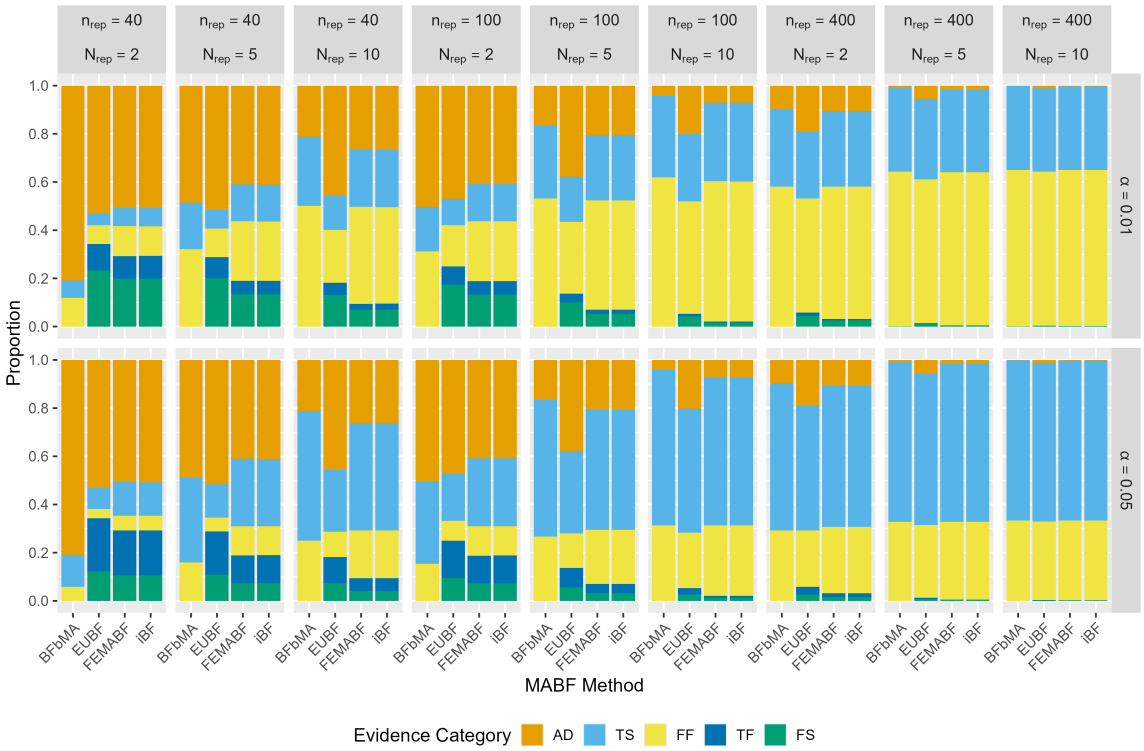

Figure S54 Proportions of anecdotal evidence, true success, false failure, true failure, and false success produced by the MABF methods (one-sided test) across different replication designs separated by original study significance levels when $\theta = 0.2$.

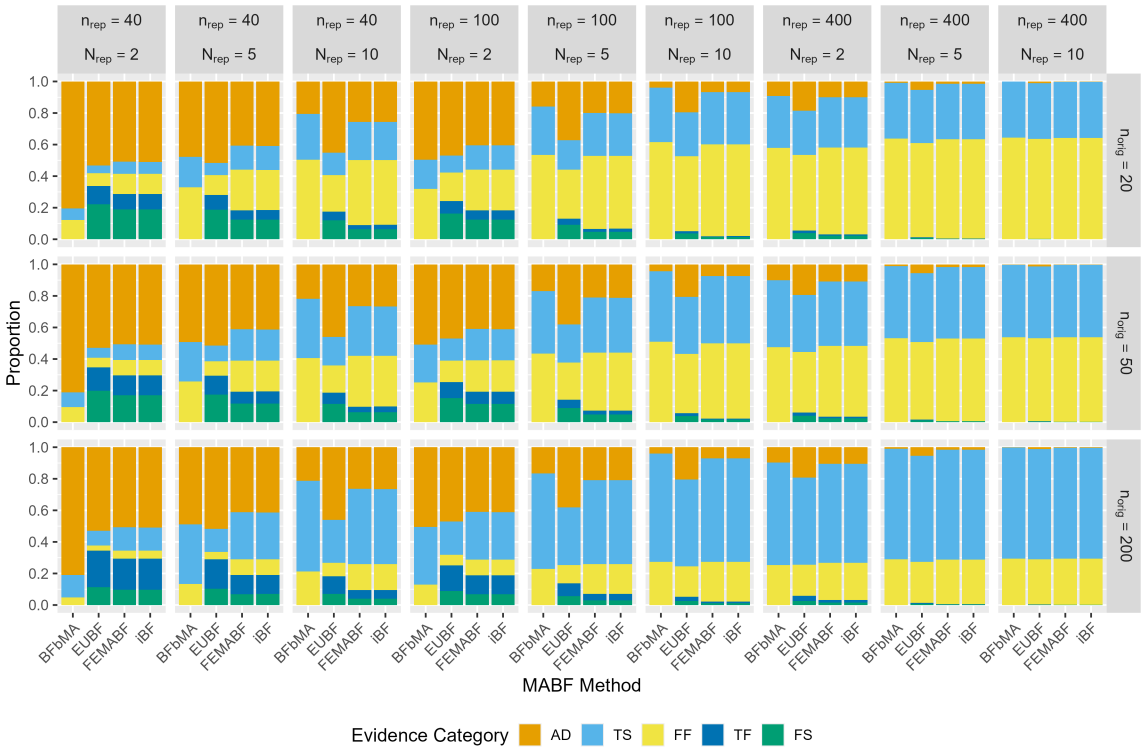

Figure S55 Proportions of anecdotal evidence, true success, false failure, true failure, and false success produced by the MABF methods (one-sided test) across different replication designs separated by original study sample sizes when $\theta = 0.2$.

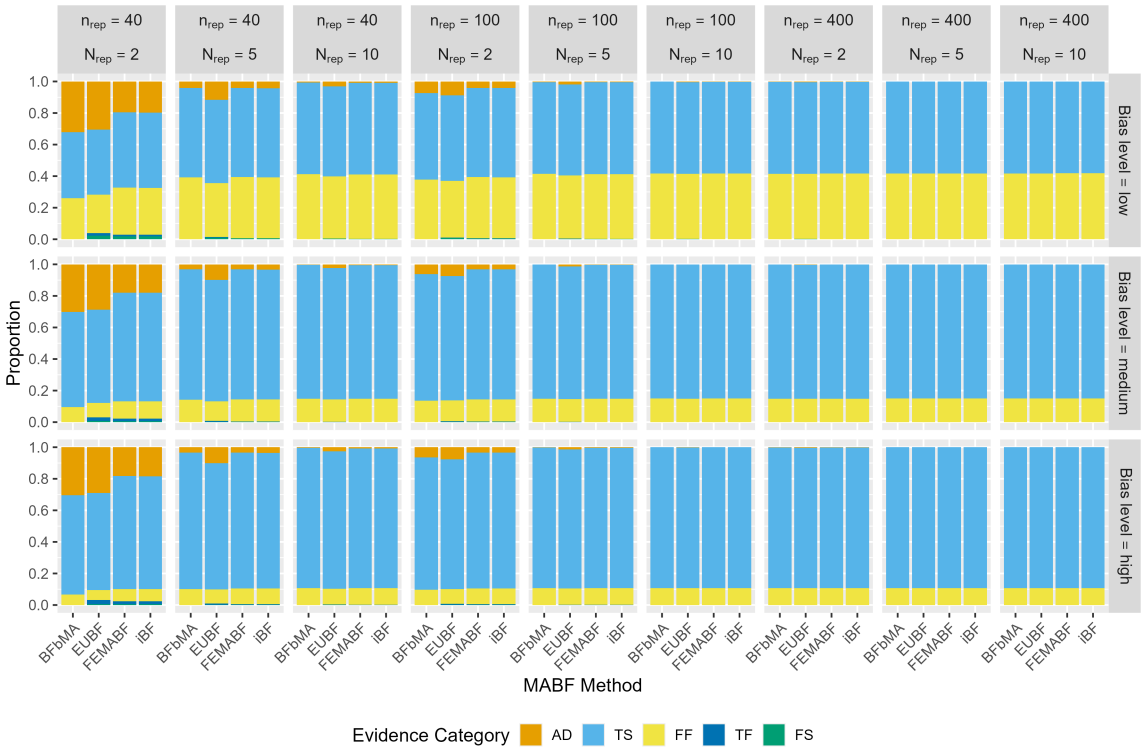

Figure S56 Proportions of anecdotal evidence, true success, false failure, true failure, and false success produced by the MABF methods (one-sided test) across different replication designs separated by levels of biasing mechanisms when $\theta = 0.5$.

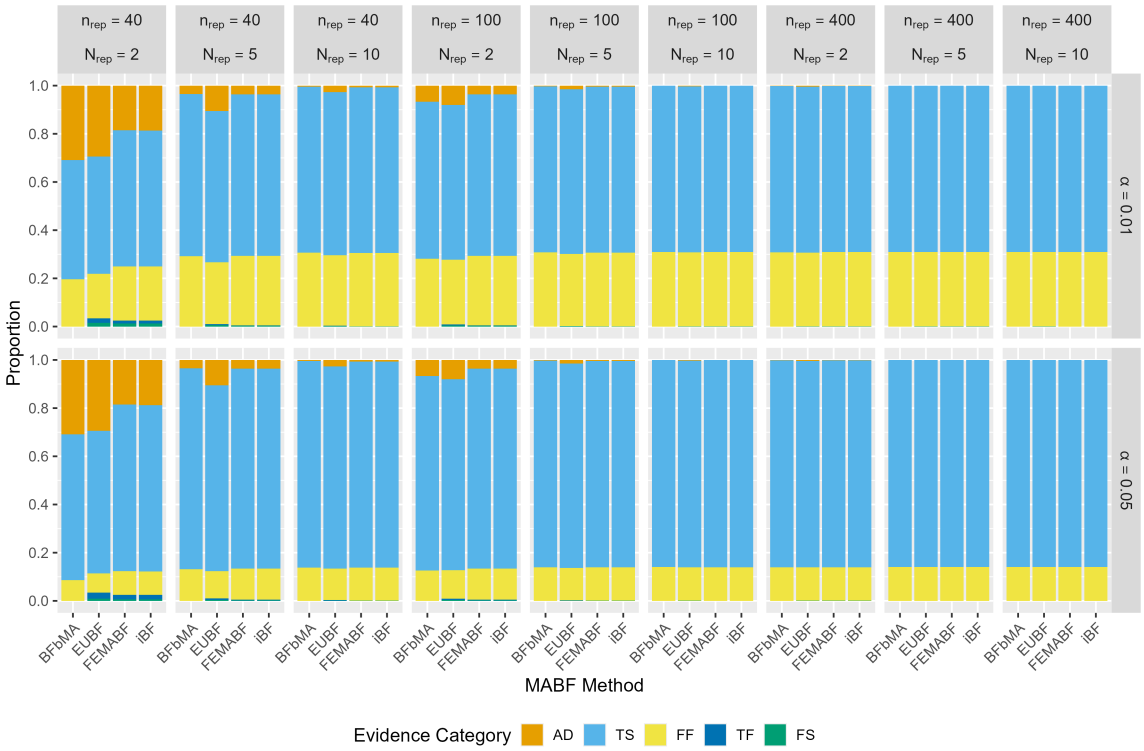

Figure S57 Proportions of anecdotal evidence, true success, false failure, true failure, and false success produced by the MABF methods (one-sided test) across different replication designs separated by original study significance levels when $\theta = 0.5$.

Figure S58 Proportions of anecdotal evidence, true success, false failure, true failure, and false success produced by the MABF methods (one-sided test) across different replication designs separated by original study sample sizes when $\theta = 0.5$.

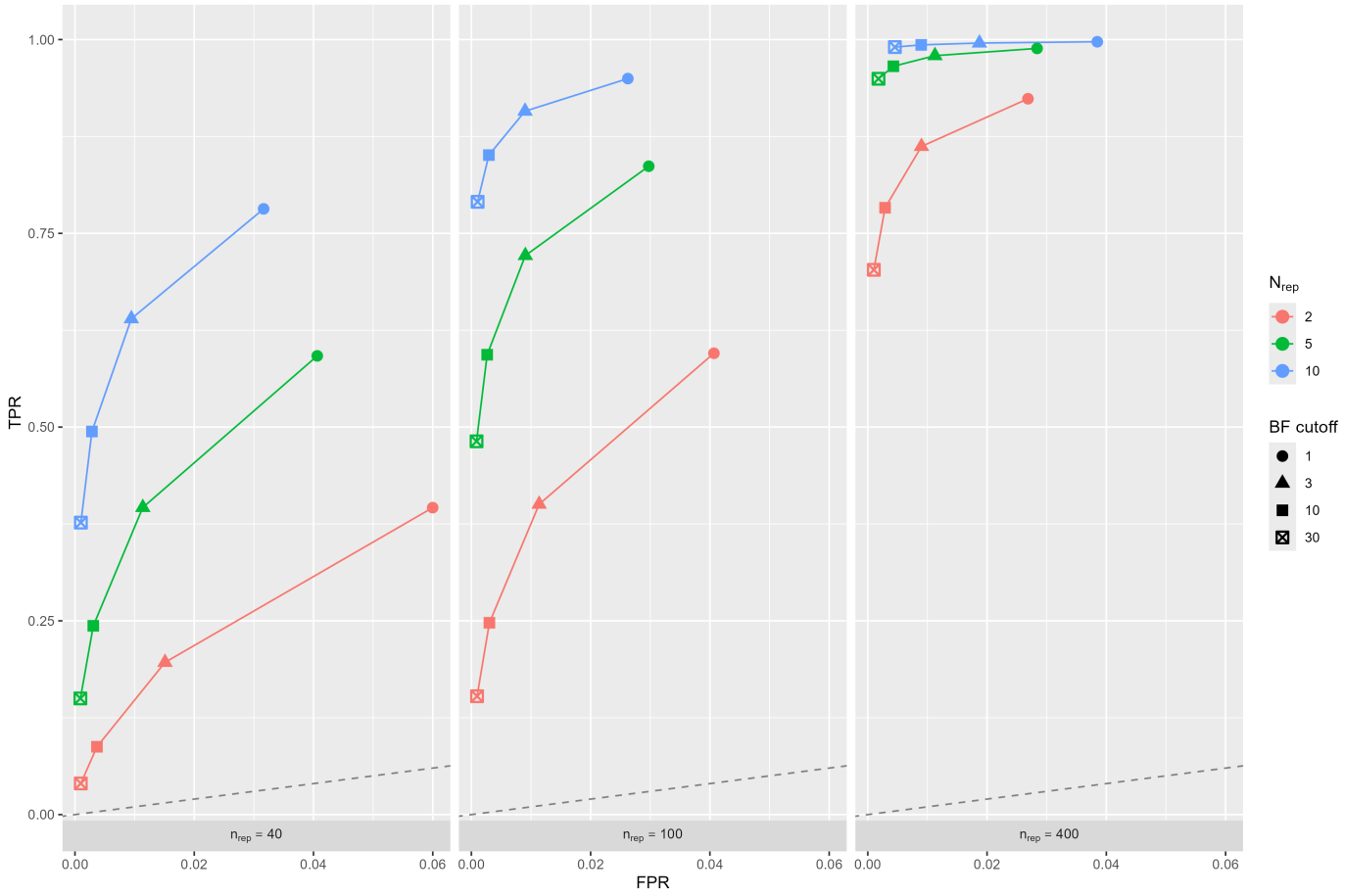

Figure S59 True positive rates (TPR) plotted against false positive rates (FPR) for the inclusion Bayes factor (iBF) method across varying replication sample sizes ($n_{rep} = 40, 100, 400$). Each panel corresponds to a replication sample size. The number of replications ($N_{rep} = 2, 5, 10$) is distinguished by line color (red, green, blue), and Bayes factor cutoffs (1, 3, 10, 30) are distinguished by point symbols. The dashed line represents random classification, where the true positive rate equals the false positive rate.

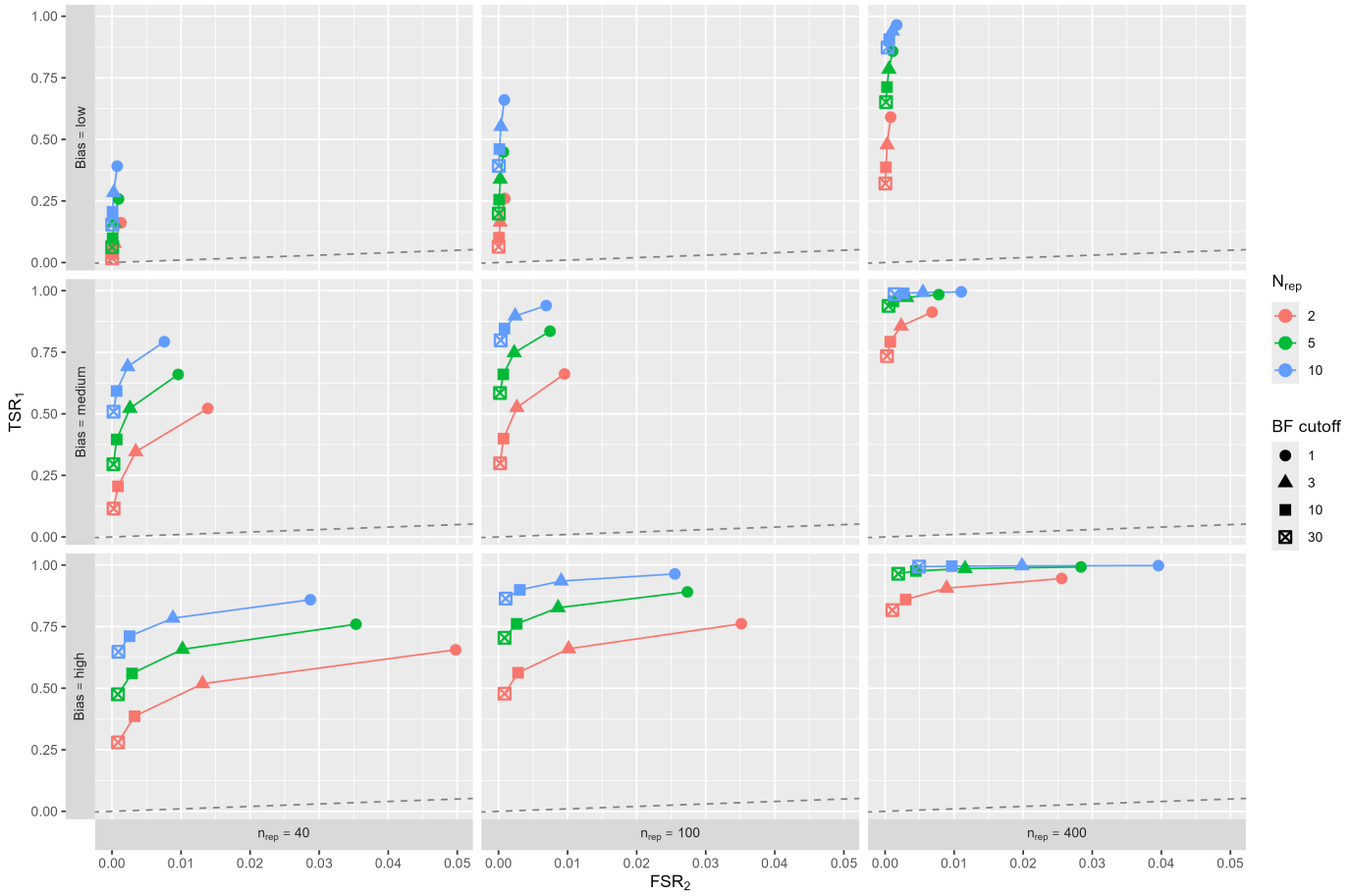

Figure S60 True success rates (TSR_1) plotted against false success rates (FSR_2) for the one-sided iBF method across different bias mechanism levels (low, medium, high) when $\theta = 0.2$. Each row corresponds to a bias mechanism level, and each column corresponds to a replication sample size ($n_{rep} = 40, 100, 400$). The number of replications ($N_{rep} = 2, 5, 10$) is distinguished by line color (red, green, blue), and Bayes factor cutoffs (1, 3, 10, 30) are distinguished by point symbols. The dashed line represents random classification, where the true success rate equals the false success rate.

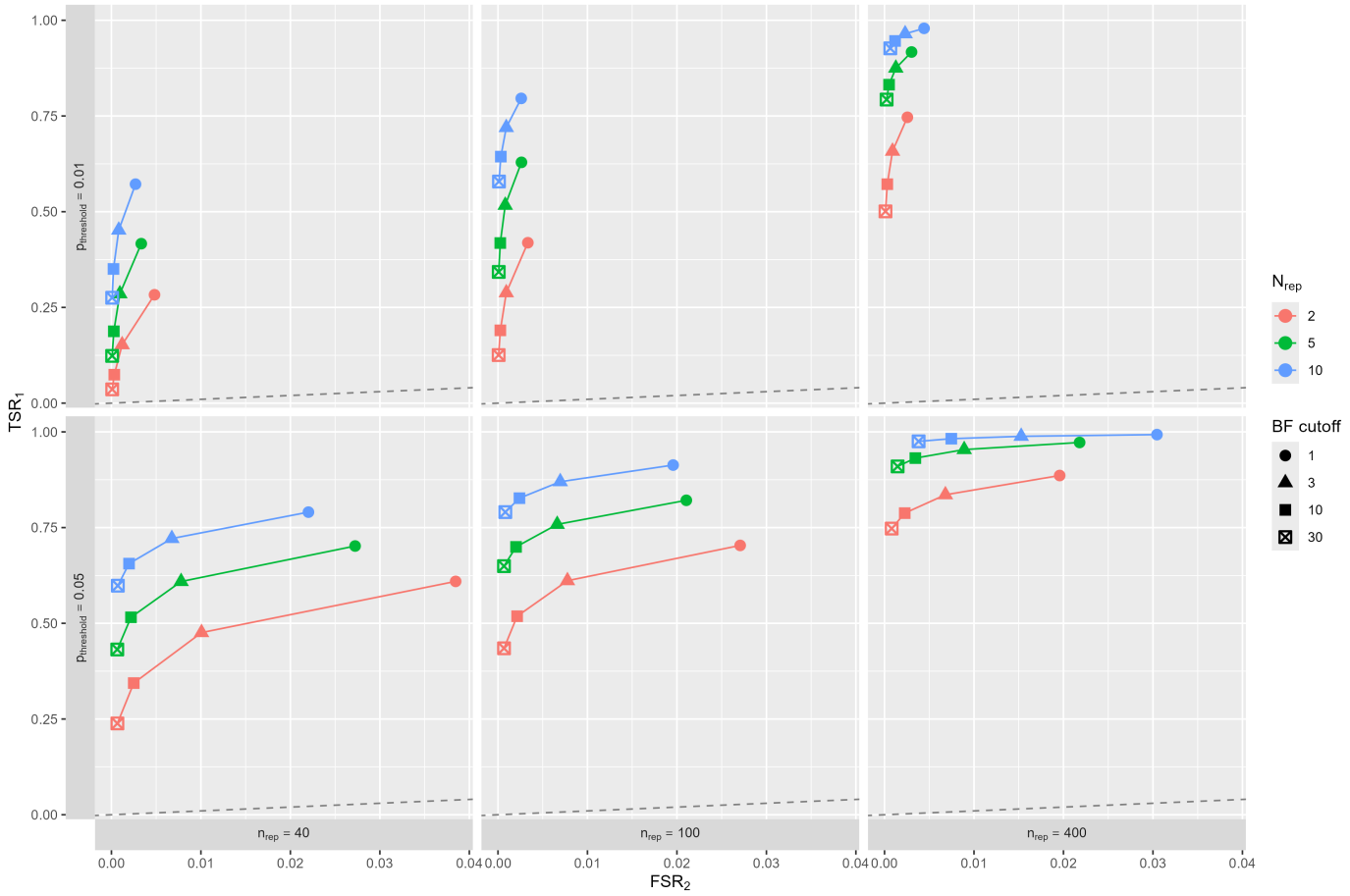

Figure S61 True success rates (TSR_1) plotted against false success rates (FSR_2) for the one-sided iBF method $\theta = 0.2$ under different original study's nominal significance levels ($\alpha = .01$ and .05). Each row corresponds to a nominal significance level, and each column corresponds to a replication sample size ($n_{rep} = 40, 100, 400$). The number of replications ($N_{rep} = 2, 5, 10$) is distinguished by line color (red, green, blue), and Bayes factor cutoffs (1, 3, 10, 30) are distinguished by point symbols. The dashed line represents random classification, where the true success rate equals the false success rate.

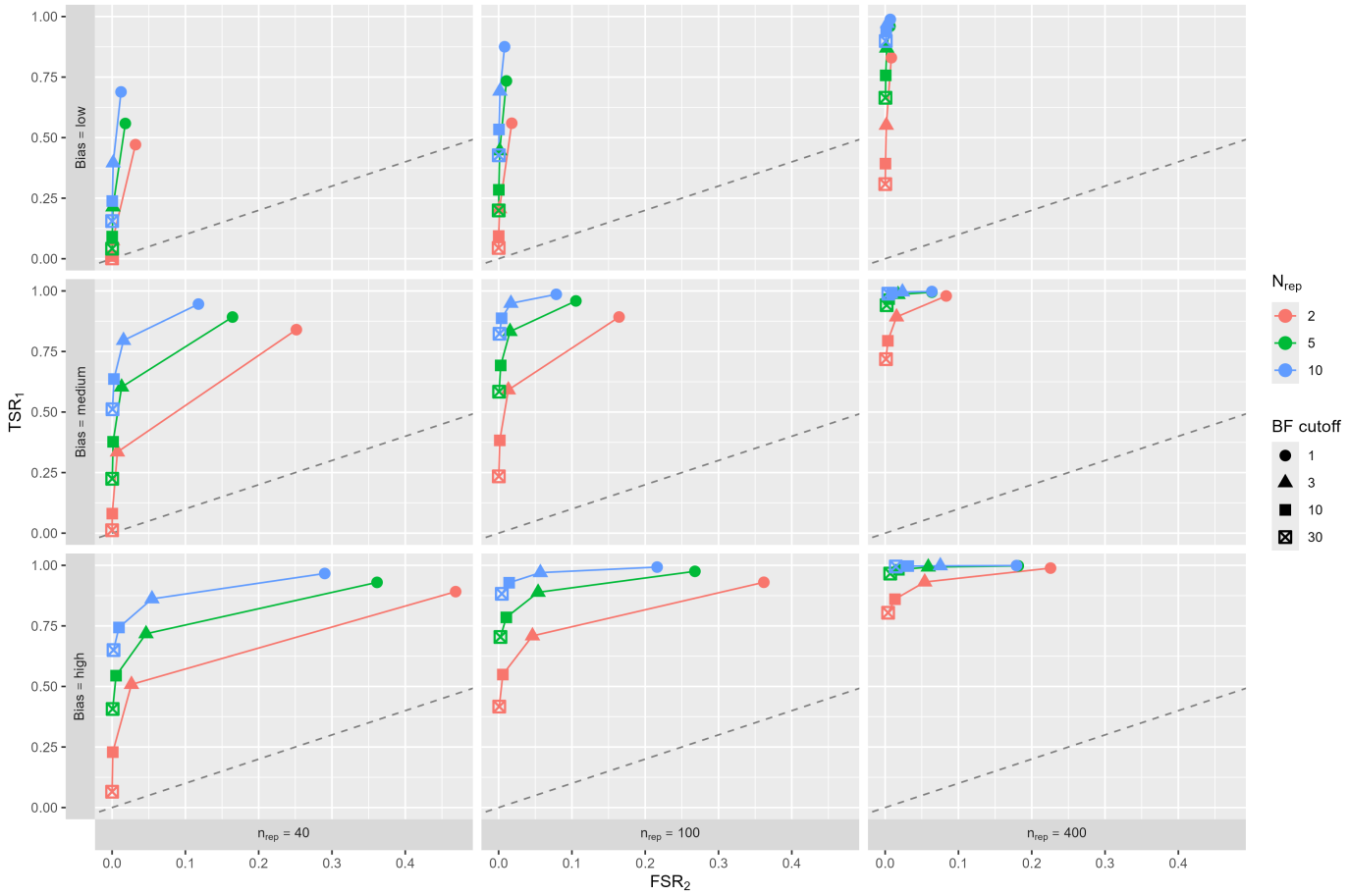

Figure S62 True success rates (TSR_1) plotted against false success rates (FSR_2) for the Bayes factor based on meta-analysis (BFbMA) method when $\theta = 0.2$ under different bias mechanism levels (low, medium, high) across varying replication sample sizes, numbers of replications, and Bayes factor cutoffs.