Quantifying Genomic Variation at the Individual Level: Putting the “Person” in Personalized Medicine

Nathan J. Markward, PhD, MPH
Pennington Biomedical Research Center, Baton Rouge, LA

Abstract

This project evaluates how the Rasch measurement framework can be used 1) to construct and interpret individualized genomic measures and fit statistics from nominal single nucleotide polymorphism (SNP) data and 2) to identify interaction patterns that are relevant to person-specific disease risk assessment, case management, and prevention.

Introduction

Recent advances in human genome epidemiologic research suggest that inferences based on single-locus analyses may be of limited utility due to their inability to incorporate information about the person-specific genomic context in which a hypothesized causal allele is enmeshed. Refined construct theories and mathematical methods are needed for 1) analyzing multiple genomic loci; 2) assessing their complex interactions; and 3) relating these dynamics to health and disease outcomes of interest.

This study utilizes a Rasch measurement model to evaluate different SNP/phenotype scoring schemes and, based on these results, to develop a composite “genometric” scale that simultaneously captures information about the persons, loci, and phenotypes included in the analysis. Rasch measures and model fit statistics are evaluated to assess their utility for detecting and interpreting different species of interaction (person-SNP, SNP-SNP, SNP-phenotype) that may be relevant to complex disease risk assessment, prevention, and case management.

Materials and Methods

Publicly available data were downloaded from the Cardiogenomics Program for Genomic Applications (PGA) website at Harvard. The data set included genotype and phenotype data for 12 families (N = 93 individuals, 2N = 186 haplotypes) and 43 SNP variants located in the angiotensin I-converting enzyme (ACE) gene region. 24 of the 93 (25%) study participants were diagnosed with cardiovascular disease (CVD).

Each individual’s data string was scored using methods described elsewhere. Once scores were assigned, a genomic measurement scale was constructed using a dichotomous Rasch model

\[ \ln\left(\frac{P_{ni}}{1 - P_{ni}}\right) = B_n - D_i \]

where

\[ B_n = \ln\left(\frac{P_n}{1 - P_n}\right) \]

is the measure estimated for person \( n \), and

\[ D_i = \ln\left(\frac{P_i}{1 - P_i}\right) \]

is the estimated calibration for each locus \( i \). Data analysis was conducted using WINSTEPS®, a popular Rasch measurement software package.

Results

Figure 1 is a Wright (variable) map depicting the relative locations of persons, loci, and CVD on the measurement scale. Table 2 and Table 3 summarize measures, model fit statistics, and measurement precision for persons and loci. Table 4 reports calibrations and fit statistics for each locus. Figure 2 and Figure 3 illustrate the person-specific diagnostic output generated for two unrelated study subjects.

Conclusion

A theoretical and mathematical framework does, in fact, exist to guide the assessment of genotype-phenotype relationships at the level of individuals. If coupled with high-throughput genotyping and computer adaptive assessment technologies, the approach might become a powerful analytical complement to conventional human genome epidemiologic, statistical, and bioinformatic tools that are used to estimate disease risk.

References