# allelic test in case-control association studies.

**Mickaël Guedj, E. Della-Chiesa, K. Forner, J. Wojcik, G. Nuel**

*guedj@genopole.cnrs.fr*

A popular method to analyze case-control association data is to compare allele frequencies using the Pearson statistic. The $p$-value is computed comparing the value of the statistic to a one degree-of-freedom chi-square distribution. However this strategy has been shown not to be always appropriate, in particular when the Hardy-Weinberg equilibrium (HWE) does not hold in the combined population of cases and controls, what is restrictive. Two alternatives have been proposed: the **Cochran-Armitage test for trends** and an **unbiased allelic test**. Since the **biased allelic test** is still widely used in the community, we illustrate its pitfalls in the context of genome-wide association studies and particularly in the case of low-level tests. To guide the choice between the two alternatives available, we propose to study their **power** in various genetic scenarios and focus on situations where genotype frequencies does not comply with HWE that has not been exhaustively studied and hence deserve further attention.

## Allele-based tests

Simple-marker strategies treat one marker at a time. They are generally used as a first step in the analysis process. Individuals are organized into contingency tables according to their marker and disease status; various approaches are then proposed to test for association based on either genotypes or alleles.

| | $ss$ | $sS$ | $SS$ | total |
|---|---|---|---|---|
| diseased | $D_0$ | $D_1$ | $D_2$ | $n_D$ |
| control | $C_0$ | $C_1$ | $C_2$ | $n_C$ |
| total | $n_0$ | $n_1$ | $n_2$ | $n$ |

| | $s$ | $S$ | total |
|---|---|---|---|
| | $2D_0 + D_1$ | $2D_2 + D_1$ | $2n_D$ |
| | $2C_0 + C_1$ | $2C_2 + C_1$ | $2n_C$ |
| total | $2n_0 + n_1$ | $2n_2 + n_1$ | $2n$ |

### The biased allelic test

It aims at identifying significant differences in allelic frequencies between cases and controls. Alleles are supposed to be sampled from the case and control populations according to binomial distributions:

$$2D_0 + D_1 \sim \mathcal{B}(2n_D, p_D),$$
$$2C_0 + C_1 \sim \mathcal{B}(2n_C, p_C),$$

where $p_D$ and $p_C$ correspond to the proportions of the susceptibility allele in the case and control populations respectively. Under the null hypothesis of no association, alleles are supposed to be sampled from the same general population and $p_D$ and $p_C$ to be equal to $p$. To test this hypothesis we classically consider the statistic:

$$\mathcal{Z}_A = \frac{\hat{p}_D - \hat{p}_C}{\sqrt{N\hat{p}(1-\hat{p})}} \sim \mathcal{N}(0,1),$$

with $\hat{p}_D = \frac{2D_2+D_1}{2n_D}$, $\hat{p}_C = \frac{2C_2+C_1}{2n_C}$ and $\hat{p} = \frac{2n_2+n_1}{2n}$, or the equivalent Pearson statistic applied to the allelic table:

$$\mathcal{S}_A = (\mathcal{Z}_A)^2 \sim \chi^2(1).$$

The validity of this test has been discussed (Sasieni 1997): the $H_0$ hypothesis makes actually the further assumption that alleles are sampled independently. If it is the case for genotypes, it is however clear that alleles are not since they are sampled by two at a time. Consequently, the way alleles are paired (corresponding to a departure from HWE) deviates $\mathcal{S}_A$ from its wrongly supposed $H_0$ distribution and hence dramatically biased the estimation of the $p$-value.

### The unbiased allelic test

This test is also based on the Pearson statistic ($\mathcal{S}_A$) but on the multinomial sampling of genotypes instead of alleles taken independently, what is more consistent with the reality:

$$(D_0, D_1, D_2) \sim \mathcal{M}(n_D, p_{D0}, p_{D1}, p_{D2}),$$
$$(C_0, C_1, C_2) \sim \mathcal{M}(n_C, p_{C0}, p_{C1}, p_{C2}).$$

### The trend test

It is based on the genotypic contingency table. In practice, the trend statistic $\mathcal{S}_T$ measures a linear trend in proportions weighted by a dose effect score $x_i$ associated to each column of the contingency table. When $x_i$ quantifies the number of high-risk allele, this test is equivalent to the score test in the logistic regression model where each SNP is coded according to its count of high-risk allele. In our case:

$$\mathcal{S}_T = \frac{n.\left[n.(D_1 + 2D_2) - n_D.(n_1 + 2n_2)\right]^2}{n_D n_C.\left[n.(n_1 + 4n_2) - (n_1 + 2n_2)^2\right]} \sim \chi^2(1)$$

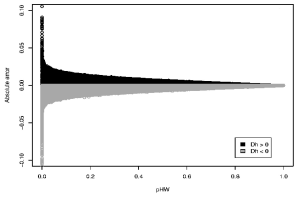
## Consequences of using the biased test

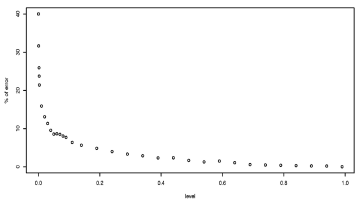
### Data

We applied the biased and unbiased exact tests on genome-wide data concerning the multiple sclerosis. The data set consists in 66,990 SNPs. DNA from 279 patients and 301 controls were genotyped with this marker set using the 100K Affymetrix chip. The algorithm used for making genotype calls has been previously described by Affymetrix.

### Results

Figures display **(i)** the impact on the p-value computation with respect to the strengh of departure from HWE and **(ii)** the impact on predictions with respect to the level of the test.



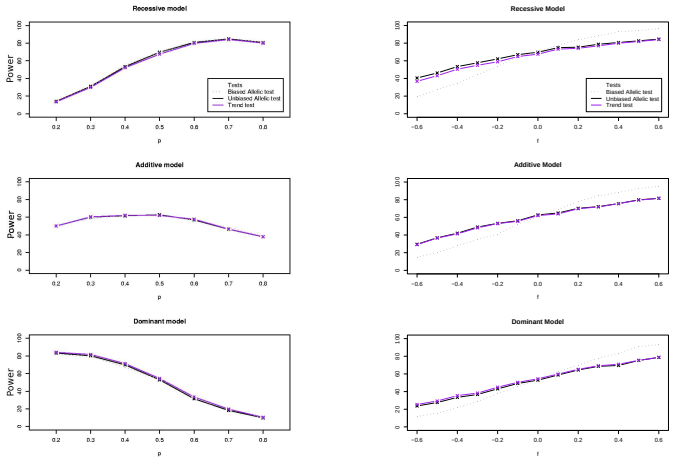




## Which alternative

### Computing power

Power is estimated using Monte-Carlo simulations. Each estimate is done on the basis of 10,000 simulations first with respect to the susceptibility allele frequency ($p$), and then according to the coefficient of consanguinity ($f$) introduce in the general population. All situations are considered for a prevalence $K_p = 0.05$, and three Mode Of Inheritance (recessive, additive and dominant).

### Results




## Conclusions

- The classical allelic test is biased by the allele matching (Sasieni 1997).
- An excess of heterozygotes ($f < 0$) decreases the type-I error-rate leading to a more conservative test whereas a deficit inflates it (Schaid and Jacobsen 1999).
- The impact on predictions can be substantial since the threshold chosen to decide for the acceptance or the rejection of the $H_0$ hypothesis is generally low, commonly set to 5 or 1% and even lower when one takes the multiple-testing into account.
- Besides the impact on predictions, let us add that the exactness of $p$-values can be really important when one applies methods based on $p$-value distributions such as FDR-based approaches.
- Two alternatives have been proposed and the three are strictly identical when HWE holds in the combined population.
- The three are strictly equivalent when HWE holds in the combined population.
- Despites small differences for recessive and dominant models, the trend and unbiased tests present comparative efficiency.

## References

[1] Sasieni (1997). From genotypes to genes: doubling the sample size. *Biometrics*, 53, 1253–1261.

[2] Schaid and Jacobsen (1999). Biased tests of association: comparisons of allele frequencies when departing from Hardy-Weinberg proportions. *Am. J. of Epidemiology*, 149, 706–711.

[3] Guedj, Wojcik, Nuel and Forner. A fast, unbiased and exact allelic test for case-control association studies. [submitted]

[4] Guedj, Della-Chiesa, Nuel. On the power of allele-based tests in case-control association studies. [submitted]