Genetic Epidemiology 17(Suppl 1): S61–S66 (1999)

The Effect of Phenotype Variation on Detection of Linkage in the COGA Data

Gunther Birznieks, Soumitra Ghosh, Richard M. Watanabe, and Braxton D. Mitchell

Genetics and Molecular Biology Branch (G.B., S.G.), National Human Genome Research Institute, Bethesda, Maryland; Department of Biostatistics (R.M.W.), University of Michigan School of Public Health, Ann Arbor, Michigan; Department of Genetics (B.D.M.), Southwest Foundation for Biomedical Research, San Antonio, Texas

Error in phenotypic measurement can significantly compromise ability to detect linkage. We assessed the impact of introducing phenotypic measurement error on our ability to detect a quantitative trait locus in the Collaborative Study on the Genetics of Alcoholism (COGA) data. The impact of introducing three different types of errors was evaluated: 1) errors generated by sampling from a normal distribution; 2) errors generated by permuting phenotype values between subjects; and 3) errors generated by sampling from a uniform error distribution.

Key words: alcoholism, phenotypes, sensitivity analysis, quantitative traits

INTRODUCTION

The development of new analytic methods for quantitative trait linkage analysis will greatly facilitate efforts to detect genes for complex diseases. The appeal of these methods is their applicability over a range of genetic models, thus obviating the need to specify model parameters. However, the ability to detect linkage is influenced by the quality of the genotypic and phenotypic information. Relatively little is known about the sensitivity of some of these methods to variability in these data. The goal of this study is to investigate the sensitivity in our ability to detect linkage in the face of phenotype error. Such errors may occur for a variety of factors, including circadian rhythm, measurement error or because of assay errors resulting from variability of reagents required to quantitate the trait in question. In addition, collection errors such as data entry, sample ID switching, or sample switches in the lab may cause problems with the phenotypes.

Address reprint requests to Dr. S. Ghosh, Room 1W108, Building 9, 9000 Rockville Pike, Bethesda, MD 20892.

In this paper, we investigated the effect of introducing phenotypic measurement error on our ability to detect linkage to a quantitative trait in the Collaborative Study on the Genetics of Alcoholism (COGA) data. In an initial genome scan of the COGA data, we detected significant evidence for linkage to a quantitative trait using a variance component approach [Amos, 1994; Almasy and Blangero, 1998]. In addition, evidence for linkage for the same trait and chromosome was also previously found in this data set [Begleiter et al., 1998]. To address our study question, we performed a sensitivity analysis in which we introduced three different types of measurement error, in varying degrees, and then re-evaluated evidence for linkage at that locus.

METHODS

We analyzed the trait Cz for the P300 event-related potential (ERP) data for 604 people in a total dataset of 1,214 individuals. Linkage analyses were performed using the variance components approach of Almasy and Blangero [1998], in which the covariance of trait values in pairs of relatives is decomposed into major gene, polygenic and environmental components. Analyses were conducted using the SOLAR software program (Sequential Oligogenic Linkage Analysis Routines). Multipoint analysis revealed a maximum lod score of 3.54 at 198 cM on chromosome 6 using the information from all 15 markers from the COGA dataset. All subsequent analyses were then performed at this locus.

We performed a sensitivity analysis to evaluate the effect of introducing error in phenotype measurement on our ability to detect linkage. Our strategy involved replacing the Cz values of a subset of individuals with values taken from a pre-defined error distribution. Following replacement of the phenotype values, we recomputed the lod score at position 198.

We considered three different types of error distributions, each corresponding to a different source of error. For each type of error we introduced, we also evaluated the effect of varying the proportion of individuals whose phenotypes were altered. Specifically, for each type of error we evaluated the impact of altering phenotypes in 1%, 5%, 20%, 50%, and 100% of the total population. Each of these five sampling proportions was generated randomly.

The first type of error we considered corresponded to variability associated with imprecision in the measurement. Errors corresponding to this source were generated by replacing the observed phenotype value with a value drawn from an "error" distribution with mean equal to that of the observed value and variance, σ^2_e. To evaluate the effect of different degrees of imprecision, we assessed sensitivity for values of σ^2_e corresponding to 1%, 5%, 20%, 50%, 100%, 120%, 150%, 175%, 200%, 225%, 250%, 275%, and 300% of the original population variance of Cz P300 ($\sigma^2 = 74.96$). We then generated 500 replicates for each unique combination of sampling proportion (ranging from 1% to 100%) and variance (ranging from 1% to 300%) and linkage analysis was then performed on each replicate data set. Thus, a total of 27,500 (500 x 5 x 13) analyses were performed.

The second type of error we considered corresponded to errors introduced by sample swapping, as, for example, might occur by incorrectly matching the ID numbers to their corresponding phenotype measurements. We simulated this type of error by permuting the phenotypic values between individuals. Again, 500 replicate data sets were generated across each of the five sampling proportions, for a total of 2,500 separate analyses.

The third type of error we considered corresponded to errors introduced by an invalid measurement that was uncorrelated with the observed measurement. Such a

measurement could be obtained in a number of ways, including data recording error and instrument malfunction. To simulate this type of error, we replaced the phenotypic value with a value obtained from a uniform distribution, whose minimum and maximum values were defined by the corresponding values taken from the observed distribution of Cz P300 measurements. As before, we generated 500 replicate data for each of the five sampling proportions, for a total of 2,500 separate analyses.

RESULTS

The results obtained from introducing measurement imprecision into the data are shown in Figure 1, which shows the mean and standard error of the lod scores across the 500 replicates for each combination of generating parameters. As expected, for each of the five sampling proportions, the lod scores decreased as the variability in the replaced phenotypic measurements increased. Replacing only 5% of the sample reduces the lod score from 3.54 to 1.75 if the variability associated with the replaced values is 300% of the original population variance. Conversely, if measurement error (at 100% of the population variance) is introduced to 50% of the sample, lod score decreased to approximately 1.5. Interestingly, if the variance surrounding the replaced values is 50% of the population variance or less, the ability to detect linkage remained relatively high, even if phenotypes for large proportions of the population were replaced. In fact, if variability equaling 50% of the population variance was introduced for every single individual in the entire population, the lod score decreased to only 2.2.

We also examined the minimum and maximum lod scores out of the replicates for each sample proportion and variance combination. The minimum lod score was often below 2, suggesting that even for a small phenotype change, there may be a small subset of individuals whose change of phenotype led to a large drop in lod score even if the mean of all the replicates remains fairly stable. The maximum lod score often increased with small variance and sample changes but showed a general decline when both these factors became larger.

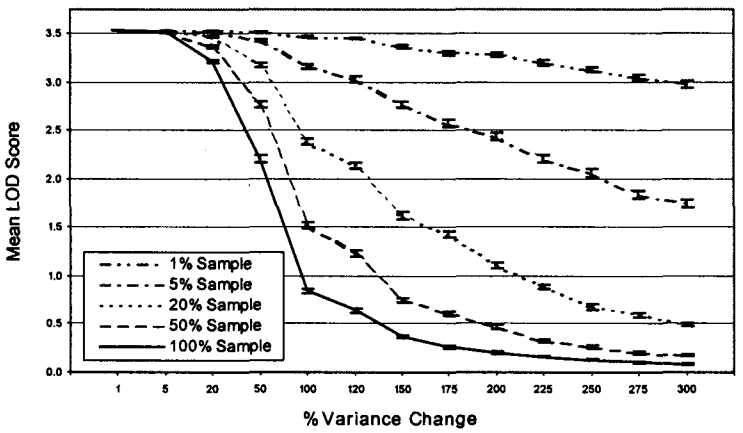

Fig. 1. Mean (and standard error) lod scores associated with introducing phenotypic imprecision for sampling proportions of 1%, 5%, 20%, 50%, and 100%.

The effects of introducing measurement error attributable to permutation errors and invalid measurements are shown in Figure 2. As expected, in both cases the mean lod scores decreased as the proportion of the individuals whose values were replaced increased. At each sampling proportion, ability to detect linkage was lower for invalid measurement errors than for permutation-based errors. Compared with the errors introduced by measurement imprecision, errors introduced by both permutation errors and invalid measurements tended to have a more severe effect on the ability to detect linkage. For example, if phenotypic values of only 20% of the population were replaced, the mean lod scores were 1.4 and 0.7 for the permutation and invalid measurements groups, respectively, compared to 2.45 to 0.5 for the imprecise measurement group, for phenotypic variability ranging from 100% to 300% of the original population variance.

Finally, we examined the proportion of replicates whose lod score appeared above a given threshold (Table I). As expected, an increase in the percent sample change led to fewer replicates above a specific threshold.

DISCUSSION

Measurement error is a ubiquitous and unavoidable feature of any scientific study. The impact of measurement error on quantitative trait linkage analysis has been left relatively unexplored. We have investigated this issue in the COGA data by examining the robustness of a linkage observed with Cz P300 values. We modeled three different types of errors that could occur in the measurement of this phenotype separately. Other sources of error are, of course, also possible, and in the "real world," measurement errors accrue not from a single source, but rather from multiple sources. Nevertheless, evaluation

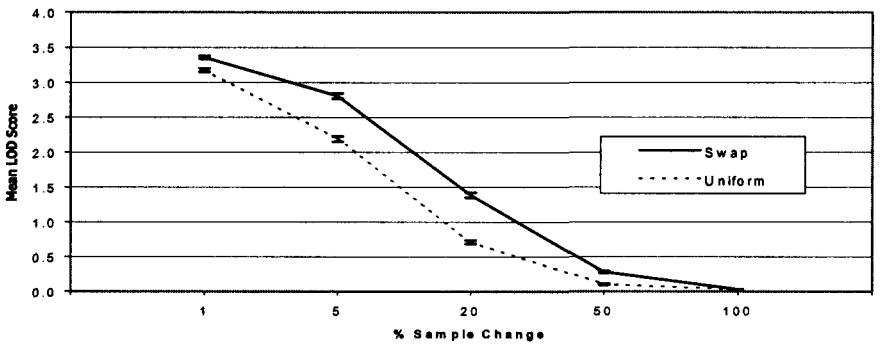

Fig. 2. Mean (and standard error) of lod scores associated with introducing phenotype measurement error attributable to permutation erros and invalid measurement for samplings of 1%, 5%, 20%, 50%, and 100%.

TABLE I. Percent of Replications Above the Lod Score Cutoff for Different Phenotype Errors

Lod cutoff	lod ≥ 2					lod ≥ 3				
% Sample	1	5	20	50	100	1	5	20	50	100
Uniform	94.8	57.8	23.6	3.8	0.0	67.0	16.8	0.4	0.0	0.0
Swap	98.4	84.4	21.4	0.2	0.0	86.6	44.2	4.6	0.0	0.0
100% Variance	100.0	98.8	66.4	24.0	4.4	96.4	61.8	21.0	3.4	0.0

of the errors arising from these three broad classes can provide insights into the overall robustness of the observed Cz P300 linkage in the COGA data set.

Some general patterns are evident. First of all, as expected, ability to detect linkage declined as the proportion of the population to be replaced increased. Secondly, sources of error resulting in replacement of observed phenotypic value with an uncorrelated value (for example, errors attributable to permutation or to invalid measurement) tended to reduce the ability to detect linkage to a greater extent than replacement of the phenotypic value with a correlated error measurement (such as that attributable to an imprecise measurement).

In general, ability to detect linkage to a locus influencing Cz P300 values remained excellent if phenotypes were altered for only 1% of the population and fairly good if phenotypes were altered for 5% of the sample. This was true even for relatively drastic changes in the phenotype (e.g., permutation changes or replacement with a value sampled from the uniform distribution). However, if phenotypes were replaced with uncorrelated values in more than 5% of the sample, ability to detect linkage declined rather quickly (Fig. 2). In contrast, ability to detect linkage was somewhat preserved if relatively small amounts of variability were introduced to the phenotypic measure in even a large proportion of subjects (Fig. 1).

It is also noteworthy that the minimum lod score in each replication set dropped significantly with replacement of even a very small proportion of phenotypes. One possible explanation for this is that in these replications, the values of a few critical individuals who were highly informative for linkage were changed. In this context, the work of Mitchell et al. [this issue] may be relevant. In the analysis described by these authors, individuals were systematically removed from the Cz P300 COGA data set to see what effect this would have on the resulting lod score at 198 cM on chromosome 6. A small subset of individuals were identified who contributed disproportionately to the linkage because the lod score was substantially reduced if they were removed from the analysis. We expect that the influential individuals identified by Mitchell et al. were probably more likely in these analyses to have had their phenotypes replaced in the replications resulting in low lod scores.

Because an increase in the proportion of people whose phenotypes were changed led to lower lod scores, we also looked at the variance as well as the coefficient of variation among the replicates. In general, while the variance decreased with a greater proportion change of phenotype, the coefficient of variation actually increased. Therefore, it is difficult to use these measures to examine whether a few critical individuals affect the lod scores significantly.

The analyses presented in this paper focus on the effects of measurement error on the ability to detect a true linkage. The p-value for the lod score of 3.54 is 2.75×10^{-5} under asymptotic assumptions, so there is a possibility, albeit very small, that the result is a false positive. In addition, all three classes of error could, in combination, impact phenotypic variation. This requires further exploring.

In summary, we wish to emphasize that accurate and precise phenotype measurement is essential to detect linkage and to possibly guard against false positive results (work in progress). Incorrect phenotype measurement of certain critical individuals due to various causes can impact the lod score substantially.

ACKNOWLEDGMENTS

This work was supported in part by National Institutes of Health grants R01-AR43351 and P01-HL45522. The work of Richard M. Watanabe was supported by the ADA Career Development Award.

REFERENCES

Amos CI (1994): Robust variance-components approach for assessing genetic linkage in pedigrees. Am J Hum Genet 54:535-543.

Almasy L, Blangero J (1998): Multipoint quantitative trait linkage analysis in general pedigrees. Am J Hum Genet 62:1198-1211.

Begleiter H, Porjesz B, Reich T, Edenberg HJ, Goate A, Blangero J, Almasy L, Foroud T, Van Eerdewegh P, Polich J, Rohrbaugh J, Kuperman S, Bauer LO, O'Connor SJ, Chorlian DB, Li TK, Conneally PM, Hesselbrock V, Rice JP, Schuckit MA, Cloninger R, Nurnberger J Jr, Crowe R, Bloom FE (1998): Quantitative trait loci analysis of human event-related brain potentials: P3 voltage. Electroencephalogr Clin Neurophysiol 108:244-250.