

Population Stratification and Use of Genetic Ancestry to Assess Risk for Cancer

Abstract:

Population stratification (PS) refers to differences in allele frequencies between cases and controls due to systematic differences in ancestry rather than association of genes with disease. Much variation in genetic ancestry can exist within or between racial/ethnic groups thereby causing significant PS to be present. The main consequence of PS is bias in the estimate of genetic associations, which can lead to incorrect inferences as well as inconsistency across reports. When race/ethnicity can be accurately described in terms of actual ancestry and there is ancestral homogeneity in a study population, standard epidemiological approaches of matching or statistical adjustment by race/ethnicity may be sufficient to remove or reduce bias due to PS. Controlling for self reported race has generally been thought to suffice, however, self-reported race/ethnicity and/or ancestry can be quite unreliable. Recent data show that matching on ancestry is more robust. However, in many populations, whether recently admixed or not, individuals cannot accurately report their precise ancestry. Therefore, multiple methods have been developed to assess and control for PS within association studies, many utilizing individual or group ancestry estimates. These methods will be described and examples shown as related to estimating cancer risk.