Supplemental Tables and Figures

The Weighting Is The Hardest Part: On The Behavior of the Likelihood Ratio Test and the Score Test Under a Data-Driven Weighting Scheme in Sequenced Samples

Camelia C. Minică, Giulio Genovese, Christina M. Hultman, René Pool, Michael C. Neale, Conor V. Dolan[☯], Benjamin M. Neale[☯]
[☯] These authors contributed equally to this work.

S1 Table. Comparison of 4 programs implementing genetic similarity/kernel based variance component tests

For the sake of comparison, we analyzed one simulated sample of 5000 individuals by using 4 independent programs implementing genetic similarity/kernel-based variance component tests. Data were simulated in R using the MASS package. We simulated a gene harboring 50 active variants (with a minor allele frequency ranging between 5% and .05%, explaining 10% of variance) and a continuous phenotype. Association analyses were perfomed in the nlme R-package, the software Genome-wide Complex Trait Analysis (GCTA; [2]), the software FaST-LMM-set [4] and the software OpenMx [5]. Results are presented in S1 Table below.

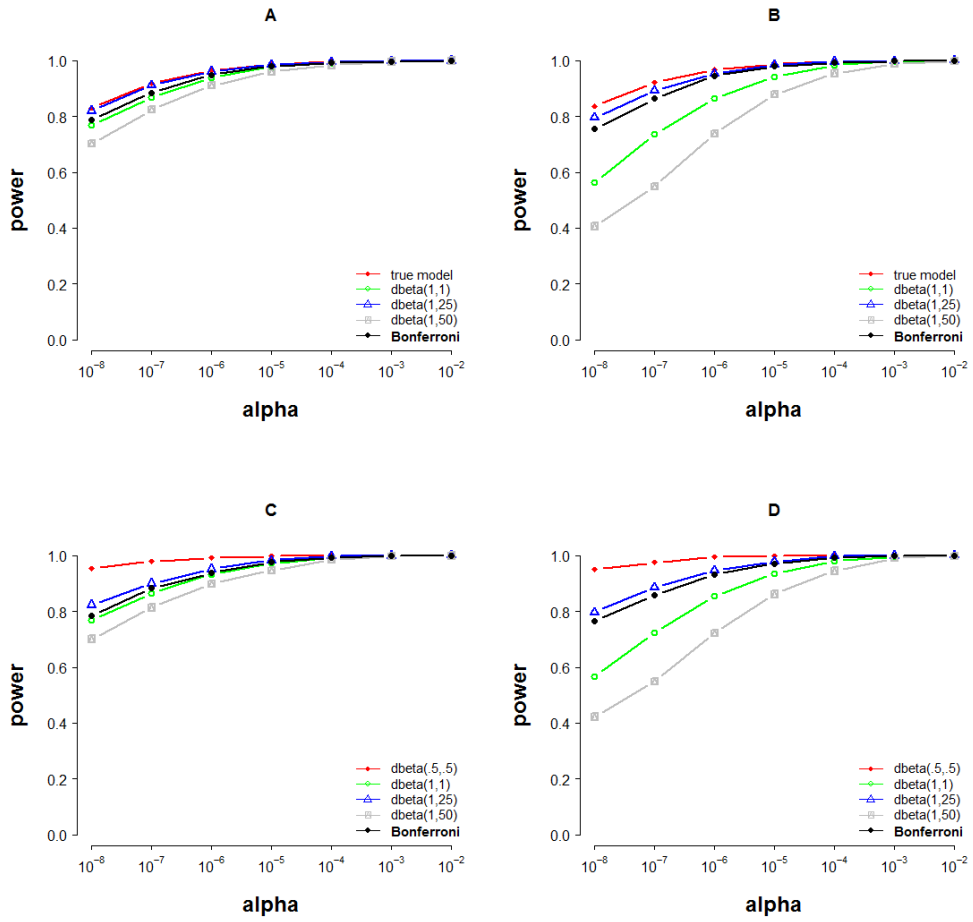
Results of a kernel(set)-based association test between a gene and a continuous phenotypic score, available in a simulated sample of 5,000 individuals. Analyses were performed in: the R-nlme package, the software Genome-wide Complex Trait Analysis (GCTA), the software FaST-LMM-set, and the software OpenMx . We report the log restricted likelihood (excluding the constant in GCTA) under the null model (LL_0), the log restricted likelihood under the alternative model (LL_1), the chi-square test with 1 degree of freedom ($\chi^2(1)$), the variance attributable to the 50 genetic variants (V(G)).

Software	$\mathbf{LL_0}$	$\mathbf{LL_1}$	$\boldsymbol{\chi^2(1)}$	**V(G)**	$\mathbf{V(G)/V_{phenotype}}$
GCTA	-2542.12	-2329.29	425.65	0.107	0.106
R-nlme	-7135.89	-6923.01	425.7	0.107	0.106
FaST-LMM-set	-7131.63	-6918.75	425.7	-	0.1055
OpenMx	-7140.57	-6927.91	425.31	0.107	0.1055

The values for the LRT and the estimate for the variance component obtained by the 4 programs were almost identical (the small differences are likely due to numerical precision of calculation/optimization), indicating that these implement equivalent approaches.

S1 Figure. The power of the likelihood ratio test (LRT; A and C) and the score test (B and D) to detect a gene harboring 50 low-frequency variants: all functional (A and B) or a mixture of 30 functional and 20 neutral variants (C and D). We randomly sampled MAFs ranging from 0.5% to 5% from the uniform distribution.The gene

explains 1% of the phenotypic variance. Power was evaluated in 1000 datasets consisting of 10,000 individuals. Note that while the variants within the set explain the same amount of variance across all scenarios considered, the true individual variant weights increase as the proportion of functional variants in the set decreases. Data were simulated and analyzed in R using the R-packages R-nlme and SKAT. Following Visscher [6] we used a .5:.5 mixture of χ^2_0 and χ^2_1 distributions to compute the p-value.The inclusion of neutral variants dilutes the power of both tests. However, the differences in power between the LRT and the score test follow the same pattern as that observed when the target region includes only functional variants.



S2 Table. The performance of the likelihood ratio test and the score test in association studies involving variants in linkage disequilibrium

We evaluated the performance of the likelihood ratio test and of the score test in association analyses involving target sites harboring rare variants in LD. Genotypic data were simulated using the software Cosi2 (`http://www.broadinstitute.org/mpg/cosi2` [3]). We used a coalescent model, 20,000 chromosomes of length 5,000, a population of size 100,000, and we set the mutation rate to equal 1.5E-08. The simulated chromosomes were then randomly paired up to form 10,000 diploids. We randomly selected target sites of 50 variants with MAF ranging from 0.5 to 5%; based on these we simulated continuous phenotypic scores as

described in the section *Data driven search for optimal weights: exploring the misspecification space*. The 50 variants were all active and jointly explained on average (across the 1000 simulations) 0.94% of the phenotypic variance. The site's LD structure varied across simulations, mimicing LD in real data.

As our proposed data-driven weighting scheme renders thresholding unnecessary (i.e., the use of alternative weights is equivalent to applying variable freqeuncy thresholds), we also considered the behavior of the two tests in association analyses involving target sites harboring both common and rare variants (with MAF between 0.5 to 50 %). As above, all variants were functional (50% deleterious), and jointly explained 1.57% of the phenotypic variance (on average, across the simulated samples). The type I error rate is given in S2 Table , and the power results are displayed in S2 Figure .

The 95% confidence intervals for the type I error for the restricted likelihood ratio test (LRT) and the score test, given genotypic data simulated under the null model of no association between the target site and the phenotype. Type I error was evaluated at $\alpha = 0.01$. The restricted LRT and the score tests were computed for four weights beta in each of the 1000 simulated samples of 10,000 individuals with genotypes at 50 variants. The LD structure varied across simulations.

Abbreviations: MAF - minor allele frequency.

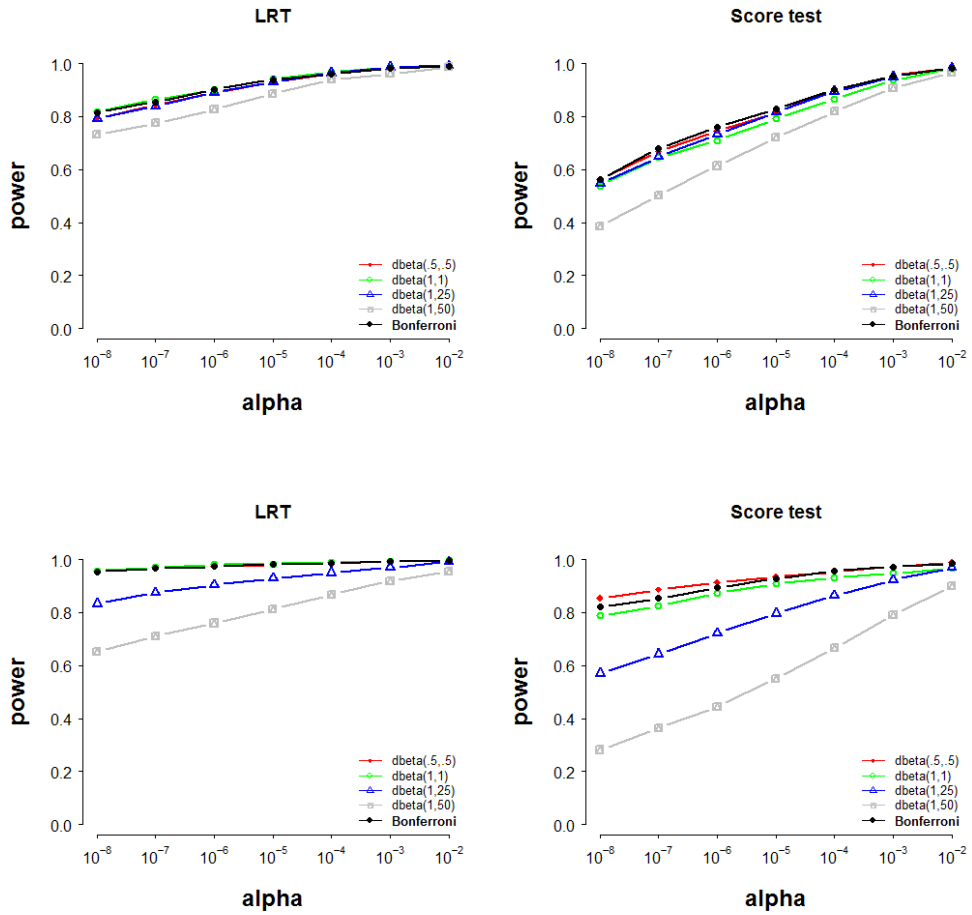
weights dbeta	LRT	Score test
MAF=0.5%-5%,		
(.5,.5)	[*0.0024, 0.0137*]	[*0.0037, 0.0163*]
(1,1)	[*0.0043, 0.0176*]	[*0.0024, 0.0137*]
(1,25)	[*0.0030, 0.0150*]	[*0.0030, 0.0150*]
(1,50)	[*0.0054, 0.0183*]	[*0.0024, 0.0137*]
Bonferroni	[*0.0018, 0.0123*]	[*0.0012, 0.0109*]
MAF=0.5%-50%		
(.5,.5)	[*0.0057, 0.0202*]	[*0.0072, 0.0227*]
(1,1)	[*0.0087, 0.0252*]	[*0.0087, 0.0252*]
(1,25)	[*0.0079, 0.0239*]	[*0.0072, 0.0227*]
(1,50)	[*0.0094, 0.0264*]	[*0.0087, 0.0252*]
Bonferroni	[*0.0054, 0.0183*]	[*0.0057, 0.0202*]

S2 Table indicates that the type I error rates of the two tests is correct also when the association analysis targets sites harboring rare variants, or rare and common variants in linkage disequilibrium. The type I error rate is correct, regardless of the weighting scheme used.

S2 Figure S2 Figure demonstrates that the powers of the two tests follow the pattern observed when the variants within the target site are in linkage equilibrium. The likelihood ratio test has more power of detection than the score test, and is generally more robust to misspecification. In these scenarios dbeta (1,25) and dbeta(1,50) are the least optimal weighting schemes, as with these weights a large proportion of (more) common causal variants are assigned zero weights and are therefore discarded from the test. Note that in performing these simulations, we no longer controlled the effect size, as this now depends on the LD structure (thence, from here the differences in power between the top and bottom figures).

The power of the likelihood ratio test and the score test to detect a gene harboring 50 rare variants with MAF=0.5%-5% (top figures), and 50 common and rare variants

with MAF=0.5%-50% (bottom figures). The simulated effect size varied with LD. Power was evaluated in 1000 datasets consisting of 10,000 individuals.



S3 Table. Association results for the single variant analysis of the SNPs with minor allele frequency $> 0.5\%$ located in:(a) the *PRRC2A* gene, (b) the *AKT3* gene and (c) the *VARS2* gene. Association between each SNP and the schizophrenia disease status was tested in a logistic regression model using Plink. Two principal components explaining the larges amount of variance in the sample and reflecting the Finish and Northern/Southern Swedish ancestry (see Extended Data Figure 1 in [1]) were included as covariates. Abbreviations: CHR - chromosome; BP = physical position; A1 - minor allele name (tested allele); OR - odds ratio; STAT - coefficient t-statistic; P - asymptotic p-value for t-statistic; MAF - minor allele frequency.

(a) *PRRC2A*

SNP	CHR	BP	A1	OR	STAT	P	MAF
6:31598489:A:G	6	31598489	G	0.8374	-4.934	8.074E-07	0.1792
6:31604591:C:T	6	31604591	T	0.8385	-4.9	9.596E-07	0.1793
6:31603770:A:G	6	31603770	G	0.8394	-4.866	1.137E-06	0.1792
6:31602967:G:A	6	31602967	A	0.892	-4.023	5.737E-05	0.3739
6:31604010:C:G	6	31604010	G	1.126	3.982	6.837E-05	0.3049
6:31590898:C:T	6	31590898	T	1.102	3.352	0.0008012	0.3507
6:31601735:G:A	6	31601735	A	1.095	1.583	0.1134	0.05965
6:31593265:G:T	6	31593265	T	1.091	1.518	0.129	0.05956
6:31595882:C:A	6	31595882	C	1.028	0.9038	0.3661	0.2667
6:31607050:C:T	6	31607050	C	1.03	0.8721	0.3832	0.1839
6:31601344:T:C	6	31601344	T	1.027	0.8615	0.389	0.2667
6:31594628:T:A	6	31594628	T	1.028	0.7508	0.4528	0.1622
6:31604044:T:G	6	31604044	T	0.9851	-0.3038	0.7613	0.08361
6:31601520:T:C	6	31601520	C	1.011	0.3036	0.7614	0.1746
6:31602489:T:C	6	31602489	C	1.011	0.2983	0.7655	0.1807

S4 Table.

(b) *AKT3*

SNP	CHR	BP	A1	OR	STAT	P	MAF
1:243777066:G:A	1	243777066	A	0.8965	-3.052	0.002276	0.1796

S5 Table.

(c) *VARS2*

SNP	CHR	BP	A1	OR	STAT	P	MAF
6:30886729:C:T	6	30886729	T	1.186	4.288	1.801E-05	0.1371
6:30893941:G:A	6	30893941	A	0.8646	-4.113	3.904E-05	0.1888
6:30890871:T:C	6	30890871	T	0.9119	-3.307	0.0009422	0.4308
6:30888161:T:C	6	30888161	T	0.9137	-3.234	0.001221	0.4318
6:30889120:A:G	6	30889120	G	1.077	2.632	0.008496	0.3866
6:30892322:C:T	6	30892322	T	1.076	2.609	0.009088	0.3867
6:30890055:T:C	6	30890055	C	1.076	2.602	0.009265	0.3874
6:30893428:G:A	6	30893428	A	1.076	2.599	0.009349	0.3865
6:30882781:G:T	6	30882781	T	1.075	2.561	0.01044	0.3877
6:30882634:C:T	6	30882634	T	1.075	2.56	0.01047	0.3877
6:30890206:T:C	6	30890206	C	1.075	2.56	0.01047	0.3877
6:30893831:G:A	6	30893831	A	1.075	2.557	0.01056	0.3875
6:30890483:G:T	6	30890483	T	1.075	2.551	0.01074	0.3876
6:30890569:G:T	6	30890569	T	1.074	2.547	0.01088	0.3874
6:30882856:A:G	6	30882856	G	1.074	2.546	0.01091	0.3876
6:30883920:C:T	6	30883920	T	1.074	2.546	0.01091	0.3876
6:30888169:G:A	6	30888169	A	1.072	2.476	0.01328	0.3867
6:30887988:C:T	6	30887988	T	1.072	2.461	0.01385	0.3875
6:30893127:G:A	6	30893127	A	1.037	1.032	0.3021	0.1813
6:30887972:C:T	6	30887972	T	1.035	0.9699	0.3321	0.1814
6:30878919:T:C	6	30878919	C	1.034	0.9465	0.3439	0.1894
6:30877202:T:C	6	30877202	C	1.032	0.8932	0.3717	0.1896
6:30878579:A:T	6	30878579	T	1.032	0.8882	0.3745	0.1883
6:30889389:T:C	6	30889389	C	1.032	0.8672	0.3858	0.1748
6:30883878: AAAGCAACC:A	6	30883878	AAAGC AACC	1.012	0.365	0.7151	0.2194
6:30890789:G:A	6	30890789	A	1.015	0.2901	0.7718	0.07364
6:30890195:A:G	6	30890195	G	1.014	0.2587	0.7959	0.0736
6:30893728:C:T	6	30893728	C	0.9985	-0.0475	0.9621	0.242

Acknowledgments

We thank to the Swedish cohort participants whose data we analyzed in this study. Camelia C. Minica and Michael C. Neale are supported by the National Institute on Drug Abuse grant DA-018673.

References

1. Shaun M Purcell, Jennifer L Moran, Menachem Fromer, Douglas Ruderfer, Nadia Solovieff, Panos Roussos, Colm O'Dushlaine, Kimberly Chambert, Sarah E Bergen, and Anna Kähler. A polygenic burden of rare disruptive mutations in schizophrenia. *Nature*, 506(7487):185–190, 2014.

2. Jian Yang, S Hong Lee, Michael E Goddard, and Peter M Visscher. Gcta: a tool for genome-wide complex trait analysis. *The American Journal of Human Genetics*, 88(1):76–82, 2011.

3. Ilya Shlyakhter, Pardis C Sabeti, Stephen F Schaffner Cosi2: an efficient simulator of exact and approximate coalescent with selection. *Bioinformatics*, 3427–3429, Oxford Univ Press, 2014.

4. Jennifer Listgarten, Christoph Lippert, Eun Yong Kang, Jing Xiang, Carl M Kadie, and David Heckerman. A powerful and efficient set test for genetic markers that handles confounders. *Bioinformatics*, 29(12):1526–1533, 2013.

5. Michael C Neale, Michael D Hunter, Joshua N Pritikin, Mahsa Zahery, Timothy R Brick, Robert M Kirkpatrick, Ryne Estabrook, Timothy C Bates, Hermine H Maes, Steven M Boker OpenMx 2.0: Extended structural equation and statistical modeling. *Psychometrika*, 1–15, Springer, 2015.

6. Peter M Visscher. A note on the asymptotic distribution of likelihood ratio tests to test variance components. *Twin research and human genetics*, 9(04):490–495, 2006.