Simulations to assess the contribution of survival bias to protective associations of body mass index and Parkinson's disease

2016-09-21

## Simulation strategy

We construct a model whereby Parkinson's disease (PD) is not biologically related to body mass index (BMI), but BMI is associated with mortality, and PD is associated with age

BMI ~ snp(s)  
mortality ~ age + BMI  
PD ~ age

We simulated a large population () where each individual has alleles at 77 BMI associated SNPs (Locke et al. 2015) (using only LD independent SNPs out of the reported 97, r-square cutoff 0.001 within 10Mb windows). PD status, age values, BMI values, and alive/dead status are also simulated.

Age values are generated to match the reported age distributions in (M. A. Nalls et al. 2014). BMI SNPs are generated as a function of their allele frequencies, such that for individual at SNP their genotype value is where is the allele frequency of SNP . The BMI values are a function of the BMI SNPs, such that

where , where the genetic variance and residual variance . The phenotypic variance, , is the variance of BMI that was used to obtain the effect sizes.

PD status was simulated as a function of age, based on age related incidence obtained from (Driver et al. 2009). Alive/dead status was modelled as a function of age and BMI values. The baseline survival function was generated from the Gompertz-Makeham law of mortality, with age related hazard function

which has CDF

giving the baseline survival function:

The influence of BMI on survival is then incorporated into the full survival model as

where is the BMI value and is a function that uses external data to relate BMI with mortality. Following (Davey Smith et al. 2009), a causal effect on all-cause mortality hazard ratios of 1.16 per standard deviation increase in BMI was simulated.

Once these simulated data are generated, only 'alive' individuals are retained, and individuals are then selected for MR analysis. 13,708 PD individuals and 95,282 non-PD individuals are sampled based on the distribution of ages of cases and controls in the original PD GWAS (M. A. Nalls et al. 2014), such that the mean age of the cases are 60.6 years old, and of controls are 53.1 years old. Finally, two sample MR analysis is applied to the selected individuals from these simulated data.

This procedure is repeated 1000 times to obtain an empirical distribution of frailty effects for the given parameters.

## References

Davey Smith, George, Jonathan AC Sterne, Abigail Fraser, Per Tynelius, Debbie A Lawlor, and Finn Rasmussen. 2009. “The association between BMI and mortality using offspring BMI as an indicator of own BMI: large intergenerational mortality study.” *BMJ* 339.

Driver, Jane A., Giancarlo Logroscino, J. Michael Gaziano, and Tobias Kurth. 2009. “Incidence and remaining lifetime risk of Parkinson disease in advanced age.” *Neurology* 72 (5): 432–38. doi:[10.1212/01.wnl.0000341769.50075.bb](http://dx.doi.org/10.1212/01.wnl.0000341769.50075.bb).

Locke, Adam E., Bratati Kahali, Sonja I. Berndt, Anne E. Justice, Tune H. Pers, Felix R. Day, Corey Powell, et al. 2015. “Genetic studies of body mass index yield new insights for obesity biology.” *Nature* 518 (7538): 197–206. doi:[10.1038/nature14177](http://dx.doi.org/10.1038/nature14177).

Nalls, Mike A, Nathan Pankratz, Christina M Lill, Chuong B Do, Dena G Hernandez, Mohamad Saad, Anita L DeStefano, et al. 2014. “Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson’s disease.” *Nature Genetics* 46 (9). Nature Publishing Group: 989–93. doi:[10.1038/ng.3043](http://dx.doi.org/10.1038/ng.3043).