Research Article   **Open Access**

The Influence of Errors Inherent in Genome Wide Association Studies (GWAS) in Relation To Single Gene Models

Philip Cooley[1]*, Robert F. Clark[1] and Grier Page[2]

[1]*RTI International, Research Triangle Park, North Carolina*
[2]*RTI International, Atlanta, Georgia*

Abstract

Nearly one thousand human genome wide association studies (GWAS) have examined over 210 diseases and traits and found over 1,200 SNP associations. With improved genotyping technologies and the growing number of available markers, case-control Genome Wide Association Studies (GWAS) have become a key tool for investigating complex diseases. This study assesses the influence of genotype and diagnosis errors present in GWAS by analyzing a synthetic gene dataset incorporating factors known to influence association measurement. Monte Carlo methods were used to generate the synthetic gene data, which incorporated factors including gene inheritance, relative risk levels, disease penetrance, genotype distribution, sample size, as well as the two error factors that are the focus of this study. The resulting dataset provides a truth set for assessing statistical method performance and association sensitivity.

While previously understood, these results quantify and document the extent of the relationship between genotype and diagnosis error measures and statistical power loss. Our results also demonstrate that for low risk non-recessive loci, sample sizes in the range of 1,000 - 2,000 cases will achieve 80% power thresholds for error type I error levels of 10^{-8} even with realistic genotype and phenotype error assumptions. Nevertheless, compensating for power loss due to the presence of genotype and diagnosis errors by increasing sample size should not be underestimated. Our estimates indicate that sample size increase requirements are in the range of 20% to 40%, depending on the gene inheritance model assumed.

Introduction

Over 900 human genome wide association studies (GWAS) have examined over 210 diseases and traits and found over 1,200 SNP associations [1]. With improved genotyping technologies and the growing number of available markers, case-control GWAS have become a key tool for investigating complex diseases. Because GWAS have become a standard primary investigative tool, researchers need to be aware of how errors influence their studies and how to overcome or compensate for them. The initial step in a GWAS is to apply univariate statistical tests for each single nucleotide polymorphism (SNP) in the dataset. Applying the tests is statistically straightforward and done with several methods (e.g., χ^2 tests, regression methods) that are standard approaches.

Studies on the consequences of genotype error have led to a modest number of investigations in the statistical genetics literature. Gordon [2] and colleagues investigated the effects of three published models of genotyping errors on the 2df genotype chi-square test. In another study, Gordon et al. [3] described a statistical power calculator (PAWE-3D) that produces power and sample size calculations that can support study designs for GWAS, and computes power and/or sample size requirements for a specified significance level. Zheng et al. [4] and Edwards et al. [5] contributed to the development of PAWE. Gordon et al. [6] analyzed the influence of both random phenotype and genotype misclassification errors on statistical power contrasting the Cochran Armitage Trend test (CA-A) with the 2 df genotype test and concluded that the CA-A is more powerful.

Ahn et al. [7] addressed the effect of different types of genotyping errors on statistical power in GWAS. While their prior work focused on non-differential genotype error rates, this study considered errors in each of the three bi-allelic genotypes differentially. The methods were based on a Taylor-series expansion of a non-centrality parameter of the asymptotic distribution of the trend test. In a follow-up study, Ahn et al. [8] extended their work by developing a closed form analytic procedure for both the 2df genotype and the Cochran Armitage trend tests. They reported that misclassifying the heterozygote genotype is particularly detrimental when using the recessive trend test (CA-R) on data from a recessive mode of inheritance (MOI) model.

While the accuracy of the genotyping process has improved, data errors still occur. Hao et al. [9] reported an overall 0.5% error rate imputation process but they also reported a 2% error rate in underrepresented subpopulations. Miclaus et al. [10] examined genotype calling algorithms on HapMap samples and found that different algorithms can produce genotyping errors that influence downstream genotype calls. They reported a 2-3% error estimate attributable to the genotype-calling algorithm. Laurie et al. [11] estimated genotyping error rates from duplicate sample discordance rates from Addiction and Lung Cancer projects genotyped on Illumina Human1Mv1_c and HumanHap550-2v3_b arrays by the Center for Inherited Disease Research (CIDR). The investigators calculated genotyping error rates on the order of 10^{-4}, which corresponds to

*Corresponding author: Philip Cooley, Senior Fellow, Research Computing Division, RTI International, 3040 Cornwallis Road, Research Triangle Park, NC 27709, E-mail: pcc@rti.org

mean completion call rates of 99.7% and 99.8% respectively for the two projects. If study samples are not duplicated, as in the Type-2 Diabetes project, but with multiple replicates of the HapMap control sample, discordant rates of 1-4 x 10^{-3} lead to completion rates of 99.6-99.7%.

Phenotypic misclassification errors are also a source of bias and can reduce the power to detect a statistical association between a phenotype and a specific allele [12-15]. Edwards et al. [5] presented a quantification of the effect of phenotypic error on power and sample size calculations for case-control genetic association studies between a marker locus and a disease phenotype [5]. Barendse et al. [13] described the effect of investigator measurement-error on the phenotypes – the error was significant when looking at quantitative traits. When the traits were coded as affected or unaffected the error effect sizably decreased (14.5% down to 5.3%). For many diseases, the interrater agreement for disease diagnosis can be quite low. For example, the sensitivity and specificity of pre-death Alzheimer's diagnosis with post mortem autopsies can be as low as 0.83 and 0.84 respectively [16].

Gene traits that influence prediction accuracy have been also reported in other studies. For example, Sasieni [17] and Freidlin, et al. [18] demonstrated that the phenotype mode of inheritance (MOI) model was a major influence on association prediction accuracy.

To help provide additional insight into the influence of genotype and diagnosis errors affecting the accuracy of the phenotype measure in a GWAS, we ran simulations with synthetically generated data. We focused on assessing the impact on statistical power caused by the influence of these two often overlooked errors. Our simulations demonstrated that genotype (even at low error rates) and phenotype (diagnosis) errors produce substantial power losses for all MOIs, with significant power losses for recessive MOIs. Because GWAS involving recessive loci have additional power requirements relative to other MOI types, researchers need to address these requirements in developing appropriate sample sizes for their studies.

Methods

Our approach is based entirely on a simulation framework. In the following sections, we identify the data generation method we used to produce our synthetic data and how we used these data to assess the influence of errors on statistical power loss. While a similar version of the data generation process was presented in Cooley et al. [19] the description below adds specific material on how the non-differential errors were incorporated into the synthetic gene data and the impact of genotype and diagnosis errors on statistical power performance.

We developed our assessments by analyzing a dataset of synthetic gene data that incorporates factors known to influence association measurements in GWAS. These include phenotypic errors (i.e., due to improper disease diagnosis) and genotype errors (due to incorrect genotype calls). We employed Monte Carlo methods to generate simulated gene data that we analyzed to assess the influence of the individual factors on statistical power in the context of GWAS. There are two advantages to using simulated data. First, the association-affecting factors are isolated and can be linked to the affecting locus. Second, we can choose any specific statistical method to perform the association assessment. The simulated dataset provides a truth set for assessing the role of statistical methods on association sensitivity and highlights the particular role of errors in disease diagnosis and incorrect genotype assignments.

Generating the synthetic SNP data

We derived our data generation method from a study by Iles [20] and from Mendelian concepts of inheritance, and specifically incorporated autosomal dominant, recessive, additive and multiplicative inheritance patterns. These data incorporate factors known to influence the association of the measurements in the context of GWAS. We began the generation process with disease penetrance. Because our simulation process assumed Mendelian behaviors, we expected that the findings would apply to genes that exhibit these types of patterns. Thus, the findings might not be typical for unknown (poly-gene) diseases.

Penetrance is defined as the proportion of individuals without the risk allele that has a specific trait (phenotype). In other words, it is a genotype-specific probability of being affected with disease. We designated **a** as the risk allele, and **A** as the allele without risk. Generating the synthetic dataset was straightforward by using the relationships between penetrance and risk for different MOI categories (see Cooley et al. [19] for further detail).

Initially, we identified:

- n_j = the target number of cases and controls in a given experiment,

- P_j = the disease penetrance,

- $Perr_j$ = the misclassification error rate contained in the phenotype data, (0, 2, 5%),

- $Gerr_j$ = the misclassification error rate contained in the genotype data, (0, .5, 1%),

- Φ_j = the relative risk, (1.0, 1.15, 1.3, 1.45), and

- g0, g1, g2 the distribution of genotypes which were drawn at random from a master set of genotype distributions obtained from real SNP data, Schymick et al. [21].

In screening samples from the master set, Chan et al. [22] recommends not applying a minor allele frequency (MAF) threshold as a filter. They argue that filtering MAFs out of the process because of low frequencies or to maintain Hardy–Weinberg equilibrium (HWE) deviation has little effect on the overall false positive rate and in some cases, filtering on MAF only serves to exclude SNPs. The effect of this step is to select a specific genotype distribution (at random) from the master distribution.

From the selected relative risk (Φ_j), penetrance (P_j) and MOI assumptions, the formulas in Table 1 are used to assign a case or control code (1,0). This step converts the relative risk ratio, Φ_j into the probability of a case (disease) given the MOI gene model assumed. This genotype specific process can be represented by the following logic.

Major Homozygote (AA): If the AA (non disease) genotype is selected, the probability of a case equals the disease penetrance P_j.

Minor Homozygote (aa): Ψ_{aa} is the ratio of two probabilities: the probability of a case for a minor homozygote divided by the probability of a case for a major homozygote, i.e.,

$$\Psi_{aa} = \text{Prob(case/aa)} / \text{Prob(case/AA)} = x/P_j. \qquad (1)$$

Thus the probability of a case (x) given the minor genotype is:

$$x = \Psi aa * Pj \qquad (2)$$

where Ψ_{aa} = an assumed risk factor.

Heterozygote (aA): By the same argument, the phenotype risk given a heterozygote is:

$$\Psi aA = \text{Prob(case/aA)} / \text{Prob(case/AA)} = y/P_j, \qquad (3)$$

Thus, the probability of a case (y) given the heterozygote genotype is:

$$y = \Psi aA * Pj \qquad (4)$$

where ΨaA is the assumed risk factor and P_j is the assumed penetrance.

Note that implicit in equations (1), (2) and (3) is a consistent definition of penetrance defined as the proportion of cases that are present in the major genotype AA.

Using the estimate of x from equation (2) and y from equation (4), we assigned a case or control at random using the four different MOI models in Table 1 below. For the MOI models that assume an elevated risk from the minor and the heterozygote genotypes, we would expect a higher proportion of cases to be more easily identified via the statistical procedures. Specifying risk depends on specific and unknown disease mechanisms. A relative risk of 1.7 is considered strong and is associated with positive replication [23], and a risk of 1.3 is considered by Ziegler et al. [24] to be a realistic assumption for complex diseases. We limited our focus to a relative risk range of 1.15 to 1.45 and were particularly interested in cases with low relative risk.

Errors were introduced into both genotype and the phenotype assignments. The first pass (described above) generated an error free record. During the second pass, the "correct" data was altered to reflect diagnostic errors and/or errors made by the genotyping platform. Our data included examples in which random changes in both phenotype and genotype data were introduced at three error rates, including a zero error rate. To introduce phenotype errors, we selected a record at random and changed its designation from case to control or from control to case, depending on the original assignment. In a separate step, we introduced genotype errors into the data. We selected a record and a genotype at random and altered the genotype code that corresponded to mistakes made by either the chip or human recorder during the genotype assignment. In the genotype error simulation, the error is assumed to involve any of the three genotypes and is non-differential. Accordingly, the first step selected a genotype code at random. For example, if a minor homozygote genotype was selected, that genotype code was changed to either a heterozygote genotype code or to a major homozygote genotype, also selected at random.

- This process continued until we generated n1 cases and n2 controls (note in this example n1 = n2, but that can be tailored to specific n1 – n2 targets).

- We then applied a set of statistical methods to predict associations and record the results.

- For each set of 3,456 factor combinations (i.e., 3 penetrance levels by 8 sample sizes by 9 error combinations by 4 relative risk levels by 4

	Major homozygote	Minor homozygote	Heterozygote
MOI	Ψ_{aa}	Ψ_{AA}	Ψ_{aA}
Recessive	1	Φ	1
Dominant	1	Φ	Φ
Additive	1	$2*\Phi-1$	Φ
Multiplicative	1	$\Phi*\Phi$	Φ

Where:

$$\Psi AA = \frac{Pr(case / aa)}{Pr(case / AA)}$$

$$\Psi aA = \frac{Pr(case / aA)}{Pr(case / AA)}$$

Table 1: Relative Risk Assumptions by Mode of Inheritance [18].
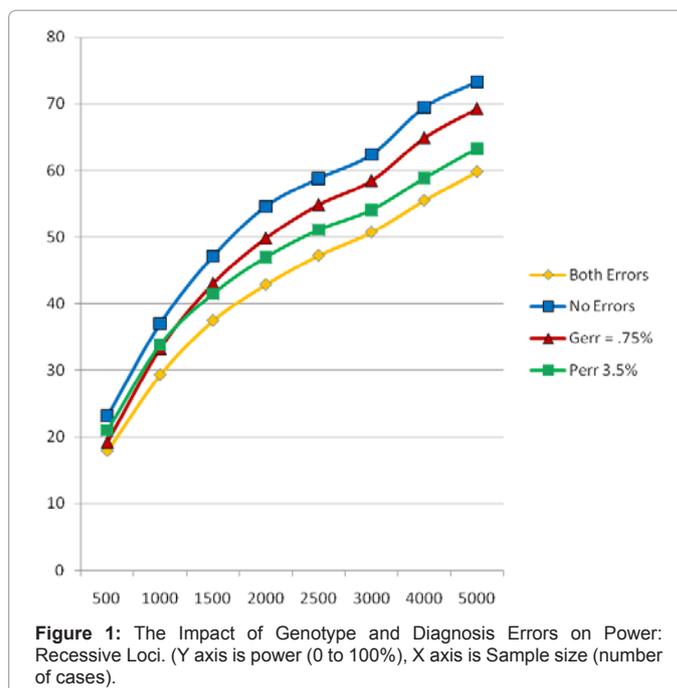

Figure 1: The Impact of Genotype and Diagnosis Errors on Power: Recessive Loci. (Y axis is power (0 to 100%), X axis is Sample size (number of cases).

MOI categories) we generated 1,000 replicate experiments.

Results

Dataset summary

Using the methods described above, we generated a synthetic gene simulated dataset with the following characteristics:

- The proportion of cases (controls) that are Major Homozygotes = 50.3 (63.0) %.

- The proportion of cases (controls) that are Heterozygotes = 39.2 (31.3) %.

- The proportion of cases (controls) that are Minor Homozygotes = 10.5 (5.7) %.

  - With MOI distribution:

o Recessive = 25%,
o Dominant = 25%,
o Additive = 25%, and
o Multiplicative = 25%,

While the distribution of MOI traits above used in our analysis is not an accurate representation of a "true" distribution, we currently know of no accurate way to obtain this distribution. Consequently, even though we gave each of the four MOI traits equal representation in the simulated data, we confined our examinations to within-MOI assessments.

Error analyses

To simulate the association estimation process in a GWA experiment, we applied three variations of the Cochran Armitage (CA) Trend test to each of the 1,000 replicates of the 3,456 possible data subsets. Each of the variations of the CA tests used a distinct genotype score vector: [0,0,1] for recessive, [0,0.5,1] for additive and [0,1,1] for dominant. These tests are described by several researchers [25,26]. We applied each of the tests to all of the replicates in each of the data

subsets. This process allowed us to show that the optimal strategy for maximizing statistical power is MOI specific. This strategy posits that the recessive version (CA-R) be used to estimate associations involving recessive loci data, the dominant version (CA-D) to the dominant loci data, and the additive version (CA-A) to both additive and multiplicative loci data. This strategy is cited by others for single gene models [23]. Cooley et al. provided a similar assessment and identified a multiple test strategy that combines the three tests into an overall score that has merit if the MOI of the causative loci is not known [17]. For this assessment, we assumed that the MOI is known and selected the best statistical method to measure the association. Consequently, our results tended to be optimistic.

Error rates of 0%, 2% and 5% are incorporated into the simulated data for the phenotype and 0% .5% and 1% for the genotype. Our approach combines the three risk levels (mean risk = 1.3), three penetrance levels (mean penetrance = 0.4), and groups the data into a "with error" (mean phenotype error = 3.5%, and mean genotype
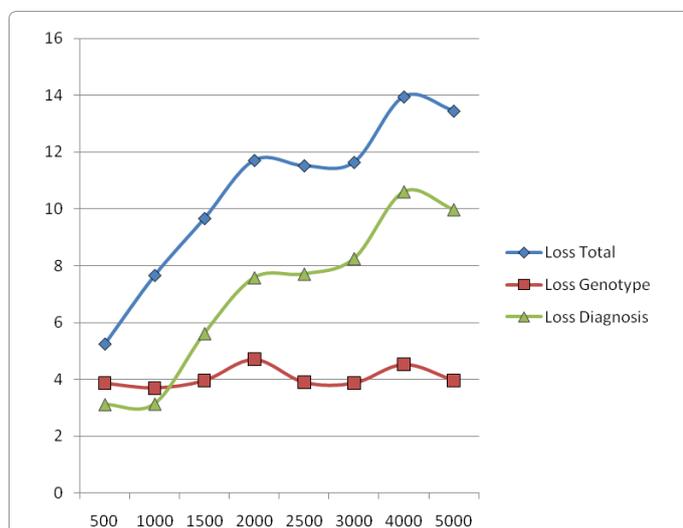

Figure 1a: Power Loss: Total, Genotype and Diagnosis: Recessive Loci. (Y axis is power (0 to 16%), X axis is Sample size (number of cases).
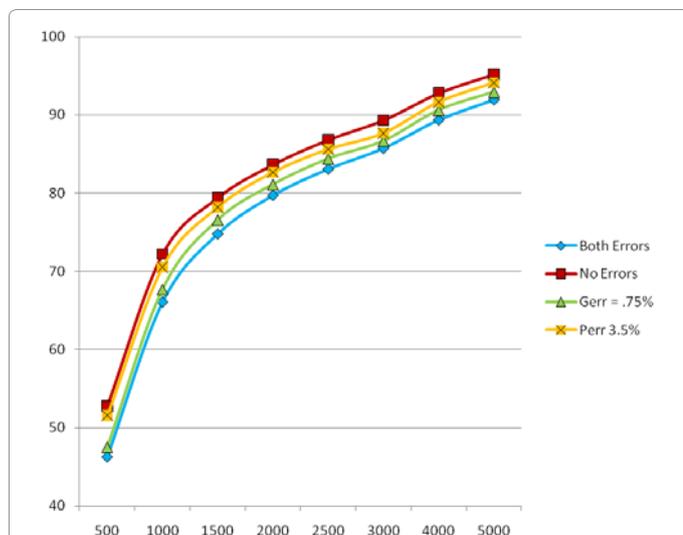

Figure 2: The Impact of Genotype and Diagnosis Errors on Power: Dominant Loci. (Y axis is power (40 to 100%), X axis is Sample size (number of cases)).

MOI	Genotype N = 500	Errors N = 5,000	Diagnosis N = 500	Errors N = 5,000	Both N = 500	Errors N = 5,000
R	3.86	3.96	3.11	9.98	5.23	13.45
D	5.23	3.43	1.21	1.06	6.46	3.26
A	4.57	1.27	1.97	1.36	6.51	2.69
M	4.09	0.98	1.81	1.40	5.58	2.35

Table 2: Max Power Loss due to Genotype or Diagnosis Error.

0.75%) and "without error" strata. Also, we also stratified the analysis by MOI. Figures 1 thru 4 identify the 4 MOI specific results. Each figure includes a 0.75% genotype error curve, a 3.5% phenotype error curve, a curve that includes both error sources and a curve generated without either source of error. Figure 1 presents the recessive loci analysis. The impact of a 0.75% average genotype error rate and a 3.5% average diagnosis error rate with respect to power loss for recessive loci is non-trivial. However, each profile exhibits distinct behavior. The effect of the phenotype error increases with N and peaks at N = 4,000 cases, whereas the genotype error effect is constant across all N. Also observed at the peak is a genotype impact of 6.04% per 1.0% genotype error and a power loss of 3.03% per 1.0% phenotype error. Note that with $\alpha < 10^{-8}$ as the significance threshold, an 80% power target is far from being realized even with N = 5,000 cases and controls.

Figures 2, 3, and 4 present the error effects of the dominant, additive and multiplicative loci respectively. All three figures indicate that power loss is non-trivial for the MOI categories they represent but that the effect is substantially less than recessive modes.
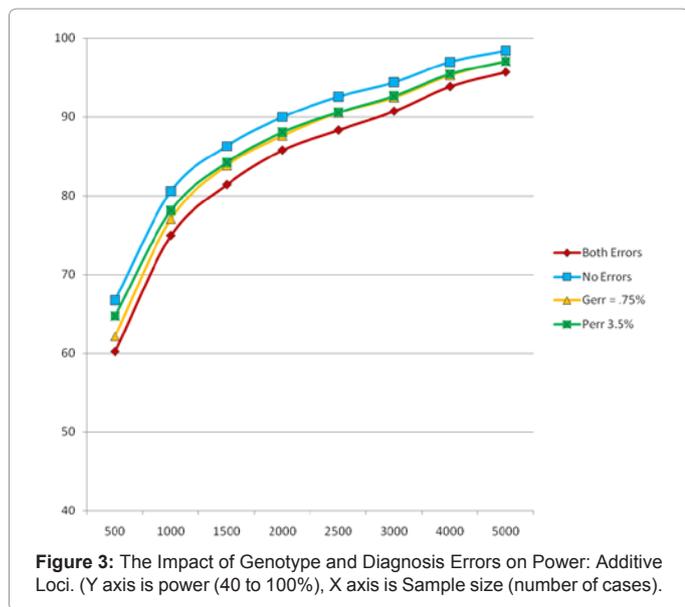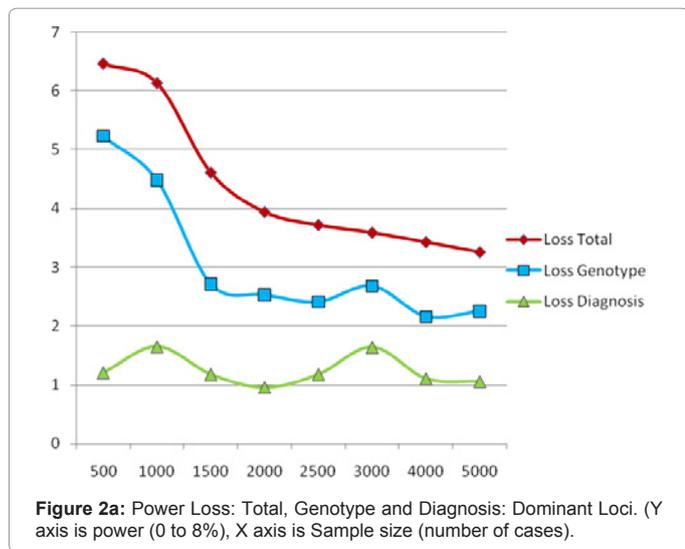
The pattern of the power profile for dominant loci is in sharp contrast to the recessive loci profile. The diagnosis error pattern is constant across N for dominant loci – the recessive loci show an increasing pattern. The genotype patterns are also different. As N increases, the power differences decline for the dominant loci whereas the patterns are constant for recessive loci.

The additive and the multiplicative loci show similar error profiles. The impact of a 3.5% diagnosis error and a 0.75% genotype error has a similar quantitative impact. Both have a declining power loss as N increases.

In summary, the four figures indicate that the genotype error versus the diagnosis errors effects vary by MOI. For the recessive MOI, a 3.5% diagnosis error has a larger impact than a 0.75% genotype error. This result is reversed in the dominant MOI scenarios (Figure 2). The additive and the multiplicative MOI scenarios represented in Figure 3 and Figure 4 indicate that a 0.75% genotype error is comparable in effect to a 3.5% diagnosis error with respect to power loss.

These results are summarized in Table 2, which displays the power loss for the smallest sample size (N = 500 cases) and the largest sample size (N = 5,000 cases). For example, row R (recessive) of Table 2 illustrates that error loss due to genotype errors at N = 500 and N = 5,000 is flat, but that error loss due to diagnosis error increases dramatically from N = 500 to N = 5,000 and dominates the total error profile. The power loss pattern changes for the other 3 MOIs where error loss patterns for both genotype and diagnosis sources decline from N = 500 to N = 5,000.

We also examined the simultaneous influence of relative risk and error effects on statistical power. As above, we analyzed our dataset using all three penetrance levels (mean penetrance = 0.4), but we also stratified the curves by the low (1.15), the medium (1.3) and the high risk (1.45) categories. Figure 5 displays the combined (genotype plus diagnosis) error effects for the three risk categories using the CA-A

Figure 2a: Power Loss: Total, Genotype and Diagnosis: Dominant Loci. (Y axis is power (0 to 8%), X axis is Sample size (number of cases).

Figure 3: The Impact of Genotype and Diagnosis Errors on Power: Additive Loci. (Y axis is power (40 to 100%), X axis is Sample size (number of cases).

indicate that both sources of error can adversely affect power levels. This outcome is more pronounced for recessive MOI and low risk loci, which is common knowledge. What our study shows is that the error magnitude depends on a variety of factors in addition to MOI, especially relative risk and sample size; our study quantifies this magnitude and indicates the significance of this impact. This loss can be compensated for by increasing sample sizes. Gordon et al. [2] reported that a 1% increase in genotype error rates requires an increase in sample size of 2-8%, which they also noted depends on the MOI scenario. Our estimates are much higher than those reported by Gordon et al. [2] and are based on achieving a power threshold of 80%. Using the additive model, results at N = 1,000 (assuming no genotype errors) exceeds the 80% threshold (80.6%). Introducing a 1% genotype error, power drops to 75%. An additional 405 cases are needed to compensate for this loss to restore an 80.6% power level, which is a 40.5% increase

Figure 3a: Power Loss: Total, Genotype and Diagnosis: Additive Loci. (Y axis is power (0 to 8%), X axis is Sample size (number of cases).

Figure 4: The Impact of Genotype and Diagnosis Errors on Power: Multiplicative Loci. (Y axis is power (40 to 100%), X axis is Sample size (number of cases).
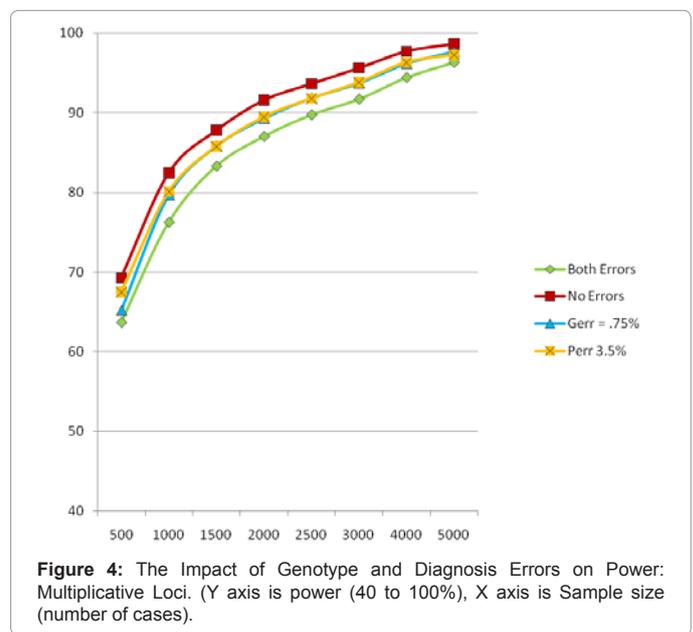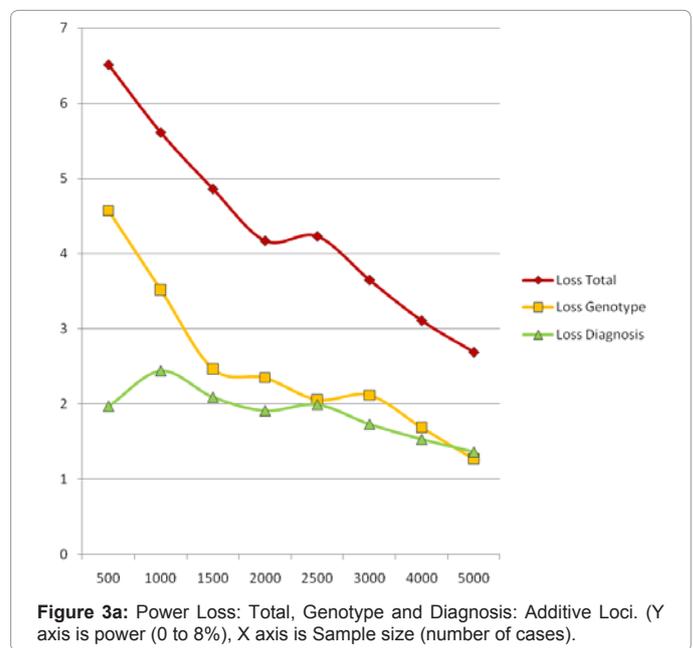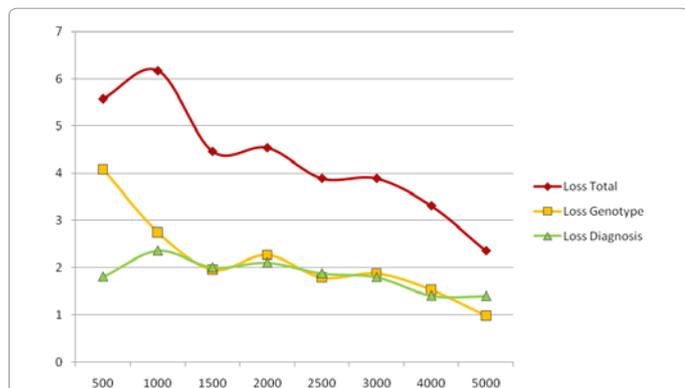
(additive) method applied to the additive MOI data. Similar curves can be generated for the dominant and multiplicative scenarios. This figure suggests that for additive inheritance scenarios, researchers can predict associations in the context of GWAS with a type I error threshold of $\alpha < 10^{-8}$ and still achieve a power level greater than 80%. This statement applies to low risk loci even when diagnosis errors are 3.5% and genotype errors are 0.75%.

Figure 6 shows the same results for the recessive scenario. In this recessive scenario, the likelihood of achieving an 80% power level is low and is only possible for high risk loci in the absence of genotype and diagnosis error with a sample size N larger than attempted by our simulation experiments.

Summary/Discussion

We examined the influence of genotype and diagnosis errors that affect the accuracy of association predictions in a GWAS and focused on assessing the effect on statistical power loss caused by the influence of these two sources of error. Our findings are MOI specific and

Figure 4a: Power Loss: Total, Genotype and Diagnosis: Multiplicative Loci. (Y axis is power (0 to 8%), X axis is Sample size (number of cases).
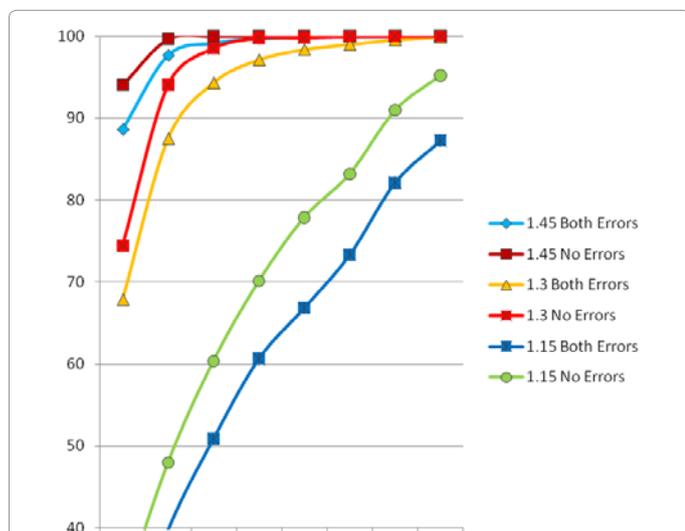

Figure 5: Genotype error, Recessive MOI by Risk level and N. (High = Relative Risk = 1.45; Med = Relative Risk = 1.30; Low = Relative Risk = 1.15. Y axis is power (0 to 100.0), X axis is Sample size (number of cases).

Figure 6: Phenotype error, Recessive MOI by Risk level and N. (High = Relative Risk = 1.45; Med = Relative Risk = 1.30; Low = Relative Risk = 1.15.). Y axis is power (0 to 100.0), X axis is Sample size (number of cases).
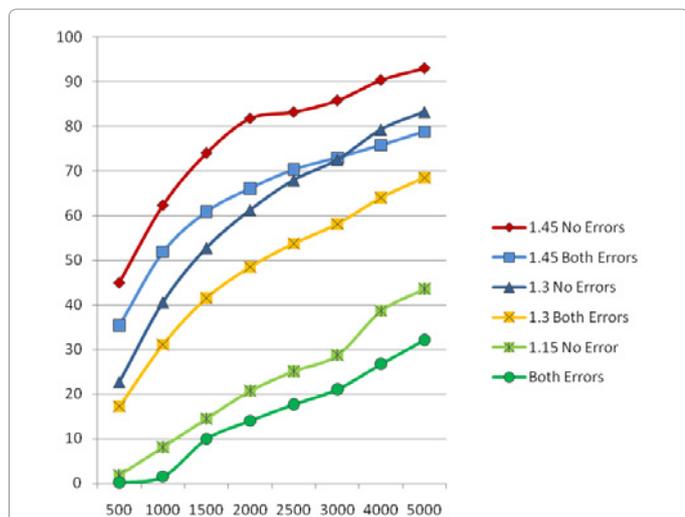
MOI	Genotype %	Diagnosis %
R	57.2	35.9
D	40.1	19.7
A	40.5	20.9
M	40.2	18.9

Table 3: % Sample Size Increase to Restore Power caused by a 1% Genotype or Diagnosis Error.

in sample size. Note: we are not suggesting that 1% error is standard operating procedure. In fact, genotype errors are improving with the introduction of each new technology, and currently are likely below 0.5%. Table 3 presents these results for all MOIs for both Genotype and Diagnosis errors.

In summary, our results quantify the relationship between genotype and diagnosis error measures and statistical power loss. These relationships are understood, but we document their extent. Our results also assume that the MOI of the locus being analyzed is known; therefore, our results will understate the true power loss and the compensating sample size increases. Our results also demonstrate that for low risk non-recessive loci, sample sizes in the range of 1,000 – 2,000 cases will achieve 80% power thresholds for type I error levels of 10^{-8} even with realistic genotype and phenotype error assumptions.

However, the recessive loci model remains problematic. Desirable power thresholds for moderate risk levels can only be realized with sample sizes in the tens of thousands, which is further complicated by accounting for power loss as a result of genotype and diagnosis errors.

References

1. Hindorff LA, Junkins HA, Hall PN, Mehta JP, and Manolio TA (2011) A Catalog of Published Genome-Wide Association Studies.

2. Gordon D, Finch SJ, Nothnagel M, Ott J (2002) Power and sample size calculations for case-control genetic association tests when errors present: application to single nucleotide polymorphisms. Hum Hered 54: 22-33.

3. Gordon D, Haynes C, Blumenfeld J, Finch SJ (2005) PAWE-3D: visualizing power for association with error in case/control genetic studies of complex traits. Bioinformatics 21: 3935-3937.

4. Zheng G, Tian X (2005) The impact of diagnostic error on testing genetic association in case-control studies. Stat Med 24: 869-882.

5. Edwards BJ, Haynes C, Levenstien MA, Finch SJ, and Gordon D (2005) Power and sample size calculations in the presence of phenotype errors for case/control genetic association studies, BMC Genetics. 6:18.

6. Gordon D, Haynes C, Yang Y, Kramer PL, Finch SJ (2007) Linear trend tests for case-control genetic association that incorporate random phenotype and genotype misclassification error. Genet Epidemiol 31: 853-870.

7. Ahn K, Haynes C, Kim W, Fleur RS, Gordon D, Finch SJ (2007) The effects of SNP genotyping errors on the power of the Cochran-Armitage linear trend test for case/control association studies, Ann Hum Genet 2007: 249-61.

8. Ahn K, Gordon D, Finch SJ (2009) Increase of rejection rate in case-control studies with differential genotype error rates. Stat Appl Genet Mol Biol 8.

9. Hao K, Chudin E, McElwee J, Schadt EE (2009) Accuracy of genome-wide imputation of untyped markers and impacts on statistical power for association studies. BMC Genet 10: 27.

10. Miclaus K, Chierici M, Lambert C, Zhang L, Vega S, et al. (2010) Variability in GWAS analysis: the impact of genotype calling algorithm inconsistencies, Pharmacogenomics J 10: 324-335.

11. Laurie CC, Doheny KF, Mirel DB, Pugh EW, Bierut LJ, et al. (2010) Quality control and quality assurance in genotypic data for genome-wide association studies. Genetic Epidemiol 34: 591-602.

12. Gordon D (2006) Gene mapping: balance among quality, quantity and cost of data in the era of whole-genome mapping for complex disease. Eur J Hum Genet 14: 1147-1148.

13. Barendse, William (2011) The effect of measurement error of phenotypes on genome wide association studies. BMC Genomics 12: 232.

14. Hughes AJ, Daniel SE, Kilford L, Lees AJ (1992) Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. J Neurol Neurosurg Psychiatry 55: 181-184.

15. Samuels DC, Burn DJ, Chinnery PF (2009) Detecting new neurodegenerative disease genes: does phenotype accuracy limit the horizon? Trends Genet 25: 486-488.

16. Blacker D, Albert MS, Bassett SB, Go RCP, Harrell LH, et al. (1994) Reliability and validity of NINCDS/ADRDA criteria for Alzheimer's disease: the NIMH Genetics Initiative. Arch Neurol 51: 1198.

17. Sasieni PD (1997) From genotypes to genes: doubling the sample size. Biometrics 4: 1253-1261.

18. Freidlin B, Zheng G, Li Z, Gastwirth JL, et al. (2002) Trend tests for case-control studies of genetic markers: power, sample size and robustness. Hum Hered 53: 146-152.

19. Cooley P, Clark R, Folsom R, Page G (2010) Genetic Inheritance and Genome Wide Association Statistical Test Performance. J Proteomics Bioinform 3: 321-325.

20. Iles MM (2002) Effect of mode of inheritance when calculating the power of a transmission/disequilibrium test study. Hum Hered 53: 153-157.

21. Schymick JC, Scholz SW, Fung H, et al. (2007) Genome-wide genotyping in amyotrophic lateral sclerosis and neurologically normal controls: first stage analysis and public release of data. Lancet Neurol 6: 322-328.

22. Chan EKF, Hawken R, Reverter A (2009) The combined effect of SNP-marker and phenotype attributes in genome-wide association studies. Anim Genet 40: 149-156.

23. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, et al. (2007) A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 445: 881-885.

24. Ziegler A, Inke RK, Thompson JR (2008) Biostatistical aspects of genome-wide association studies. Biom J 50: 1-21.

25. Kuo C, Feingold E (2010) What's the best statistic for a simple test of genetic association in a case-control study? Genet Epidemiol 34: 246-253.

26. Zheng G, Gastwirth JL (2006) On estimation of the variance in Cochran–Armitage trend tests for genetic association using case–control studies. Statist Med 25: 3150-3159.