The Impact of Longer Schooling on Brain Ageing: Evidence from a Quasi-Experiment in the UK Biobank

Abstract (Long)

Pavithra Manoj, London School of Economics, London, UK

Mengyao Wang, Yale University, New Haven, US

Stella Tsoli, London School of Economics, London, UK

James Cole, University College London, London, UK

Emilie Courtin, London School of Economics, London, UK

Motivation-

A vast literature has investigated the effect of educational attainment on cognitive ageing and brain health (Lövdén et al., 2020; Stern et al., 2020). Education has been proposed to contribute to cognitive reserve, akin to a mental resilience which would allow individuals to better withstand age-related neural changes, potentially delaying cognitive impairment or dementia (Stern, 2009). Recent studies have challenged this hypothesis.

One of the earliest empirical studies to support the cognitive reserve hypothesis explored the relationship between education and brain size using MRI in older adults (Coffey et al., 1999). They found that higher education was associated with greater total brain volume, suggesting a structural advantage linked to educational attainment. However, they also observed that age-related brain changes occurred regardless of education level, implying that if education may be linked to larger brain reserve, it does not prevent brain aging itself. A longitudinal study examining the link between education and brain ageing using neural biomarkers showed no evidence of educational attainment influencing the rate of brain ageing over time (Nyberg et al., 2021). However, cross-sectional analyses revealed a modest positive association between education and regional cortical volume. This suggests that although education may be associated with greater initial brain reserve, it does not necessarily slow down brain ageing processes. These studies investigate the association between educational attainment and brain ageing, making them potentially prone to selection bias and reverse causality.

To overcome these limitations, we are proposing to use the 1972 Raising the Schooling leaving Age (ROSLA) reform as a natural experiment to assess the causal relationships between duration of schooling and brain age.

The 1972 ROSLA Order increased the minimum leaving age from 15 to 16. This reform applied to students born on or after 1st September 1957, thereby extending compulsory education by one additional year for those born after that date. By exploiting the policy-induced variation in schooling duration, we aim to isolate the effects of education on brain ageing from confounding factors. We have three aims in this paper. First, we will calculate and validate innovative Brain Ageing clocks in our sample of the UK Biobank. Second, we will describe the association between educational attainment and Brain Ageing clocks in our sample. Finally, we will overcome the methodological limitations of previous studies by assessing the causal impact of a change in the duration of schooling induced by ROSLA on the rate of brain ageing as measured by these clocks.

Data:

The UK Biobank (UKB) is a large population-based prospective cohort study comprising of almost 500,000 participants aged 40 to 70 at baseline, recruited between 2006-2010 (Bycroft et al., 2018). Approximately, 9.2 million individuals living within 40km of one of 22 assessment centres across England, Scotland and Wales were invited to participate, of whom 5.5% enrolled. At recruitment, participants completed detailed questionnaires, underwent physical measurements, and provided biological samples. The study was further expanded to include enhanced follow-up data, including genotyping, biochemical assays, web-based questionnaires, physical activity tracking, and multimodal imaging (including brain MRI data critical for brain ageing research).

For our analysis, we will focus on a subsample of participants born in England, Scotland, or Wales within ten years of the September 1957 school-leaving age reform threshold. This restriction is aligned with previous studies using the same natural experiment design. This subsample includes 69,553 individuals. All participants attended an initial assessment centre to undergo MRI scanning in 2014. A proportion were invited several years later in 2019 to repeat the assessment. The second criteria of selection are the viability of the MRI brain scans. Some of the participants had empty scans, scans not in the NIFTI format and corrupted files. These participants were excluded from our study. Our final subsample comes down to 53,503 brain scans (including both assessments) for 49,448 individuals.

Exposure: educational attainment and ROSLA

In the UK Biobank (UKB) dataset, educational attainment is measured using two variables: the ages at which participants completed full-time education, and their highest educational qualification.

These measures provide insight into both the duration and level of formal education attained by participants. For our analysis, we will primarily use the first variable as a proxy for years of schooling. This variable enables us to detect whether individuals were affected by the policy change, as those compelled to remain in school for an extra year due to the reform are expected to report a higher leaving age compared to similar individuals born just prior to the cutoff.

The following covariates will be incorporated in our models: gender and age (age at recruitment as well as age at the time of the MRI).

Methods:

Aim 1: Calculation of Brain Age

Our first aim is to calculate, validate and select a measure of brain age in our sample. Brain Age Gap Estimation (Brain AGE) or Brain PAD (Predicted age difference) calculate the difference between the predicted brain age and chronological age. Brain age can be used a biomarker for brain health (Kalc et al., 2024). A positive value may indicate an older looking brain compared to their chronological brain and a negative value may indicate a brain age younger than chronological age (Cole, 2020). Brain Age will be used as our primary outcome in all other analyses.

We will use two pre-trained neural networks for building the clocks for a small subsample of our cohort, namely- DeepBrainNet (Bashyam et al., 2020) and MIDI (Wood et al., 2024). These models take minimally pre-processed data (skull-stripped and linearly registered through the automated pipeline from the FMRIB (Jenkinson et al., 2012)) pass them through a convolutional neural network with inception-resnet-v2 and DenseNet201 architecture respectively and give us brain age estimates (Jenkinson et al., 2012). To evaluate the accuracy and efficiency of the models, we employ a range of optimisation metrics, including AUC, ROC, R^2, Pearson's correlation coefficient, mean squared error (MSE), Mean Absolute Error (MAE) and system runtime. The model that best aligns with our performance criteria will be selected. This selected model will subsequently be applied to the entire population. employ a range of optimisation metrics, including AUC, ROC, R^2, Pearson's correlation coefficient, mean squared error (MSE), Mean Absolute Error (MAE) and system runtime. The model that best aligns with our performance criteria will be selected. This selected model will subsequently be applied to the entire population.

Aim 2. Association between education and brain ageing clocks

In our second aim, we ask whether educational attainment is associated with measures of brain aging. We hypothesize that individuals with more years of schooling will exhibit healthier brain aging profiles, as measured by a lower brain age gap. Gender may modify the association between education and brain aging. To explore this, we will first estimate models for the full sample and then stratify by gender. In each case we regress the brain age gap on years of education, adjusting for additional covariates including age, age², sex, height, volumetric MRI scaling (T1-weighted), and head motion during imaging. We also include interaction terms between gender and years of schooling to formally test whether the educational gradient in brain aging differs between men and women.

Aim 3. Impact of longer schooling on Brain Ageing

To estimate the causal effect of education on brain ageing, we employ a Regression Discontinuity Design (RDD) leveraging the 1972 Raising of School Leaving Age (ROSLA) reform in Great Britain. The ROSLA policy increased the minimum school leaving age from 15 to 16 for individuals born on or after September 1, 1957. In this framework, individuals born just after the cut-off serve as treatment group and those born just before are control group (Barcellos et al., 2018). Under the assumptions described below, this design enables us to establish a causal relationship between schooling duration and brain ageing.

We estimate two key parameters. First, we estimate the impact of eligibility to the reform on brain ageing, regardless of whether the individual actually stayed in school longer, akin to an intent-to-treat estimate. Potentially not all individuals complied and would have stayed at school until age 16 (or more) in the absence of the reform. Whilst some individuals stayed at school until age 16 because of the reform (the 'compliers'). Second, we also document the effect of taking up additional schooling on brain ageing using a two-stage least-squares. The coefficient of interest is interpreted as a Local Average Treatment Effect (LATE) and should be interpreted as the average effect of the additional year of compulsory schooling on brain ageing for those who would have left school at 15 in the absence of the ROSLA reform. We employ a 2SLS as follows-

Stage 1: $Educ_i = \alpha_o + \alpha_1 T_i + f(R_i) + \epsilon_i$

Stage 2: $BrainAgeGap_i = \beta_0 + \beta_1 \widehat{Educ_i} + f(R_i) + \delta X_i + \epsilon_i$

$BrainAgeGap_i$ is the outcome of interest for individual i; T_i is the indicator variable taking value 1 for individual born after the cut-off for eligibility (treated group) and 0 if born before that threshold (control group); $R_i$is an individual's birth cohort relative to the cut-off and measured in months (it can be linear, quadratic or cubic); $X_i$is a vector individual characteristic: age, age squared, , gender and month of birth. We will use robust standard errors in all specifications.

We will select the mean squared error optimal bandwidths to empirically establish our bandwidth size on each side of the cut-off (Calonico et al., 2014). We will conduct robustness checks by estimating models with alternative bandwidths of 3, 5, and 10 years. For graphical analysis, we will use small bins at the month of birth level to allow for fine-grained visualisation of the discontinuity cut-off.

In addition, focusing on average effects may conceal heterogeneous effects on specific parts of the health distribution (Barcellos et al., 2023). In additional models, we will estimate quantile treatment effects (QTE) across the distributions of our outcomes (Bitler et al., 2006; Frandsen et al., 2012).

We will also conduct a series of robustness check as part of supplementary analysis. First, we will check there is no manipulation or discontinuity at the threshold in the density of the running variable (date of birth) using the McCrary graph. We will also check the covariate balance at the cut-off to ensure that the treated and control groups are comparable at the cut-off. Higher order polynomials and triangular kernels, and additional adjustments will be included. We will also run placebo tests by shifting the cut-off to earlier or later than the actual threshold (e.g. on September 1st 1954-56 and then September 1st 58-59) to check for discontinuity. We will also check if our findings are sensitive to different bandwidth sizes (ranging from 0 to 10 years).

Expected findings:

We expect that eligibility for the 1972 ROSLA reform, which increased compulsory schooling by one year, will be associated with healthier brain ageing. Specifically, individuals affected by the reform are anticipated to show a lower brain age gap, as illustrated by Figure 1, indicating brains that appear younger relative to chronological age. This would provide causal evidence that extended education contributes to improved brain health later in life. In preliminary analyses, we tested the DeepBrainNet neural network model for estimating brain age and obtained a Mean Absolute Error (MAE) of 5.64 years, indicating strong predictive accuracy and reliability for use in our cohort. We also anticipate heterogeneous effects across the distribution of brain ageing measures, with stronger benefits among individuals who would otherwise have left school earlier. Together, these results would suggest that social policies extending compulsory education can have lasting effects on brain structure and ageing, contributing to the reduction of health inequalities across the life course.

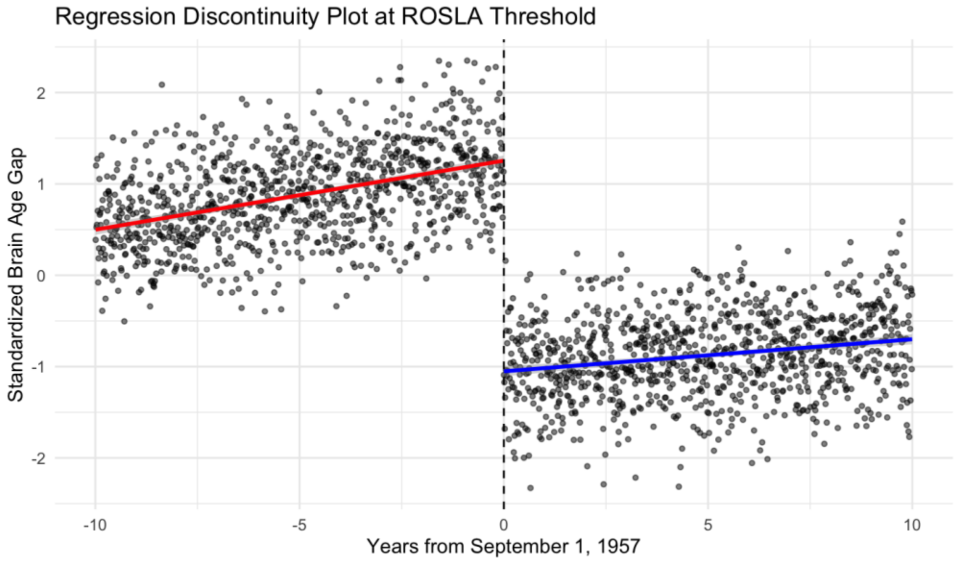

Figure 1: The plot illustrates with simulated data the populations we will consider in our analyses.

Figure 1

Bibliography-

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). https://doi.org/10.1073/pnas.1802909115

Barcellos, S. H., Carvalho, L. S., & Turley, P. (2023). Distributional Effects of Education on Health. *Journal of Human Resources*, *58*(4), 1273–1306. https://doi.org/10.3368/jhr.59.2.0720-11064R1

Bashyam, V. M., Erus, G., Doshi, J., Habes, M., Nasrallah, I. M., Truelove-Hill, M., Srinivasan, D., Mamourian, L., Pomponio, R., Fan, Y., Launer, L. J., Masters, C. L., Maruff, P., Zhuo, C., Völzke, H., Johnson, S. C., Fripp, J., Koutsouleris, N., Satterthwaite, T. D., … Davatzikos, C. (2020). MRI signatures of brain age and disease over the lifespan based on a deep brain network and 14 468 individuals worldwide. *Brain*, *143*(7), 2312–2324. https://doi.org/10.1093/brain/awaa160

Bitler, M. P., Gelbach, J. B., & Hoynes, H. W. (2006). What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiments. *American Economic Review*, *96*(4), 988–1012. https://doi.org/10.1257/aer.96.4.988

Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J., Cortes, A., Welsh, S., Young, A., Effingham, M., McVean, G., Leslie, S., Allen, N., Donnelly, P., & Marchini, J. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature*, *562*(7726), 203–209. https://doi.org/10.1038/s41586-018-0579-z

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica*, *82*(6), 2295–2326. https://doi.org/10.3982/ECTA11757

Coffey, C. E., Saxton, J. A., Ratcliff, G., Bryan, R. N., & Lucke, J. F. (1999). Relation of education to brain size in normal aging. *Neurology*, *53*(1), 189–189. https://doi.org/10.1212/WNL.53.1.189

Cole, J. H. (2020). Multimodality neuroimaging brain-age in UK biobank: relationship to biomedical, lifestyle, and cognitive factors. *Neurobiology of Aging*, *92*, 34–42. https://doi.org/10.1016/j.neurobiolaging.2020.03.014

Frandsen, B. R., Frölich, M., & Melly, B. (2012). Quantile treatment effects in the regression discontinuity design. *Journal of Econometrics*, *168*(2), 382–395. https://doi.org/10.1016/j.jeconom.2012.02.004

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, *62*(2), 782–790. https://doi.org/10.1016/j.neuroimage.2011.09.015

Kalc, P., Dahnke, R., Hoffstaedter, F., & Gaser, C. (2024). BrainAGE : Revisited and reframed machine learning workflow. *Human Brain Mapping*, *45*(3). https://doi.org/10.1002/hbm.26632

Lövdén, M., Fratiglioni, L., Glymour, M. M., Lindenberger, U., & Tucker-Drob, E. M. (2020). Education and Cognitive Functioning Across the Life Span. *Psychological Science in the Public Interest*, *21*(1), 6–41. https://doi.org/10.1177/1529100620920576

Nyberg, L., Magnussen, F., Lundquist, A., Baaré, W., Bartrés-Faz, D., Bertram, L., Boraxbekk, C. J., Brandmaier, A. M., Drevon, C. A., Ebmeier, K., Ghisletta, P., Henson, R. N., Junqué, C., Kievit, R., Kleemeyer, M., Knights, E., Kühn, S., Lindenberger, U., Penninx, B. W. J. H., … Fjell, A. M. (2021). Educational attainment does not influence brain aging. *Proceedings of the National Academy of Sciences*, *118*(18). https://doi.org/10.1073/pnas.2101644118

Stern, Y. (2009). Cognitive reserve☆. *Neuropsychologia*, *47*(10), 2015–2028. https://doi.org/10.1016/j.neuropsychologia.2009.03.004

Stern, Y., Arenaza-Urquijo, E. M., Bartrés-Faz, D., Belleville, S., Cantilon, M., Chetelat, G., Ewers, M., Franzmeier, N., Kempermann, G., Kremen, W. S., Okonkwo, O., Scarmeas, N., Soldan, A., Udeh-Momoh, C., Valenzuela, M., Vemuri, P., & Vuoksimaa, E. (2020). Whitepaper: Defining and investigating cognitive reserve, brain reserve, and brain maintenance. *Alzheimer's & Dementia*, *16*(9), 1305–1311. https://doi.org/10.1016/j.jalz.2018.07.219

Wood, D. A., Townend, M., Guilhem, E., Kafiabadi, S., Hammam, A., Wei, Y., Al Busaidi, A., Mazumder, A., Sasieni, P., Barker, G. J., Ourselin, S., Cole, J. H., & Booth, T. C. (2024). Optimising brain age estimation through transfer learning: A suite of pre-trained foundation models for improved performance and generalisability in a clinical setting. *Human Brain Mapping*, *45*(4). https://doi.org/10.1002/hbm.26625