Does compulsory schooling interact with genetic predisposition to improve cognition? Evidence from the ELSA study

Yan Liu[1], Giorgio Di Gessa[2], Jennifer B. Dowd[1]

[1]Department of Global Health and Social Medicine, King's College London

[2]Epidemiology and Public Health, University College London

Abstract (150 Words)

Quasi-experimental methods in social and health sciences have been increasingly used to improve causal inference in the impact of social factors, such as education and job loss, on health outcomes. This approach is beginning to be used to better identify gene-environment interactions with "exogenous" changes in the environment that are uncorrelated with family upbringing. We exploit two changes in the minimum school leaving age from 14 to 15 from March 1947 and 15 to 16 from September 1972 in England, which have been shown to have strong impacts on educational attainment. Data, including Polygenic Risk Scores (PGS) for cognitive function and dementia, come from the English Longitudinal Study of Ageing (ELSA). A fuzzy regression discontinuity design (RDD) will be used to identify the causal effect of the policy-induced additional year of education change and how this interacts with genetic risk for cognitive decline.

Background

There is a high burden of cognitive impairment and dementia facing ageing societies around the world. There are approximately 50 million people living with dementia all over the world and nearly 10 million new dementia cases every year [1]. In the United Kingdom, there were 850,000 people with dementia in 2015 and the prevalence is projected to rise to over one million by 2025 and over two million by 2051 [2].The worldwide costs of dementia have increasing rapidly to US $818 billion in 2015 with an increase of 35% since 2010, [3, 4] 86% of which occur in high-income countries. Further research is needed to understand cognitive ageing and to identify what drives variability in disease development, which will offer crucial insights to better target prevention efforts.

Positive correlations between educational attainment and cognition in older age are well established in the literature [5, 6, 7], though how much of this association is causal is still subject to debate, due to unobservable factors associated with both education and cognition. To address this problem, recent studies have explored the causality between education and cognitive function using exogenous policy shocks. Glymour (2008) [8] et al. used state compulsory schooling laws (CSLs) in the US during 1907-1961 and used an instrumental variable (IV) approach to estimate the impact of education on memory. Banks and Mazzonna (2012) [9] exploited the compulsory schooling reform in England in 1947, which raised the minimum school-leaving age from 14 to 15 for individuals born after a specific cut-off in March 1933, and used a fuzzy regression discontinuity (FRD) design to compare individuals born before and after the cut-off. Despite adopting different designs, the two studies both take

advantage of CSLs as exogenous policy shocks to education to identify the causal effect of education attainment on cognitive function. These two studies both found that education attainment can protect against cognitive decline among older people. Most recently, Courtin et al. (2019) [10] provided new quasi-experimental evidence leveraging a French reform of schooling duration and offered a more complex picture: while men performed better in cognitive function in older age as a result of longer schooling, women did not improve their cognitive function and instead showed more depressive symptoms. Overall, these studies suggest that higher education and additional schooling may protect against cognitive decline, but there are significant differences across countries, cohorts and reforms in the consistency of this finding [11].

Gene– environment interactions for cognition

The availability of genetic data in large population surveys has resulted in an explosion of research on gene-environment interactions (G×E) [12, 13, 14, 15]. G×E occurs when "difference genotypes lead to variation in responses to social and/or physical environments (or vice versa)" [16]. Gene-environment interactions for cognition have been reported in several studies [17, 18, 19, 20, 21, 22] among which the ε4 allele of ApoE (APOE4) is one of the crucial genetic risk factors for cognitive decline [23]. Few other risk genes of cognitive decline have been explored systematically to examine the G×E interaction.

The G×E research in cognition has been dominated by two hypotheses [24]: On the one hand, the "social trigger" model poses that the risky genotype will have a more deleterious effect on cognition when triggered by the most adverse environments. On the other hand, the "social push" model poses that the phenotype which refers to the expression of the trait will be pushed by the most harmful environments, while the risky genotype will become effective on cognition when the environments turns to be favourable.

Some studies suggested that education can moderate the relationship between APOE4 and cognitive decline [25, 26] or dementia [27]. Using the HRS data, McCardle and Prescott (2010) [28] found steeper memory declines for APOE4 carriers compared to the non- APOE4 carriers only among those with 8 years of education or less, which is consistent with the social trigger model [29]. More research is needed to identify the gene-education interaction on cognitive decline in different populations and involving multiple genes [30].Considering other social factors such as socioeconomic status (SES), Boardman (2012) [62] combined the social characteristics of older adults' neighbourhoods with APOE genotype to predict how cognitive function varies over time. This study shows that the effect of APOE4 varies significantly across neighbourhoods and it displays stronger effects on cognitive function among those in the most adverse neighbourhoods, which supports the "social push" model.

Challenges of traditional GxE research and new area of quasi-experimental GxE designs

One key challenge in GxE research is the conceptual issue of modelling interactions between variables that are themselves correlated (gene-environment correlation [rGE]). One the one hand, rGE can emerge due to parents and children sharing both genes and environments [31, 32]. Alternatively, rGE can otherwise occur when children's characteristics which are inherited from parents cause responses from the social environment [33, 34]. The issue of rGE would be overcome by using research designs that leverage exogenous environmental variation. A few studies have recently contributed to incorporating quasi-experimental

designs into a G×E framework to reduce the influence of rGE and strengthen the ability to demonstrate causal inferences.

Polygenic risk scores, or polygenetic scores (PGS), have been recently employed in the gene-environment interaction research. PGS are generally calculated from a weighted quantity of allelic count and are presented as continuous scores [35, 36, 37]. They are specific to each individual to measure the propensity to a phenotype [38]. Domingue [39] conducted a polygenic score (PGS) study to test whether genetic predispositions to better subjective well-being can buffer against the risk of the development of depression following a stressful life event: spousal death. Utilizing an RD design, Domingue found that having a higher PGS for subjective well-being, which may reflect the opposite end of depression, buffered against increased depressive symptoms following spousal death.

To estimate the causal effect of genes and education on health, two studies in the UK exploited the raising of the minimum school leaving age in 1972 from 15 to 16, which induced sharp across-cohort differences in educational attainment. Using data from the UK Biobank, Barcellos [40] combining an RD design and PGS found that the additional year of schooling affected body mass index (BMI) in middle age and can reduce the differences in unhealthy body size related to genetic risk of obesity, which, specifically, can reduce from 20 to 6 percentage points of obesity between the top and bottom PGS terciles. Also using a RD design, Davies [41] employed the educational attainment genome-wide score to examine whether participants influenced by the school reform have more genetic variants associated with educational attainment. The findings are consistent with prior studies, indicating that remaining in school after 15 years of age has a causal effect on decreased risks of diabetes and mortality. Those who remained in school tend to have more single nucleotide polymorphisms (SNPs) associated with higher educational attainment.

Fletcher [42] examined the effects of early-life exposure to pneumonia – one leading cause of infant death in the early 20th century – on cognitive outcomes among older people in one working paper. Leveraging the introduction of sulfonamide antibiotics in 1937 – which led to dramatic reductions in pneumonia morbidity and mortality – along with state-level differences in baseline disease rate – Fletcher used an instrumental variable (IV) model and found that infant exposure led to faster cognitive decline in adulthood. These effects were largest for individuals with higher genetic endowments (as measured by polygenetic scores (PGS) for cognition), and null for those with lower endowments.

Using the Vietnam-era draft lottery when many young men were called to military service, Schmitz and Conley incorporated polygenic scores into an instrumental variable design to investigate the effects of compulsory military service and genotype on schooling performances [43] and smoking behaviours [44]. Their findings suggested that conscription may interact with low PGSs for education to reduce veterans' educational attainment, resulting in fewer years of schooling and less likelihood of obtaining a postsecondary degree. They also found that the interaction between compulsory military service and a high genetic predisposition for smoking increase the risk of smoking, smoking heavily and being diagnosed with hypertension or cancer in later life.

Quasi-natural experimental designs are able to more effectively isolate exogenous variation in observational data, the results from which can be used as a stepping stone for future GxE research. To examine how social and environmental factors impact cognition and how this

might vary by genetic susceptibility, the most rigorous way is combining the GxE and quasi-experimental designs. However, to our knowledge there is no published GxE research using quasi-experimental exposures for cognition, except the working paper by Fletcher reviewed above. [82] The current study will identify the impact of education on cognitive function in the ELSA sample and how this interacts with genetic risk for cognitive decline and dementia.

Methods

Data

Data will come from the English Longitudinal Study of Ageing (ELSA) [45], which include an original sample of more than 18 000 people aged 50 and older living in England. Respondents have been interviewed at two-yearly interviews and the sample has been refreshed periodically. This project will use data up to wave 8 (2002-2003 to 2016–2017). Meanwhile, this project will use the ELSA Special License data which includes the "month-of-birth" variable and the Genetic data which includes polygenic scores for cognition.

Polygenic scores for general cognition

The PGSs for general cognition in this project were created by ELSA using results from a 2015 Genome Wide Association Study (GWAS) across 31 cohorts by the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium [46]. A total of 2,473,946 SNPs was included in the CHARGE meta-analysis summary statistics. Of these, the sample size for general cognition is 7223, and 795,327 SNPs overlapped with the ELSA genetic database and were included in the PGS for the general cognition phenotype.

Cognitive function in ELSA

Cognitive function was assessed at each wave covering three major cognitive domains: memory, executive function and basic skills. Following standard practice, we will compute an overall score by adding the points of selected questions in each domain, including 20 in word list learning and recall, 64 in letter cancellation and 60 in word finding. Then we will analyze each domain and the overall score as the cognitive outcomes in the three studies below.

Analytical strategy

Building off of work by Banks [25] and Clark [47], we will exploit two changes in the minimum school leaving age from 14 to 15 from March 1947 and 15 to 16 from September 1972 in England. These school reforms have been shown to have strong impacts on educational attainment among a large fraction of the UK population after the two changes. A fuzzy regression discontinuity design (RDD) [48, 49] will be used to identify the causal effect of the policy-induced additional year of education change and how this interacts with genetic risk for cognitive decline. All regressions will use pooled waves (wave1-wave8) of the ELSA.

Specifically, this study will estimate equation (1) and (2) for each school reform including a linear function of month of birth, the relevant reform dummy, an interaction term of month of birth and the relevant reform dummy, a quadratic term of age (in year) and dummy variables for gender and survey wave:

$$E_{icw} = \alpha_0 + \alpha_1 D_{ic} + \alpha_2 D_{ic} \times PGS_{ic} + f(R_{ic}) + X_{icw}\alpha_3 + \mu_{icw} \quad (1)$$

$$E_{icw} \times PGS_{ic} = \gamma_0 + \gamma_1 D_{ic} + \gamma_2 D_{ic} \times PGS_{ic} + f(R_{ic}) + X_{icw}\gamma_3 + \theta_{icw} \quad (2)$$

where E_{icw} indicates years of schooling for individual i in birth cohort c at wave w; D_{ic} is a dummy variable indicating whether a respondent belongs to a post-reform cohort; PGS_{ic} is polygenic scores for cognitive function; RR_{ic} is an respondent's birth month cohort relative to the relevant cut-offs (March 1933 for the 1947 reform, September 1957 for the 1972 reform). R_{ic} is positive when the respondent is born after the reform and negative when born before the reform; X includes predetermined characteristics. E_{icw} and $E_{icw} \times PGS_{ic}$ terms are endogenous and D_{ic} and $D_{ic} \times PGS_{ic}$ are used as instruments. Then this study adds an outcome equation (3) to identify the effect of education on cognitive function:

$$CF_{icw} = \beta_0 + \beta_1 E_{ic} + \beta_2 PGS_{ic} + \beta_3 E_{ic} \times PGS_{ic} + g(R_{ic}) + X_{icw}\beta_4 + \omega_{icw} \quad (3)$$

where CF_{icw} is a measure of cognitive function.

Conclusion

This paper will examine how social and environmental exposures impact cognition and interact with genetic risk for cognitive decline in older age, focusing on an important social exposure early in the life course–education. This project will advance the GxE research by incorporating quasi-experimental designs and offer insights as to how policy effects might be modified by genetic predisposition.

References:

[1] Dementia. (2019). Retrieved from https://www.who.int/news-room/fact-sheets/detail/dementia

[2] Prince, M., Knapp, M., Guerchet, M., McCrone, P., Prina, M., Comas-Herrera, A., ... & Rehill, A. (2014). Dementia UK: UpdateSecond Edition.

[3] Wimo, A., Jönsson, L., Bond, J., Prince, M., Winblad, B., & International, A. D. (2013). The worldwide economic impact of dementia 2010. Alzheimer's & Dementia, 9(1), 1-11.

[4] Wimo, A., Guerchet, M., Ali, G. C., Wu, Y. T., Prina, A. M., Winblad, B., ... & Prince, M. (2017). The worldwide costs of dementia 2015 and comparisons with 2010. Alzheimer's & Dementia, 13(1), 1-7.

[5] Zahodne, L. B., Glymour, M. M., Sparks, C., Bontempo, D., Dixon, R. A., MacDonald, S. W., & Manly, J. J. (2011). Education does not slow cognitive decline with aging: 12-year evidence from the Victoria Longitudinal Study. Journal of the International Neuropsychological Society, 17(6), 1039-1046.

[6] Arvanitakis, Z., Wilson, R. S., Bienias, J. L., Evans, D. A., & Bennett, D. A. (2004). Diabetes mellitus and risk of Alzheimer disease and decline in cognitive function. Archives of neurology, 61(5), 661-666.

[7] McArdle, J. J., Ferrer-Caja, E., Hamagami, F., & Woodcock, R. W. (2002). Comparative longitudinal structural analyses of the growth and decline of multiple intellectual abilities over the life span. Developmental psychology, 38(1), 115.

[8] Glymour, M. M., Kawachi, I., Jencks, C. S., & Berkman, L. F. (2008). Does childhood schooling affect old age memory or mental status? Using state schooling laws as natural experiments. Journal of Epidemiology & Community Health, 62(6), 532-537.

[9] Banks, J., & Mazzonna, F. (2012). The effect of education on old age cognitive abilities: evidence from a regression discontinuity design. The Economic Journal, 122(560), 418-448.

[10] Courtin, E., Nafilyan, V., Glymour, M., Goldberg, M., Berr, C., Berkman, L. F., ... & Avendano, M. (2019). Long-term effects of compulsory schooling on physical, mental and cognitive ageing: a natural experiment. J Epidemiol Community Health, 73(4), 370-376.

[11] Galama, T. J., Lleras-Muney, A., & van Kippersluis, H. (2018). The Effect of Education on Health and Mortality: A Review of Experimental and Quasi-Experimental Evidence (No. w24225). National Bureau of Economic Research.

[12] Bearman, P. (2008). Exploring genetics and social structure. American Journal of Sociology, 114(S1), v-x.

[13] Freese, J., & Shostak, S. (2009). Genetics and social inquiry. Annual Review of Sociology, 35, 107-128.

[14] Boardman, J. D., Blalock, C. L., & Pampel, F. C. (2010). Trends in the genetic influences on smoking. Journal of health and social behavior, 51(1), 108-123.

[15] Schnittker, J. (2008). Happiness and success: Genes, families, and the psychological effects of socioeconomic position and social support. American Journal of Sociology, 114(S1), S233-S259.
[16] Seabrook, J. A., & Avison, W. R. (2010). Genotype–environment interaction and sociology: Contributions and complexities. Social Science & Medicine, 70(9), 1277-1284.
[17] Harden, K. P., Turkheimer, E., & Loehlin, J. C. (2007). Genotype by environment interaction in adolescents' cognitive aptitude. Behavior genetics, 37(2), 273-283.
[18] Rhemtulla, M., & Tucker-Drob, E. M. (2012). Gene-by-socioeconomic status interaction on school readiness. Behavior genetics, 42(4), 549-558.
[19] Rowe, D. C., Jacobson, K. C., & Van den Oord, E. J. (1999). Genetic and environmental influences on vocabulary IQ: Parental education level as moderator. Child development, 70(5), 1151-1162.
[20] Scarr-Salapatek, S. (1971). Race, social class, and IQ. Science, 174(4016), 1285-1295.
[21] Tucker-Drob, E. M., Rhemtulla, M., Harden, K. P., Turkheimer, E., & Fask, D. (2011). Emergence of a gene× socioeconomic status interaction on infant mental ability between 10 months and 2 years. Psychological Science, 22(1), 125-133.
[22] Tucker-Drob, E. M., Rhemtulla, M., Harden, K. P., Turkheimer, E., & Fask, D. (2011). Emergence of a gene× socioeconomic status interaction on infant mental ability between 10 months and 2 years. Psychological Science, 22(1), 125-133.
[23] Davies, G., Harris, S. E., Reynolds, C. A., Payton, A., Knight, H. M., Liewald, D. C., ... & Henderson, R. (2014). A genome-wide association study implicates the APOE locus in nonpathological cognitive ageing. Molecular Psychiatry, 19(1), 76.
[24] Boardman, J. D., Barnes, L. L., Wilson, R. S., Evans, D. A., & de Leon, C. F. M. (2012). Social disorder, APOE-E4 genotype, and change in cognitive function among older adults living in Chicago. Social Science & Medicine, 74(10), 1584-1590.
[25] McArdle, J. J., & Prescott, C. A. (2010). Contemporary modeling of gene× environment effects in randomized multivariate longitudinal studies. Perspectives on Psychological Science, 5(5), 606-621.
[26] Seeman, T. E., Huang, M. H., Bretsky, P., Crimmins, E., Launer, L., & Guralnik, J. M. (2005). Education and APOE-e4 in longitudinal cognitive decline: MacArthur Studies of Successful Aging. The Journals of Gerontology Series B: Psychological sciences and social sciences, 60(2), P74-P83.
[27] Arenaza-Urquijo, E. M., Gonneaud, J., Fouquet, M., Perrotin, A., Mézenge, F., Landeau, B., ... & Chételat, G. (2015). Interaction between years of education and APOE ε4 status on frontal and temporal metabolism. Neurology, 85(16), 1392-1399.
[28] McArdle, J. J., & Prescott, C. A. (2010). Contemporary modeling of gene× environment effects in randomized multivariate longitudinal studies. Perspectives on Psychological Science, 5(5), 606-621.
[29] Shanahan, M. J., & Hofer, S. M. (2005). Social context in gene–environment interactions: Retrospect and prospect. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 60(Special_Issue_1), 65-76.
[30] Smith, J. A., Kho, M., Zhao, W., Yu, M., Mitchell, C., & Faul, J. D. (2018). Genetic effects and gene-by-education interactions on episodic memory performance and decline in an aging population. Social Science & Medicine.
[31] Kendler, K. S. (1996). Parenting: A genetic-epidemiologic perspective. The American journal of psychiatry, 153(1), 11.
[32] Horwitz, B. N., & Neiderhiser, J. M. (2011). Gene–environment interplay, family relationships, and child adjustment. Journal of Marriage and Family, 73(4), 804-816.
[33] McGue, M., Elkins, I., Walden, B., & Iacono, W. G. (2005). Perceptions of the parent-adolescent relationship: a longitudinal investigation. Developmental psychology, 41(6), 971.
[34] Horwitz, B. N., & Neiderhiser, J. M. (2011). Gene–environment interplay, family relationships, and child adjustment. Journal of Marriage and Family, 73(4), 804-816.
[35] Dudbridge, F. (2013). Power and predictive accuracy of polygenic risk scores. PLoS genetics, 9(3), e1003348.
[36] Purcell, S. M., Wray, N. R., Stone, J. L., Visscher, P. M., O'donovan, M. C., Sullivan, P. F., & Sklar, P. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature, 460(7256), 748-752.
[37] Choi, S. W., Mak, T. S. H., & O'reilly, P. (2018). A guide to performing Polygenic Risk Score analyses. BioRxiv, 416545.
[38] So, H. C., & Sham, P. C. (2017). Improving polygenic risk prediction from summary statistics by an empirical Bayes approach. Scientific reports, 7, 41262.

[39] Domingue, B. W., Liu, H., Okbay, A., & Belsky, D. W. (2017). Genetic heterogeneity in depressive symptoms following the death of a spouse: Polygenic score analysis of the US Health and Retirement Study. American Journal of Psychiatry, 174(10), 963-970.
[40] Barcellos, S. H., Carvalho, L. S., & Turley, P. (2018). Education can reduce health differences related to genetic risk of obesity. Proceedings of the National Academy of Sciences, 115(42), E9765-E9772.
[41] Davies, N. M., Dickson, M., Smith, G. D., Van Den Berg, G. J., & Windmeijer, F. (2018). The causal effects of education on health outcomes in the UK Biobank. Nature human behaviour, 2(2), 117.
[42] Jason M. Fletcher, Atheendar S. VenkataramaniEarly Life Exposures, Gene-Environment Interactions, and Cognition in Old Age. (working paper)
[43] Schmitz, L. L., & Conley, D. (2017). The effect of Vietnam-era conscription and genetic potential for educational attainment on schooling outcomes. Economics of education review, 61, 85-97.
[44] Schmitz, L., & Conley, D. (2016). The long-term consequences of Vietnam-era conscription and genotype on smoking behavior and health. Behavior genetics, 46(1), 43-58.
[45] Steptoe, A., Breeze, E., Banks, J., & Nazroo, J. (2012). Cohort profile: the English longitudinal study of ageing. International journal of epidemiology, 42(6), 1640-1648.
[46] Davies, G., Armstrong, N., Bis, J. C., Bressler, J., Chouraki, V., Giddaluru, S., ... & van der Lee, S. J. (2015). Genetic contributions to variation in general cognitive function: a meta-analysis of genome-wide association studies in the CHARGE consortium (N= 53 949). Molecular psychiatry, 20(2), 183.
[47] Clark, D., & Royer, H. (2013). The effect of education on adult mortality and health: Evidence from Britain. American Economic Review, 103(6), 2087-2120.
[48] Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. Journal of economic literature, 48(2), 281-355.
[49] Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. Journal of econometrics, 142(2), 615-635.