A Progress Report on the GWAS of Educational Attainment

Niels Rietveld

Erasmus School of Economics
Erasmus University Rotterdam

3rd SSGAC Workshop • 26 May 2012

Background

- First initiative SSGAC
  - Official started with first SSGAC workshop
- Educational attainment:
  - Moderately heritable
    - $h^2 \sim 0.40$ (e.g. Taubman 1976, AER, Miller et al. 2001, EER)
  - Relevant in medicine and social sciences
  - Widely measured
  - Measures can be harmonized (ISCED)
  - However, it's biologically distal

Progress

- Discovery stage
  - Recruiting cohorts / Distribution analysis plan Feb 2011
  - Deadline uploading results June 2011
  - QC discovery stage finished Jan 2012
- Replication stage
  - Recruiting cohorts / Distribution analysis plan Nov 2011
  - Deadline uploading results Dec 2011
  - QC discovery stage finished May 2012
- Overall:
  - 8 conference calls, 2 Working Group meetings
  - Preliminary results at 2nd SSGAC workshop

Analyses

- Two model specifications
  - OLS on educational attainment in US schooling years (ISCED)
  - Logit on college degree
- Sample sizes
  - ~100,000 in discovery stage
  - ~30,000 in replication stage

Meta-analysis

- Analysts:
  - Sarah Medland (Queensland Institute of Medical Research)
  - Jamie Derringer (Psychology, U Colorado)
  - Niels Rietveld (Economics, Erasmus U Rotterdam)
- Quality control:
  - *MAF* > 1%
  - Imputation quality R^2 > 40% (MACH and Impute)
  - Visual inspection Cohort-specific QQ-plots
  - Single Genomic Control
- Analysis plan:
  - Replication for SNPs with $p < 10^{-6}$ in discovery stage, GWAS hit when:
    - $p < 5x10^{-8}$ in discovery stage, nominal significance in replication stage
    - $p < 5x10^{-6}$ in discovery stage, $p < 5x10^{-8}$ in combined stage

Summary

- Genome-wide significant SNPs for:
  - College females
  - College & EduYears pooled
- Results got stronger after replication phase