

Statistical Issues in Linkage Analysis

Nicholas Schork, Ph.D.
University of California, San Diego

Abstract:

Some time ago, Risch and Botstein (1997) wrote a fairly scathing critique of the experiences and practices of genome-wide linkage analysis studies of complex traits and diseases. In particular, Risch and Botstein emphasized what they felt to be highly discouraging results of applications of linkage analysis to neuropsychiatric diseases. The Risch and Botstein critique was, in many respects, completely justified at the time, as there were many published linkage studies in the neuropsychiatric field reporting highly conflicting results, non-replicated findings, or no evidence for linkage at all. As linkage analysis is being applied to many multifactorial diseases in an attempt to identify relevant genes in extremely large and expensive consortium-like initiatives such as the Consortium on the Genetics of Schizophrenia (COGS) and the Family Blood Pressure Program (FBPP), it is important to consider factors that might be contributing to an historical lack of overt and consistent success in the application of linkage analysis. In the following, we consider eleven factors that pertain to the way in which traditional non-parametric linkage analysis models are both formulated and implemented that may impact their successful application to complex traits and diseases. We consider in greater detail six of the eleven factors with a focus on an inherent bias in multipoint non-parametric linkage towards the null hypothesis of no linkage. The actual degree to which any of these factors ultimately influences the successful application of linkage analysis, as well as the degree to which they can be overcome, is open to question and should promote further research in these areas. However, we argue that many of the problems associated with these eleven factors can indeed be overcome at some level, thereby potentially increasing the power and utility of linkage analyses. In addition, we provide evidence that some simplified parametric linkage analysis models, when combined with appropriate modeling of recombination, marker informativity, and other phenomena, have the potential to be orders of magnitude more powerful than traditional non-parametric and/or variance component-based linkage analysis models.