European Journal of Human Genetics (2006) 14, 1223–1230

www.nature.com/ejhg

ARTICLE

A test of homogeneity of Hardy-Weinberg disequilibrium across strata

Xiao-Lin Yin[1], Wen-Qing Ma[1], Man-Lai Tang[2] and Jianhua Guo*,[1]

[1]*Key Laboratory for Applied Statistics of MOE, School of Mathematics and Statistics, Northeast Normal University, ChangChun, China;* [2]*Department of Mathematics, Hong Kong Baptist University, Hong Kong, China*

For genotype data being sampled from several strata with different allele frequencies, it is necessary to verify the assumption of homogeneity of Hardy–Weinberg disequilibrium across strata before testing Hardy–Weinberg law across strata. In practice, disequilibrium can be measured via fixation coefficients (ie, ratios of genotypic frequencies) or disequilibrium coefficients (ie, differences of genotypic frequencies). Test for homogeneity of Hardy–Weinberg disequilibrium using data from several populations has been derived according to fixation coefficients. In this article, using the likelihood score theory extended to nuisance parameters, we derive a homogeneity score test for comparing disequilibrium coefficients across several independent strata. Simulation results demonstrate that the homogeneity score test performs satisfactorily in the sense that its empirical size seldom exceeds the pre-chosen nominal level by more than 10% even for small sample sizes. Corresponding power and sample size formulae are provided as well. We illustrate our test with a real glyoxalase genotype data set.
European Journal of Human Genetics (2006) **14,** 1223–1230. doi:10.1038/sj.ejhg.5201689; published online 26 July 2006

Keywords: disequilibrium coefficient; Hardy–Weinberg law; homogeneity; score test

Introduction

The law of Hardy–Weinberg equilibrium (HWE) states that in a large random mating population that is not affected by the evolutionary processes of mutation, migration, or selection, both the allele frequencies and the genotype frequencies are constant from generation to generation.[1,2] Furthermore, the genotype frequencies are related to the allele frequencies by the square expansion of those allele frequencies. In other words, the law of HWE states that under a restrictive set of assumptions, it is possible to calculate the expected frequencies of genotypes in a population if the frequency of the different alleles in a population is known. The original descriptions of HWE become an important landmark in the history of population genetics,[3] and it is now a common practice to verify whether observed genotypes conform to Hardy–Weinberg expectations.[4,5]

In a diallelic locus with alleles A_1 and A_2 across K strata, let the genotypic array of the kth ($k=1, \ldots, K$) stratum be

$$p_{11k}A_1A_1 + p_{12k}A_1A_2 + p_{22k}A_2A_2$$

Let p_k be the allelic frequency of A_1 in the kth stratum and $q_k = 1-p_k$ ($k=1, \ldots, K$). Populations with genotypic frequencies satisfying $p_{11k}=p_k^2$, $p_{12k}=2p_kq_k$, and $p_{22k}=q_k^2$ ($k=1, \ldots, K$) are said to be in HWE at the locus under consideration. In studies of HWE, there are two widely used coefficients, namely the fixation and disequilibrium coefficients.[6] For stratum k ($k=1, \ldots, K$), the fixation and disequilibrium coefficients are defined by $f_k = 1 - p_{12k}/(2\sqrt{p_{11k}p_{22k}})$ and $D_k = p_kq_k - p_{12k}/2$, respectively. Hence, the problem of testing HWE when individuals are sampled from several strata is equivalent to testing one of the following hypotheses:

$$\begin{aligned} &\mathrm{H}_0' : \theta_k = 0 \quad \text{for all } k = 1, \ldots, K \quad \textit{versus} \\ &\mathrm{H}_1' : \theta_k \neq 0 \quad \text{for some } k, \end{aligned} \tag{1}$$

where $\theta_k = f_k$ or D_k. For statistical tests based on disequilibrium coefficient, one can refer to the work of Haldane[7] and Smith.[8]

*Correspondence: Professor J Guo, School of Mathematics and Statistics, Northeast Normal University, 5268 Renmin Street, ChangChun 130024, Jilin Province, P.R. China.
Tel: +86 431 5098576; Fax: +86 431 5098237;
E-mail: jhguo@nenu.edu.cn
Received 13 January 2006; revised 29 May 2006; accepted 1 June 2006; published online 26 July 2006

For test procedures based on functions of fixation coefficients (eg, $(1-f_k)^2$), one can consult the work of Emigh,[9] Troendle and Yu,[10] and Nam.[11]

It is noteworthy that any statistical procedure for testing the null hypothesis in (1) assumes that the measure of disequilibrium (ie, θ_k) is constant across the strata. In this regard, it is important that one should consider testing the assumption of homogeneity of the measure of disequilibrium across strata before any testing of the null hypothesis in (1). For this purpose, we consider the following hypotheses:

$$\mathrm{H}_0 : \theta_1 = \ldots = \theta_k \quad \textit{versus} \quad \mathrm{H}_1 : \text{Not all } \theta_k' \text{s are equal} \qquad (2)$$

Olson and Foley[12] proposed a large-sample test and an exact test for verifying the null hypothesis H_0 via a function of fixation coefficients, $(1-f_k)^2$. They also approximate the *P*-value of the exact test using a Markov chain Monte Carlo approach. Although the use of fixation coefficients to describe departures from HWE has some merit, it has the disadvantage that these parameters are estimated as ratios of genotypic frequencies. It is difficult to study sampling properties of ratio statistics.[4,6] Besides, functions of fixation coefficients such as $(1-f_k)^2$ may possess infinite upper bound. On the other hand, there are advantages in working with a composite kind of quantity such as the disequilibrium coefficient. This is simply the difference between a frequency and its values expected when there are no association between alleles. Moreover, it is easy to show that disequilibrium coefficient D_k satisfies $\max\{-p_k^2, -q_k^2\} \leqslant D_k \leqslant p_k q_k$. Unfortunately, test of homogeneity of disequilibrium coefficients across several strata has not been considered in the literature yet. Therefore, the objective of this study is to develop a new homogeneity score statistic for testing the null hypothesis in (2) based on disequilibrium coefficients. We first develop the theory and method and then demonstrate the advantage of our method over the method proposed by Olson and Foley[12] via Monte Carlo simulation studies. We also derive the approximate power and sample size formulae, which are necessary in design of studies. Finally, we illustrate our test with a real glyoxalase genotype data set.

Method

Homogeneity test

Let X_{ijk} ($i \leqslant j = 1,\ 2$ and $k = 1, \ldots, K$) be the number of individuals with genotype A_iA_j in the kth population with $n_k = X_{11k} + X_{12k} + X_{22k}$. Let $M(n_k, \{p_{ijk}\})$ denote the trinomial distribution with parameter vector $(p_{11k}, p_{12k}, p_{22k})$. Hence, we have $\{X_{ijk}: i, j = 1, 2;\ i \leqslant j\} \sim M(n_k, \{p_{ijk}\})$ for $k = 1, \ldots, K$. In this article, we are interested to test the homogeneity hypothesis in (2) with $\theta_k = D_k$. That is,

$$\mathrm{H}_0 : D_1 = \cdots = D_k \quad \textit{versus} \quad H_1 : \text{Not all } D_k' \text{s are equal},$$

where $D_k = p_k q_k - p_{12k}/2$. All subsequent results are obtained under the assumptions that K is fixed and n_k is sufficiently large for $k = 1, 2, \ldots, K$.

Note that $p_{11k} = p_k^2 + D_k$, $p_{12k} = 2(p_k q_k - D_k)$ and $p_{22k} = q_k^2 + D_k$, the log-likelihood for the kth strata can be expressed in terms of D_k and p_k ($k = 1, \ldots, K$) as

$$\begin{aligned} l_k(D_k, p_k) =& x_{11k} \ln(p_k^2 + D_k) + x_{12k} \ln(2(p_k q_k - D_k)) \\ &+ x_{22k} \ln(q_k^2 + D_k) \end{aligned}$$

Let D denote the common disequilibrium coefficient under H_0 and $\mathbf{p} = (p_1, \ldots, p_K)'$ the nuisance parameter vector. Under H_0, the total log-likelihood for all K strata is given by

$$l(D, \mathbf{p}) = \sum_{k=1}^{K} l_k(D, p_k)$$

Hence, the efficient scores for the kth stratum (ie, the first-order derivatives of $l_k(D,p_k)$ with respect to D and p_k) are given by

$$H_{kD}(D, p_k) = \frac{\partial l_k(D, p_k)}{\partial D} = \frac{x_{11k}}{p_k^2 + D} - \frac{x_{12k}}{p_k q_k - D} + \frac{x_{22k}}{q_k^2 + D},$$

$$H_k p_k(D, p_k) = \frac{\partial l_k(D, p_k)}{\partial p_k} = \frac{2x_{11k} p_k}{p_k^2 + D} + \frac{x_{12k}(1 - 2p_k)}{p_k q_k - D} - \frac{2x_{22k} q_k}{q_k^2 + D}$$

Let $\hat{D}$ and $\hat{\mathrm{p}}$ be the maximum-likelihood estimates (MLEs) of D and p under the null hypothesis H_0. In this case, $\hat{D}$ and $\hat{\mathrm{p}}$ must satisfy the following $K+1$ equations:

$$\sum_{k=1}^{K} H_{kD}(\hat{D}, \hat{p}_k) = 0,$$

and

$$H_k p_k(\hat{D}, \hat{p}_k) = 0, \quad k = 1, 2, \ldots, K$$

Denote

$$\begin{aligned} V_{kDD}(D, p_k) =& \frac{\partial H_{kD}(D, p_k)}{\partial D} = -\frac{x_{11k}}{(p_k^2 + D)^2} \\ &- \frac{x_{12k}}{(p_k q_k - D)^2} - \frac{x_{22k}}{(q_k^2 + D)^2}, \end{aligned}$$

$$\begin{aligned} V_{kD} p_k(D, p_k) =& \frac{\partial H_{kD}(D, p_k)}{\partial p_k} = -\frac{2x_{11k} p_k}{(p_k^2 + D)^2} \\ &+ \frac{x_{12k}(1 - 2p_k)}{(p_k q_k - D)^2} + \frac{2x_{22k} q_k}{(q_k^2 + D)^2}, \end{aligned}$$

and

$$\begin{aligned} V_k p_k p_k(D, p_k) =& \frac{\partial H_k p_k(D, p_k)}{\partial p_k} = \frac{2x_{11k}(D - p_k^2)}{(p_k^2 + D)^2} \\ &+ \frac{x_{12k}(2D - p_k^2 - q_k^2)}{(p_k q_k - D)^2} + \frac{2x_{22k}(D - q_k^2)}{(q_k^2 + D)^2} \end{aligned}$$

In addition, denote $I_{kD}|p_k = I_{kDD} - I_{kDp_k}^2 / I_{kp_kp_k}$, where

$$I_{kDD} = -E(V_{kDD}) = \frac{n_k(p_k q_k + D)}{(p_k^2 + D)(p_k q_k - D)(q_k^2 + D)},$$

$$I_{kD} p_k = -E(_{VkD} p_k) = \frac{2n_k(2p_k - 1)D}{(p_k^2 + D)(p_k q_k - D)(q_k^2 + D)},$$

and

$$I_k P_k P_k = -E(V_k P_k P_k) = 2n_k \frac{(p_k^2+D)(q_k^2+D)^2 + (p_k q_k - D)^3 + (p_k^2+D)^2(q_k^2+D) - 4D^2}{(p_k{}^2+D)(p_k q_k - D)(q_k^2+D)}.$$

Hence, the likelihood score test for testing H_0: $D_1 = \cdots = D_K$ is given by

$$X^2 = \sum_{k=1}^{K} \frac{H_{kD}^2(\hat{D}, \hat{p}_k)}{I_{kD}|P_k(\hat{D}, \hat{p}_k)},$$

which is asymptotically distributed as a χ^2 variate with $K-1$ degrees of freedom under H_0. Unfortunately, we note that $\hat{D}$ and $\hat{p}$ cannot be expressed in closed form and this makes the likelihood score test X^2 less appealing in real applications. To over this issue, using the theory of homogeneity score test extended to nuisance parameters,[13] we consider the following modified score statistic:

$$X^{2*} = \sum_{k=1}^{K} \frac{H_{kD}^2(D^*, p_k^*)}{I_{kD|p_k}(D^*, p_k^*)} - \frac{\left[\sum_{k=1}^{K} H_{kD}^2(D^*, p_k^*)\right]^2}{\sum_{k=1}^{K} I_{kD|p_k}(D^*, p_k^*)}, \qquad (3)$$

where D^* and $\mathbf{p}^*$ are any consistent estimators of D and $\mathbf{p}$, respectively. To this end, we choose D^* to be $\sum_{k=1}^{k}(4x_{11k}x_{22k}/x_{12k}^2 - 1)/\sum_{k=1}^{k}(4n_k^2/x_{12k}^2)$ and $p^*{}_k$ be the solution to the following equation:

$$H_{kp_k}(D^*, p_k) \equiv \frac{2x_{11k}p_k}{p_k^2 + D^*} + \frac{x_{12k}(1-2p_k)}{p_k q_k - D^*} - \frac{2x_{22k}q_k}{q_k^2 + D^*} = 0,$$

or equivalently the following quintic polynomial equation,

$$a_0 + a_1 p_k + a_2 p_k^2 + a_3 p_k^3 + a_4 p_k^4 + a_5 p_k^5 = 0,$$

where $a_0 = x_{12k}D^*(1+D^*) + 2x_{22k}(D^*)^2$, $a_1 = -2(n_k D^*(1+D^*) + x_{12k}D^*)$, $a_2 = 6n_k D^* + 2x_{11k} + x_{12k}$, $a_3 = -2(2n_k D^* + n_k + 2x_{11k} + x_{12k})$, $a_4 = 4n_k + 2x_{11k} + x_{12k}$, and $a_5 = -2n_k$. Here, D^* is analogous to the Mantel–Haenszel estimator[14] and is a consistent estimator to D. However, it is not an efficient estimator to D in general. The proof of consistency and the condition to attain asymptotic efficiency for D^* is given in Appendix A. We note that the calculation of $I_{kD|p_k}$ in (3) could be tedious. Nonetheless, it is easy to show that $I_{kD|p_k}$ is simply given by $n_k/w_k(D, p_k)$ with $w_k(D, p_k) = (p_k^2+D)(q_k^2+D)^2 + 2(p_k q_k - D)^3 + (p_k^2+D)^2(q_k^2+D) - 4D^2$ (see Appendix B for the proof). Similarly, X^{2*} is asymptotically distributed as a χ^2 variate with $K-1$ degrees of freedom under H_0. Therefore, the homogeneity hypothesis H_0 is rejected at level α if $X^{2*} \geqslant \chi^2_{K-1,(1-\alpha)}$, where $\chi^2_{K-1,(1-\alpha)}$, is the $100 \times (1-\alpha)$ percentile point of the χ^2 distribution with $K-1$ degrees of freedom. Finally, it is noteworthy that if the consistent estimators of D and $\mathbf{p}$ are the constrained maximum-likelihood estimators under H_0, then the second term of (3) vanishes, since $\sum_{k=1}^{K} H_{kD}(D^*, p_k^*) = 0$, and (3) reduces to the likelihood score statistic.

Asymptotic power and sample size formulae

In this section, we aim to derive the asymptotic power and sample size formulae[15] based on X^{2*}. For these purposes, we assume $n_k = nb_k$ for some n and $b_k > 0$. Let $\bar{D}_k$ and $\bar{p}_k$ be the true parameter values for D_k and p_k under the alternative H_1, where $k = 1, 2, \ldots, K$ and $\bar{D}_k \neq \bar{D}_j$ for some $k \neq j$. Hence, the asymptotic power of the homogeneity score test X^{2*} at α level is given by

$$\Pr(X^{2*} > \chi^2_{K-1,(1-\alpha)}|H_1) = Pr(\chi^2_{K-1}(\delta) \geqslant \chi^2_{K-1,(1-\alpha)}), \qquad (4)$$

where $\chi^2_{K-1}(\delta)$ denotes the non-central χ^2 distribution with $K-1$ degrees of freedom and the non-centrality parameter δ is equal to

$$\delta = n\left\{\sum_{k=1}^{K} \frac{\left[b_k\left(\frac{\bar{p}_k^2+\bar{D}_k}{\bar{p}_k^2+D} - \frac{2(\bar{p}_k\bar{q}_k - D_k)}{p_k q_k - D} + \frac{\bar{q}_k^2+\bar{D}_k}{q_k^2+D}\right)\right]^2}{b_k/w_k(D, p_k)} - \frac{\left[\sum_{k=1}^{K} b_k\left(\frac{\bar{p}_k^2+\bar{D}_k}{\bar{p}_k^2+D} - \frac{2(\bar{p}_k\bar{q}_k - \bar{D}_k)}{p_k q_k - D} + \frac{\bar{q}_k^2+\bar{D}_k}{q_k^2+D}\right)\right]^2}{\sum_{k=1}^{K}[b_k/w_k(D, p_k)]}\right\},$$

with $\bar{q}_k = 1 - \bar{p}_k$,

$$D = \sum_{k=1}^{K}\left(\frac{(\bar{p}_k^2 + \bar{D})(\bar{q}_k^2 + \bar{D})}{(\bar{p}_k\bar{q}_k - \bar{D}_k)^2} - 1\right) \Big/ \sum_{k=1}^{K}\left[1/(\bar{p}_k\bar{q}_k - \bar{D}_k)^2\right],$$

and p_k is the solution to the following equation:

$$\bar{a}_0 + \bar{a}_1 p_k + \bar{a}_2 p_k^2 + \bar{a}_3 p_k^3 + \bar{a}_4 p_k^4 + \bar{a}_5 p_k^5 = 0,$$

where $\bar{a}_0 = 2(\bar{p}_k\bar{q}_k - \bar{D}_k)D(1+D) + 2(\bar{q}_k^2 + \bar{D}_k)D^2$, $\bar{a}_1 = -2D(1+D) + 4(\bar{p}_k\bar{q}_k - \bar{D}_k)D$, $\bar{a}_2 = 6D + 2\bar{p}_k$, $\bar{a}_3 = -2(2D + 1 + \bar{p}_k)$, $\bar{a}_4 = 4 + \bar{p}_k$, and $\bar{a}_5 = -2$.

As a result, the desirable sample size n required to attain the power at $1-\beta$ with $\bar{D}_k$ and $\bar{p}_k$ being the true parameter values for D_k and p_k under the alternative H_1 at nominal level α can be determined from the following equality:

$$\chi^2_{K-1,\beta}(\delta) = \chi^2_{K-1,(1-\alpha)},$$

where $\chi^2_{K-1,\beta}(\delta)$ is the $100 \times \beta$ percentage point of the non-central χ^2 distribution with $K-1$ degrees of freedom with non-centrality parameter being δ. The value of n can be readily obtained by solving the equation given in (4).

Simulation

We evaluate the performance of our proposed homogeneity score test in terms of type I error rate and power. For type I error rate, we include the homogeneity test proposed by Olson and Foley[12] in our comparison study. In their case, they adopted a function of fixation coefficients as the measure for Hardy–Weinberg disequilibrium. Specifically, they were interested to test the homogeneity hypothesis in (2) with $\theta_k = (1-f_k)^2$. That is,

H_0^*: $(1-f_1)^2 = \cdots = (1-f_K)^2$ *versus* H_1^*: Not all $(1-f_k)^2$ are equal, and their proposed statistic for testing the above

hypotheses is given by

$$T^2_{homog} = \sum_{k=1}^{k} \frac{h_k^2(\hat{\theta})}{\hat{V}ar[h_k(\hat{\theta})]}, \quad (5)$$

where $\theta = (\sum_{k=1}^{K}((x^2_{12k}-x_{12k})/(2(2n_k-1)))/(\sum_{k=1}^{K}((2x^2_{11k}-x_{22k})/(2n_k-1)))$, $h_k(\theta) = x^2_{12k}-x_{12k}-4\theta x_{11k}x_{22k}$, and $\hat{Var}[h_k(\theta)] = 4(x^3_{11k}-3x^2_{11k}+2x_{11k})(1-\theta)+2(x^2_{11k}-x_{11k})(2n_k\theta-3\theta+2)$ for $k=1, \ldots, K$. We would like to point out here that our proposed homogeneity score test (ie, X^{2*}) and Olson and Foley's test (ie, T^2_{homog}) can be fairly comparable only when $\theta_k=0$ for $k=1, \ldots, K$ in the null hypotheses H_0 and H_0* (ie, H'_0 in (1)). In the present comparisons, we consider both the asymptotic (denoted as $T^2_{homog,a}$) and exact (denoted as $T_{homog,e}$) versions of T^2_{homog}. For the implementation of $T^2_{homog,e}$, one can refer to Olson and Foley (1996, p 975). Here, we investigate type I error rates of X^{2*} and $T_{homog,a}$ for small (eg, $n_k=20$ and 30) to large sample sizes (eg, $n_k=50$–200) when $\theta_k=0$ for $k=1, \ldots, K$. As $T^2_{homog,e}$ is computationally intensive for large sample sizes, we consider its small-sample behavior only. Results of Monte Carlo experiments with 5 000 repetitions for different designed allele probabilities $p'_k s$ with $k=1, \ldots, K$ and $K=3$ and 5 at 0.05 nominal level are summarized in Tables 1 (for small sample sizes) and 2 (for moderate to large sample sizes).

As expected, the exact test $T^2_{homog,e}$ is always conservative (ie, its type I error rates are always less than the pre-assigned nominal level). The empirical type I error rates of our asymptotic homogeneity score test X^{2*} are satisfactorily close to the nominal 0.05 level for allelic probabilities being bounded away from 0 and 1, whereas those of the $T^2_{homog,a}$ are generally liberal (eg, more than 11 times of the given nominal level) even for large sample sizes. It is noteworthy that X^{2*} appears to be conservative than $T^2_{homog,e}$, for small allele probabilities (eg, p_k's being 0.1). However, the conservativeness of X^{2*} vanishes with an increase in sample sizes and the computation of X^{2*} is much more simpler than $T^2_{homog,e}$.

In view of the above observations, we prefer the proposed homogeneity score test X^{2*} (based on disequilibrium coefficients) to the existing homogeneity tests based on function of fixation coefficients (ie, $T^2_{homog,a}$ and $T^2_{homog,e}$). Hence, we exclude $T^2_{homog,a}$ and $T^2_{homog,e}$ in all subsequent evaluation and discussion. Table 3 further summarizes the type I error rate of X^{2*} for some non-zero (common) disequilibrium coefficients (ie, $D\neq 0$) under different settings. Again, the propose homogeneity score test performs satisfactorily in the sense that its empirical type I error rates are close to the pre-chosen nominal level and seldom exceed the nominal level by more than 10%.

For power performance, the parameters and sample size are quite similar to those adopted in Table 3, except that $\{D_k\}$ are now specifically designed under H_1. For this purpose, we set $D_k = D_0 + \Delta(k-1)$. For $K=3$, we consider: (i) $D_0=-0.03$, $\Delta=0.03$ and (ii) $D_0=-0.05$, $\Delta=0.05$. For $K=5$, we consider: (i) $D_0=-0.06$, $\Delta=0.03$ and (ii) $D_0=-0.1$, $\Delta=0.05$. The results are reported in Table 4. From the simulation results, the power of X^{2*} increases with the sample size n or Δ. For those settings with the same $\{D_k\}$, the one with varied allele probabilities across

Table 1 Empirical type I error rates for X^{2*}, $T^2_{homog,a}$ and $T^2_{homog,e}$ under H'_0 when $K=3$ and $K=5$

n	*p*	X^{2*}	$T^2_{homog,a}$	$T^2_{homog,e}$
20, 20, 20	0.5, 0.5, 0.5	0.058	0.133	0.041
	0.5, 0.4, 0.3	0.047	0.141	0.041
	0.5, 0.3, 0.1	0.046	0.215	0.033
	0.3, 0.3, 0.3	0.029	0.144	0.035
	0.3, 0.2, 0.1	0.023	0.228	0.024
	0.1, 0.1, 0.1	0.006	0.420	0.010
30, 30, 30	0.5, 0.5, 0.5	0.055	0.104	0.046
	0.5, 0.4, 0.3	0.047	0.113	0.046
	0.5, 0.3, 0.1	0.045	0.144	0.041
	0.3, 0.3, 0.3	0.033	0.107	0.041
	0.3, 0.2, 0.1	0.026	0.141	0.027
	0.1, 0.1, 0.1	0.004	0.229	0.011
20, 20, 20, 20, 20	0.5, 0.5, 0.5, 0.5, 0.5	0.059	0.192	0.048
	0.5, 0.4, 0.3, 0.2, 0.1	0.041	0.294	0.049
	0.5, 0.3, 0.1, 0.3, 0.5	0.047	0.280	0.047
	0.3, 0.3, 0.3, 0.3, 0.3	0.021	0.223	0.047
	0.1, 0.3, 0.5, 0.3, 0.1	0.032	0.363	0.048
	0.1, 0.1, 0.1, 0.1, 0.1	0.008	0.581	0.022
30, 30, 30, 30, 30	0.5, 0.5, 0.5, 0.5, 0.5	0.054	0.151	0.047
	0.5, 0.4, 0.3, 0.2, 0.1	0.040	0.190	0.046
	0.5, 0.3, 0.1, 0.3, 0.5	0.046	0.178	0.048
	0.3, 0.3, 0.3, 0.3, 0.3	0.028	0.156	0.047
	0.1, 0.3, 0.5, 0.3, 0.1	0.033	0.221	0.049
	0.1, 0.1, 0.1, 0.1, 0.1	0.006	0.330	0.026

Table 2 Empirical type I error rates for X^{2*} and $T^2_{homog,,a}$ under H'_0 when $K=3$ and $K=5$

n	*P*	X^{2*}	$T^2_{homog,,a}$
50, 50, 50	0.5, 0.5, 0.5	0.053	0.083
	0.5, 0.4, 0.3	0.046	0.083
	0.5, 0.3, 0.1	0.043	0.105
	0.3, 0.3, 0.3	0.041	0.086
	0.3, 0.2, 0.1	0.034	0.098
	0.1, 0.1, 0.1	0.003	0.124
100, 100, 100	0.5, 0.5, 0.5	0.051	0.065
	0.5, 0.4, 0.3	0.051	0.070
	0.5, 0.3, 0.1	0.046	0.075
	0.3, 0.3, 0.3	0.048	0.070
	0.3, 0.2, 0.1	0.036	0.075
	0.1, 0.1, 0.1	0.009	0.073
200, 200, 200	0.5, 0.5, 0.5	0.050	0.058
	0.5, 0.4, 0.3	0.051	0.060
	0.5, 0.3, 0.1	0.047	0.061
	0.3, 0.3, 0.3	0.050	0.061
	0.3, 0.2, 0.1	0.043	0.064
	0.1, 0.1, 0.1	0.023	0.063
50, 100, 200	0.5, 0.5, 0.5	0.050	0.072
	0.5, 0.4, 0.3	0.050	0.070
	0.5, 0.3, 0.1	0.046	0.068
	0.3, 0.3, 0.3	0.045	0.074
	0.3, 0.2, 0.1	0.042	0.071
	0.1, 0.1, 0.1	0.010	0.097
100, 100, 100, 100, 100	0.5, 0.5, 0.5, 0.5, 0.5	0.050	0.079
	0.5, 0.4, 0.3, 0.2, 0.1	0.042	0.096
	0.5, 0.3, 0.1, 0.3, 0.5	0.045	0.091
	0.3, 0.3, 0.3, 0.3, 0.3	0.046	0.086
	0.1, 0.3, 0.5, 0.3, 0.1	0.036	0.102
	0.1, 0.1, 0.1, 0.1, 0.1	0.008	0.106
50, 75, 100, 125, 150	0.5, 0.5, 0.5, 0.5, 0.5	0.048	0.082
	0.5, 0.4, 0.3, 0.2, 0.1	0.043	0.094
	0.5, 0.3, 0.1, 0.3, 0.5	0.042	0.096
	0.3, 0.3, 0.3, 0.3, 0.3	0.043	0.092
	0.1, 0.3, 0.5, 0.3, 0.1	0.034	0.117
	0.1, 0.1, 0.1, 0.1, 0.1	0.008	0.125

strata usually have power greater than that with equal allele probabilities across strata.

Real example

Ghosh reported genotype frequencies of red cell glyoxalase 1 (GLO) polymorphism from several populations.[16] We consider the data, reproduced in Table 5, from four populations in the Western Pacific Area. The gene frequencies of four populations highly vary from 0.0455 in the Eastern Carolines to 0.3611 in the Tokelau Islands, Samoa and Fuji in between. The estimated disequilibrium coefficients (ie, $\hat{D}_k$) in the four populations are ranging from 0.0145–0.019, which are close to zero. This seems to suggest that the homogeneity of HWE across the four populations, although the gene frequencies vary appreciably. Our proposed homogeneity score test yields $X^{2*}=2.33$ with *P*-value being 0.51. Hence, it is now safe to assume that the HWE is simultaneously valid across the four populations. We apply the Olson and Foley's test to the same glyoxalase genotype data set in the Western Pacific Area. Function of fixation coefficients (ie, $(1-f_K)^2$) was adopted and the corresponding homogeneity test yields $T^2_{homog,a}=2.78$ with *P*-value being 0.43. In this case, both tests reach the same conclusion.

Discussion

In practice, one is tempted to test the Hardy–Weinberg law across several independent populations without verifying the underlying assumption of homogeneity of Hardy–Weinberg disequilibrium across populations. Verification of the latter assumption is critical in genotype data analysis. Olson and Foley proposed a homogeneity test for this purpose. Unfortunately, our simulations show that their asymptotic version test is not reliable (ie, inflated type I error rates) even in large sample size. Although an exact version test was also proposed to overcome the liberty issue, such a test is however always conservativeness and computationally intensive for large sample sizes.

In this paper, we consider a homogeneity score test based on disequilibrium coefficients. Empirical results from our

simulation studies support that our homogeneity score test is a reliable asymptotic testing procedure even for small sample sizes. However, our test may suffer the drawback that it may be quite conservative for rare allelic probabilities (eg, $\leqslant 0.1$). In this case, one may require larger sample sizes to overcome the conservativeness issue. In this regard, we also provide a sample size formula for design purpose. We have implemented the test procedures described in this manuscript in a Matlab program, which can be downloaded from the web site: http://math.nenu.edu.cn/jhguo/program.htm.

We also applied the Kolmogorov–Smirnov test to study the asymptotic behaviors of our test (ie, X^{2*}). Briefly, for allele frequency greater than or equal to 0.1, we find that the asymptotic χ^2 sampling distribution property follows for moderate sample sizes (eg, $n_k \geqslant 50$). For rare allele frequency (ie, <0.1), larger sample sizes are required. In fact, after some straightforward algebra, we observe that H_{kD} (D^*, p_k^*) has larger variance for rare p_k. This may explain the severe conservativeness of X^{2*} for rare p_k. We are now undertaking an investigation of possible modification of X^{2*} for conservative correction.

We note that exact (conditional) method works in Olson and Foley[12] as they considered fixation coefficient f's which in turn are odds ratio. In their case, sufficient statistics for those nuisance parameters exist and can be eliminated by conditioning on their sufficient statistics. On the contrary, we consider the disequilibrium coefficient D's, which are actually rate differences. In our case, sufficient statistics do not exist for the corresponding nuisance parameters and the exact conditional method hence is not applicable.[17]

Table 3 Empirical type I error rates for X^{2*} under H_0

n	*D*	*p*	*Empirical size*
30, 30, 30	0.03	0.5, 0.5, 0.5	0.061
50, 50, 50			0.055
100, 100, 100			0.051
30, 30, 30	−0.03	0.5, 0.5, 0.5	0.048
50, 50, 50			0.049
100, 100, 100			0.050
30, 30, 30	0.03	0.5, 0.4, 0.3	0.045
50, 50, 50			0.048
100, 100, 100			0.049
30, 30, 30	−0.03	0.5, 0.4, 0.3	0.040
50, 50, 50			0.042
100, 100, 100			0.049
30, 30, 30, 30, 30	0.03	0.5, 0.5, 0.5, 0.5, 0.5	0.063
50, 50, 50, 50, 50			0.057
100, 100, 100, 100, 100			0.051
30, 30, 30, 30, 30	−0.03	0.5, 0.5, 0.5, 0.5, 0.5	0.049
50, 50, 50, 50, 50			0.050
100, 100, 100, 100, 100			0.050
30, 30, 30, 30, 30	0.03	0.5, 0.4, 0.3, 0.4, 0.5	0.047
50, 50, 50, 50, 50			0.049
100, 100, 100, 100, 100			0.049
30, 30, 30, 30, 30	−0.03	0.5, 0.4, 0.3, 0.4, 0.5	0.038
50, 50, 50, 50, 50			0.042
100, 100, 100, 100, 100			0.046

Table 4 Empirical power for X^{2*}

n	*D*	*p*	*Power*
30, 30, 30	−0.03, 0.0, 0.03	0.5, 0.5, 0.5	0.124
50, 50, 50			0.177
100, 100, 100			0.310
30, 30, 30	−0.03, 0.0, 0.03	0.5, 0.4, 0.3	0.133
50, 50, 50			0.196
100, 100, 100			0.363
30, 30, 30	−0.05, 0.0, 0.05	0.5, 0.5, 0.5	0.270
50, 50, 50			0.420
100, 100, 100			0.730
30, 30, 30	−0.05, 0.0, 0.05	0.5, 0.4, 0.3	0.297
50, 50, 50			0.473
100, 100, 100			0.790
30, 30, 30, 30, 30	−0.06, −0.03, 0.0, 0.03, 0.06	0.5, 0.5, 0.5, 0.5, 0.5	0.348
50, 50, 50, 50, 50			0.546
100, 100, 100, 100, 100			0.886
30, 30, 30, 30, 30	−0.06, −0.03, 0.0, 0.03, 0.06	0.5, 0.4, 0.3, 0.4, 0.5	0.353
50, 50, 50, 50, 50			0.560
100, 100, 100, 100, 100			0.887
30, 30, 30, 30, 30	−0.1, 0.05, 0.0, 0.05, 0.1	0.5, 0.5, 0.5, 0.5, 0.5	0.806
50, 50, 50, 50, 50			0.969
100, 100, 100, 100, 100			1.0
30, 30, 30, 30, 30	−0.1, 0.05, 0.0, 0.05, 0.1	0.5, 0.4, 0.3, 0.4, 0.5	0.811
50, 50, 50, 50, 50			0.972
100, 100, 100, 100, 100			1.0

Table 5 Glyoxalase genotype data in Western Pacific Area

Population	*Genotype counts*					
	n_k	*1–1*	*1–2*	*2–2*	$\hat{p}_k$	$\hat{D}_k$
Eastern Carolines	748	3	62	683	0.0455	0.0019
Tokelau Islands	961	118	458	385	0.3611	−0.0076
Samoa	101	4	39	58	0.2327	−0.0145
Fiji	137	4	38	95	0.1679	0.0010

Finally, the theories developed in this paper can be readily extended to genotype data with multiple alleles.

Acknowledgements

This research is supported by the NSFC (Grant Numbers 10431010, 10329102, and 10371015), the Science and Technology Keystone Fund of MOE, P.R. China (Grant Numbers 104070 and 00041), NCET-04-0310, EYTP, the Jilin Distinguished Young Scholars Program (Grant Number 20030113), and the Science Foundation for Young Teachers of NENU (Grant Number 20060101). The work of ML Tang was fully supported by a grant from the Research Grant Council of the Hong Kong Special Administration (Project Number CUHK4371/04 M) and Hong Kong Baptist University Grants FRG/04-05/II-01 and FRG/04-05/II-20.

References

1 Hardy GH: Mendelian proportions in a mixed population. *Science* 1908; **28**: 49–50.
2 Weinberg W: *Papers on Human Genetics*. Englewood Cliffs, NJ: Prentice-Hall, 1908, In (translation by Boyer SH) On the Demonstration of Heredity in Man 1963.
3 Crow JE: Eighty years ago: the beginnings of population genetics. *Genetics* 1988; **119**: 473–476.
4 Hernandez JL, Weir BS: A disequilibrium coefficient approach to Hardy–Weinberg testing. *Biometrics* 1989; **45**: 53–70.
5 Olson JM: Testing the Hardy–Weinberg law across strata. *Ann Hum Genet* 1993; **57**: 291–295.
6 Weir BS: *Genetic Data Analysis II*. Sunderland, MA: Sinauer Associates, Inc., 1996, pp 91–140.
7 Haldane JBS: An exact test for randomness of mating. *J Genet* 1954; **52**: 631–635.
8 Smith CAB: A note on testing the Hardy–Weinberg law across strata. *Ann Hum Genet* 1970; **33**: 377–383.
9 Emigh TH: A comparison of tests for Hardy–Weinberg equilibrium. *Biometrics* 1980; **36**: 627–642.
10 Troendle JJ, Yu KF: A note on testing the Hardy–Weinberg law across strata. *Ann Hum Genet* 1994; **58**: 397–402.
11 Nam JM: Testing a genetic equilibrium across strata. *Ann Hum Genet* 1997; **61**: 163–170.
12 Olson JM, Foley M: Testing for homogeneity of Hardy–Weinberg disequilibrium using data sampled from several populations. *Biometrics* 1996; **52**: 971–979.
13 Tarone RE: Homogeneity score tests with nuisance parameters. *Commun Statist Theory Meth* 1988; **17**: 1549–1556.
14 Mantel N, Haenszel W: Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 1959; **22**: 719–748.
15 Guo JH, Ma YP, Shi NZ, Lau TS: Testing for homogeneity of relative difference under inverse sampling. *Comput Statist Data Anal* 2004; **44**: 613–624.
16 Ghosh AK: Polymorphism of red cell glyoxalase 1 with special reference to South and Southeast Asian and Oceania. *Hum Genet* 1977; **39**: 91–95.
17 Agresti A: Exact inference for categorical data: recent advances and continuing controversies. *Statist Med* 2001; **20**: 2709–2722.

Appendix A
Consistency and the condition to attain asymptotic efficiency for D^*

Let $n_k = nb_k$, with $b_k > 0$ and $k = 1, 2, \ldots, K$. The asymptotic property of D^* is obtained under the assumption that K is fixed and n approaches infinity (ie, sufficiently large). Following the notation above, the Mantel–Haenszel-type estimator of D^* can be rewritten as

$$D^* = \sum_{k=1}^{K} (n_k/x_{12k})^2 \hat{D}_k / \sum_{k=1}^{K} (n_k/x_{12k})^2$$

By the Central Limit Theorem, $\sqrt{n}(y_k - g_k)$ has an asymptotic normal distribution, $N(0, \sum_k/b_k)$. By delta method, we obtain $\sqrt{n}(\hat{D}_k - D_k)$ has an asymptotic normal distribution with mean 0 and variance $w_k(D_k, p_k)/b_k$. As $D_k = D$ under H_0 for $k = 1, 2, \ldots, K$, we can derive that D^* is a consistent estimate of D. Let $w_k = w_k(D, p_k)$, $v_k = 1/(p_k q_k - D)$. Hence, the asymptotic variance of D^* under H_0 is given by

$$AsyVar(D^*) = \frac{\sum_{k=1}^{K} w_k v_k^4 / b_k}{n(\sum_{k=1}^{K} v_k^2)^2}$$

Denote the information matrix with respect to D and $\mathbf{p}$ under H_0 by

$$I = \begin{pmatrix} \sum_{k=1}^{K} I_{kDD} & I_{1Dp_1} & \cdots & I_{KDp_K} \\ I_{1Dp_1} & I_{1p_1p_1} & \cdots & I_{Kp_1p_K} \\ \vdots & \ddots & & \vdots \\ I_{KDp_K} & \cdots & \cdots & I_{Kp_Kp_K} \end{pmatrix}$$

By inverting I, we can obtain the asymptotic variance of $\hat{D}$, which is given by

$$AsyVar(\hat{D}) = \frac{1}{s}\left(\sum_{k=1}^{K} b_k/w_k\right)^{-1}$$

By the Cauchy–Schwarz inequality

$$\left(\sum_{k=1}^{K} v_k^2\right)^2 \leqslant \left(\sum_{k=1}^{K} w_k v_k^4/b_k\right)\left(\sum_{k=1}^{K} b_k/w_k\right)$$

Thus, $AsyVar(\hat{D}) \leqslant AsyVar(D^*)$ and we get the sufficient and necessary condition for the asymptotic efficiency of D^*, that is, $w_k v_k^2 = c$, $k = 1, 2, \ldots, K$, where c is a constant

independent of all the parameters. The condition is satisfied if $D=0$. From the above discussion, we know that D^* is inefficient in general case.

Appendix B

Simple expression for I_{kD/P_k}

For the kth stratum, we denote the information matrix with respect to D_k and p_k by

$$I_k = \begin{pmatrix} D_k & I_{kD_kp_k} \\ I_{kp_kD_k} & I_{kp_kp_k} \end{pmatrix}$$

By the property of inverse matrix, $I_{\mathrm{kD/pk}}(D_k,p_k)$ is equal to the reciprocal of the (1,1) element of I_k^{-1}. On the one hand, according to the property of MLEs, we have

$$\sqrt{n_k}(\hat{D}_k - D_k, \hat{p}_k - p_k)' \xrightarrow{D} N(0, n_k I_k^{-1}(D_k, p_k)),$$

where $\hat{D}_k = (4x_{11\mathrm{k}}x_{22k} - x_{12k}^2)/(4\mathrm{n}_k^2)$ is the MLE of D_{k}. Therefore, the asymptotic variance of $\sqrt{n_k}\hat{D}_k$ is $n_k/I_{kD|p_k}(D_k,p_k)$. On the other hand, let $y_k = (x_{11k},\ x_{12k},\ x_{22k})/n_k$, by the Central Limit Theorem, $\sqrt{n_k}(y_k - g_k)$ has an asymptotic normal distribution, $N(0,\ \sum_k)$, where $g_k = (p_{11k},\ p_{12k},\ p_{22k})'$, $\sum_k = diag(g_k) - g_k g'_k$. Let $c_k = \frac{\partial \hat{D}_k}{\partial y_k}|_{y_k = g_k}$. By delta method, we obtain $\sqrt{n_k}(\hat{D}_k - D_k)$ has an asymptotic normal distribution with mean 0 and variance $c'_k \sum_k c_k$. After simple calculation, we have $c'_k \sum_k c_k = w_k(D_k,p_k)$. Hence, we can give the exact expression $I_{kD|p_k}(D_k,p_k) = n_k/w_k(D_k,p_k)$. Naturally, the expression of $I_{kD|p_k}(D,p_k)$ is just $I_{kD|p_k}(D_k,p_k)$ by substituting D for D_k.