

Correcting for Genetic and Cultural Nonindependence (Galton's Problem) when Testing Genetic Associations with Substance Abuse

Luke J. Matthews

RAND Corporation, Full Behavioral and Social Scientist

Genetic and cultural inheritance mean that individuals and populations violate a core assumption of most statistical methods because the data points do not reflect independent measurements of the thing being studied. Rather, individuals and populations reflect the same underlying inherited genetic or cultural information in proportion to their degree of shared genetic and cultural ancestry. This statistical issue, identified over 128 years ago by Francis Galton and hence named "Galton's Problem", is a substantial and potentially under-appreciated problem in research on genetic associations with substance abuse. The problem may be particularly acute for alcohol use because it exhibits widely varying cultural norms across human populations, and because there are a number of functional genes linked to alcohol use disorder that vary greatly in frequency across human populations. Correcting for nonindependence requires empirical knowledge of the processes that generate it and statistical methods that can appropriately transform the data. I present results from agent based simulation models to assess the statistical performance of currently proposed methods to correct for nonindependence. The methods are found to vary in how well they control false positive associations while maintaining statistical power. I also discuss the empirical data architecture that would be needed for researchers to routinely correct for Galton's Problem in substance abuse research without themselves having to be experts on cultural and genetic ancestry.