Supplementary Figures

Figure S1: Type I error rates by fixed expected minor allele count (MAC) for different sample sizes.
Analytically calculated type I error rates ($\alpha = 5\times10^{-8}$) for joint analysis in: balanced studies (A - C), unbalanced studies (D - F), and very unbalanced studies (G - I). Variant allele frequencies are selected so that the expected MAC remains constant across studies with total sample size N = 2000, 20000 and 50000 individuals respectively. The horizontal dotted line denotes the corresponding nominal significance threshold ($\alpha = 5\times10^{-8}$).
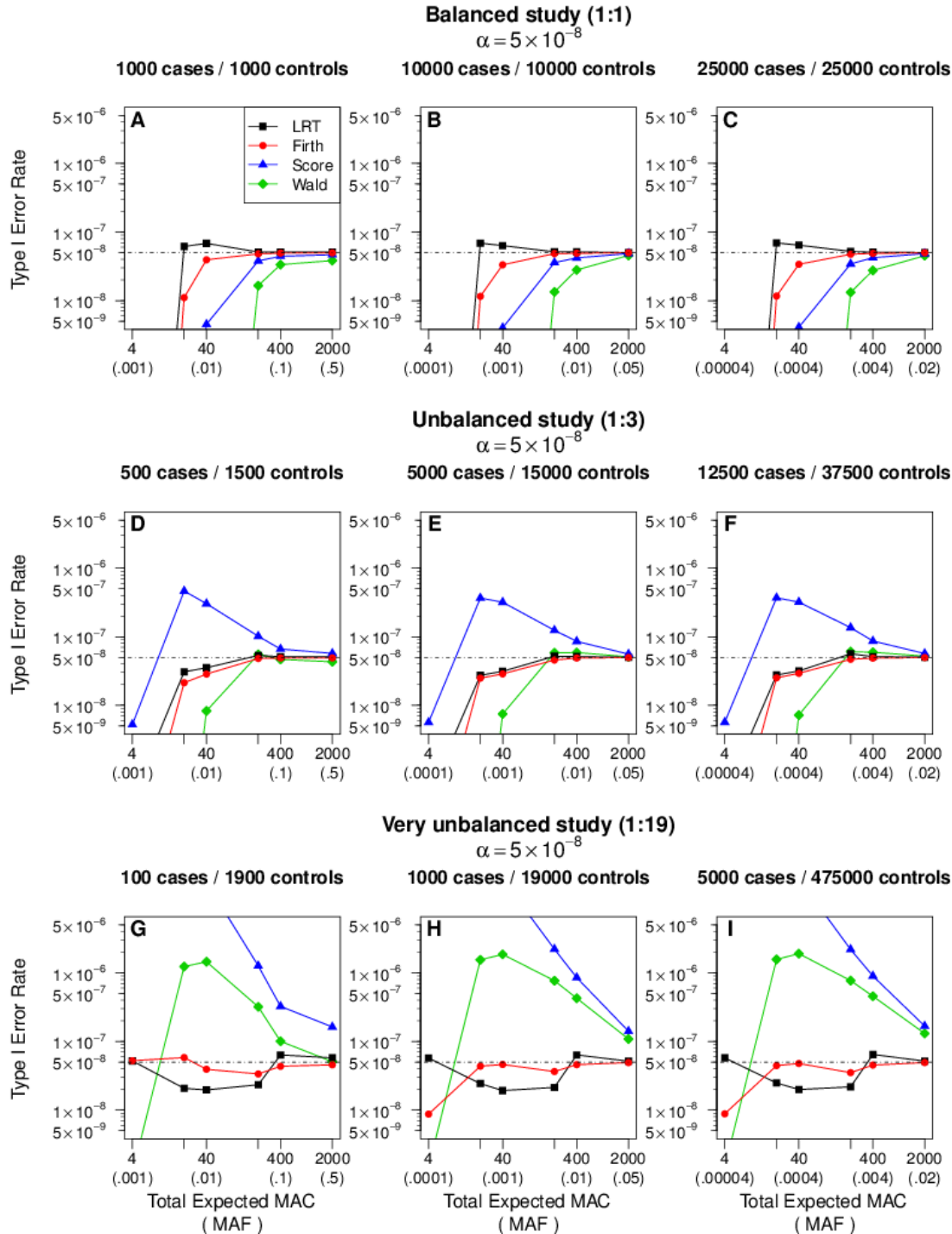
Figure S2: Meta-analysis type I error rates by sample size for fixed expected minor allele count (MAC). Simulation-based sample-size weighted meta-analysis type I error rates ($\alpha = 5\times10^{-4}$) for balanced (case-control ratio 1:1), unbalanced (1:3), and very unbalanced studies (1:19) with of various sample sizes. For each study, variant allele frequencies are selected so that the expected MAC = 2000 (A - C), 400 (D - F), or 40 (G - I). The horizontal dotted line denotes the corresponding nominal significance threshold ($\alpha = 5\times10^{-4}$). Very conservative or anti-conservative tests with type I error rates that exceed the vertical axis limits are not displayed.
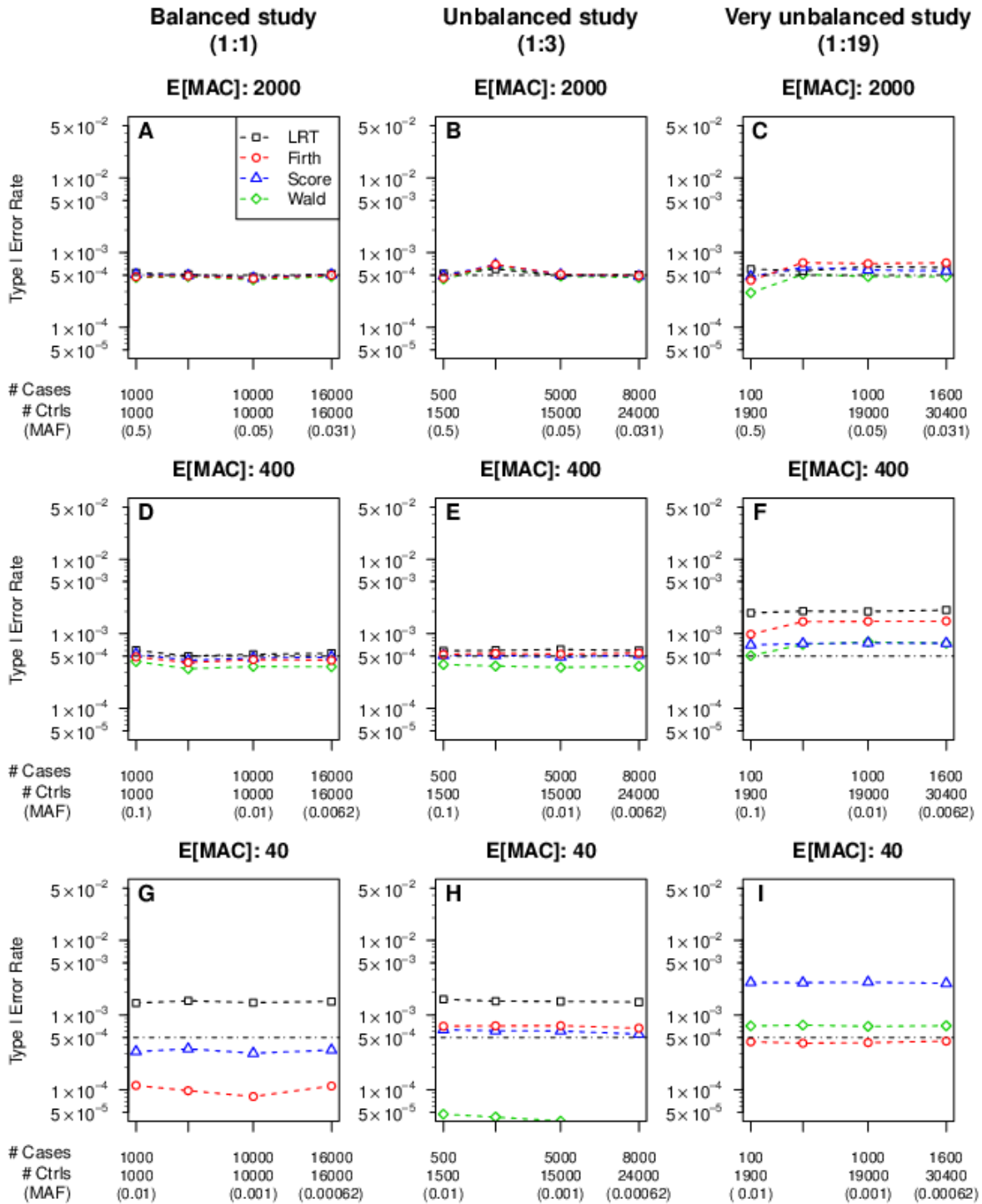
Figure S3: Comparison of score and Firth test association p-values in the GoT2D study. For different minor allele count (MAC) categories, comparison of score and Firth test-based (A-D) joint analysis p-values and (E-H) meta-analysis p-values.
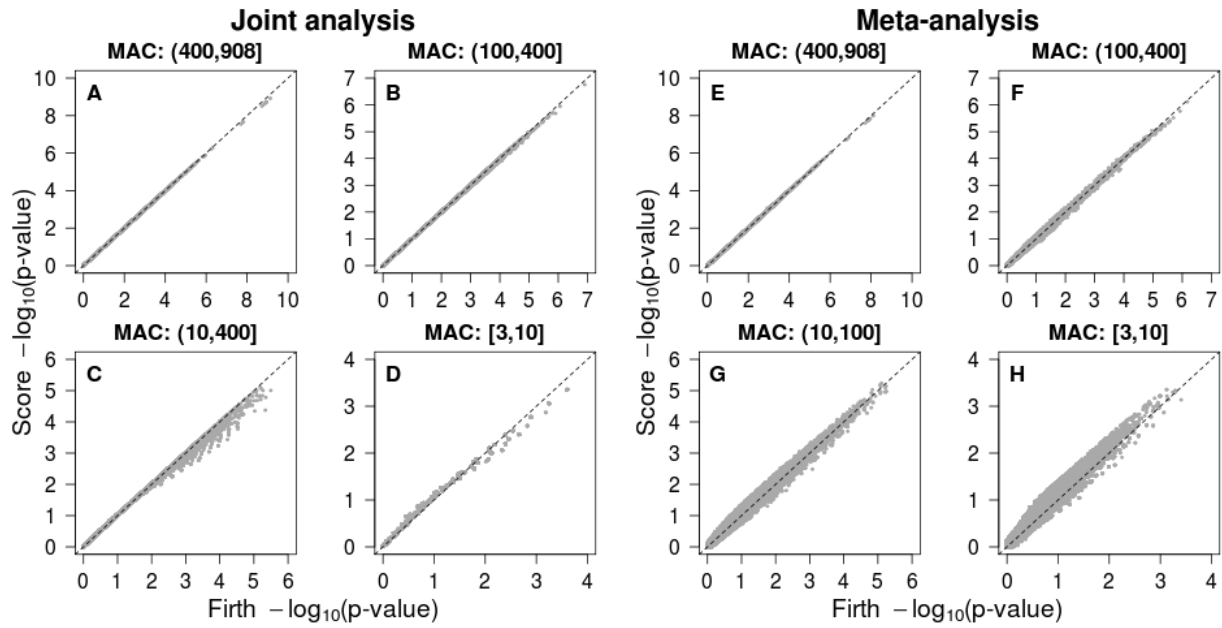
Figure S4: Comparison of joint and meta-analysis p-values in the GoT2D study. For different minor allele count (MAC) categories, comparison of joint and meta-analysis p-values using the (A-D) Firth test and (E-H) score test.
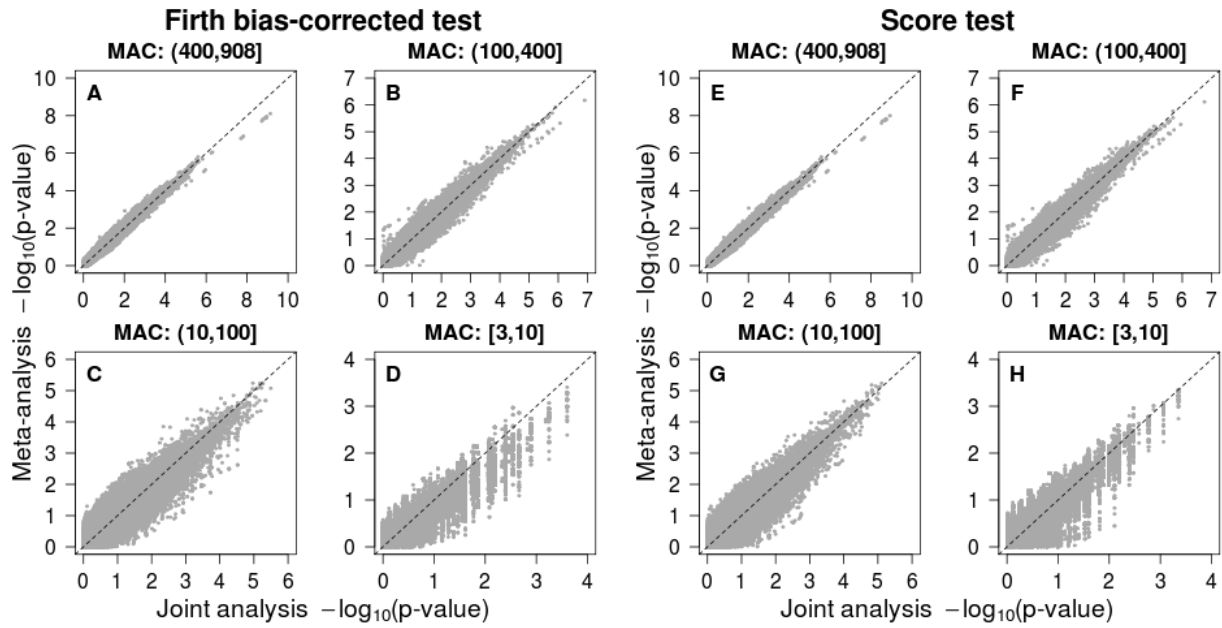
Figure S5: Score test type I error rate and power with study-level minor allele count (MAC) filters. (A) Empirical type I error rates ($\alpha = 5\times10^{-5}$) for score test-based joint and sample-size weighted meta-analysis, with varying degrees of study-level MAC filters. Type I error rates for joint analysis are estimated for studies with 10000/10000 total cases and controls; meta-analysis is based on partitioning the full dataset into 10 equal-sized sub-studies. The horizontal dotted line denotes the corresponding nominal significance threshold. (B - C) Simulated power at $\alpha = 5\times10^{-8}$ for a variant with: expected MAC = 40 (MAF = 0.001); and E[MAC] = 20 (MAF = 0.0005), for the same study design.
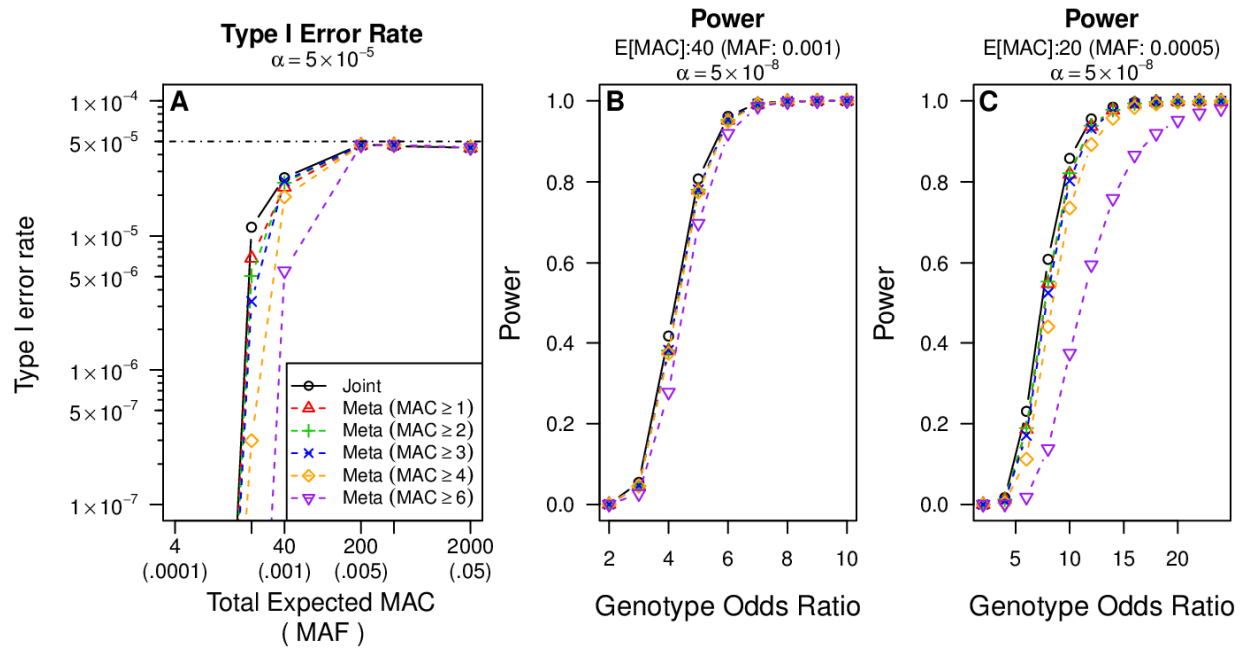
Figure S6: **Score test type I error rate and power curves for meta-analysis of *K* = 10 and 50 sub-studies.** (A) Empirical type I error rates ($\alpha = 5\times10^{-5}$) for score test-based joint analysis with 10000/10000 total cases and controls (black); sample-size weighted meta-analysis with *K* = 10 sub-studies of 1000/1000 cases and controls (red); and *K* = 50 sub-studies of 200/200 cases and controls (green). (B) Simulated power ($\alpha = 5\times10^{-8}$) for a variant with expected minor allele count = 40 (MAF = 0.001) for the same study design.
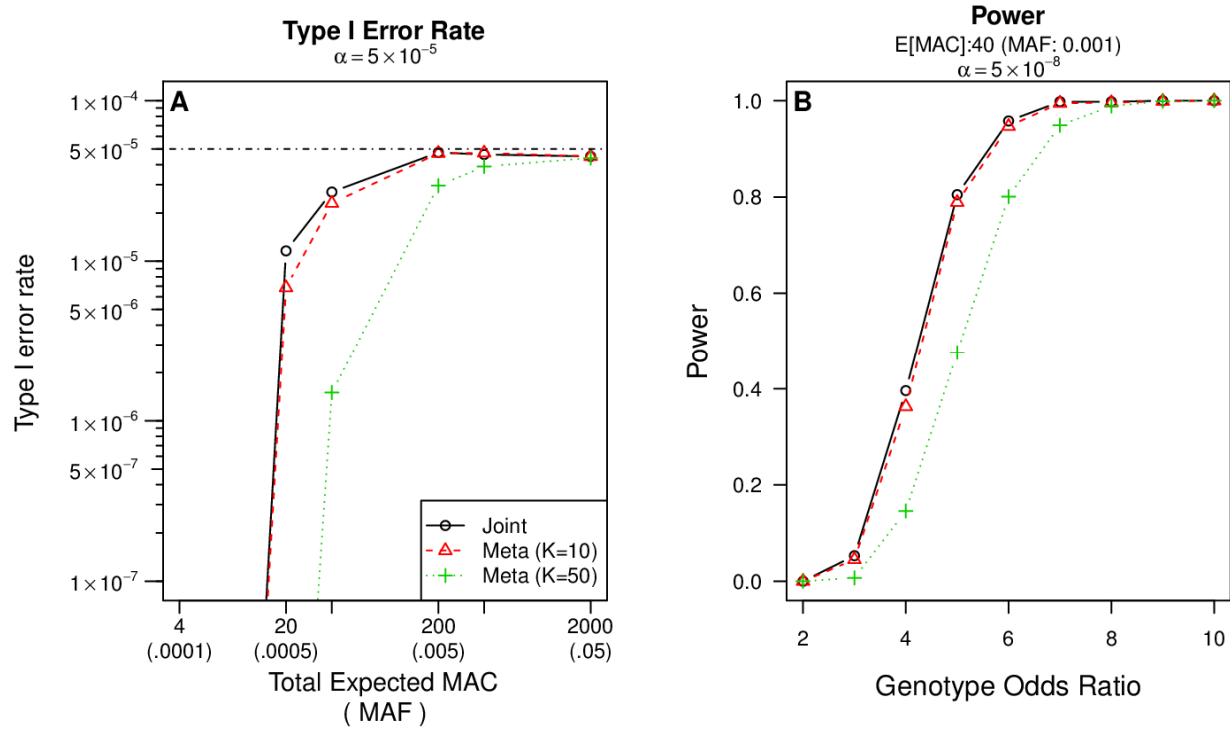
Figure S7: Type I error rates by minor allele count (MAC) for logistic regression tests and Fisher's exact test in joint and meta-analysis. (A - C) Analytically calculated type I error rates ($\alpha = 5\times10^{-8}$) for joint analysis; (D - F) empirical type I error rates ($\alpha = 5\times10^{-5}$) for joint analysis; and (G - I) empirical type I error rates ($\alpha = 5\times10^{-5}$) for sample-size weighted (dashed) and inverse-variance weighted (dotted) meta-analysis. Type I error rates for joint analysis are estimated for studies with 10000/10000, 5000/15000 and 1000/19000 total cases and controls; meta-analysis is based on partitioning the full dataset into 10 equal-sized sub-studies. The horizontal dotted line denotes the corresponding nominal significance threshold.
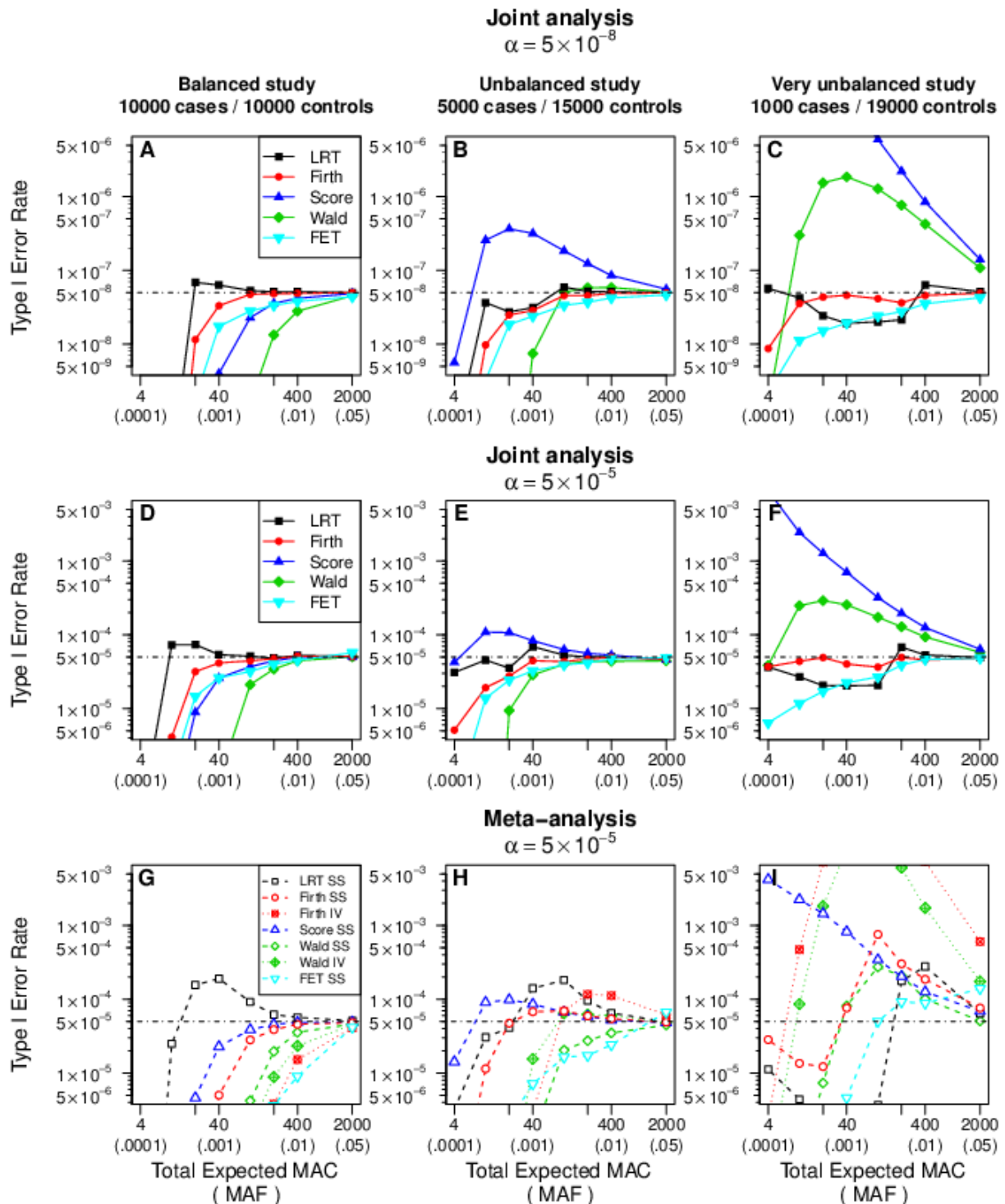
Figure S8: Type I error rates by case-control ratio for logistic regression and Fisher's exact tests in joint and meta-analysis. (A, B) Analytically calculated type I error rates (α = 5x10^{-8}) for joint analysis; (C, D) empirical type I error rates (α = 5x10^{-5}) for joint analysis; and (E, F) empirical type I error rates (α = 5x10^{-5}) for sample-size weighted (dashed) and inverse-variance weighted (dotted) meta-analysis. Type I error rates are estimated for a high count (expected MAC = 2000; MAF = 0.05), and low count (E[MAC] = 40; MAF = 0.001) variant, in studies with *N* = 20000 individuals with varying case-control ratios. The horizontal dotted line denotes the corresponding nominal significance threshold.
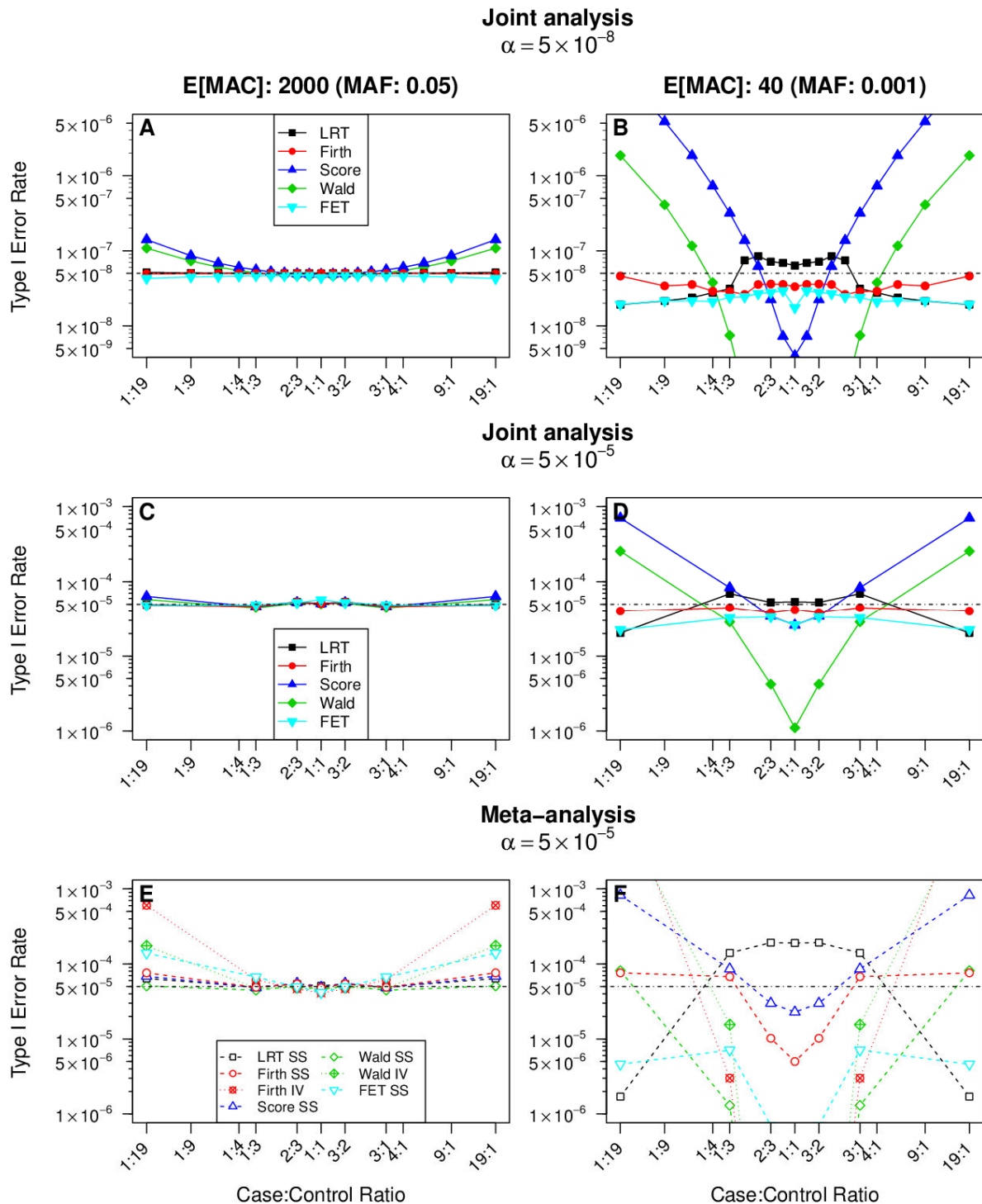
Figure S9: Simulated power curves for joint and meta-analysis. Simulated power (α = 5×10^{-8}) in joint analysis (solid), sample-size weighted (dashed) and inverse-variance weighted (dotted) meta-analysis for a variant with: (A - C) expected MAC = 2000 (MAF = 0.05); (D - F) expected MAC = 400 (MAF = 0.01); and (G - I) expected MAC = 40 (MAF = 0.001). Power for joint analysis is estimated for studies with 10000/10000, 5000/15000, and 1000/19000 total cases and controls; meta-analysis is based on partitioning the full dataset into 10 equal-sized sub-studies.