The MAX Statistic is Less Powerful for Genome Wide Association Studies Under Most Alternative Hypotheses

Benjamin Shifflett[1,2], Rong Huang[3] and Steven D. Edland[2,4,*]

[1]*Department of Physics, University of California San Diego, La Jolla CA, USA*

[2]*Division of Biostatistics, Department of Family and Preventative Medicine, University of California San Diego, La Jolla, CA, USA*

[3]*Department of Mathematics, University of California San Diego, La Jolla, CA, USA*

[4]*Department of Neurosciences, University of California San Diego, La Jolla, CA, USA*

Abstract: Genotypic association studies are prone to inflated type I error rates if multiple hypothesis testing is performed, e.g., sequentially testing for recessive, multiplicative, and dominant risk. Alternatives to multiple hypothesis testing include the model independent genotypic χ^2 test, the efficiency robust MAX statistic, which corrects for multiple comparisons but with some loss of power, or a single Armitage test for multiplicative trend, which has optimal power when the multiplicative model holds but with some loss of power when dominant or recessive models underlie the genetic association. We used Monte Carlo simulations to describe the relative performance of these three approaches under a range of scenarios. All three approaches maintained their nominal type I error rates. The genotypic χ^2 and MAX statistics were more powerful when testing a strictly recessive genetic effect or when testing a dominant effect when the allele frequency was high. The Armitage test for multiplicative trend was most powerful for the broad range of scenarios where heterozygote risk is intermediate between recessive and dominant risk. Moreover, all tests had limited power to detect recessive genetic risk unless the sample size was large, and conversely all tests were relatively well powered to detect dominant risk. Taken together, these results suggest the general utility of the multiplicative trend test when the underlying genetic model is unknown.

Keywords: Armitage test, case-control study, efficiency robust statistics, MAX statistic, multiple comparisons;, Type I error.

1. INTRODUCTION

Testing for association between genotype and disease in case-control studies is arguably one of the most important statistical analyses performed by genetic epidemiologists. However, there is no consensus on the most appropriate test procedure to use in this situation. If the mode of inheritance is known and is recessive, multiplicative, or dominant, then association can be tested optimally using the Armitage trend test corresponding to the appropriate underlying genetic model [1]. When the mode of inheritance is not known, testing multiple hypotheses is also common. For example, some investigators test and report p-values for association using both a genotypic χ^2 test and an allelic χ^2 test, or test and report p-values for association using each of the recessive, multiplicative, and dominant Armitage tests. Epidemiologists often set one genotype as the referent genotype and calculate odds ratios and test statistics for the two other genotypes relative to the referent genotype, another example of multiple hypothesis testing. Multiple hypothesis testing without correction results in inflated type I error rate, and standard corrections for multiple comparisons, such as the Bonferonni correction, are statistically inefficient. Therefore, many investigators propose using so called efficiency robust tests, which perform multiple statistical tests, but with appropriate statistical correction for multiple comparisons [2].

Table **1** lists some of the properties of various tests commonly used to test genetic associations with disease. Guidelines as to which test to use in a given situation are limited. If the underlying genetic model is known and is recessive, dominant, or multiplicative, then the most powerful statistical test is the Armitage test using the score function corresponding to the known underlying genetic model [1, 3]. This may occur, for example, when performing a validation study testing a clear *a priori* hypothesis based on a previously reported association or on laboratory observations.

If there is little prior information on the likely underlying genetic model to guide the choice of which statistic to use when testing for association, it is tempting to apply more than one statistical test. Performing more than one test without statistical correction for multiple comparisons is not valid because the probability of a false positive finding under the null hypothesis of no association will exceed the nominal

*Address correspondence to this author at the Division of Biostatistics, University of California San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0949, USA; Tel: 858-246-1250; Fax: 858-622-5876; E-mail: sedland@ucsd.edu

E-ISSN: 1929-6029/17

Table 1: Approaches to testing a biallelic single nucleotide polymorphism for association with disease

Test	Comments
Genotypic χ^2 test	The only 2 degree of freedom test considered here.
Allelic χ^2 test	Invalid when data are not in HWE [1].
Allelic χ^2 test adjusted for lack of HWE [10]	Equivalent to Armitage test for multiplicative trend [11].
Armitage Test – Recessive (with score function x = (0, 0, 1))	The optimal test when $RR_{aA} = 1$.
Armitage Test – Multiplicative (with score function x = (0, 1, 2))	The optimal test when $RR_{aA} = (RR_{AA})^{1/2}$.
Armitage Test – Dominant (with score function x = (0, 1, 1))	The optimal test when $RR_{aA} = RR_{AA}$.
Genotypic and allelic χ^2 tests without correction for multiple comparisons	Commonly applied, but invalid (type I error rate > nominal rate).
Separate 1 degree of freedom tests of the odds ratio of the *aA* genotype and *AA* genotype relative to the *aa* genotype without correction for multiple comparisons	The "referent group" method. Commonly applied, but invalid (type I error rate > nominal rate).
Separate recessive, multiplicative, and dominant Armitage tests without correction for multiple comparisons	Commonly applied, but invalid (type I error rate > nominal rate).
MAX statistic of the recessive, multiplicative, and dominant Armitage tests [3]	I.e., the most significant of the three tests with adjustment for multiple testing.

type I error rate of the individual tests. Bonferroni correction for multiple hypothesis testing can be used to ensure the type I error rate does not exceed the nominal rate. However, Bonferroni correction is inefficient in these analyses because the results of the separate test statistics are correlated. Statistics that efficiently correct for multiple hypothesis testing are therefore preferred. For example, the results of multiple statistical tests applied to the same data but under different models can be combined using *efficiency robust* test statistics [4]. Freidlin, *et al*. describe applying two such methods to the genotypic association problem [3]. These are the MAX statistic, defined simply as the most significant of the separate test statistics, and the MERT statistic, the linear combination of the statistics that maximizes the minimum efficiency under the models considered. When combining results of the three separate Armitage tests (using the recessive, multiplicative, and dominant score functions), the MAX statistic tends to outperform the MERT statistic, and was generally recommended by Freidlin, *et al*. for this application [3].

An alternative to applying multiple tests with statistical correction is to settle on a single test for the primary analysis. Examples of this approach are the two degree of freedom genotypic χ^2 test and the Armitage test for multiplicative trend. The Armitage test for multiplicative trend has optimal power under the multiplicative genetic model and reasonable power under other genetic models [5]. Moreover, the multiplicative trend test has been shown to be algebraically equivalent to an allele frequency test appropriately corrected for deviations from HWE in sample data [6].

There is limited information on the relative performance of these various approaches to statistical analysis when the genetic model is unknown. The performance of the MAX statistic combining the recessive, multiplicative, and dominant Armitage tests has been compared to the performance of the multiplicative trend test by Friedlin, *et al*. [3]. In a limited series of computer simulations, they found that the MAX statistic was more powerful than the multiplicative trend test when the risk allele was recessive and was less powerful when the risk allele had a multiplicative effect. Among the simulations performed, the difference in performance was most dramatic when the risk allele had a recessive effect, and based on this observation Freidlin, *et al*. recommended the MAX statistic for general use [3].

Freidlin, *et al*. investigated a constrained set of scenarios. Specifically, Freidlin, *et al*. simulated three scenarios: 1) recessive data with an effect size that assured 80 percent power if the data were tested with the Armitage test for recessive trend, 2) multiplicative data with an effect size that assured 80 percent power if the data were tested with the Armitage test for multiplicative trend, and 3) dominant data with an effect size that assured 80 percent power if the data were tested with the Armitage test for dominant trend [3]. Moreover, Friedlin, *et al*. did not investigate the relative performance of the familiar genotypic χ^2 test.

To address these concerns, we have used Monte Carlo simulations to more completely characterize the performance of various approaches to testing for genetic association. Type I error rates under the null hypothesis of no genotypic association are estimated for each of the statistics described in Table **1**. In addition, the power of various statistics that meet their nominal type I error rates is estimated. Statistical power is estimated for the MAX statistic pooling all three Armitage tests [3], for the two degree of freedom χ^2 statistic, and for the Armitage test for multiplicative trend.

2. METHODS

Type I error rates under the null were estimated by Monte Carlo simulation using 100,000 simulated samples for each sample size considered. This number of simulations assures that error rates will be estimated with a standard error of between 0.0001 and 0.0003 when the true type I error rate is between 0.01 and 0.10. For type I error rate simulations, the control genotype frequencies and case genotype frequencies were simulated as multinomials assuming HWE. Letting *a* indicate the wild type allele, *A* the risk allele, and *p* the frequency of the putative risk allele, the frequencies of the genotypes *aa*, *aA*, and *AA* were simulated as a multinomial with parameters $(u_1, u_2, u_3) = ((1-p)^2, 2(1-p)p$, and $p^2)$ for both controls and cases.

Statistical power was estimated by Monte Carlo simulation using 100,000 simulated samples for each alternative scenario considered. This number of simulations assures that power will be estimated with a standard error of 0.002 or better. Case and control genotype frequencies were simulated as previously described [5]. For a given allele frequency *p*, heterozygote rate ratio RR_{aA}, homozygote rate ratio RR_{AA}, and population disease prevalence *K*, case genotype frequencies were simulated as a multinomial with parameters $((f_0g_0, f_1g_1, f_2g_2)/ \Sigma f_ig_i)$ and control frequencies were simulated as a multinomial with parameters $(((1 - f_0)g_0, (1 - f_1)g_1, (1 - f_2)g_2))/ \Sigma(1 - f_i)g_i)$, where HWE population genotype frequencies $(g_0, g_1, g_2) = ((1 - p)^2, 2p(1 - p), (1 - p)^2)$ and genotype specific disease penetrances $(f_0, f_1, f_2) = (K/(g_0 + RR_{aA}g_1 + RR_{AA}g_2), RR_{aA}f_0, RR_{AA}f_0))$. All simulations were performed with $K = 0.01$.

The Armitage trend test was performed as previously described [1]. Letting $i = 0, 1, 2$ index the genotypes *aa*, *aA*, and *AA*, and given a sample of *S* controls and *R* cases, $N = S + R$, define (s_0, s_1, s_2), $\Sigma s_i = S$, as the number of controls with the genotypes *aa, aA, and AA*, define (r_0, r_1, r_2), $\Sigma r_i = R$, as the number of cases with genotypes *aa, aA, and AA*, and define $(n_0, n_1, n_2) = (s_0, s_1, s_2) + (r_0, r_1, r_2)$. Then, the Armitage test is performed by referring the Armitage statistic

$$Z^2 = \frac{N^3\left[\sum_i x_i\left(\frac{S}{N}r_i - \frac{R}{N}s_i\right)\right]^2}{RS\left[N\sum_i x_i^2 n_i - \left(\sum_i x_i n_i\right)^2\right]} \quad (1)$$

to a one degree of freedom χ^2 distribution [5]. The vector $\boldsymbol{x} = (x_0, x_1, x_2)$ is a measure of exposure dosage. $\boldsymbol{x} = (0, 0, 1)$ is used to test an underlying recessive genetic model, $\boldsymbol{x} = (0, 1, 2)$ to test an underlying multiplicative genetic model, and $\boldsymbol{x} = (0, 1, 1)$ to test an underlying dominant genetic model.

P-values for the MAX statistic were calculated as described by Freidlin, *et al*. An empirical reference distribution for the MAX statistic was generated for each sample using 10,000 simulations of the asymptotic joint distribution of the three Armitage test statistics given the allele frequency observed in the sample [3]. All simulations were performed in the statistical programming language R (the Free Software Foundation Inc., Boston, MA).

3. RESULTS

3.1. Type I Error Rates Under the Null

Type I error rates when there is no signal in the data are summarized for various sample sizes in Table **2**. The Armitage tests were slightly anti-conservative when the nominal α was 0.05, consistent with previous investigations of this statistic [7], but did nearly achieved their nominal α error rates (Table **2**). Actual α error rates were approximately equal to the nominal error rates (of 0.05 or 0.01) when testing for multiplicative or recessive genetic traits, while the Armitage test for dominant trend did not perform as well, with type I error rates as much as 10 percent higher than the nominal rates (Table **2**). Hence the dominant test, while asymptotically valid [1], did not meet its nominal α error rate for the allele frequency and sample sizes considered here. The MAX statistic and genotypic χ^2 statistic achieved their nominal α error rates to within 0.001 (Table **2**).

As expected, multiple hypothesis testing without statistical correction was anti-conservative (Table **2**). When testing both genotype and allele frequencies

Table 2: Type I error rates of various tests of genotypic association assuming a sample size of 200, 600, or 1000, equal allocation of cases and controls, a risk allele frequency of *p* = 0.3, and a population in HWE. The *"Armitage-rec. and mult. and dom."* method refers to the common practice of testing under all models and reporting statistically significant models without correction for multiple comparisons. The *"genotypic χ^2 and allelic χ^2"* method refers to the common practice of testing genotypic and allelic association without correcting for multiple comparisons. The *"referent group"* method refers to the common practice of testing heterozygous and homozygous mutant genotypes separately without correcting for multiple comparisons. Observed type I error rates are reported separately for hypothesis testing using nominal type I error rates α of 0.05 and 0.01

Test	*N* = 200		*N* = 600		*N* = 1,000	
	α = 0.05	α = 0.01	α = 0.05	α = 0.01	α = 0.05	α = 0.01
Valid tests						
Genotypic χ^2 (2 df)	0.049	0.010	0.049	0.010	0.050	0.009
Armitage-recessive	0.052	0.010	0.051	0.009	0.050	0.010
Armitage-multiplicative	0.050	0.010	0.050	0.010	0.050	0.010
Armitage-dominant	0.055	0.009	0.052	0.010	0.054	0.010
MAX statistic [3]	0.049	0.009	0.050	0.009	0.050	0.010
Uncorrected multiple comparison testing						
Armitage-(rec. and mult. and dom.)	0.109	0.022	0.106	0.022	0.106	0.023
Genotypic χ^2 and allelic χ^2	0.065	0.013	0.069	0.014	0.071	0.014
Referent group method	0.060	0.011	0.072	0.014	0.078	0.015

without statistical correction the actual type I error rate was 30 to 40 percent higher than the nominal error rate α when α = 0.05 and when α = 0.01. The type I error rate varied as a function of the sample size for the "referent group" method, and was from 10 to 50 percent higher than the nominal error rate in the scenarios considered here (Table **2**). Multiple testing for recessive, multiplicative, and dominant traits without correcting for multiple comparisons increased the type I error rate by more than 100 percent for all scenarios in Table **2**.

3.2. Statistical Power for Recessive, Multiplicative, and Dominant Data

Three accepted valid statistics are the 2 degree of freedom genotypic χ^2 test, the MAX statistic, and the Armitage test for multiplicative trend. The performance of these tests is summarized in Appendix for a range of allele frequencies, underlying genetic models (recessive, multiplicative, and dominant), and effect sizes. As a point of reference, the power of the Armitage test for dominant trend is included when the underlying genetic model is dominant, and the power of the Armitage test for recessive trend is included when the underlying genetic model is recessive. As confirmed empirically (Appendix), the appropriate Armitage test is optimal and should be the preferred test if the underlying genetic model is known.

For data generated by the recessive genetic model, the pattern of relative performance of the various model robust statistics is clear. For the range of examples we considered, the 2 degree of freedom χ^2 test and the MAX statistic performed comparably, and both tests consistently outperformed the Armitage test for multiplicative trend (Appendix). For the range of sample sizes where the MAX and χ^2 statistics have between 60 percent and 90 percent power, the multiplicative test had at least 10 percent less power than the χ^2 or MAX tests (Appendix). This decrease in power for the multiplicative test relative to the χ^2 and MAX statistic is seen for a range of allele frequencies and sample sizes (Appendix).

For data generated by an underlying multiplicative model, the Armitage test for multiplicative trend was consistently most powerful, as expected. For sample sizes and effect sizes where the multiplicative test has between 60 and 90 percent power, the MAX statistic had between four and nine percent less power than the multiplicative test (Appendix). The performance of the χ^2 statistic tracked that of the MAX statistic, although the χ^2 statistic consistently had slightly less power.

For data generated by an underlying dominant genetic model the relative performance of the various statistics is less consistent (Appendix). The dominant test, included in Appendix for reference, is clearly the

optimal test for dominant data. When the allele frequency is small, the performance of the multiplicative test approaches the (optimal) performance of the dominant test, and the MAX and χ^2 tests have less power across a range of sample sizes (Appendix). When the allele frequency is 0.3, there is little practical difference in the performance of the multiplicative, MAX, and χ^2 statistics (Appendix). Finally, when the allele frequency is 0.5 the multiplicative test performs more poorly than the χ^2 and MAX statistics (Appendix).

3.3. Statistical Power under other Alternative Hypotheses

The recessive, multiplicative, and dominant tests are optimal for a discrete list of three possible underlying genetic effects. Genotypic effects intermediate between recessive, multiplicative, and dominant may also be encountered. For example, a model that is additive on the rate ratio scale (RR_{aA} = (RR_{AA} + 1)/2) is often considered in practice. More generally, the heterozygote rate ratio may be anywhere between RR_{aA} = 1 (recessive) and RR_{aA} = RR_{AA} (dominant). To further characterize the performance of the various model robust statistics, we performed additional simulations holding the sample size fixed and modifying the underlying model by letting RR_{aA} vary between 1 and RR_{AA}. Figure **1** summarizes the relative power of the χ^2, multiplicative, and MAX statistics when the control risk allele frequency is 10 percent, a scenario where the multiplicative test performs well for dominant data, while Figure **2** summarizes the relative power of the model robust statistics when the control risk allele frequency is 50 percent, a scenario that does not favor the multiplicative Armitage test when the data are dominant.

When the allele frequency was 10 percent (Figure **1**), we found that the MAX statistic and the χ^2 test outperformed the multiplicative trend test when the genetic effect was recessive or near recessive, but that otherwise the multiplicative test was superior. This is a general finding across sample sizes (Appendix), although the advantage of the MAX and χ^2 tests over the multiplicative test when data are recessive is attenuated as the sample size gets smaller than 200 and the power for all three tests approaches five percent (data not shown).

When the allele frequency was 50 percent (Figure **2**), the multiplicative test under performed relative to the MAX and χ^2 statistics when the genetic effect was recessive and when the genetic effect was dominant or nearly dominant (when RR_{aA} was greater than about 2.5 in our example). For much of the range of possible underlying genetic effects, however, the multiplicative test outperformed the MAX and χ^2 statistics (Figure **2**).

4. DISCUSSION

There are no universally accepted guidelines for testing the association between genotype and disease in a case-control study. A number of valid statistics, including the two degree of freedom genotypic χ^2 test

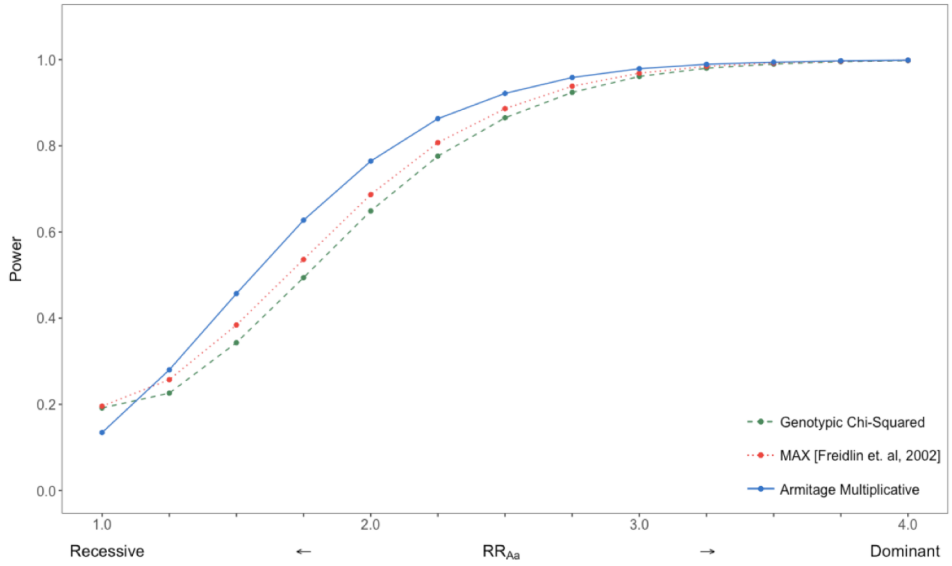

Figure 1: Statistical power as a function of the underlying genetic model when RR_{AA} = 4, risk allele frequency p = 0.1, and sample size = 100 cases plus 100 controls. The heterozygote rate ratio RR_{aA} ranges from 1 (recessive model) to RR_{AA} (dominant model).

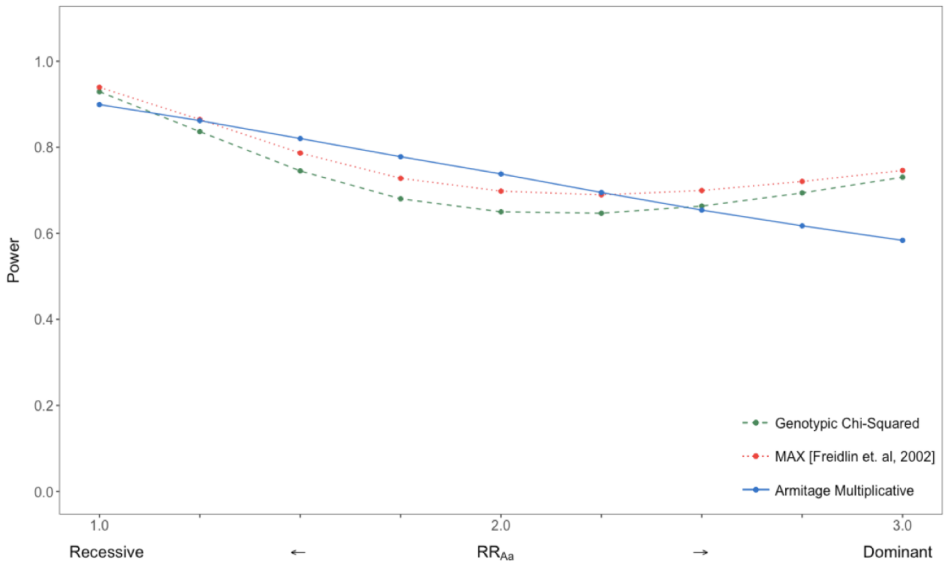

Figure 2: Statistical power as a function of the underlying genetic model when RR_{AA} = 3, risk allele frequency p = 0.5, and sample size = 100 cases plus 100 controls. The heterozygote rate ratio RR_{aA} ranges from 1 (recessive model) to RR_{AA} (dominant model).

and the Armitage test for multiplicative trend, are commonly accepted and used. It is also common, however, for multiple tests to be performed in the course of assessing a single genetic variant. Our goals in producing this report were two-fold, 1) to illustrate the potential magnitude of increase in type I error when multiple hypothesis testing is performed without correction, and 2) to compare the performance of currently available valid procedures, including valid procedures based on combining the results of multiple hypothesis tests with appropriate statistical correction for multiple comparisons.

Regarding type I error rates, we found that the practice of performing separate Armitage tests for recessive, multiplicative, and dominant trend without correcting for multiple hypothesis testing is highly anti-conservative, with true type I error rates that are twice the nominal rates. The practice of testing both genotype and allele frequencies without correction for multiple comparisons is also anti-conservative, although the problem is not as extreme in this case. Similarly, the practice of testing heterozygote and homozygote variant genotype frequencies against the referent, homozygote wild type genotype frequency without correction is anti-conservative. To our knowledge the magnitude of increase in false positive findings by these common analytic methods has not previously been quantified.

The problem of false positive findings in genetic association studies can be addressed in a number of ways, including paying more attention to study design and analysis to insure that spurious positive findings are not reported [8, 9], and reducing the number of tests performed by restricting to functional variants or variants that otherwise are high probability candidates [9]. Beyond these efforts, results reported in this paper demonstrate that the type I error rate of genetic association studies is reduced by as much as half by using a valid primary statistical analysis.

Regarding the relative performance of various valid analyses, we found that the two degree of freedom genotypic χ^2 statistic performed comparably to the MAX statistic for a range of effect sizes and underlying genetic models. The power of the two test procedures was within a few percentage points for most scenarios considered, with the MAX test consistently but only slightly outperforming relative to the χ^2 test for the scenarios we considered. Hence the MAX test is the preferred of the MAX and genotypic χ^2 tests, although the difference in performance is not dramatic.

As expected, the multiplicative trend test had more power than the genotypic χ^2 test and MAX test when the data being tested were multiplicative. Moreover, this increase in relative power held over a broad range of the possible underlying genetic models (Figures **1** and **2**). Only when the underlying genetic model was recessive or very nearly recessive, or when the underlying genetic model was dominant and the risk allele frequency high, did the multiplicative trend test perform more poorly than the genotypic χ^2 and MAX tests. Hence, the multiplicative trend test is optimal for the broadest range of likely underlying genetic effects.

On the other hand, when the underlying model is recessive the multiplicative test can under perform relative to the genotypic χ^2 and MAX tests, and there is no consistently most powerful test among the three tests considered here.

Unless there is a compelling prior likelihood that the genetic effect being tested may be recessive, we believe the Armitage test for multiplicative trend is a reasonable choice for testing candidate variants when the underlying genetic model is unknown. The Armitage test is most powerful for the broadest range of alternative hypotheses, is readily accessible using available software, and has a natural heuristic interpretation since it is algebraically equivalent to testing for a difference in allele frequencies with adjustment for departures from HWE [6]. The Armitage test also generalizes easily to the situation where one wants to control for potential confounding.

Choosing one test for the primary analysis does not rule out performing post hoc descriptive or exploratory analyses using other tests and statistics, provided the secondary analyses are identified as such. E.g. reporting subgroup analyses stratifying by age, gender, or known genetic risk factors may inform future investigators and would be available for meta-analyses testing stratum specific effects.

The practice of using multiple hypothesis tests in the course of analyzing a single candidate genetic variant without statistical correction is but one source of false positive findings in genetic association studies [8]. Nonetheless it is a potentially substantial source of false positive findings, doubling the expected number of false positive findings reported in the literature. Valid analytic approaches that obtain their nominal type I error rate are available, and should be preferred when performing genetic association studies.

ACKNOWLEDGEMENTS

This work was supported by the NIH grant AG056499, NIH grant AG005131, and the Shiley Marcos Alzheimer's Disease Research Center.

Appendix: Power to detect a genotype association as a function of the effect size, underlying genetic model (recessive, multiplicative or dominant data), the allele frequency *p*, total sample size *N*, and statistical test performed (χ^2 is the standard 2 degree of freedom genotypic χ^2 test, Mult is the Armitage test for multiplicative trend, Rec is the Armitage test for recessive trend, and Dom is the Armitage test for dominant trend); all results are for equal allocation to cases and controls

			Recessive data				Multiplicative data			Dominant data			
	p =	N =	χ^2	MAX	Mult	Rec	χ^2	MAX	Mult	χ^2	MAX	Mult	Dom
RR_{AA} = 2	0.1	200	0.045	0.041	0.057	0.055	0.133	0.138	0.200	0.464	0.446	0.549	0.568
		600	0.121	0.122	0.078	0.157	0.390	0.419	0.492	0.924	0.933	0.950	0.960
		1000	0.184	0.189	0.097	0.240	0.608	0.643	0.712	0.993	0.995	0.996	0.998
	0.3	200	0.283	0.296	0.222	0.371	0.292	0.334	0.378	0.573	0.599	0.581	0.683
		600	0.711	0.731	0.541	0.804	0.727	0.759	0.816	0.971	0.976	0.963	0.987
		1000	0.914	0.924	0.764	0.954	0.922	0.940	0.960	0.999	0.999	0.998	1.000
	0.5	200	0.527	0.551	0.499	0.635	0.323	0.369	0.415	0.388	0.401	0.324	0.493
		600	0.956	0.964	0.925	0.980	0.774	0.812	0.853	0.860	0.872	0.743	0.918
		1000	0.998	0.998	0.992	0.999	0.945	0.959	0.973	0.980	0.983	0.923	0.992
RR_{AA} = 3	0.1	200	0.087	0.085	0.083	0.140	0.335	0.360	0.454	0.897	0.872	0.930	0.944
		600	0.307	0.323	0.156	0.416	0.824	0.850	0.895	1.000	1.000	1.000	1.000
		1000	0.511	0.525	0.232	0.628	0.968	0.975	0.985	1.000	1.000	1.000	1.000
	0.3	200	0.694	0.718	0.562	0.794	0.660	0.709	0.760	0.924	0.933	0.905	0.960
		600	0.994	0.996	0.958	0.998	0.989	0.992	0.995	1.000	1.000	1.000	1.000
		1000	1.000	1.000	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.5	200	0.929	0.940	0.899	0.963	0.685	0.732	0.782	0.727	0.742	0.580	0.820
		600	1.000	1.000	1.000	1.000	0.992	0.995	0.997	0.996	0.997	0.964	0.999
		1000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	1.000

Appendix continued.

			Recessive data				Multiplicative data			Dominant data			
	p =	N =	χ^2	MAX	Mult	Rec	χ^2	MAX	Mult	χ^2	MAX	Mult	Dom
RR_{AA} = 4	0.1	200	0.156	0.159	0.118	0.246	0.545	0.576	0.668	0.989	0.963	0.993	0.996
		600	0.545	0.569	0.274	0.676	0.969	0.976	0.987	1.000	1.000	1.000	1.000
		1000	0.800	0.813	0.422	0.881	0.999	0.999	1.000	1.000	1.000	1.000	1.000
	0.3	200	0.918	0.930	0.819	0.958	0.864	0.892	0.923	0.989	0.991	0.979	0.996
		600	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		1000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.5	200	0.994	0.995	0.987	0.998	0.874	0.901	0.929	0.878	0.888	0.716	0.932
		600	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.993	1.000
		1000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

REFERENCES

[1] Sasieni PD. From genotypes to genes: doubling the sample size. Biometrics 1997; 53(4): 1253-61. https://doi.org/10.2307/2533494

[2] Zheng G, Li Q, Yuan A. Some Statistical Properties of Efficiency Robust Tests with Applications to Genetic Association Studies. Scandinavian Journal of Statistics 2014; 41(3): 762-74. https://doi.org/10.1111/sjos.12060

[3] Freidlin B, Zheng G, Li ZH, Gastwirth JL. Trend tests for case-control studies of genetic markers: Power, sample size and robustness. Human Heredity 2002; 53(3): 146-52. https://doi.org/10.1159/000064976

[4] Gastwirth JL, Freidlin B. On power and efficiency robust linkage tests for affected sibs. Annals of Human Genetics 2000; 64: 443-53. https://doi.org/10.1046/j.1469-1809.2000.6450443.x

[5] Slager SL, Schaid DJ. Case-control studies of genetic markers: Power and sample size approximations for Armitage's test for trend. Human Heredity 2001; 52(3): 149-53. https://doi.org/10.1159/000053370

[6] Schaid DJ, Jacobsen SJ. Re: "Biased tests of association: Comparisons of allele frequencies when departing from Hardy-Weinberg proportions" - The authors reply. American Journal of Epidemiology 2001; 154(3): 287-8. https://doi.org/10.1093/aje/154.3.287

[7] Neuhauser M. Exact tests for the analysis of case-control studies of genetic markers. Human Heredity 2002; 54(3): 151-6. https://doi.org/10.1159/000068838

[8] Edland SD, Slager S, Farrer M. Genetic association studies in Alzheimer's disease research: challenges and opportunities. Statistics in Medicine 2004; 23(2): 169-78. https://doi.org/10.1002/sim.1706

[9] Tabor HK, Risch NJ, Myers RM. Candidate-gene approaches for studying complex genetic traits: practical considerations. Nature reviews Genetics 2002; 3(5): 391-7. https://doi.org/10.1038/nrg796

[10] Schaid DJ, Jacobsen SJ. Biased tests of association: Comparisons of allele frequencies when departing from Hardy-Weinberg proportions. American Journal of Epidemiology 1999; 149(8): 706-11. https://doi.org/10.1093/oxfordjournals.aje.a009878

[11] Knapp M. Re: "Biased tests of association: Comparisons of allele frequencies when departing from Hardy-Weinberg proportions". American Journal of Epidemiology 2001; 154(3): 287-8. https://doi.org/10.1093/aje/154.3.287

Received on 30-10-2016 Accepted on 07-05-2017 Published on 08-12-2017

https://doi.org/10.6000/1929-6029.2017.06.04.2