

Principal Investigator

Dr Neil Davies

Address

University of Bristol, Barley House, Oakfield Grove, Bristol BS8 2BN

Summary of research

Key words: Education, health, mortality, morbidity, causal inference.

1a: Research question: what are the causal effects of education on morbidity and cause specific mortality?

Outcomes: all-cause and cause-specific mortality, coronary heart disease, lung cancer and income.

Exposure: education.

Synopsis: On average people with more years of education live longer and are healthier. We do not know whether this is because more educated people are less likely to engage in risky behaviours such as smoking, or if there are other differences, such as cognition or social class, which cause some people to be both healthier and more educated.

We will address this hypothesis using statistical techniques for differentiating causation and correlation.

1b: Low education is a major risk factor for a wide range of diseases.

We will use data from the UK Biobank to improve our understanding of how individuals' educational choices early in life affect their long term health outcomes.

Ultimately, knowledge from this project will help policy makers and practitioners mitigate these differences. Furthermore, our results will provide evidence about the consequences of the recent decisions to further increase the minimum school leaving age to 18 and to widen access to higher education.

1c: We will use data on all individuals from UK Biobank with information on date and country of birth and age at which they left full time education.

We will investigate differences in health outcomes by duration of education. We will present adjusted results adjusted for known confounders (such as income, cognition).

We will use the raising of the school leaving age in 1972 and genetic variants known to associate with education as a natural experiment (instrumental variable) for educational achievement to identify the causal effects of education.

1d: We will use participants born outside the UK as a negative control population to provide further validation of our methods (1). We do not expect the changing of the school leaving age to affect the educational attainment of individuals born outside of the UK.

Project extension: We will conduct a GWAS using the imputed data and will include the following covariates – sex, standardised age at the time of assessment, age squared, the 15 principal components of population stratification and indicators for genotyping array. We will restrict the analysis to unrelated white British individuals – the 112,338 individuals recommended by the imputation documentation. We do not need any further data for this extension.

Project extension:

The relationship between myopia and educational attainment.

Educated individuals are more likely to be short sighted. However, we do not know why this association occurs, it may be because educated individuals spend more time studying, and reading, or alternatively, short sighted individuals may be less likely to engage in non-academic activities like sports. We will apply a bidirectional Mendelian randomisation approach to investigate whether recently discovered genetic variants, which are known to associate with educational attainment, are associated with myopia, and whether genetic variants which are known to associate with myopia are also associated with educational attainment. This will allow us to determine whether education causes myopia or if myopia causes educational attainment.

The relationship between fatty acids and educational attainment.

Fatty acids are thought to be vital for normal brain development, in utero and in infancy. Fish consumption provides a major source of fatty acids, and has been found to be associated with IQ. However, there are substantial socioeconomic differences in fish consumption, and the observational association of fish consumption may be due to residual confounding. During this project we will use genetic variants, which have been found to be associated with fatty acid levels in metabolic studies, as instrumental variables for fatty acid levels in biobank participants. We will conduct a two sample study to maximise our statistical power.

Project extension August 2017:

“What are the causal effects of education on morbidity, mortality and other health related phenotypes?’ This amendment seeks to estimate the effects of educational attainment on a range of outcomes. We will use a range of methods to identify the effects of schooling, including (but not limited to) Mendelian randomisation and regression discontinuity design using the raising of the school leaving age in 1972 as a natural experiment.”

FURTHER UPDATED BY UK BIOBANK: 2 PROJECT EXTENSIONS – Approved
05.01.2017:

1) Visual acuity amendment

Visual acuity (VA) is a measure of an individual's clarity of vision. In the Avon Longitudinal Study of Parents and Children (ALSPAC) cohort we have detected a region of genome strongly associated with VA. This association has been replicated in Generation R. There is evidence for the same locus being associated with educational attainment in the Okbay et al. (2016) study.

Due to the high linkage-disequilibrium (LD) in the region, it has been difficult to attribute the cause of association to a single gene in the region. We would like to use UK Biobank to replicate the finding. We expect that the larger sample will give better resolution for understanding the LD structure in the region. If a clear functional mechanism of association can be established between variant and VA, we plan to use the variant as an instrument to test the causal effect of visual acuity on educational attainment.

No additional variables are required.

Reference

Okbay, Aysu, Jonathan P. Beauchamp, Mark Alan Fontana, James J. Lee, Tune H. Pers, Cornelius A. Rietveld, Patrick Turley, et al. "Genome-Wide Association Study Identifies 74 Loci Associated with Educational Attainment." *Nature* 533, no. 7604 (May 26, 2016): 539–42. doi:10.1038/nature17671.

2) Selection bias amendment

We will compare polygenic risk scores for key attributes (BMI, schizophrenia, ADHD, ASD, education, smoking, height) across cohorts where we expect different selection mechanisms to operate - from Biobank (where we expect a high degree of selection related to some characteristics, e.g. it has been reported that the level of smoking in Biobank participants is much lower than in the general population) to population-based birth cohorts (e.g. ALSPAC) where we expect a lesser degree of selection.

We will calculate the polygenic risk scores for each participant in each cohort, and then compare the distributions of the polygenic risk scores across cohorts. This will initially be by comparing the means, using linear regression.

However, it is possible that the selection mechanism has little effect on the means, but may affect just the tail(s) of the distribution - e.g. if only individuals with a very large BMI are less likely to participate, then cohorts will have similar average BMI but very different skewness and percentile values. We will thus also conduct between-cohort comparisons of:

- 1) the 5, 10 and 25 percentile values of each polygenic risk score
- 2) the skewness and kurtosis of each polygenic risk score

FURTHER UPDATED BY UK BIOBANK: PROJECT EXTENSION – Approved 04.04.2017:

"We would like to investigate the causal effect of education on fertility. Many studies have reported the association between educational attainment and fertility. However, we do not know whether educational attainment has a causal effect on fertility decisions. We would like to investigate this using the raising of the school leaving age in 1972. We will need additional data on variables 2405 and 2734.

Secondly, we would like to investigate the impact of education on primary prevention, such as taking statins or anti-hypertensives. There are strong associations between educational attainment and longevity. However, again it is not clear whether these associations are due a causal effect of education. We have published some initial results using UK Biobank that there may be causal effects of education on mortality. We would like to investigate whether these effects are being mediated via treatment decisions and interactions with health care such as statin and anti-hypertensive use. To do this we need the following variables: treatments 20003."