

ARTICLE

Recovering unused information in genome-wide association studies: the benefit of analyzing SNPs out of Hardy–Weinberg equilibrium

David W Fardo^{*1}, K David Becker², Lars Bertram³, Rudolph E Tanzi³ and Christoph Lange^{4,5}

¹Department of Biostatistics, University of Kentucky College of Public Health, Lexington, Kentucky, USA; ²TorreyPines Therapeutics Inc., La Jolla, California, USA; ³Department of Neurology, Genetics and Aging Research Unit, Center for Aging, Genetics, and Neurodegeneration, Massachusetts General Hospital, Charlestown, Massachusetts, USA;

⁴Channing Laboratory, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA;

⁵Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts, USA

Although the rapid advancements in high throughput genotyping technology have made genome-wide association studies possible, these studies remain an expensive undertaking, especially when considering the large sample sizes necessary to find the small to moderate effect sizes that define complex diseases. It is therefore prudent to utilize all possible information contained in a genome-wide scan. We propose a straightforward analytical approach that tests often unused SNP data without sacrificing statistical validity. We simulate genotype miscalls under a variety of models consistent with observed miscall rates and test for departures from HWE using the standard Pearson's χ^2 -test. We find that true disease susceptibility loci subjected to various patterns of genotype miscalls can be largely out of HWE and, thus, be candidates for removal before association testing. These loci, we demonstrate, can maintain sufficient statistical power even under extreme error models. We additionally show that random miscalls of null SNPs, independent of the phenotype, do not induce bias in case–control or cohort studies, and we suggest that a significant HWE test should not prevent a SNP from being tested when conducting genome-wide association studies in these scenarios. However, association findings for SNPs that are out of HWE must be treated more carefully than 'regular' findings, for example, by re-genotyping the SNP in the same study using a different genotyping technology.

European Journal of Human Genetics (2009) 17, 1676–1682; doi:10.1038/ejhg.2009.85; published online 3 June 2009

Keywords: genome-wide association; genotyping error; Hardy–Weinberg equilibrium

Introduction

Screening SNPs for Hardy–Weinberg equilibrium (HWE) departure is a quality control measure used frequently when performing genetic association studies. Gomes *et al*¹ state, 'quality control provided by a Hardy–Weinberg

test should be an essential part of any genome scan or other application of DNA typing.' The common convention is that when a marker is detected to depart from HWE, the reasons for this departure should be investigated (eg, see, Wittke-Thompson²) and then the marker is a candidate for retyping due to genotyping error. It can be prohibitively expensive to retype all of the markers from genome-wide scans that are found to deviate from HWE. Oftentimes, these markers are, thus, removed from the analysis and the study proceeds accordingly. Genome-wide studies using HWE as a screening tool can effectively remove tens of thousands of SNPs from an analysis.^{3–5}

^{*}Correspondence: Dr DW Fardo, Department of Biostatistics, University of Kentucky College of Public Health, 121 Washington Avenue, Lexington, KY 40536, USA.

Tel: +1 859 218 2070; Fax: +1 859 257 4665;

E-mail: david.fardo@uky.edu

Received 20 May 2008; revised 7 April 2009; accepted 23 April 2009; published online 3 June 2009

We investigate the effects of genotyping error on tests of HWE, study power (for true disease susceptibility loci) and the rate of false positives (for null SNPs) under different study designs. We perform extensive simulation studies to explore how factors such as disease allele frequency, mode of inheritance and the pattern of genotyping error affect these quantities. Many models for genotyping error have been proposed and investigated.^{6–9} Although some of these, such as the error model introduced by Sobel, Papp and Lange⁹, offer enough flexibility to accurately characterize common patterns of genotyping error in multiple settings, care should be taken when proposing values for the parameters that define each model. These patterns are dependent on both the genotyping platform and the genotype-calling algorithm employed. Research carried out before the creation and widespread use of the high-throughput genotyping platforms must be adapted to fit patterns consistent with the genotyping error specific to the platform and calling algorithm pair.

In the context of genome-wide association studies, the number and type of genotyping platforms is restricted to the few available commercial assays. Most of these technologies are typically associated with a particular preferred, or in-house, calling algorithm, but some alternate genotype calling strategies are becoming popularized.¹⁰ Rabbee and Speed,¹¹ among others, have found that these genome-wide platforms can be differentially susceptible to heterozygote miscalls, for example, when using the DM calling algorithm for Affymetrix genotyping platforms. It is with the goal of realistically modeling the error in a genome-wide study that we chose to use empirical genotyping error rates from a recent genome-wide study in our simulations. These simulations demonstrate the negative impact of genotyping error on study power, but they also confirm that testing SNPs out of HWE does not affect the overall false-positive rate, and, more importantly, can lead to the identification of ‘real/true’ associations.

Methods

The empirical genotyping error rates used as a model in the simulations come from a recent genome-wide study in which SNPs were initially genotyped using the Affymetrix GeneChip Human Mapping 500K Array Set and then later resequenced. The re-sequenced genotypes were treated as the ‘true’ genotypes to determine the error rates from the genome-wide assay. A total of 1502 subjects were genotyped and, of these, 1439 were successfully called from both the large-scale assay and the resequencing. Table 1 shows the concordance between the original scan and the resequencing for these 1439 subjects.

In addition to using the empirical error rates, or the ‘Empirical’ error model, we employed other models of

Table 1 SNPs cross-classified by genotype call from an Affymetrix 500K array (observed genotype) and the genotype call from resequencing (‘true’ genotype)

‘True’ genotype	Observed genotype			Total
	AA	AB	BB	
AA	48	6	0	54
AB	175	198	1	374
BB	0	2	1009	1011
Total	223	206	1010	1439

Table 2 Reverse genotyping error model probabilities

True genotype	Observed genotype		
	AA	AB	BB
AA	1	0	0
AB	0	0.5	0.5
BB	0	0	1

Probability of SNP cross-classification under the reverse genotyping error model. The genotype call from a large-scale array is the observed genotype and the actual genotype is the true genotype.

genotyping error for our simulations. We based an array of error models to reflect the proposed differential rate at which heterozygotes are miscalled. We call the model that miscalls heterozygotes to the major homozygote with probability one-half, ‘Reverse,’ as it behaves in an approximately opposite manner than the empirical data with regards to heterozygote misclassification. Table 2 displays the reverse model genotype-calling probabilities.

Case–control study

We calculated the power for a variety of case–control study designs. Each scenario was specified by several parameters: the number of cases, the number of controls, the significance level for the association test, the mode of inheritance, the disease prevalence and the odds ratio for disease. Each study design was simulated 10^8 times and the proportion of studies that generated a significant association determined the empirical power. In addition to assessing the power of the study design, the probability of rejecting the null hypothesis of Hardy–Weinberg equilibrium (tested in controls) was also obtained. The false-positive rate was calculated in the exact same fashion as the empirical power. It is simply the simulated power under the null hypothesis that $OR = 1$. When we calculated the association test to determine the false-positive rate, the nominal significance level was set to 5%.

Cohort study

The simulation process for cohort studies that recruit total population samples and measure a continuous trait/

outcome of interest follows similarly. The same genotyping error models used in the case–control setting was also investigated here. The alternative hypothesis, however, was specified through the locus-specific heritability which is defined as the proportion of phenotypic variance attributed to the typed marker, or $h^2 = \text{Var}(aX)/\text{Var}(Y)$, under the associated biometric regression model $E(Y|x) = m + ax$.¹² In this model, x is the coded genotype consistent with the assumed mode of inheritance (eg, assuming a recessive mode of inheritance with minor and major alleles A and B , respectively, we have that $x = 1$ for the genotype AA and $x = 0$ otherwise). Under the null hypothesis in this scenario, the genetic effect of size a is 0 and so is the heritability, h^2 . These studies were replicated 10^6 times under each scenario.

Results Power

We simulated the power for several study scenarios with different error models. For the case–control studies, we assumed a constant odds ratio of 1.5 and varied the number of probands across the three modes of inheritance. The disease prevalence was assumed to be 10% and the α -level for the Pearson's χ^2 association test in each scenario was 10^{-9} (Figure 1a–c). For the cohort studies, we examined different disease allele frequencies and modes of inheritance while keeping the number of probands, 2000, and the locus-specific heritability, 2%, constant throughout. We tested for a disease–SNP association using a significance level of 10^{-9} . A simple test for the significance of a regression coefficient not being zero was

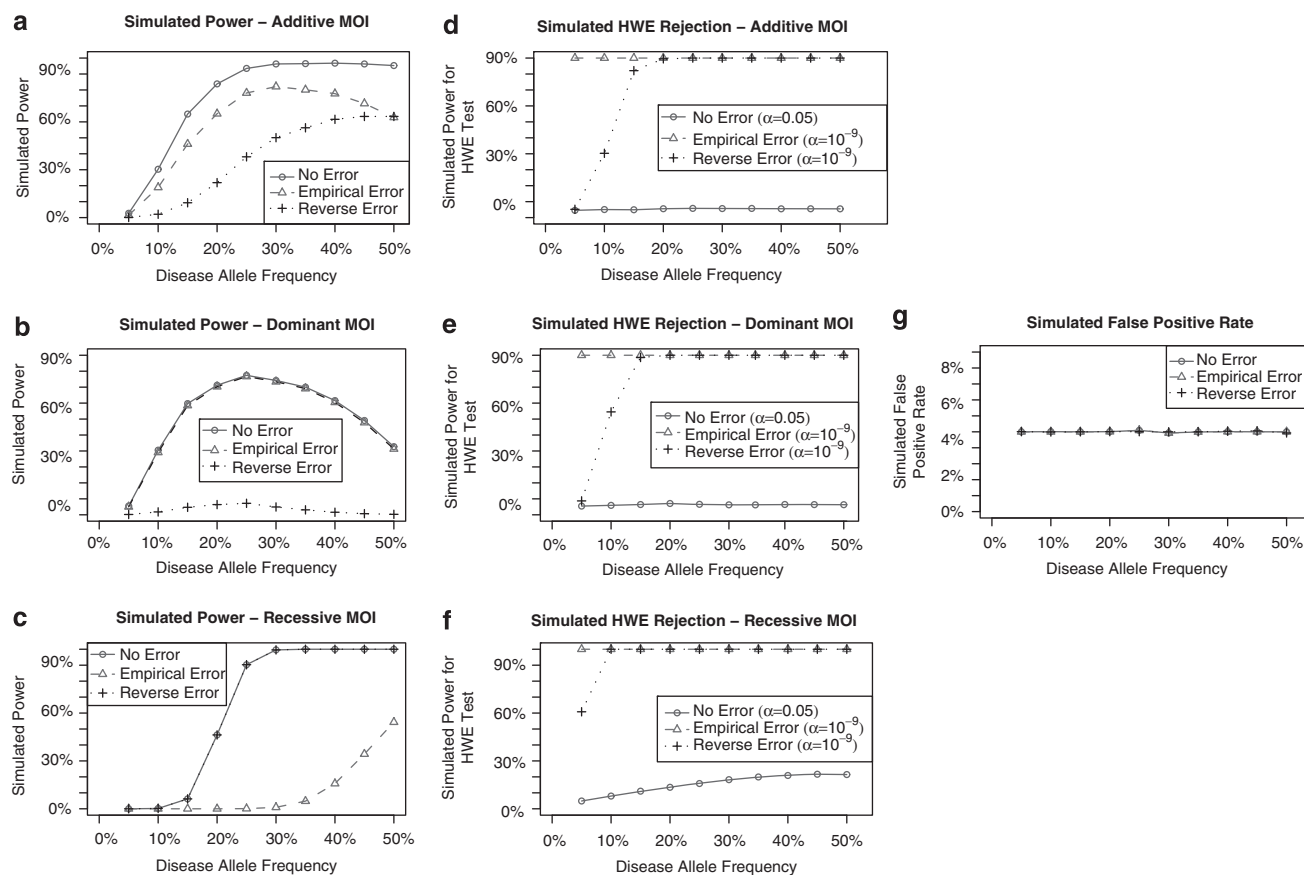


Figure 1 Case–control simulation results. Power simulations: The significance level for the association tests was $\alpha = 10^{-9}$ and 10^8 simulations were performed under each genotyping error model (no error, empirical error and reverse error). An odds ratio of 1.5 and disease prevalence of 10% was assumed. (a) Simulated power under an additive mode of inheritance (2000 cases and 2000 controls). (b) Simulated power under a dominant mode of inheritance (2500 cases and 2500 controls). (c) Simulated power under a recessive mode of inheritance (10 000 cases and 10 000 controls). HWE simulations: The significance level for the Hardy–Weinberg tests was $\alpha = 10^{-9}$ under the two error models and $\alpha = 0.05$ under no genotyping error. Simulations (10^8) were performed under each genotyping error model (no error, empirical error and reverse error). An odds ratio of 1.5 and disease prevalence of 10% was assumed. (d) Simulated probability of rejecting Hardy–Weinberg equilibrium under an additive mode of inheritance (2000 cases and 2000 controls). (e) Simulated probability of rejecting Hardy–Weinberg equilibrium under a dominant mode of inheritance (2500 cases and 2500 controls). (f) Simulated probability of rejecting Hardy–Weinberg equilibrium under a recessive mode of inheritance (10 000 cases and 10 000 controls). False-positive rate simulations: the significance level for the association test was $\alpha = 10^{-9}$ and 10^8 simulations were performed under each genotyping error model (no error, empirical error and reverse error). (g) Simulated false-positive rate (5000 study subjects).

used (Figure 2a–c). The same two models for genotyping error were used under all scenarios.

The difficulty of detecting rare disease susceptibility loci and the sensitivity of study power to the disease allele frequency are readily apparent. Intuitively, statistical power under the dominant mode of inheritance was less affected by the empirical genotyping error model where the most common error is calling a true heterozygote a homozygote minor. Similarly, power under the recessive mode of inheritance was less affected by the reverse genotyping error model. Data generated under the additive mode of inheritance, however, were fairly robust to both patterns of genotyping error. For example, examining an additive disease allele with a frequency of 35% and recruiting 2000 cases and controls, the study power of 96.5% was only reduced to 80.1% under the empirical genotyping error model and 56.3% under the reverse genotyping error model. Power loss patterns akin to these resulting from various genotyping error models have been previously documented using both a 2×3 χ^2 -test of independence^{13,14} and the Cochran–Armitage trend test.¹⁵

Probability of departure from HWE

A standard χ^2 -test for departure from Hardy–Weinberg equilibrium was performed in each of the above power simulations (results using exact tests for Hardy–Weinberg were nearly identical). Although current studies employ varying levels of significance for these tests, we tested using an unusually stringent α -level of 10^{-9} . Only control data was used to test for HWE in the case–control study simulations (Figure 1d–f) and all probands were used in the cohort study simulations (Figure 2d). We determined the empirical probability of rejecting HWE using the 10^8 replicates for each case–control scenario and the 10^6 replicates for each cohort scenario.

The probability of rejecting HWE was very high under most scenarios and was virtually one for allele frequencies greater than 20%. Only when the disease allele frequency was low, and, thus, the power to detect an association was low, did the probability of rejecting HWE fall. For the empirical genotyping error model, every test rejected HWE. Only using the theoretical reverse genotyping error model did the probability of rejecting HWE fall below one. Relating this to a study using conventional quality control methods, we see that a disease-associated risk locus can be removed from analysis without making it to the testing phase even though the disease–SNP association test is still powered.

False-positive rate

To determine how genotyping error affects the false-positive rate, we simulated the same scenarios as above under the null hypothesis of no genetic association (ie, OR = 1 for the case–control studies and $h^2 = 0$ for the

cohort studies). Using a nominal α -level of 0.05, we simulated each case–control scenario 10^8 times (Figure 1g) and each cohort scenario 10^6 times (Figure 2e).

The nominal significance level of 5% is maintained in every scenario. That is, the disease–SNP association testing is still valid in the presence of either model for genotyping error. This means that although the loci subjected to genotyping errors will be less powered than their correctly genotyped counterparts, they will not increase the previously determined acceptable rate of false-positive associations. The difference between the nominal 5% and the highest simulated false-positive rate, 5.07%, can be attributed to random sampling variability. It is worth mentioning that when the tested allele frequency is very low, the test can be slightly conservative. This could be because the distributional assumptions of the χ^2 -test statistic are more difficult to satisfy under the recessive mode of inheritance and with smaller sample sizes (note that the additive mode of inheritance scenarios use the fewest cases and controls in the simulations and are the most affected). This conservative nature is not severe but would likely be amplified when considering smaller sample sizes and/or rare variants.

Data application

To examine the effects of varying HWE test filtering criteria in the quality control steps of a real study, we applied different QC rules to a genome-wide association study for Alzheimer's disease. Because this study population is comprised of related individuals, we randomly sampled a single individual from each family to mimic a traditional population-based, case–control study. This resulted in a sample of 346 cases and 88 controls with genotypes available for 484 422 SNPs.

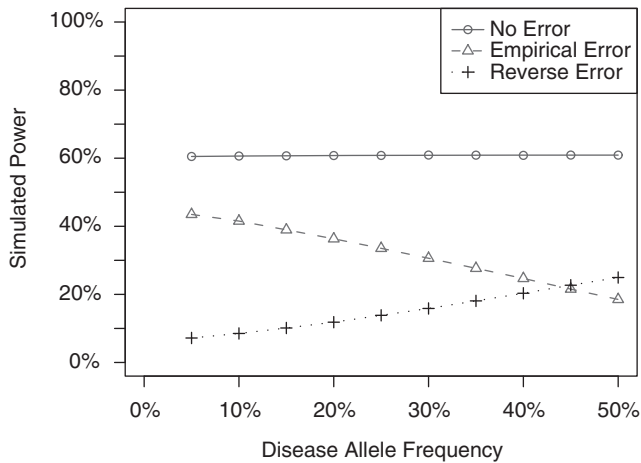
Before any analysis was performed, two individuals with more than 5% missingness were removed from consideration. SNPs with more than 5% missingness were also removed, but the specific number varied based on other QC filtering rules, which were applied before this missingness filter. The minimum minor allele frequency varied between 1 and 5%, and HWE testing was either omitted or performed using a χ^2 -test at a significance level of $\alpha = 10^{-6}$. In addition, only SNPs with each genotype represented by at least 1 affected and 1 unaffected individual were tested using a 2-d.f. genotype-by-affection status χ^2 -test.

The effects of the QC filtering are illustrated in Table 3. For example, over 50 000 SNPs have minor allele frequencies between 1 and 5%, and the more stringent criterion for MAF results in not testing these SNPs. An interesting note is that none of these fairly low frequency SNPs result in a statistically significant association with Alzheimer's disease as evidenced by the fact that the same numbers are found significant at the $\alpha = 10^{-3}$ level for both QC rules. More illuminating is that at significance levels of $\alpha = 10^{-4}$ and

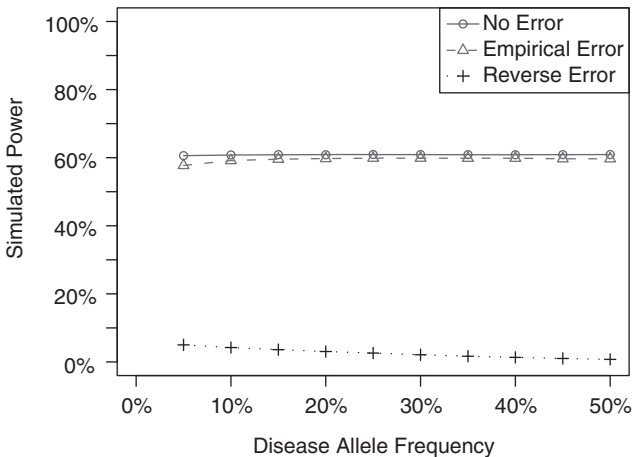
more stringent results employing HWE testing as a QC screen are identical to those omitting it. Specifically, after QC filtering, 30 SNPs are found to be significant at $\alpha = 10^{-4}$

regardless of how departure from Hardy–Weinberg equilibrium is treated. In addition, the difference in the number of SNPs that get to the testing stage is minimal. This means

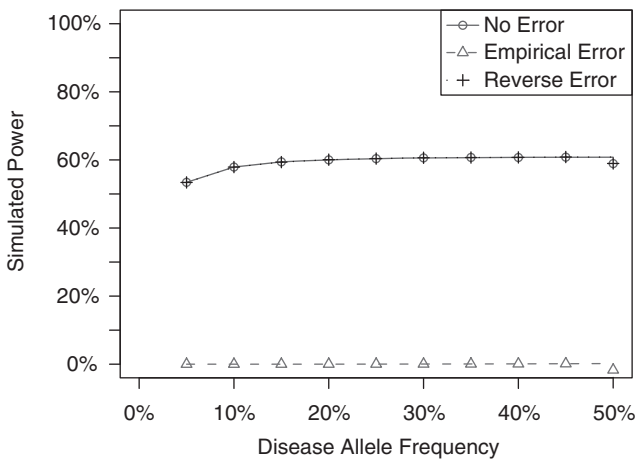
a Simulated Power – Additive MOI



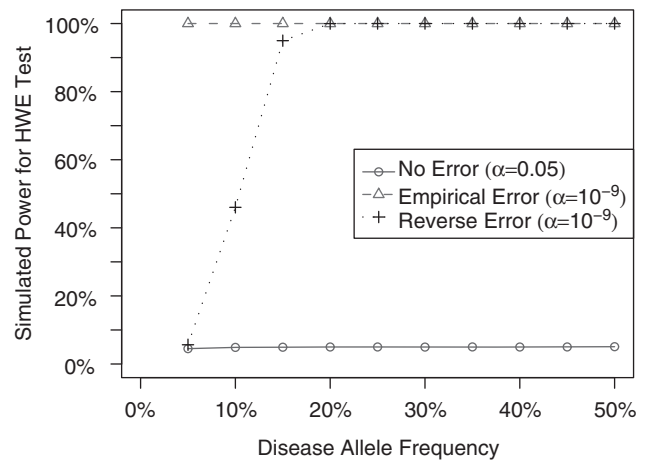
b Simulated Power – Dominant MOI



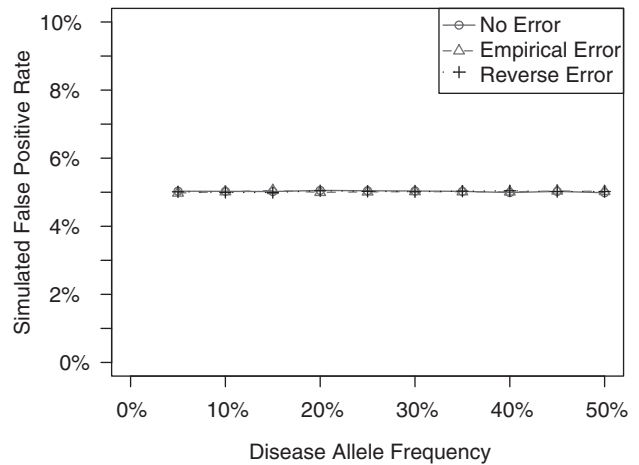
c Simulated Power – Recessive MOI



d Simulated HWE Rejection



e Simulated False Positive Rate



that any adjustments to the study-specific significance level will be minimal.

Discussion

Before the advent of high-throughput genotyping technologies, removing SNPs out of HWE was an algorithmic step in a genetic association study's quality control process used to detect genotyping error. Recently, some genome-wide association studies have continued to follow this practice and exclude SNPs out of HWE,^{3,16,17} but others have tested SNPs found to deviate from HWE.^{18,19} We have shown that in the presence of realistic genotyping error that is independent of the phenotype, common tests for SNP–disease association can remain powerful without increasing the false-positive rate. SNPs out of HWE that are significantly associated with disease should then be considered as secondary findings behind those SNPs that are both associated with the disease and that pass all quality control criteria. These secondary SNPs must be confirmed through an alternative genotyping technology or resequenced altogether.

The degree of genotyping error in any study is a function of both the genotyping platform and the algorithm used to make genotype calls, and the number of SNPs that demonstrate departure from HWE in recent studies has varied widely. Generally, studies using Affymetrix chips^{3–5,13,20} have had a higher percentage of SNPs out of HWE than studies employing Illumina arrays.^{16,18,19,21–23} It can only be economically practical to resequence all secondary

SNPs when the number found significantly associated with disease does not require a major portion of a study's budget.

Based on our results, we recommend against the use of Hardy–Weinberg testing as an agnostic quality control technique. We advise that SNPs be screened/tested as outlined by the particular study design and that SNPs deemed of interest/significant should only then be tested for HWE departures to determine if any of the significant SNPs need to be re-genotyped. Although the SNPs that have been subjected to genotyping errors are less powerful than their error-free counterparts, they do not increase the rate of finding false-positive associations. This amendment to the current practice will have the effect of increasing the probability of finding a true disease risk-conferring allele in genome-wide association studies without increasing the number of false-positive findings.

It is common to share genotyping between centers/studies, especially that of control data. In these settings, subset-based data quality artifacts can arise due to differential genotyping error, lab-specific error, etc. As such, particular caution should be practiced concerning all QC metrics, including HWE testing. That is, although we advocate for testing SNPs that violate HWE, the potential reasons for being out of HWE must be explored. If any subject subset (eg, cases, controls, subjects genotyped in a specific lab, etc) exhibits a disproportionate lack of genotyping quality as compared with the entire subject population, the subset must be examined more carefully before any statistical testing can be reliable.

Table 3 Alzheimer's disease case–control results

HWE-based QC?	min MAF	No. of SNPs		No. of SNPs with $-\log_{10}(\text{P-value}) <$					
		Passing QC	Tested	1	2	3	4	5	6
No	1%	417 531	313 554	29 927	3056	320	30	4	2
Yes	1%	415 531	312 747	29 819	3037	319	30	4	2
	Difference	2000	807	108	19	1	0	0	0
No	5%	362 228	310 571	29 798	3044	320	30	4	2
Yes	5%	361 778	310 395	29 734	3027	319	30	4	2
	Difference	450	176	64	17	1	0	0	0

Count of SNPs passing varying QC criteria, and subsequently achieving various levels of statistical significance. HWE screening is either omitted or performed using a χ^2 -test at $\alpha = 10^{-6}$. The minimum minor allele frequency (min MAF) varies between 1 and 5%. Listed are the number of SNPs passing QC and also satisfying the minimum cell count requirement of at least 1 affected and 1 unaffected subject in each genotype category. The numbers of these tested SNPs satisfying varying levels of significance are displayed. The effects of HWE screening are presented as the SNP count differences.

Figure 2 Cohort simulation results. Power simulations: the significance level for the association test was $\alpha = 10^{-9}$ and 10^6 simulations were performed under each genotyping error model (no error, empirical error and reverse error). A sample size of 2000 and a locus-specific heritability of 2% was assumed. (a) Simulated power under an additive mode of inheritance. (b) Simulated power under a dominant mode of inheritance. (c) Simulated power under a recessive mode of inheritance. HWE simulations: the significance level for the Hardy–Weinberg tests was $\alpha = 10^{-9}$ under the two error models and $\alpha = 0.05$ under no genotyping error. Simulations (10^8) were performed under each genotyping error model (no error, empirical error and reverse error). A sample size of 2000 was assumed. (d) Simulated probability of rejecting Hardy–Weinberg equilibrium. False-positive rate simulations: The significance level for the association test was $\alpha = 10^{-9}$ and 10^6 simulations were performed under each genotyping error model (no error, empirical error and reverse error). A sample size of 2000 was assumed. (e) Simulated false-positive rate.

The relationship between genotyping error and Hardy–Weinberg equilibrium has been investigated previously.^{24–28} Some recent papers have addressed this issue in genome-wide association scans from other angles. Li and Li²⁹ similarly point out that study ascertainment can lead to an inflated type I error rate for HWE testing. They develop a likelihood-based methodology that incorporates SNP–disease association into the testing for HWE departure. Our approach differs fundamentally from theirs as we do not attempt to differentiate the reason for HWE departure nor do we require estimation of any model parameters. Rather, we simply classify disease-associated SNPs into two classes: those that are in HWE and should be followed up as priority ‘hits’ and those out of HWE which are treated as second-tier ‘hits.’

An important reminder is that the results presented in this article apply only to population-based studies. The discerning characteristic of the considered genotyping error not introducing bias into association testing does not hold true in family-based designs. Although we recommend that a SNP’s departure from HWE not warrant its removal from the pool of variants to be tested for association in a population-based study, the same is not true of a family-based study; that is, genotyping error causes bias in a family-based association testing, and, thus, SNPs that depart from HWE should not be tested for association in a family-based genome-wide association study.

Acknowledgements

We thank three anonymous reviewers for helpful comments and suggestions. This work was supported in part by grant no. HL07427 (to DWF) from the National Institutes of Health, Bethesda, MD, USA.

References

- Gomes I, Collins A, Lonjou C *et al*: Hardy–Weinberg quality control. *Ann Hum Genet* 1999; **63**: 535–538.
- Wittke-Thompson JK, Pluzhnikov A, Cox NJ: Rational inferences about departures from Hardy–Weinberg equilibrium. *Am J Hum Genet* 2005; **76**: 967–986.
- Samani NJ, Erdmann J, Hall AS *et al*: Genomewide association analysis of coronary artery disease. *N Engl J Med* 2007; **357**: 443–453.
- Winkelmann J, Schormair B, Lichtner P *et al*: Genome-wide association study of restless legs syndrome identifies common variants in three genomic regions. *Nat Genet* 2007; **39**: 1000–1006.
- Zanke BW, Greenwood CMT, Rangrej J *et al*: Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. *Nat Genet* 2007; **39**: 989–994.
- Gordon D, Heath SC, Liu X, Ott J: A transmission/disequilibrium test that allows for genotyping errors in the analysis of single-nucleotide polymorphism data. *Am J Hum Genet* 2001; **69**: 371–380.
- Gordon D, Ott J: Assessment and management of single nucleotide polymorphism genotype errors in genetic association analysis. *Pac Symp Biocomput* 2001; **18**: 29.
- Douglas JA, Skol AD, Boehnke M: Probability of detection of genotyping errors and mutations as inheritance inconsistencies in nuclear-family data. *Am J Hum Genet* 2002; **70**: 487–495.
- Sobel E, Papp JC, Lange K: Detection and integration of genotyping errors in statistical genetics. *Am J Hum Genet* 2002; **70**: 496–508.
- Carvalho B, Bengtsson H, Speed TP, Irizarry RA: Exploration, normalization, and genotype calls of high-density oligonucleotide SNP array data. *Biostatistics* 2007; **8**: 485–499.
- Rabbee N, Speed TP: A genotype calling algorithm for affymetrix SNP arrays. *Bioinformatics* 2006; **22**: 7–12.
- Falconer DS, Mackay TFC: Introduction to quantitative genetics. New York, Longman 1997.
- Kang SJ, Finch SJ, Haynes C, Gordon D: Quantifying the percent increase in minimum sample size for SNP genotyping errors in genetic model-based association studies. *Hum Hered* 2004; **58**: 139–144.
- Kang SJ, Gordon D, Finch SJ: What SNP genotyping errors are most costly for genetic association studies? *Genet Epidemiol* 2004; **26**: 132–141.
- Ahn K, Haynes C, Kim W *et al*: The effects of SNP genotyping errors on the power of the cochrane-armitage linear trend test for case/control association studies. *Ann Hum Genet* 2007; **71**: 249–261.
- Cronin S, Berger S, Ding J *et al*: A genome-wide association study of sporadic ALS in a homogenous Irish population. *Hum Mol Genet* 2008; **17**: 768–774.
- Wallace C, Newhouse SJ, Braund P *et al*: Genome-wide association study identifies genes for biomarkers of cardiovascular disease: serum urate and dyslipidemia. *Am J Hum Genet* 2008; **82**: 139–149.
- Hunter DJ, Kraft P, Jacobs KB *et al*: A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet* 2007; **39**: 870–874.
- Yeager M, Orr N, Hayes RB *et al*: Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. *Nat Genet* 2007; **39**: 645–649.
- Saxena R, Voight BF, Lyssenko V *et al*: Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 2007; **316**: 1331–1336.
- Gudmundsson J, Sulem P, Manolescu A *et al*: Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. *Nat Genet* 2007; **39**: 631–637.
- Scott LJ, Mohlke KL, Bonnycastle LL *et al*: A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* 2007; **316**: 1341–1345.
- Sladek R, Rocheleau G, Rung J *et al*: A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 2007; **445**: 881–885.
- Zou GY, Donner A: The merits of testing Hardy–Weinberg equilibrium in the analysis of unmatched case-control data: a cautionary note. *Ann Hum Genet* 2006; **70**: 923–933.
- Hosking L, Lumsden S, Lewis K *et al*: Detection of genotyping errors by Hardy–Weinberg equilibrium testing. *Eur J Hum Genet* 2004; **12**: 395–399.
- Leal SM: Detection of genotyping errors and pseudo-SNPs via deviations from Hardy–Weinberg equilibrium. *Genet Epidemiol* 2005; **29**: 204–214.
- Cox DG, Kraft P: Quantification of the power of Hardy–Weinberg equilibrium testing to detect genotyping error. *Hum Hered* 2006; **61**: 10–14.
- Teo YY, Fry AE, Clark TG, Tai ES, Seielstad M: On the usage of HWE for identifying genotyping errors. *Ann Hum Genet* 2007; **71**: 701–703.
- Li M, Li C: Assessing departure from Hardy–Weinberg equilibrium in the presence of disease association. *Genet Epidemiol* 2008; **32**: 589–599.