Testing for Genetic Association With Constrained Models Using Triads

James F. Troendle, **Kai F. Yu**, and **James L. Mills**[*,†]

Abstract

Wang and Sheffield (2005) showed that it is preferable to use a robust model that incorporated constraints on the genotype relative risk rather than rely on a model that assumes the disease operates in a recessive or dominant fashion. Wang and Sheffield's method is applicable to case-control studies, but not to family based studies of case children along with their parents (triads). We show here how to implement analogous constraints while analyzing triad data. The likelihood, conditional on the parents genotype, is maximized over the appropriately constrained parameter space. The asymptotic distribution for the maximized likelihood ratio statistic is found and used to estimate the null distribution of the test statistics. The properties of several methods of testing for association are compared by simulation. The constrained method provides higher power across a wide range of genetic models with little cost when compared to methods that restrict to a dominant, recessive, or multiplicative model, or make no modeling restriction. The methods are applied to two SNPs on the methylenetetrahy-drofolate reductase (MTHFR) gene with neural tube defect (NTD) triads.

Keywords

conditional distribution; genetic risk model; likelihood ratio test; power

1 Introduction

Two common problems in case-control genetic association studies are: (a) it is often difficult, a priori, to determine the proper model of risk (dominant, recessive, etc.) and (b) controls may not be a good match for the cases. Wang and Sheffield (2005) developed a method to deal with (a). They showed how likelihood methods can be used to fit models that only assume a monotone relationship between the genotypic relative risks without needing to correctly specify the exact risk model. The use of triads (case, mother, father) avoids the problems of control selection, (b). However, Wang and Sheffield's method does not apply to studies based on triads.

Triad studies are important because the results are more robust than those of case-control studies. This is because case-control studies require the case and control groups to be representative of the diseased and non-diseased population. There are many ways that the case or control groups can be skewed. Ascertainment and selection bias are two of the most common (Schlesselman, 1982). Another problem for case-control studies is confounding due to population stratification. Case control studies often face the problem that several ethnic/racial groups are studied. If the case group contains different proportions of the ethnic

groups than the control group, false positive results may appear. That is to say that differences between the groups that are actually due to ethnic differences in gene frequencies may be mistakenly considered to be risk factors for disease. While race is an obvious example of the problems in matching cases to controls, other more subtle differences may be difficult, or even impossible, to detect. If in a genetic association study, the population consists of strata with different allele frequencies and different disease risks, then the cases and controls will appear different in allele frequency if the stratification variable is not controlled for (Lee and Wang, 2008). Triad studies eliminate the danger that ethnic, or other non-disease related, case control differences will mistakenly be interpreted as risk factors for the disease being studied because the comparisons in triad studies are made between the mathematically expected transmission rates and those observed in the case families. Triad studies have the additional advantages on the practical level that it is often simpler to identify case parents than to find and match appropriate controls. Identifying case mothers has the benefit that in pregnancy studies possible maternal risk factors can be investigated as well. In this paper, we develop a method of testing for use in triad studies that is based on the same model considered by Wang and Sheffield. We show that in triad studies likelihood methods can be used to fit models that only assume a monotone relationship between the genotypic relative risks, thereby yielding tests that are powerful across a wide range of genetic risk models.

Powerful methods to test for association between disease and the number of risk alleles using triads are obtained by specifying a genetic risk model and using the likelihood ratio test (conditional on the parental mating genotype) for the null hypothesis that all genotypes have the same risk. Examples are the transmission/disequilibrium test (TDT) of Spielman (1993), the unrestricted likelihood ratio test conditional on the parental genes of Schaid and Sommer (1993), and tests based on a dominant or recessive genetic risk model. The unrestricted test is the most general and uses two parameters (ψ_1 and ψ_2) to model the genotypic relative risk for one or two copies of the risk allele compared to no copies. The null hypothesis of no genetic association is $\psi_1 = \psi_2 = 1.0$. The unrestricted test makes no restriction on the model of genetic risk. The TDT uses a multiplicative model, where risk of two copies is the square of the risk of one copy ($\psi_2 = \psi_1^2$). A dominant model is obtained by setting $\psi_1 = \psi_2$ and a recessive model is obtained by setting $\psi_1 = 1.0$. The dominant and recessive models are quite powerful when those models are correctly specified, but suffer greatly when the true model is different than that used in likelihood ratio test construction.

A monotone restricted model is obtained by enforcing that either $1.0 \leq \psi_1 \leq \psi_2$ or $\psi_2 \leq \psi_1 \leq 1.0$. Although this restriction is not true for all genetic risk models, there are relatively few known exceptions and the investigator may know ahead of time if the true model has the potential to contradict this restriction. As will be seen, the power gained by using a model that uses this restriction may make it worthwhile even for investigators uncertain about its truth. Moreover, the test based on the monotone restriction is more robust to that assumption than the TDT, which is based on a model even more restrictive ($\psi_2 = \psi_1^2$ implies either $1.0 \leq \psi_1 \leq \psi_2$ or $\psi_2 \leq \psi_1 \leq 1.0$).

The paper proceeds as follows. Section 2 reviews the previous methods of association testing with triads and describes the monotone constrained likelihood method. Section 3 presents simulations comparing the type I error rates and power under various true disease models and risk allele proportions. Section 4 applies the tests to two SNPs on the MTHFR gene with NTD triads. Finally, section 5 contains recommendations.

2 Genetic Association Tests Using Triads

Assume that one is interested in testing for association between disease and the number of risk alleles at a single bi-allelic locus. It is not important which allele actually confer risk; the name risk allele is merely for reference. Triad data can be represented as an ordered triple, (M, F,C), where

M = # of risk alleles held by the mother
F = # of risk alleles held by the father
C = # of risk alleles held by the case child.

Each of the methods considered in this paper can be written as a likelihood ratio test (LRT) for a certain parameterization of the likelihood conditional on the mating type (MT), specified by knowledge of M and F without knowledge of which is which. Thus, MT is a set of possible (M, F). The derivation of the distribution of case genotype given MT is given in Schaid and Sommer (1993) and summarized in Table I.

The conditional log-likelihood is

$$\mathscr{L}(\psi_1,\psi_2) \approx n_{22} \log\left(\frac{\psi_2}{\psi_1+\psi_2}\right)+n_{21} \log\left(\frac{\psi_1}{\psi_1+\psi_2}\right)+n_{42} \log\left(\frac{\psi_2}{\psi_2+2\psi_1+1}\right)+n_{41} \log\left(\frac{2\psi_1}{\psi_2+2\psi_1+1}\right)$$
$$+n_{40} \log\left(\frac{1}{\psi_2+2\psi_1+1}\right)+n_{51} \log\left(\frac{\psi_1}{\psi_1+1}\right)+n_{50} \log\left(\frac{1}{\psi_1+1}\right)$$

(1)

The unrestricted test of Schaid and Sommer (1993) is a LRT based on (1). Maximization of (1) can be achieved by numerical means quite reliably. Finding a numerical solution is facilitated by noticing that the second score equation leads to a quadratic equation in ψ_2. Thus, ψ_2 can effectively be eliminated from the numerical search by considering for any give ψ_1 the two possible roots for ψ_2. This enables use of reliable univariate root finding algorithms to solve for the maximum likelihood estimates (MLE) of (1). Let $\hat{\psi}_1$ and $\hat{\psi}_2$ be the MLE of ψ_1 and ψ_2. The test statistic for a LRT of association is then

$$UNR=2\left\{\mathscr{L}\left(\widehat{\psi}_1,\widehat{\psi}_2\right)-\mathscr{L}(1,1)\right\}.$$

The *UNR* test statistic is asymptotically distributed as chi-squared with 2 degrees of freedom (DF) under the null hypothesis (Kendall and Stuart, 1973). A critical value for a 5% level test is therefore the 95% quantile of the chi-squared distribution with 2 DF.

If a dominant disease model is desired, one imposes the restriction $\psi_1 = \psi_2 = \psi$ The resulting likelihood, $\mathscr{L}_D(\psi)$, is then a function of only one parameter, and maximization is achieved in closed form (the solution is a root of a quadratic equation). Let $\tilde{\psi}$ be the maximizer of $\mathscr{L}_D$. The test statistic for a LRT of association is then

$$DOM=2\left\{\mathscr{L}_D\left(\widehat{\psi}\right)-\mathscr{L}_D(1)\right\}.$$

The *DOM* test statistic is asymptotically distributed as chi-squared with 1 DF under the null hypothesis. A critical value for a 5% level test is therefore the 95% quantile of the chi-squared distribution with 1 DF.

If a recessive disease model is desired, one imposes the restriction $\psi_1 = 1$. The resulting likelihood, $\mathscr{L}_R(\psi_2)$, is then a function of only one parameter, and maximization is achieved

in closed form (the solution is a root of a quadratic equation). Let $\bar{\psi}_2$ be the maximizer of $\mathcal{L}_R$. The test statistic for a LRT of association is then

$$REC = 2\left\{\mathcal{L}_R\left(\tilde{\psi}_2\right) - \mathcal{L}_R(1)\right\}.$$

The *REC* test statistic is asymptotically distributed as chi-squared with 1 DF under the null hypothesis. A critical value for a 5% level test is therefore the 95% quantile of the chi-squared distribution with 1 DF.

If a multiplicative model is desired, one may impose the restriction $\psi_2 = \psi_1^2$. This model has the nice feature that the effect of the risk allele is linear on the log scale, analogous to what one would get in a case-control analysis when using a logistic regression model with a continuous term for the number of risk alleles. The resulting likelihood, $\mathcal{L}_{MULT}(\psi_1)$, is then a function of only one parameter, and maximization is achieved in closed form. Let ψ_1 be the maximizer of $\mathcal{L}_{MULT}$. The test statistic for a LRT of association is then

$$MULT = 2\left\{\mathcal{L}_{MULT}\left(\tilde{\psi}_1\right) - \mathcal{L}_{MULT}(1)\right\}.$$

This test is equivalent to the TDT test of Spielman (1993). The *MULT* test statistic is asymptotically distributed as chi-squared with 1 DF under the null hypothesis. A critical value for a 5% level test is therefore the 95% quantile of the chi-squared distribution with 1 DF.

If a model is desired that assumes a monotone risk relationship, one may impose the restriction that either

$$1.0 \leq \psi_1 \leq \psi_2 \text{ or } \psi_2 \leq \psi_1 \leq 1.0. \tag{2}$$

This model has the flexibility to be correctly specified in cases between dominant, recessive, or multiplicative models, without allowing for unlikely non-monotone patterns. To obtain a LRT for association under this model, one must obtain restricted MLE of $\mathcal{L}(\psi_1, \psi_2)$. The possible maximizers of $\mathcal{L}$ under the restriction (2) are the critical points for the unrestricted problem or those from either the dominant or recessive problems described above. Thus, it is quite easy to compare the likelihood (1) at the critical points from the dominant and recessive problems to the likelihood at the maximizer of the unrestricted problem. The maximizer of the monotone restricted problem, $(\breve{\psi}_1, \breve{\psi}_2)$, is the critical point with maximum likelihood. The test statistic for a LRT of association is then

$$MONO = 2\left\{\mathcal{L}\left(\breve{\psi}_1, \breve{\psi}_2\right) - \mathcal{L}(1,1)\right\}.$$

The asymptotic distribution of the *MONO* test statistic under the null hypothesis is derived in the appendix.

3 Simulations

Monte Carlo simulations were used to assess the type I error rate and power of the testing procedures. Triads were generated assuming Hardy-Weinberg equilibrium in the population and random mating. One obtains the genotype for a given triad (*M, F, C*) according to

$$Pr\left[M{=}x,\ F{=}y,\ C{=}z\right] \quad = \quad Pr\left[M{=}x\right]\cdot Pr\left[F{=}y\right]$$
$$\cdot \quad Pr\left[C{=}z\,\middle|\,MT{=}\{(x,y),(y,x)\}\right],$$

where the first two terms are genotype frequency functions and the last term is given as $Pr\{C|MT\}$ in Table I. For reference, let p denote the frequency of the risk allele in the population.

Table II shows the results of simulations under the null hypothesis ($\psi_1 = \psi_2 = 1.0$) for testing at 5% significance level. The frequency of the risk allele was taken as .2 or .5. Each simulation used 100 or 500 triads and consisted of 100,000 replications, yielding a simulation error of $\pm 0.14\%$. The results indicate that the tests are all controlling the type I error at the desired level except for a few exceptions with $n = 100$ and $p = .2$. In this case, the method based on a recessive model is conservative and the unrestricted method has an inflated significance level. This departure from the target significance level does not greatly affect the conclusions of the power comparisons below as the unrestricted method is dominated by the monotone restricted test in the region of highest interest.

Table III shows the results of the simulations under four different alternative models. Each simulation used 100 or 500 triads and consisted of 10,000 replications, yielding a simulation error bounded by $\pm 1.0\%$. The first three alternative configurations correspond to dominant, recessive, and multiplicative disease inheritance patterns. The fourth alternative was chosen as an example of a non-monotone risk pattern, where $\psi_2 = 1/\psi_1$. The alter-native configurations were chosen so that the highest power among the methods included in Table III was approximately 80%. One can see that the methods based on dominant, recessive, and multiplicative models do indeed have maximal power when those models are correctly specified. However, all three methods suffer extreme power loss if the model is mis-specified. One can also see that the power advantage of the monotone restricted method is greater in cases favoring the monotone restricted method than the power advantage of the multiplicative method in cases favoring the multiplicative method. Furthermore, in cases where the true model is multiplicative the monotone restricted method compares favorably with the unrestricted method, and this is a more likely occurrence in actual biological applications than a non-monotone model is. Also, simulation (results not shown here) at other points inside the monotone region but different than the multiplicative model show that the monotone restricted method has higher power than the unrestricted method across the entire monotone interval.

4 Examples

We illustrate the testing methods by applying them to two SNPs of the 5,10-methylenetetrahydrofolate reductase (MTHFR) gene obtained for neural tube defect (NTD) triads in the Republic of Ireland. The MTHFR variant (677C→T) has been reported to be a risk factor for NTDs in many populations including the Irish (Parle-McDermott et al., 2003; Kirke et al., 2004; Botto and Yang, 2000). Another variant (1298A→C) of the MTHFR gene has also been reported as a risk factor for NTDs, although the association is not consistent. Since the SNPs are in linkage disequilibrium, if 677C→T is a risk factor for NTDs then 1298A→C will appear to be as well to the extent that the two are in linkage disequilibrium.

The NTD triads were recruited through various branches of the Irish Association for Spina Bifida and Hydrocephalus. For this comparison we genotyped 449 complete triads for 677C→T and 450 complete triads for 1298A→C. The numbers of informative triads are given in Table IV along with the p-values from the various tests of association.

All of the tests of association for variant 677C→T except the test based on a dominant model give a moderate level of evidence. The unrestricted maximum likelihood estimates of the genotype relative risks are $\hat{\psi}_1 = 1.13$ and $\hat{\psi}_2 = 1.61$. The evidence for association seems to come primarily from the triads of MT=4, where both parents are heterozygotes. Cases from this MT yield 24 who inherited both T alleles and only 11 who inherited both C alleles, whereas without association one would expect equal numbers. This is consistent with the dominant model giving the least evidence for association. Kirke et al. (2004) found from case-control study of the same population somewhat higher risks (odds ratio of disease with CT of 1.52 and odds ratio of disease with TT of 2.56).

The tests of association for variant 1298A→C are less consistent across models. The unrestricted maximum likelihood estimates of the genotype relative risks are $\hat{\psi}_1 = 0.67$ and $\hat{\psi}_2 = 0.72$. The evidence for association seems to come primarily from the triads of MT=2 and MT=5, where one parent is heterozygous and one parent homozygous. In this case however, if the homozygous parent has allele C, more cases are found who inherited the C allele from the heterozygote (16 to 9), whereas if the homozygous parent has allele A, more cases are found who inherited the A allele from the heterozygote (58 to 36). This creates the estimated non-monotone risk pattern, leading to similar p-values for all of the tests except for the test based on a recessive model which yields virtually no evidence of association. The non-monotone pattern of risk does not seem biologically plausible for 1298A→C, and as stated before the apparent association with NTDs seems to be coming through its linkage disequilibrium with 677C→T.

5 Discussion

We have shown how to implement a restricted model based on a monotone relationship for the genotype relative risks in genetic association testing using triads. The appendix derives an approximate asymptotic distribution under the null hypothesis for the test statistic of this new test. The resulting procedure is quite powerful across all commonly encountered genetic risk models. Furthermore, the procedure has more power than tests based on more restrictive models, even outside of the monotone risk setting. Moreover, tests based on assuming a dominant, recessive, or multiplicative risk model suffer substantial power loss when the model is incorrectly specified. We therefore recommend the monotone restricted test for genetic association studies using triads.

Triad studies do not eliminate all of the problems seen in case control studies. Genotyping errors can still produce false positive results (Morris and Kaplan, 2004); however, in triad studies they are more likely to be identified because of impossible inheritance patterns. Those subjects can then be dropped. Cases who are phenocopies will make both triad and case control studies less likely to identify risk factors because they dilute out the population with the true genetic risk factor. The same is true when the etiology of the disease is heterogeneous. If the causes are genetic, a large study may be able to overcome the problem and identify several true associations. If the causes are not all genetic, however, the inclusion of non-genetic cases will reduce the power in both triad and case control studies. Despite the fact that triad studies cannot overcome all the limitations of case control studies, the monotone restricted test has important advantages and is recommended for triad studies of genetic association.

References

1. Wang K, Sheffield VC. A Constrained Likelihood Approach to Marker-Trait Association Studies. Am J Hum Genet. 2005; 77:768–780. [PubMed: 16252237]

2. Schlesselman, JJ. Case-Control Studies: Design, Conduct, Analysis. New York: Oxford University Press; 1982.

3. Lee W-C, Wang L-Y. Simple Formulas for Gauging the Potential Impact of Population Stratification Bias. Am J Epi. 2008; 167:86–89.

4. Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and Insulin-Dependent Diabetes Mellitus (IDDM). Am J Hum Genet. 1993; 52:506–516. [PubMed: 8447318]

5. Schaid DJ, Sommer SS. Genotype relative risks: methods for design and analysis of candidate-gene association studies. Am J Hum Genet. 1993; 53:1114–1126. [PubMed: 8213835]

6. Kendall, MG.; Stuart, A. The Advanced Theory of Statistics: Volume 2, Inference and Relationship. Third Edition. Hafner Publishing Company: New York; 1973.

7. Parle-McDermott A, Mills JL, Kirke PN, O'Leary VB, Swanson DA, Pangilinan F, Conley M, Molloy AM, Cox C, Scott JM, Brody LC. Analysis of the MTHFR 1298A→C and 677C→T polymorphisms as risk factors for neural tube defects. J Hum Genet. 2003; 48:190–193. [PubMed: 12730722]

8. Kirke PN, Mills JL, Molloy AM, Brody LC, O'Leary VB, Daly L, Murray S, Conley M, Mayne PD, Smith O, Scott JM. Impact of the MTHFR C677T polymorphism on risk of neural tube defects: case-control study. BMJ. 2004; 328:1535–1536. [PubMed: 15155469]

9. Botto LD, Yang Q. 5,10-Methylenetetrahydrofolate reductase gene variants and congenital anomalies: a HuGE review. Am J Epi. 2000; 151:862–877.

10. Morris RW, Kaplan NL. Testing for Association With a Case-Parents Design in the Presence of Genotyping Errors. Gen Epi. 2004; 26:142–154.

11. van der Vaart, AW. Asymptotic Statistics. Cambridge: Cambridge University Press; 1998.

6 Appendix: Asymptotic Distribution of CPGMONO

6.1 Conceptual Experiment

Let k_2, k_4 and k_5 be positive integers. Assume ψ_1 and ψ_2 are constants. Define

$$p_{21} = \psi_1 / (\psi_1 + \psi_2), \ p_{40} = 1 / (2\psi_1 + \psi_2 + 1), \ p_{41} = 2\psi_1 / (2\psi_1 + \psi_2 + 1), \ p_{50} = 1 / (\psi_1 + 1).$$

Let X_{21}, X_{40}, X_{41} and X_{50} have marginal distributions

$$P[X_{21} = k_{21}] = \frac{k_2!}{k_{21}!(k_2 - k_{21})!} p_{21}^{k_{21}} (1 - p_{21})^{(k_2 - k_{21})}$$

$$P[X_{40} = k_{40}, X_{41} = k_{41}] = \frac{k_4!}{k_{40}! k_{41}! (k_4 - k_{40} - k_{41})!} p_{40}^{k_{40}} p_{41}^{k_{41}} (1 - p_{40} - p_{41})^{(k_4 - k_{40} - k_{41})}$$

$$P[X_{50} = k_{50}] = \frac{k_5!}{k_{50}!(k_5 - k_{50})!} p_{50}^{k_{50}} (1 - p_{50})^{(k_5 - k_{50})},$$

and their joint distribution be

$$P[X_{21} = k_{21}] P[X_{40} = k_{40}, X_{41} = k_{41}] P[X_{50} = k_{50}].$$

This statistical experiment can also be reparametrized by $\theta_1 = \log(\psi_1)$ and $\theta_2 = \log(\psi_2)-\log(\psi_1)$. With this θ-parametrization, the Fisher information matrix evaluated at $\theta_1 = \theta_2 = 0$, is

$$I_0 = \frac{1}{16} \begin{pmatrix} 8k_2+7k_4+4k_5 & k_4 \\ k_4 & 4k_2+3k_4 \end{pmatrix}.$$

If there are K replicates, the total information is KI_0. To test $H_0 : \theta_1 = \theta_2 = 0$ versus $H_1 : \theta_1\theta_2 \geq 0$ with $\max(|\theta_1|, |\theta_2|) > 0$, we consider the log-likelihood ratio statistic. By Theorem 16.7 of van Der Vaart (1998), under the null hypothesis the asymptotic distribution of 2(log-likelihood ratio statistic) is that of T, where

$$T = X^t I_0 X - \inf_{\theta \in H_1} (X - \theta)^t I_0 (X - \theta),$$

(3)

and X is normally distributed with mean vector zero and covariance matrix I_0^{-1}.

6.2 Replications of the Conceptual Experiment

For any positive constant K, let $I_{K0} = KI_0$; then the distribution of T_K is the same as that of T, where

$$T_K = Y^t I_{K0} Y - \inf_{\theta \in H_1} (Y - \theta)^t I_{K0} (Y - \theta)$$

and Y is normally distributed with mean vector zero and covariance matrix I_{K0}^{-1}.

6.3 The Actual Experiment

Let p_i be the probability of $MT = i$, $i = 1, \ldots, 6$. Choose m large enough so that the probabilities p_i can be represented to a satisfactory degree in a sample of size m. Conceptually subdivide the n triads into $K = n/m$ subsets of size m as follows. In each subset, if possible, put $k_i = p_i m$ triads of mating type i, $i = 1, \ldots, 6$. With n small this may not be possible, but as $n \to \infty$, we will be able to do this for a fraction of the subsets approaching 1. Each subset can be viewed as having fixed numbers of each MT as described in 6.1. Thus, according to 6.2, the asymptotic distribution of the CPG_{MONO} test statistic under the null hypothesis is approximately that of T in (3) where k_i is replaced by n_i, $i = 2, 4, 5$.

The distribution of T is a mixture of a chi-squared with 2 DF and the maximum of the squared components of a bivariate normal with zero means, unit variances, and correlation ρ.

The mixing probability of the chi-squared is $\cos^{-1}(\rho)/\pi$, and ρ is the correlation from I_0^{-1}. This correlation can be found as

$$\rho = \frac{-n_4}{\sqrt{(4n_2+3n_4)(8n_2+7n_4+4n_5)}}.$$

A critical value for a 5% level test is therefore the 95% quantile of the distribution of T.

Table I

Notation and Probability of Case Genotype Conditional on MT

MT	MT Number	Number of Cases	C	$Pr\{C \mid MT\}$
$\{(2, 2)\}$	1	n_{12}	2	1
$\{(2, 1), (1, 2)\}$	2	n_{22}	2	$\dfrac{\psi_2}{\psi_1+\psi_2}$
$\{(2, 1), (1, 2)\}$	2	n_{21}	1	$\dfrac{\psi_1}{\psi_1+\psi_2}$
$\{(2, 0), (0, 2)\}$	3	n_{31}	1	1
$\{(1, 1)\}$	4	n_{42}	2	$\dfrac{\psi_2}{\psi_2+2\psi_1+1}$
$\{(1, 1)\}$	4	n_{41}	1	$\dfrac{2\psi_1}{\psi_2+2\psi_1+1}$
$\{(1, 1)\}$	4	n_{40}	0	$\dfrac{1}{\psi_2+2\psi_1+1}$
$\{(1, 0), (0, 1)\}$	5	n_{51}	1	$\dfrac{\psi_1}{\psi_1+1}$
$\{(1, 0), (0, 1)\}$	5	n_{50}	0	$\dfrac{1}{\psi_1+1}$
$\{(0, 0)\}$	6	n_{60}	0	1
Total		n		

Table II

Type I error (%) of the tests.

n	p	UNR	DOM	REC	MULT	MONO
100	.2	5.60	5.07	4.06	5.04	4.76
	.5	5.22	5.15	5.21	5.08	4.90
500	.2	5.12	4.96	5.19	5.04	4.76
	.5	5.00	4.99	4.99	5.06	4.71

Table III

Power (%) of the tests for nonnull configurations.

n	p	ψ_1	ψ_2	UNR	DOM	REC	MULT	MONO
100	.2	2.2	2.2	72.1	80.8	6.0	70.3	71.7
		1.0	3.6	70.6	6.8	79.4	43.0	70.6
		1.9	3.61	70.8	68.5	39.1	79.5	73.7
		0.5	2.0	79.5	41.8	57.5	10.4	61.4
	.5	2.6	2.6	68.2	78.4	6.4	44.1	68.0
		1.0	2.2	71.8	6.9	80.5	63.9	71.8
		1.8	3.24	73.9	49.0	66.6	82.6	76.9
		0.65	1.54	79.0	12.1	78.2	30.7	68.2
500	.2	1.45	1.45	75.4	83.9	6.2	76.2	75.4
		1.0	2.0	74.8	6.2	83.0	39.1	74.6
		1.35	1.82	70.3	70.2	33.3	79.7	73.3
		0.73	1.37	81.6	53.4	55.0	14.6	65.1
	.5	1.5	1.5	71.9	81.2	6.7	52.9	71.9
		1.0	1.45	73.6	7.6	82.5	61.9	73.8
		1.3	1.69	74.9	55.5	63.9	83.7	77.5
		0.82	1.22	81.2	15.5	80.1	26.0	71.0

Table IV

Triad numbers and p-values for MTHFR variants.

Variant	n_{22}	n_{21}	n_{21}	n_{41}	n_{40}	n_{51}	n_{50}	UNR	DOM	REC	MULT	MONO
(677C→T)	27	21	24	34	11	60	59	0.12	0.30	0.05	0.06	0.10
(1298A→C)	16	9	9	27	15	36	58	0.07	0.02	0.95	0.07	0.06