

**The American Journal of Human Genetics, Volume 92
Supplemental Data**

**Genome-wide Association Analysis
for Multiple Continuous Secondary Phenotypes**

Elizabeth D. Schifano, Lin Li, David C. Christiani, and Xihong Lin

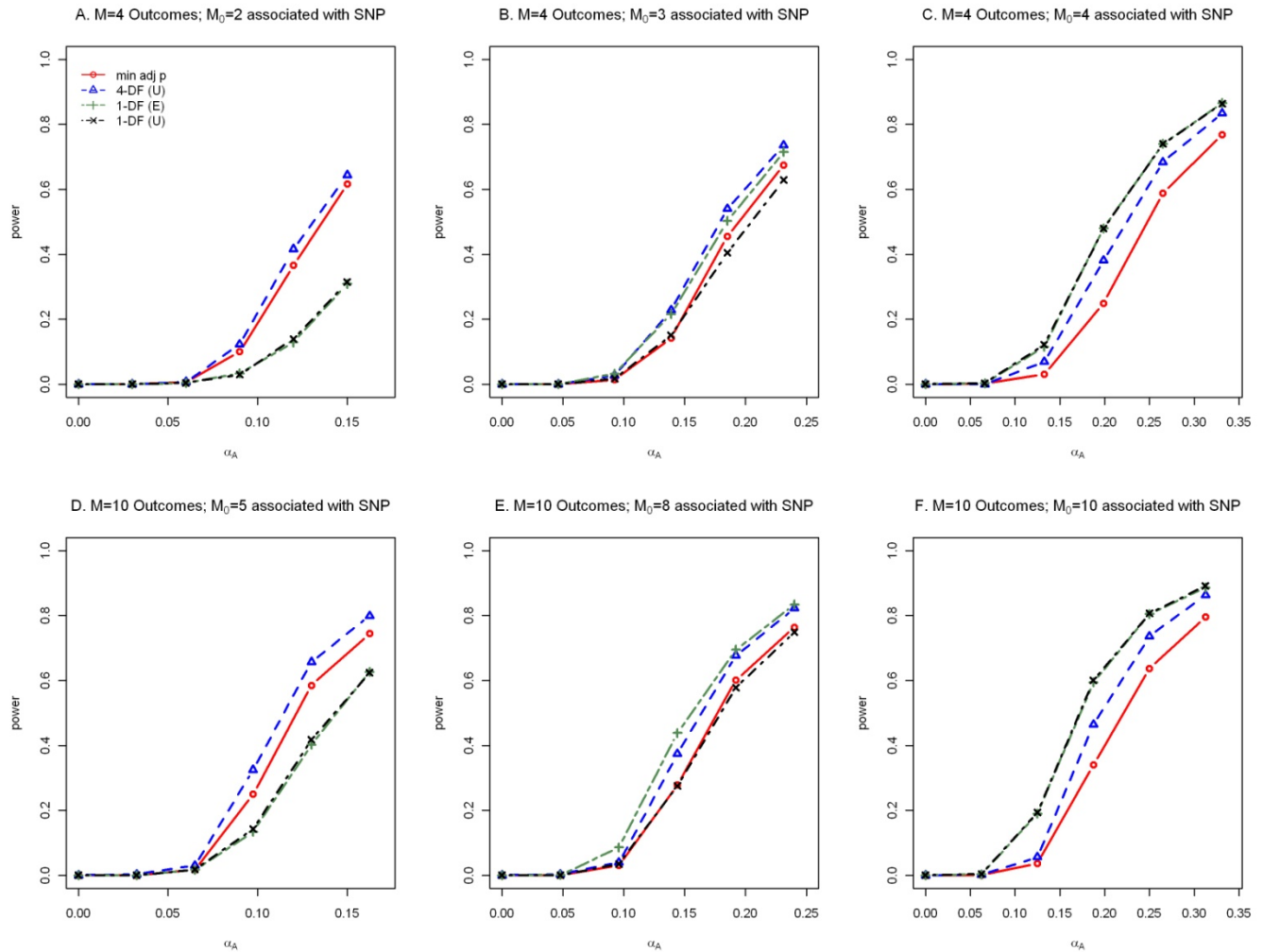


Figure S1. Power for Control + Affected Analysis under the Scaled Heterogeneous Effect Model

Power for Control + Affected analysis under the scaled heterogeneous effect model for $M=4$ phenotypes (A–C) and $M=10$ (D–F) for various numbers of phenotypes associated with SNP (M_0). Three working correlation matrix structures for the joint outcome analyses were considered: I=Independent, E=Exchangeable, U= Unstructured. Power results were nearly identical using I, E, and U for the 4-DF GEE tests, while power results were nearly identical using I and E, but not U, in the 1-DF SMAT; thus only the results for GEE (U) and SMAT (E) and (U) are included.

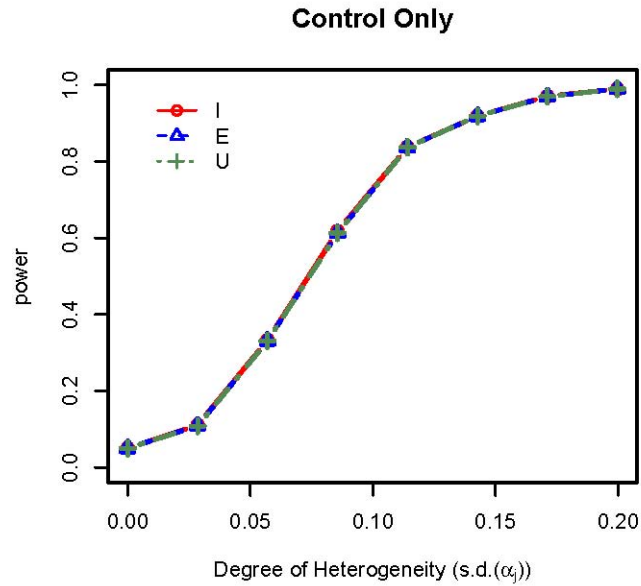


Figure S2. Power Results for Test for Scaled Homogeneity for Control Only Analysis

Power is plotted as a function of increasing effect size heterogeneity across outcomes, as measured by the standard deviation of SNP effect sizes α_j , $j=1 \dots 4$, i.e. $s.d.(\alpha_j)$; see Materials and Methods subsection *Simulation: Empirical Performance of Test for Scaled Homogeneity* for more details. While the data were generated with an unstructured correlation matrix, three working correlation matrix structures for the joint outcome analyses were considered: I=Independent, E=Exchangeable, U=Unstructured; power is similar for all three choices.

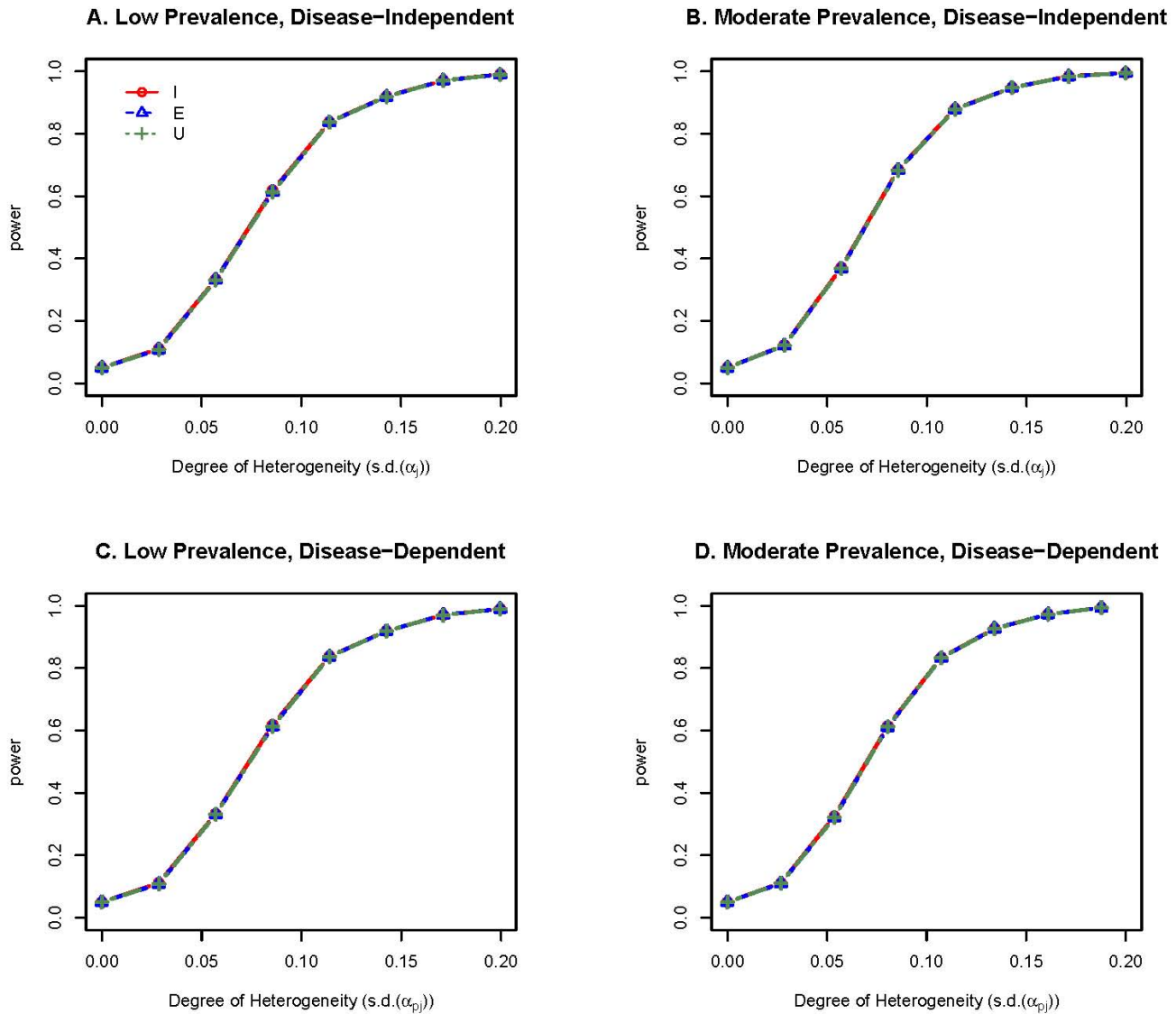


Figure S3. Power Results for Test for Scaled Homogeneity for Control + Affected Analysis

Power is plotted as a function of increasing effect size heterogeneity across outcomes, as measured by the standard deviation of SNP effect sizes α_j (s.d.(α_j)) or α_{pj} (s.d.(α_{pj})), $j=1\dots 4$, for the Disease-Independent and Disease-Dependent scenarios, respectively; see Materials and Methods subsection *Simulation: Empirical Performance of Test for Scaled Homogeneity* for more details. While the data were generated with an unstructured correlation matrix, three working correlation matrix structures for the joint outcome analyses were considered: I=Independent, E=Exchangeable, U= Unstructured; power is similar for all three choices.

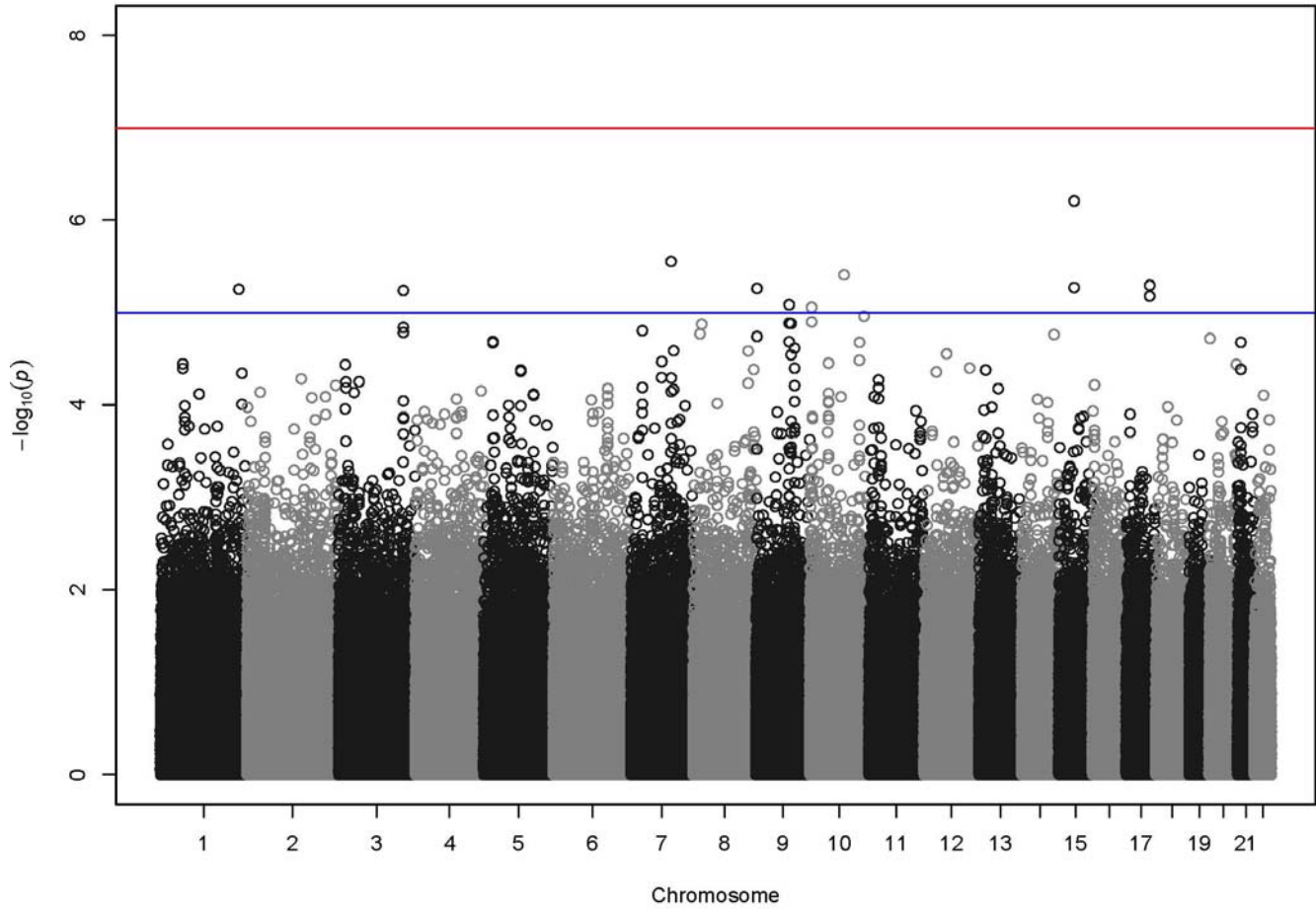


Figure S4. Manhattan Plot for Individual Outcome $\sqrt{\text{DURATION}}$

Analysis was performed on both affected ($n_1 = 696$) and control ($n_0 = 730$) ever-smokers using $\pi = 0.000745$ to determine the weights. p values are unadjusted.

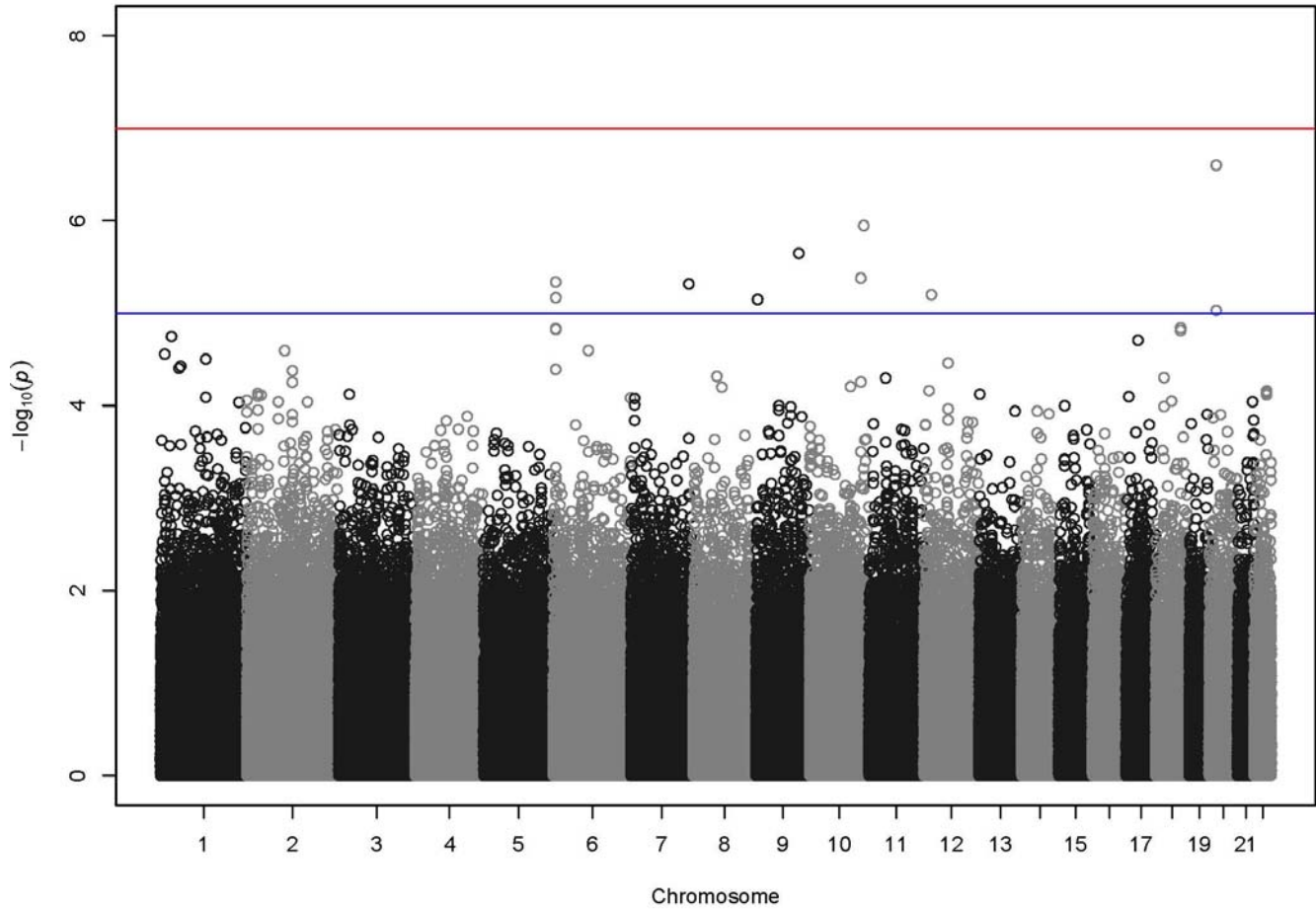


Figure S5. Manhattan Plot for Individual Outcome $\sqrt{\text{INITIATION}}$

Analysis was performed on both affected ($n_1 = 696$) and control ($n_0 = 730$) ever-smokers using $\pi = 0.000745$ to determine the weights. p values are unadjusted.

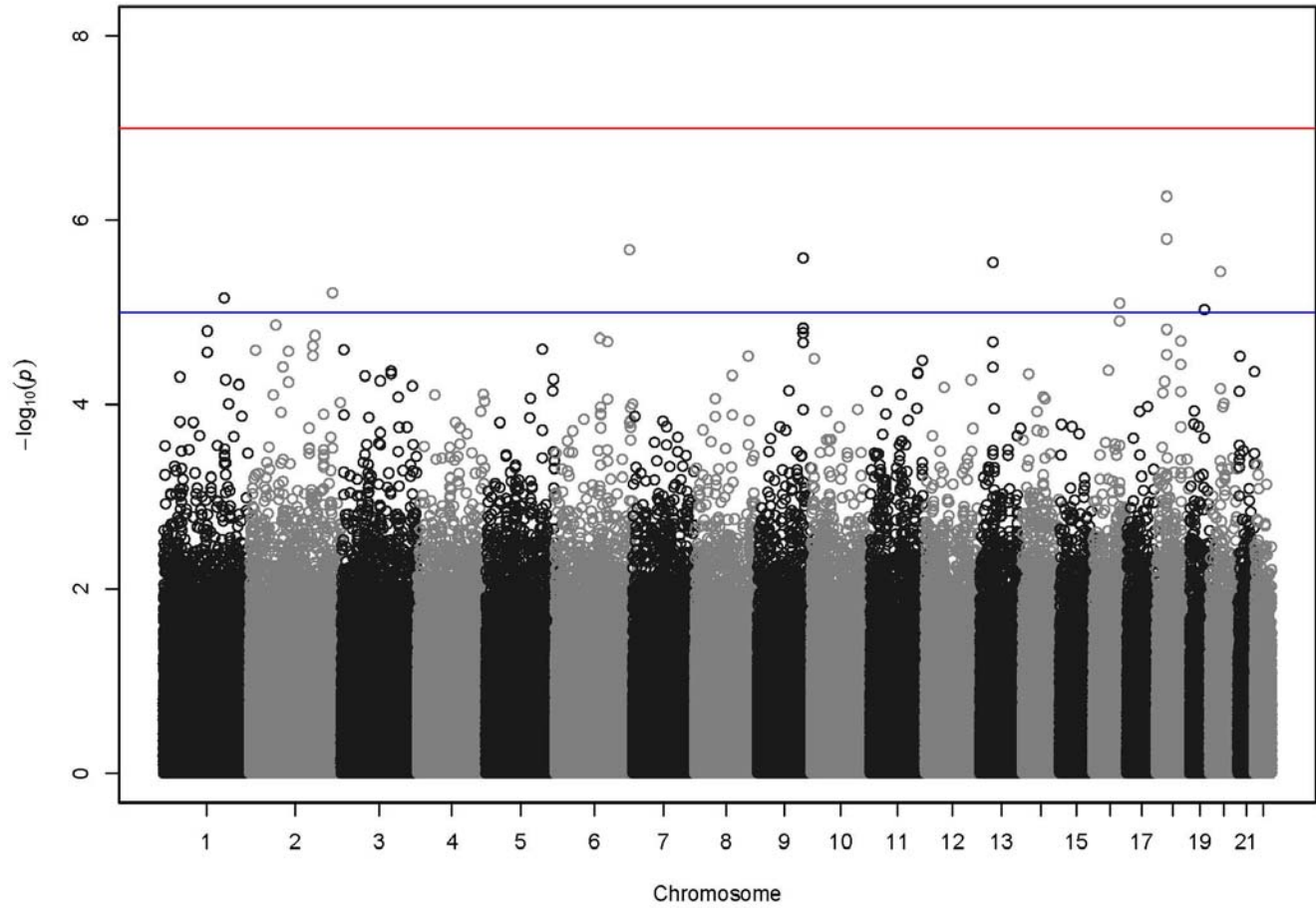


Figure S6. Manhattan Plot for Individual Outcome $\sqrt{\text{CPD}}$

Analysis was performed on both affected ($n_1 = 696$) and control ($n_0 = 730$) ever-smokers using $\pi = 0.000745$ to determine the weights. p values are unadjusted.

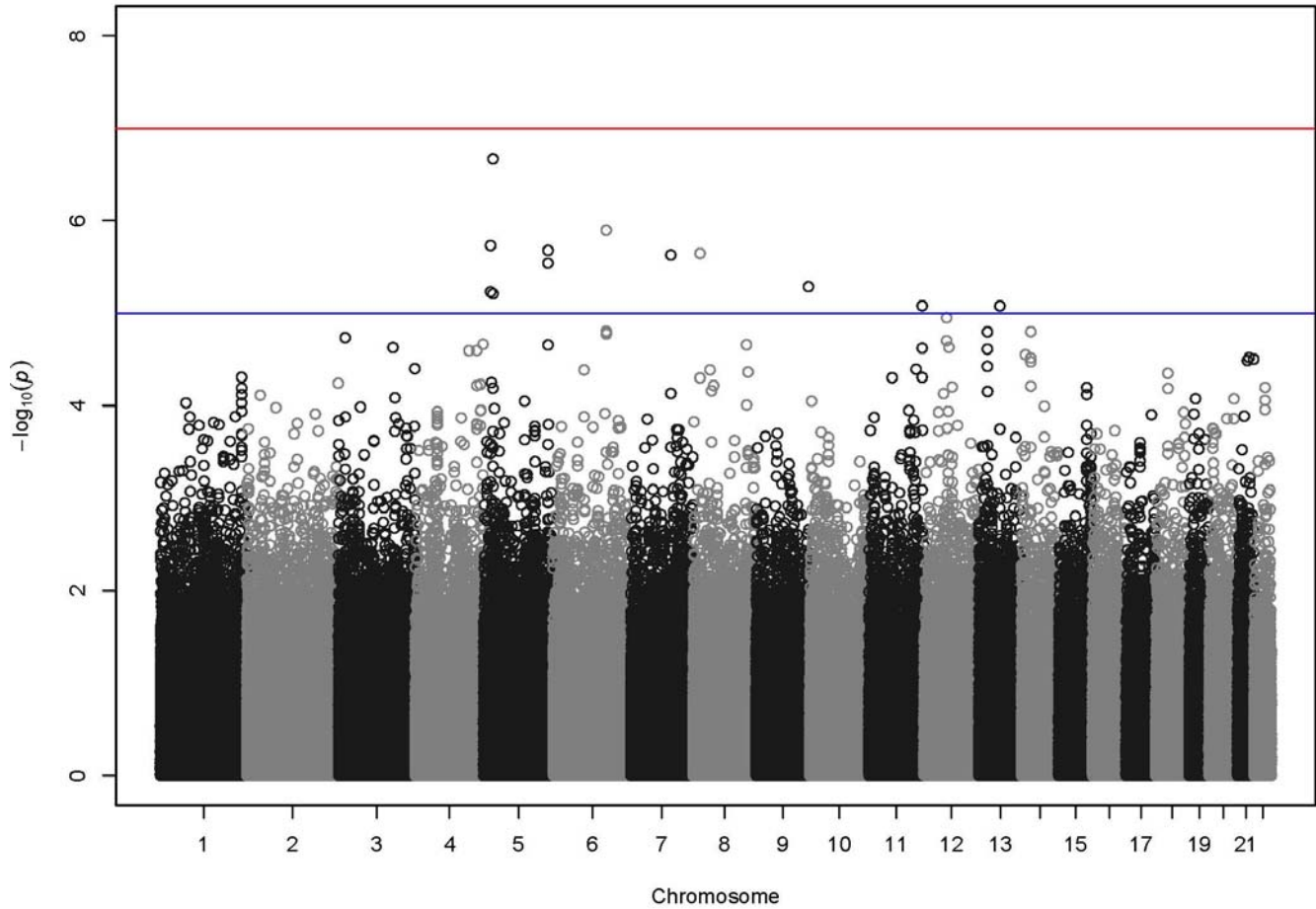


Figure S7. Manhattan Plot for Individual Outcome $\sqrt{\text{CESSATION}}$

Analysis was performed on both affected ($n_1 = 696$) and control ($n_0 = 730$) ever-smokers using $\pi = 0.000745$ to determine the weights. p values are unadjusted.

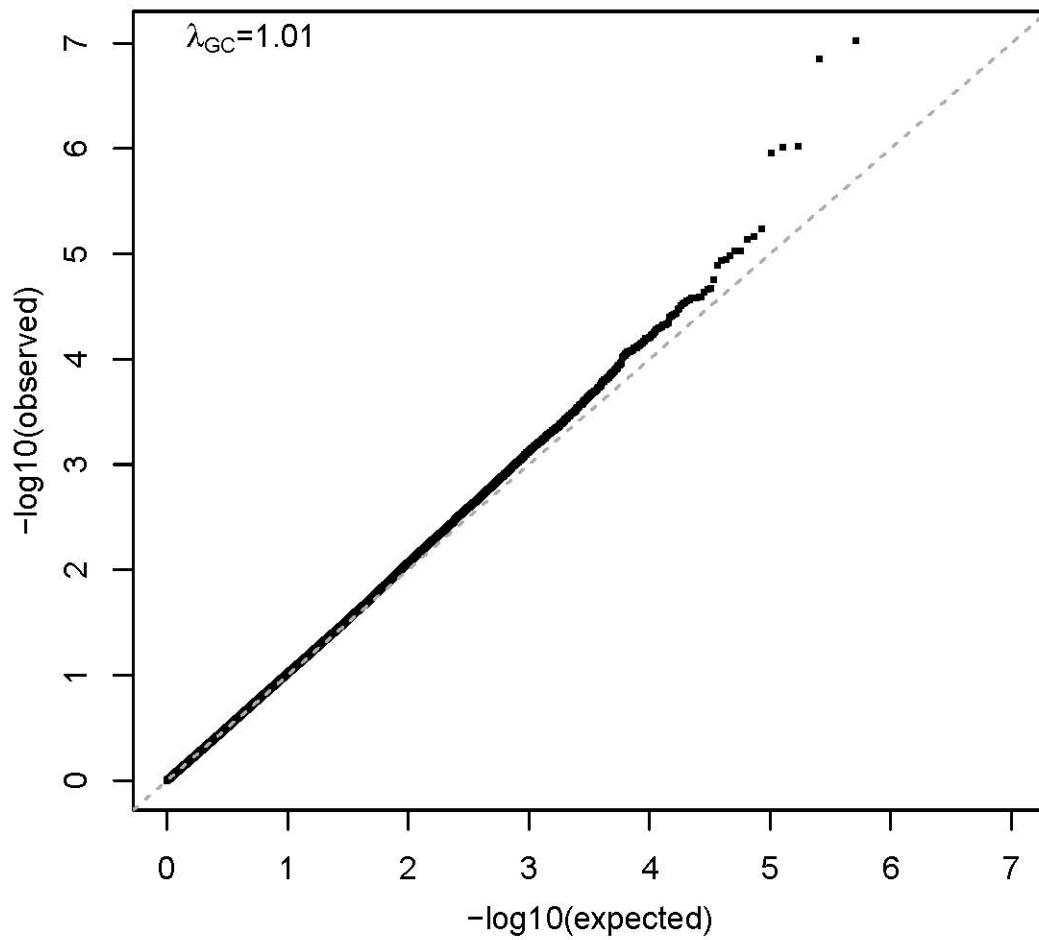


Figure S8. Quantile-Quantile Plot for 1-DF SMAT p Values

Analysis was performed on both affected ($n_1 = 696$) and control ($n_0 = 730$) ever-smokers using $\pi = 0.000745$ to determine the weights and an unstructured working correlation matrix. The estimated genomic inflation factor is indicated by λ_{GC} in the top left corner.

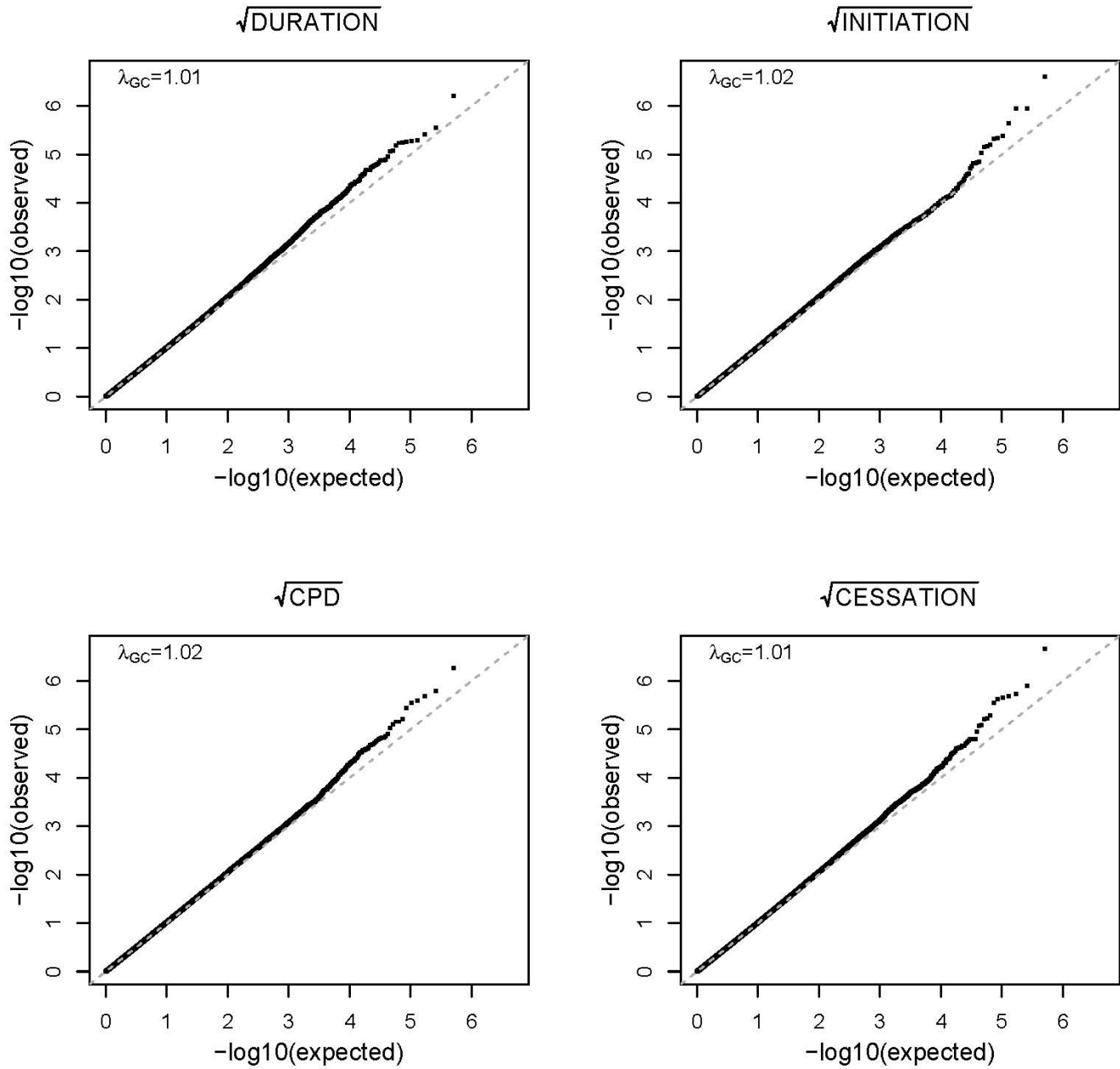


Figure S9. Quantile-Quantile Plots for Single-Outcome Tests

Analysis was performed on both affected ($n_1 = 696$) and control ($n_0 = 730$) ever-smokers using $\pi = 0.000745$ to determine the weights. The estimated genomic inflation factor for each plot is indicated by the λ_{GC} value in the top left corner. p values are unadjusted.

Table S1. Empirical Size Results for Test for Homogeneity

Working Correlation	Control Only	Control + Affected (LOW)^a	Control + Affected (MOD)^a
Disease-Independent			
I / E	0.049	0.050	0.052
U	0.049	0.049	0.051
Disease-Dependent			
I / E	0.049	0.050	0.052
U	0.049	0.049	0.052

^aLOW and MOD refer to disease prevalence $\pi=0.000745$ and $\pi=0.0745$, respectively.

Empirical size results for $B=5000$ simulated datasets and $n_0=n_1=700$ assuming the true correlation among the phenotypes is unstructured (as given in Table 1 of the main text); I=Independent, E=Exchangeable, and U= Unstructured working correlation matrices were considered. Disease-(In)dependent refers to whether the true SNP effect sizes are the same for affected and control individuals; see Materials and Methods subsection *Simulation: Empirical Performance of Test for Scaled Homogeneity* for more details. The theoretical size is 0.05.