The causal effects of education on health outcomes in the UK Biobank

Neil M. Davies1,2*, Matt Dickson3, George Davey Smith1,2, Gerard J. van den Berg1,4 and Frank Windmeijer1,4

Educated people are generally healthier, have fewer comorbidities and live longer than people with less education1–3. Much of the evidence about the effects of education comes from observational studies, which can be affected by residual confounding. Natural experiments, such as laws that increase the minimum school leaving age, are a potentially more robust source of evidence about the causal effects of education. Previous studies have exploited this natural experiment using population-level administrative data to investigate mortality, and surveys to investigate the effect on morbidity4–6. Here, we add to the evidence using data from a large sample from the UK Biobank. We exploit the raising of the minimum school leaving age in the UK in September 1972 as a natural experiment. We used a regression discontinuity design to investigate the causal effects of remaining in school. We found consistent evidence that remaining in school causally reduced the risk of diabetes and mortality in all specifications.

We do not know whether the differences in outcomes across education groups are because education directly causes these outcomes by affecting behaviours, such as smoking, or whether these differences are due to other factors, such as socioeconomic or genetic differences. Whether education causes differences in outcomes later in life has been the subject of considerable debate by epidemiologists, economists and other social scientists7–15. Economists have argued that, in addition to its effects on income, a substantial portion of the benefits of education accrue via its potential effects on mortality and morbidity1. Epidemiologists have found that people who attended university have higher fluid intelligence in adulthood16. These associations are robust to adjustment for parental social class and adolescent cognition, which has been taken by some as proof that education causes later outcomes17. Despite this, many epidemiologists and economists are acutely aware that correlations and multivariable adjusted regressions can be unreliable evidence of causation18–20. The ideal experiment to test this hypothesis, randomizing the age at which children leave school, is unlikely to be ethical, cost-effective or timely. A more feasible, and potentially robust, research design is to exploit natural experiments that affected when people left school but are not related to confounding factors21,22. One widely used natural experiment is changes to the legal minimum school leaving age. These changes forced some people to stay in school for longer than they would have otherwise chosen.

In September 1972, the school leaving age increased from 15 to 16 years of age for children in England. Before the reform, the vast majority of those who left school at 15 years of age went into the labour force and found employment. The 1971 census indicated that in April 1971, 32% of 15-year olds were non-students, of whom 87% were in the labour force. At this time, the unemployment rates in this group were 21.7% and 14.9% for males and females, respectively23. Government discussions at the time of the reform raised concerns over the impact of the immediate withdrawal of 400,000 15-year olds from the labour force as a result of the reform. School leavers at this time were strongly attached to the labour market24. Researchers have previously used this policy change to investigate the effects of forcing students to stay in school longer using administrative data and longitudinal cohort studies25–27. However, the cohort studies had relatively small samples and, as a result, produced relatively imprecise estimates of the effects of education. Previous results from administrative data lacked detailed information needed to identify people born in England who were affected by the reform, or on many outcomes of interest such as cognition or clinical measures of ageing, for example, grip strength.

In the current study, we used the raising of the school leaving age in 1972 as a natural experiment to estimate the causal effects of schooling. We used a regression discontinuity design and data from the UK Biobank28–30. We add to the literature in two ways. First, this is the largest sample with detailed individual-level information from the school years immediately before and after the reform. Second, we used genome-wide data to demonstrate that the observational associations of education and other outcomes are probably affected by genomic confounding.

Of the 502,644 participants in the UK Biobank, who were all aged between 37 and 74 years at recruitment in 2008, 390,412 were born in England (see Supplementary Fig. 1 for a flow diagram of inclusion and exclusion of participants in this study, and Supplementary Table 1 for a description of their characteristics). The youngest participants, those born between 1960 and 1971, obtained more education than those born earlier in the twentieth century (Fig. 1). This is consistent with the well-documented secular increase in the length of education over the period31. The UK Biobank includes 11,240 and 10,898 participants who turned 15-years old in the last year before and the first year after the school leaving age increased, respectively. Before the reform, 85% of participants remained in school after 15 years of age, whereas after the reform, almost 100% of participants remained in school after 15 years of age. The proportions of men and women who remained in school after 15 years of age increased over time (Supplementary Fig. 2). Participants born in July and August could still technically leave school before their 16th birthday; this is why participants born in the summer term were more likely to report leaving school before 16 years of age.

1Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, UK. 2School of Social and Community Medicine, University of Bristol, Bristol, UK. 3Department of Social and Policy Sciences, University of Bath, Bath, UK. 4Department of Economics, University of Bristol, Bristol, UK.

* e-mail: neil.davies@bristol.ac.uk

© 2018 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.
People who remained in school after 15 years of age had higher birth weights, their mothers were less likely to smoke during pregnancy, were more likely to have been breastfed, were more likely to have parents who were alive and had fewer siblings (Supplementary Table 2) than those who left school before 16 years of age. In addition, they had more genetic variants (single-nucleotide polymorphisms (SNPs)) that are known to be associated with higher educational attainment\(^1\) (Supplementary Table 2). This suggests that the association of educational attainment and later outcomes will be affected by residual genomic confounding. In comparison, there were few detectable pre-existing differences between people affected and unaffected by the reform. The only detectable difference was that the parents of participants in the first year affected by the reform were more likely to be alive when they attended the assessment centre in 2008–2010 (4.3 (95% confidence interval (95% CI): 2.5–6.1)) and 3.7 (95% CI: 2.6–4.8) percentage points for father and mother, respectively. These associations could be due to age effects, because on average, the parents of those in the first year affected by the reform will be a year younger than parents’ of those in the previous school year. Alternatively, having more educated and potentially richer offspring may increase parents’ longevity, perhaps via improved care\(^2\). There was some evidence that fewer participants in the younger cohort were breastfed. On average, participants in the cohorts before and after the reform had similar numbers of education-associated genetic variants. This suggests that associations of the reform and later outcomes are unlikely to be affected by residual genomic confounding. The participants affected by the reform are, by definition, an average of 1 year younger than those who were not affected. The raw differences above do not account for this age difference. There was little evidence of manipulation around the discontinuity (McCrary robust bias-corrected regression discontinuity manipulation test, \(P = 0.21\))\(^3\).

In this section, we report two comparisons: first, the differences between participants who chose to stay in school after 15 years of age and those who left, and second, the regression discontinuity results. The regression discontinuity results are the difference between participants not affected by the reform (those born before September 1957) and those affected by it (those born in or after September 1957).

On average, participants who chose to stay in school after 15 years of age had better outcomes later in life. They were less likely to be diagnosed with hypertension, diabetes, a stroke or a heart attack, to die, smoke or have ever smoked and were more likely to be diagnosed with depression (left columns in Table 1). Rates of cancer diagnoses were similar across education levels. Participants who remained in school had stronger grips, lower arterial stiffness, and lower systolic and diastolic blood pressure. They also reported higher incomes, were taller, thinner, achieved higher scores on the intelligence test, drank more, watched less television and exercised less. There was little difference in happiness.

Turning to the regression discontinuity results, there was little evidence that the reform affected rates of depression, diastolic blood pressure and rates of moderate and vigorous exercise (right columns in Table 1). For the other outcomes, the effect of the reform was consistent with the association of choosing to remain in school and the outcomes. We found some evidence that the reform may have had a larger effect on the likelihood of males earning more than £31,000 (\(P\) value for interaction = 0.008), but with little evidence of interactions by sex with any other outcomes (Supplementary Tables 3 and 4). There was some evidence that the reform had larger effects on participants who were predicted to leave school before 16 years of age: specifically, increasing the likelihood of earning over £18,000 or £31,000, increasing grip strength and happiness and alcohol consumption (Supplementary Table 5).

As a sensitivity analysis, we repeated the analyses reported in Table 1 using optimal bandwidths\(^4\) (reported in Supplementary Table 6, and sex stratified in Supplementary Tables 7 and 8). These bandwidths are calculated using each outcome and the running variable (the difference between the participant’s date of birth and 1 September 1957 in months). They minimize the mean squared error of the estimates. The bandwidths ranged from 24 to 65.4 months, which was greater than the 12 months used for the results above. These analyses allow for differential linear time trends either side of the reform. This substantially increased the sample size and statistical power (standard errors fell by a factor of 1.25–4). The results were consistent in accordance with the main results reported in Table 1, except for cancer, income over £100,000 and happiness. However, these differences are consistent with sampling error. Supplementary Tables 9–11 provide the results for the regression discontinuity results using a 1-year bandwidth without using inverse probability weights (see Methods).

The associations reported in Table 1 are valid tests of the null hypotheses that education does not affect the outcomes. However, these associations are not informative about the size of the effect of remaining in school. We estimated the effect of remaining in school using instrumental variable analysis. Participants affected by the reform were 23.0 (95% CI: 21.7–24.4) percentage points more likely to remain in school past 15 years of age than those who were unaffected. This suggests that these analyses are unlikely to suffer from weak instrument bias (minimum partial \(F\) statistic = 811). In Supplementary Table 12, we report instrumental variable estimates of the effect of remaining in school past 15 years of age. The instrumental variable estimates are consistent in line with the effect of the reform described above. There was evidence that the linear regression overestimated the effect of remaining in school on rates of ever or current smoking, income, intelligence, sedentary behaviour and exercise (all Hausman test for difference, \(P < 0.007\)).

The instrumental variable results imply that staying in school increases the likelihood of earning more than £18,000, £31,000 or £52,000 by 11.1 (95% CI: 8.9–13.3), 24.0 (95% CI: 21.8–26.2) and 14.6 (95% CI: 9.8–19.3) percentage points. These results exceeded the false discovery rate threshold at \(q = 0.05\) for 18 of the 25 outcomes\(^5\).
Table 1 | The associations of remaining in school after 15 years of age, and attending school after the ROSLA and outcomes

| Outcomes                          | Left school after 15 years of age | Affected by ROSLA |
|----------------------------------|-----------------------------------|-------------------|
|                                  | N       | Risk/mean difference | 95% CI            | P value | N       | Risk/mean difference | 95% CI            | P value |
|                                  |         |                     | Lower  | Upper |         |                     | Lower  | Upper |         |
| Hypertension                     | 21,768  | −0.039              | −0.057  | −0.021 | 1.9×10⁻⁴| −0.018              | −0.026  | −0.010 | 9.0×10⁻⁵|
| Diabetes                         | 22,049  | −0.019              | −0.031  | −0.008 | 0.002   | −0.008              | −0.011  | −0.005 | 3.5×10⁻⁶|
| Stroke                           | 22,110  | −0.006              | −0.011  | −0.002 | 0.009   | −0.003              | −0.005  | −0.001 | 0.001   |
| Heart attack                     | 22,110  | −0.011              | −0.017  | −0.005 | 9.5×10⁻⁴| −0.003              | −0.004  | −0.002 | 2.5×10⁻⁵|
| Depression                       | 21,085  | 0.031               | 0.017   | 0.045  | 9.7×10⁻⁵| −0.003              | −0.010  | 0.005  | 0.47    |
| Cancer                           | 22,111  | −0.006              | −0.020  | 0.008  | 0.38    | −0.005              | −0.011  | 0.001  | 0.09    |
| Died                             | 22,138  | −0.008              | −0.013  | −0.003 | 0.004   | −0.005              | −0.007  | −0.002 | 0.001   |
| Ever smoked                      | 22,086  | −0.205              | −0.228  | −0.183 | 1.9×10⁻⁵| −0.023              | −0.034  | −0.012 | 3.0×10⁻⁴|
| Currently smoke                  | 22,086  | −0.141              | −0.155  | −0.127 | 1.7×10⁻⁶| −0.009              | −0.014  | −0.003 | 0.004   |
| Income over £18,000              | 19,921  | 0.174               | 0.154   | 0.195  | 8.0×10⁻⁵| 0.024               | 0.019   | 0.029  | 2.3×10⁻⁵|
| Income over £31,000              | 19,921  | 0.296               | 0.274   | 0.318  | 4.1×10⁻⁵| 0.052               | 0.047   | 0.058  | 6.7×10⁻⁶|
| Income over £52,000              | 19,921  | 0.256               | 0.239   | 0.274  | 3.2×10⁻⁵| 0.032               | 0.020   | 0.043  | 1.1×10⁻⁵|
| Income over £100,000             | 19,921  | 0.079               | 0.071   | 0.087  | 2.5×10⁻⁶| 0.005               | −0.001  | 0.012  | 0.08    |
| Grip strength (kg)a              | 21,989  | 1.215               | 0.947   | 1.484  | 2.6×10⁻⁴| 0.551               | 0.476   | 0.626  | 1.7×10⁻⁵|
| Arterial stiffness (cm)³a         | 8,537   | −0.750              | −0.931  | −0.570 | 1.2×10⁻⁸| −0.113              | −0.223  | −0.003 | 0.04    |
| Height (cm)³a                    | 22,077  | 1.765               | 1.517   | 2.014  | 3.6×10⁻¹⁵| 0.286               | 0.193   | 0.379  | 1.7×10⁻⁶|
| BMI (kg m⁻²)a                    | 22,055  | −1.235              | −1.478  | −0.992 | 2.9×10⁻⁵| −0.252              | −0.324  | −0.179 | 2.6×10⁻⁷|
| Diastolic blood pressure (mm Hg)a| 21,494  | −0.877              | −1.377  | −0.377 | 0.001   | −0.069              | −0.291  | 0.154  | 0.53    |
| Systolic blood pressure (mm Hg)a | 21,492  | −1.688              | −2.444  | −0.933 | 1.2×10⁻⁴| −0.611              | −0.923  | −0.299 | 4.9×10⁻⁴|
| Intelligence (0-13)a              | 8,540   | 1.653               | 1.458   | 1.849  | 9.0×10⁻¹⁵| 0.148               | 0.086   | 0.210  | 5.8×10⁻⁵|
| Happiness (0-5 Likert)a           | 8,626   | 0.008               | −0.047  | 0.062  | 0.77    | −0.015              | −0.039  | 0.009  | 0.21    |
| Alcohol consumption (1 low, 5 high)a| 22,123 | 0.316               | 0.229   | 0.404  | 1.3×10⁻⁷| 0.036               | 0.009   | 0.064  | 0.01    |
| Television viewing per day (hours)a| 21,206 | −0.834              | −0.916  | −0.752 | 1.5×10⁻⁶| −0.137              | −0.172  | −0.102 | 3.0×10⁻⁸|
| Moderate exercise (days per week)a| 21,330 | −0.480              | −0.639  | −0.321 | 2.2×10⁻⁵| 0.005               | −0.040  | 0.049  | 0.84    |
| Vigorous exercise (days per week)a| 21,379 | −0.129              | −0.207  | −0.051 | 0.002   | 0.010               | −0.019  | 0.038  | 0.50    |

Participants were born between September 1956 and August 1958. Estimated using robust linear regression, with standard errors clustered by year and month of birth. All estimates adjust for the month of birth and sex. The same sample was used for both the conventional linear regression and the raising of the school leaving age (ROSLA) analyses. Inverse probability weights were used to correct for under-sampling of participants who left school at 15 years of age (weight = 1.8857). The difference in outcomes between those who remained and left school at 15 years of age is included for comparison, and may be affected by residual confounding. aDenotes mean differences.

We investigated whether the differences in the outcomes seen in the regression discontinuity results could be solely explained by the ageing process using a difference-in difference approach. We created a series of non-overlapping negative control samples, which contained participants who were born in consecutive school years in the 10 years before and after the reform. For each of these samples, we allocated the younger cohort to a ‘placebo’ reform (see Supplementary Fig. 1 for diagram and sample sizes). In each of these negative control samples, all of the participants experienced the same minimum school leaving age. Thus, any differences between the younger and the older school cohort cannot be due to the raising of the school leaving age in 1972, and are probably due to the ageing process and not an effect of education.

Forest plots of the differences in the outcomes for the negative control analyses are reported in the Supplementary Information (Supplementary Figs. 5–29). There was evidence of an effect of age. On average, younger participants in both the raising of the school leaving age and the negative control cohorts were less likely to report having had a diagnosis of hypertension, a heart attack or cancer, die during follow-up, currently smoke, report higher incomes, have higher grip strength, lower arterial stiffness, be taller and slimmer, have lower diastolic and systolic blood pressure, have higher scores on the intelligence tests, be less sedentary and do less moderate exercise. The effect of the reform on diastolic blood pressure was similar to the year-on-year differences seen before the reform, but was smaller than differences observed after the reform. The effect of the reform on the likelihood of earning over £18,000 and £52,000 was

Supplementary Figs. 3 and 4 plot the conventional linear regression and the instrumental variable point estimates and 95% CI using a 12-month bandwidth. Supplementary Tables 13 and 14 report the instrumental variable results stratified by sex. There was little evidence that the reform had larger effects on men than on women, except for the likelihood of having an income above £31,000 (P value for interaction = 0.009).
similar to the year-on-year differences observed before the reform, but was larger than the differences observed after the reform.

The effects of the reform on the outcomes after accounting for age are shown in Fig. 2. The effect of the reform exceeded the false discovery threshold for: diabetes, stroke, mortality, former smoker, current smoker, earning over £18,000 or £31,000, grip strength, body mass index (BMI), intelligence, alcohol consumption and sedentary behaviour. We report sensitivity analyses of the overall result without using inverse probability weights (see Methods) in Supplementary Figs. 30 and 31. The effects of the reform exceed the false discovery rate threshold in both the weighted and unweighted analysis for diabetes, stroke, mortality and grip strength.

This study provides some of the strongest evidence to date about the causal effects of education. We found that the raising of the school leaving age in 1972 affected some health outcomes. A conservative analysis is to focus on the effects that were consistently found across all estimation methods. We found that there was consistent evidence that the reform had generally beneficial effects on the risk of diabetes and mortality. Finally, we found molecular genetic evidence that regression discontinuity designs using raising of the school leaving age are unlikely to be affected by residual genomic confounding.

Clark and Royer found that the participants of the Health Survey for England and the General Household Survey who were affected by the reform were by 26.1 (95% CI: 23.0–29.2) percentage points more likely to stay in school after 15 years of age. After correcting for under-sampling of people who left school at 15 years of age, we found a slightly smaller difference (23.0 (95% CI: 21.7–24.4)). Clark and Royer found that people affected by the reform may have had lower mortality between 40 and 44 years of age (odds ratio = 0.95, 95% CI: 0.89–1.01), but had no detectable effects on current or ever smoking, or drinking. Figure 3 presents a sensitivity analyses using identical bandwidths and covariates, as in Clark and Royer, for mortality, current and ever smoking, and drinking alcohol (coded as a binary rather than ordinal variable in our main analysis). As with our main results, the estimates using Clark and Royer’s specification suggest that those affected by the reform had a substantially lower risk of mortality (odds ratio = 0.58, 95% CI: 0.39–0.87) (Fig. 3). Furthermore, this difference was greater than the average year-on-year difference in mortality seen before and after the reform (Supplementary Fig. 11).

The difference between the UK Biobank and Clark and Royer mortality results may be because the UK Biobank participants were almost 10 years older (mean age: 53.2 years) than those sampled by Clark and Royer. Clark and Royer sampled those 40–44 years of age and had a 5-year follow-up. The 5-year mortality rate for this age group is 0.79%56. The five leading causes of death for this age group in 2001 were cancer (22.9%), ischaemic heart disease (14.9%), alcohol-related disease (13.3%), suicides (12.1%) and accidental injuries (7.0%). By contrast, the subsample of the UK Biobank used in the study comprises individuals between 42 and 62 years of age and has a 7.78-year follow-up. The 8-year probability of mortality between 42 and 62 years of age was 3.44% in 2008. The five leading causes of death for this age group in 2008 were cancer (37.0%), ischaemic heart disease (20.0%), alcohol-related disease (9.0%), cerebrovascular diseases (5.7%) and chronic obstructive pulmonary disease (4.8%). Thus, the absolute probability of mortality is over four times as high in the UK Biobank, and the causes of death differ. In particular, the risk of mortality due to smoking-related illness, such as ischaemic heart disease, cancer (particularly lung cancer) and chronic obstructive pulmonary disease was much higher in people eligible to participate in the UK Biobank. Thus, it is possible that Clark and Royer’s sample was too young to detect any difference in mortality. Finally, Clark and Royer could not exclude immigrants, who were not affected by the reform, from their sample. This could attenuate their estimates towards the null.

In the sensitivity analysis reported in Fig. 3, our estimates of the effect of the reform on smoking and alcohol consumption were almost identical to Clark and Royer. However, we found some evidence that the reform affected alcohol consumption and smoking rates using an ordinal measure of alcohol consumption and tighter bandwidths. These effects exceeded the age effects found in the difference-in-difference analysis for the inverse probability weighted but not in the unweighted analysis. This suggests that the reform may have affected the frequency of alcohol consumption in those who drink alcohol, but had little effect on whether participants drank.

Epidemiologists have argued that education has causal effects on intelligence later in life. A previous study found that educational attainment by 26 years of age was associated with intelligence at 53 years of age77, which the authors argued was evidence that education had a causal effect on intelligence78. However, another study raised doubts about this interpretation and called for greater clarity about the assumptions underlying these analyses79. We found modest evidence of a causal effect of education on intelligence later in life from the inverse probability weighted estimates. This suggests that the raw differences in intelligence between those who remain and leave school at 15 years of age may overestimate the effect of schooling on cognition. Our results are also consistent with a previous study, which used increases in the legal school leaving ages in the United States to investigate the effects of education on the risk of dementia later in life80. They found evidence that education reduced the risk of dementia. We cannot test this hypothesis directly with the UK Biobank data because too few participants have been diagnosed with dementia.

People with more education were much less likely to smoke. However, it is not clear whether this is due to a causal effect of education. A previous study found that the association between education and smoking status was attenuated in sibling fixed effects designs81. We found evidence that participants affected by the reform were less likely to smoke or have ever smoked. Educated participants drank more heavily, but the instrumental variable estimates suggested that this was probably an overestimate of the causal effect of education on alcohol consumption. However, these effects only exceeded the false discovery rate in the weighted analysis. We found some evidence that the effects of the reform on income were greatest in participants who would otherwise have been expected to leave school at 15 years of age. Our results are consistent with those of a previous study14 that used data from the UK Biobank to investigate heterogeneity in the effects of education on BMI and blood pressure. The study used a 110-month bandwidth and a triangle kernel. The results of this study, allowing for differential linear trends before and after the reform, suggested that remaining in school caused a 0.42 kg/m² (95% CI: −0.30 to 1.14) reduction in BMI and a 2.3 (95% CI: −0.1 to 4.7) percentage point reduction in the risk of diabetes82.

A key strength of our study is that we used a natural experiment to identify the effects of education. The raising of the school leaving age in 1972 provided exogenous variation in the length of schooling. We found few pre-existing differences between participants on either side of the reform, suggesting that it can be used as a potentially valid instrumental variable83. Another strength of our study is that it uses one of the largest samples to date to investigate the effects of education on a wide range of outcomes. Our outcomes were recorded both in clinics and via linked NHS mortality registry data. This means that our outcomes are probably affected by relatively little measurement error. Furthermore, we were able to restrict our sample to people born in England who were affected by the reform. In addition, we used genome-wide data to show that this natural experiment is unlikely to be affected by residual genomic confounding. Participants unaffected and affected by the reform had very similar genome-wide scores for education. A potential limitation of our study is that our treatment group, people affected by the
reform, are 1 year younger than our control group, those born in the last school year unaffected by the reform. Many of the outcomes we investigated increase linearly or log-linearly over time. This means that it is difficult to determine whether any of the differences we observed in the regression discontinuity design with 12-month bandwidths were due to an additional year of ageing or the reform.

### Table: Risk Difference per 100 (95% CI) and P value

| Outcome                          | Risk difference per 100 (95% CI) | P value   |
|---------------------------------|----------------------------------|-----------|
| High blood pressure             | −0.86 (−1.65 to −0.08)           | 3.2 × 10⁻²|
| Diabetes                        | −0.81 (−1.09 to −0.54)           | 6.9 × 10⁻³|
| Stroke                          | −0.28 (−0.44 to −0.12)           | 6.6 × 10⁻⁴|
| Heart attack                    | −0.12 (−0.24 to −0.01)           | 4.1 × 10⁻²|
| Depression                      | −0.35 (−1.10 to 0.39)            | 3.5 × 10⁻¹|
| Cancer                          | −0.02 (−0.63 to 0.59)            | 9.4 × 10⁻¹|
| Died                            | −0.37 (−0.62 to −0.12)           | 4.0 × 10⁻³|
| Ever smoked                     |                                  |           |
| Currently smoke                 |                                  |           |
| Income over £18,000             | 1.51 (1.01 to 2.01)              | 2.8 × 10⁻⁵|
| Income over £31,000             | 3.35 (2.76 to 3.94)              | 7.8 × 10⁻⁹|
| Income over £52,000             | 1.29 (0.15 to 2.43)              | 2.7 × 10⁻²|
| Income over £100,000            | 0.19 (−0.39 to 0.78)             | 5.2 × 10⁻¹|

### Table: Mean difference (95% CI) and P value

| Outcome                                      | Mean difference (95% CI) | P value |
|----------------------------------------------|--------------------------|---------|
| Grip strength (kg)                           | 0.26 (0.18 to 0.34)      | 1.4 × 10⁻¹⁰|
| Arterial stiffness                           | −0.02 (−0.13 to 0.08)    | 6.6 × 10⁻¹|
| Height (cm)                                  | 0.09 (−0.01 to 0.18)     | 7.7 × 10⁻²|
| BMI (kg m⁻²)                                 | −0.22 (−0.30 to −0.15)   | 2.6 × 10⁻⁹|
| Diastolic blood pressure (mm Hg)             | 0.06 (−0.16 to 0.27)     | 6.1 × 10⁻¹|
| Systolic blood pressure (mm Hg)              | 0.33 (0.02 to 0.64)      | 3.6 × 10⁻²|
| Intelligence (0–13)                          | 0.12 (0.06 to 0.18)      | 1.8 × 10⁻⁴|
| Happiness (0–5 Likert)                       | −0.00 (−0.03 to 0.02)    | 8.3 × 10⁻¹|
| Alcohol consumption (1 low, 5 high)          | 0.05 (0.02 to 0.07)      | 1.2 × 10⁻⁹|
| Television viewing per day (hours)           | −0.10 (−0.13 to −0.06)   | 1.6 × 10⁻⁹|
| Moderate exercise (days per week)            | 0.04 (−0.00 to 0.09)     | 6.5 × 10⁻²|
| Vigorous exercise (days per week)            | 0.01 (−0.02 to 0.04)     | 6.6 × 10⁻¹|

### Fig. 2 | The effect of the reform on each outcome estimated using the difference-in-difference approach accounting for age effects. The units in the top panel are reported on the absolute risk difference scale (risk differences per 100 people). This is interpreted as the change in the number of events per 100 people affected by the reform. The units in the bottom panel differ by outcome and are listed in the legend on the left-hand side. All estimates control for sex and month of birth. Estimates are the difference between the year-on-year difference in outcome across the raising of the school leaving age compared to the average year-on-year difference. Estimated using robust linear regression, with standard errors clustered by month of birth and weighting. Differences and 95% CI were calculated using Bland–Altman tests. The estimates for diabetes, stroke, mortality, former and current smoking, income over £18,000 and £31,000, grip strength, BMI, intelligence, alcohol consumption and sedentary behaviour exceed false discovery rate threshold for multiple hypothesis testing. Maximum N = 262,348. Error bars indicate 95% CIs.
reduce the average outcome in the ‘treatment’ group, and be miss-
our results towards the null, because these marginal students would
school had they attended school after the reform (the compliers),
could cause less-educated people, who would have remained in
England (minimum
entire English and Welsh population) and the General Health Survey for
from the Office of National Statistics Census (summary data from the

The UK Biobank participants were between 42 and 62 years of age
risk of mortality in the 5 years between 40 and 44 years of age, whereas
the UK Biobank participants were between 42 and 62 years of age and
follow-up spanned 7.78 years (over the period 10 May 2006 to 17 February
2014). Error bars indicate 95% CIs.

We addressed this by using a difference-in-difference approach to
estimate the average effects of a year of ageing (Fig. 3), and allowed for a
differential linear time trend before and after the reform as a
sensitivity analysis using wider bandwidths (Supplementary Tables 4–6). These results suggest that ageing rather than the reform probably explain the differences observed across the regression discontinuity for outcomes such as height. However, it is probable that the reform affected outcomes where substantial effects remained in the difference-in-difference analysis.

A representative sample is not a necessary condition for making causal inferences. Nevertheless, collider (attenuation) bias could affect our results because the UK Biobank is a volunteer sample, which over-sampled more-educated people. People affected by the reform may be more likely to participate in the study. This could cause less-educated people, who would have remained in school had they attended school after the reform (the compliers), to be under-represented in the UK Biobank. This could attenuate our results towards the null, because these marginal students would reduce the average outcome in the ‘treatment’ group, and be missing from the ‘control’ group. This would improve the control group's outcomes relative to the treatment group. Despite these differences, we found little evidence that people affected by the reform were more likely to participate in the UK Biobank (see Supplementary Fig. 32). In our primary analysis, we used inverse probability weighting to account for this sampling. This requires the assumption that the participants sampled in the UK Biobank who left school at 15 years of age are representative of the population that left school at 15 years of age. However, this issue warrants further investigation in future research.

There was limited time to collect measures during the participants’ assessment centre visits; thus, our measure of intelligence is relatively coarse. Despite this, participants who remained in school had substantially higher intelligence. The instrumental variable estimates suggest that this difference substantially overestimates the causal effect. Finally, our instrumental variable results are estimates of the local average treatment effect of schooling. They can be interpreted (‘point identified’) either under the assumption that the reform had a monotonic effect on the likelihood of staying in school (monotonicity), or that the effects of schooling on the outcomes were not affected by the reform (no effect modification). Under the monotonicity assumption, our results are estimates of the causal effects of being forced to remain in school after 15 years of age on those who would otherwise have left school. These effects may not be externally valid to infer either the effects of compelling students to remain in school for longer, or of the effects of education on other populations. In particular, these results may not be valid estimates of the effect of education on ‘always takers’, that is, people who would always remain in school regardless of the reform. Under the no effect modification assumption, we identified the average effect of education on those who remained in school. At a minimum, our results are internally valid estimates of the effects of schooling on people affected by the reform.

Does education affect outcomes later in life? Yes, although education is not the panacea implied by naïve multivariable adjusted regression, in this sample, increasing the length of compulsory schooling had substantial benefits. We found robust evidence that staying in school probably has causal effects on the risk of diabetes and mortality. These results add to our understanding of the long-term consequences of educational decisions in childhood and adolescence.

Methods

Data. We used data from 502,624 participants from the UK Biobank project. The participants, aged between 37 and 74 years, were originally recruited between 2006 and 2010. In our regression discontinuity analysis, we restricted our sample to participants who were born in England in the school cohorts in the years immediately before and after the reform took place. We did this because we have a large enough sample born in these years to precisely identify the effects of schooling.

Exposure. The participants were asked whether they had a college or university degree. If they did not have a degree, they were asked what age they left full-time education. We coded participants who reported having a degree as leaving full-time education at 21 years of age. Participants who did not report having a degree and did not have data on the age at which they left education were coded as missing.

Outcomes. Health outcomes. The participants were asked whether they had ever been diagnosed by a doctor with the following health conditions: hypertension, stroke, type 2 diabetes or heart attack. They were asked whether they had ever had a whole week where they felt depressed or down. The death of the participants was defined using linked NHS mortality registry data. Follow-up for the linked mortality data started with the first death on 10 May 2006 and ended with the last recorded death on 17 February 2014. The cancer diagnoses were taken from the national cancer registries. The first recorded cancer diagnosis was on 20 September 1957 and the last was on 25 October 2013.

Height, BMI, blood pressure, arterial stiffness, grip strength and intelligence. Height and weight were measured during the participants’ visit to a UK Biobank assessment centre. Two measures of diastolic and systolic blood pressure were recorded via an electronic blood pressure monitor. The measurements were taken 2 min apart. Arterial stiffness was measured using an electronic measuring device. Grip strength was measured in kilograms using a hydraulic hand dynamometer. We residualized the measures of grip strength and arterial stiffness to control for potential between device heterogeneity. Fluid intelligence was measured via the number of 13 logic puzzles that the participants could answer correctly in 2 min.

Health behaviours and income. During their assessment centre visit, the participants were asked to report their health behaviours. They were asked about how frequently they consumed alcohol. This is coded 6 if they drank every day, 5 for 3 or 4 times a week, 4 for 1 or 2 times a week, 3 for 1–3 times a month,
2 for special occasions only, and 1 for never. They were asked if they smoked or had ever smoked. They were asked how often they moderately and vigorously exercised in a typical week. Finally, they were asked if their pre-tax income was below £18,000, between £18,000 and £30,999, between £31,000 and £51,999, between £52,000 and £100,000, or above £100,000. Participants who did not answer these questions were coded as missing.

Genotype data. The participants provided a blood sample. This sample was used to extract DNA and genotype using the Axiom and BiLEVE genome-wide arrays. These arrays genotyped around 800,000 SNPs for each participant. The genotyping data were used to impute SNPs that were not directly genotyped using the 1,000 genomes and UK10K reference panels. The imputation produced a likelihood of each participant having a specific genotype (for example, AA, 0.1; TA, 0.9; and TT, 0). This resulted in a data set of around 80,000,000 SNPs. For each participant, we created a weighted genotype score by summing over all SNPs they had that were associated with higher educational attainment. We weighted each variant by its association with education reported in a large genome-wide association study, using a version of the genome-wide association study that did not include the UK Biobank. This study reported the association of 8,259,394 genetic variants and each of education in a meta-analysis of 64 studies. We normalized the allele score to have a mean of zero and a standard deviation of one. This score only explains a minority ($r^2 = 0.32$) of the full UK Biobank sample of the variation in educational attainment explained by the genome-wide data. This is because of limited statistical power of existing genome-wide association studies for educational attainment. One concern with this is that the score is too poor a proxy for the total genetic effects on educational attainment to be used as a conventional covariate in a regression. Thus, we used the educational attainment genome-wide score to test whether on average participants affected by the reform had more genetic variants known to associate with education.

Statistical methods. We used the changes in the school leaving age to identify the effects of schooling on a range of outcomes. Our empirical strategy has five steps. First, we estimated the effect of the reforms on the proportion of participants who remained in school after 15 years of age. Second, we investigated the associations of potential confounders with educational attainment and across the cohorts affected by the reform. Third, we used a regression discontinuity design to estimate the effect of the reform on the outcomes. Fourth, we used instrumental variable estimators to estimate the effects of remaining in school. For continuous outcomes, we used conventional Wald estimators; for binary outcomes, we used semi-parametric additive structural mean models. To address concerns about multiple hypothesis testing, we report whether the instrumental variable results for each outcome exceed a Benjamini and Hochberg false discovery rate threshold at $\alpha = 0.05$ across 25 outcomes. Fifth, we conducted a difference-in-difference analyses.

Inverse probability weighting. The UK Biobank is a volunteer sample and, as a result, people who left school at 16 years of age were likely to be unrepresentative of the cohort who left school at 15 years of age. Therefore, we used inverse probability weighting to account for non-random sampling. Thus, we corrected for the non-random sampling using inverse probability weights. This is a valid test of the null hypothesis that remaining in school does not affect the outcomes. We tested whether the reform had larger effects on people who would otherwise have been expected to leave school at 15 years of age. We estimated the probability that a participant would remain in school after 15 years of age using logistic regression and data from individuals born before 31 August 1956. This model included indicators for the participants' assessment centre, year and month of birth, sex, whether the mother smoked during pregnancy, were breastfed, number of brothers and sisters, the normalized genome-wide education score, and their ethnicity. Missing data were replaced at the mean, and indicator variables for missing values were included. We estimated the following regression:

$$
H_{ist} = \beta_1 + \delta_i F_{ist} + \pi w_{ist} + \epsilon_{ist}
$$

where $F_{ist}$ is the probability of remaining in education from the logistic regression. For each outcome, we report the coefficients on the reform indicator, and the coefficient on the interaction term and the effect of the reform. The effect of the reform on participants predicted to leave is indicated by $\pi$, and the effect on those not expected to stay is indicated by $\delta_i$. As with the main results above, we adjust for sex and month of birth, and the interaction of these variables with predicted education.

As a sensitivity analysis, we used a regression discontinuity design with variable month bandwidths to investigate the robustness of our findings. In our main analysis above, we present difference in outcomes for the last school cohort of participants who were born between 1 September 1957 and the first cohort affected by the reform (those born between September 1957 and August 1958). This is a regression discontinuity analysis with a bandwidth of 1 year. This is a fuzzy regression discontinuity design, as the reform...
only increased the probability of staying in school. In the sensitivity analyses, we investigated whether our results were sensitive to the size of the bandwidth around the reform. We did this by repeating our instrumental variable analyses on a sample defined using optimal bandwidths. Analyses using these bandwidths use the same specification as the instrumental variable analyses described above and also include linear time trends, which vary either side of the reform. We estimated the optimal bandwidths using the rdselect command in Stata.

Instrumental variables. We estimated the causal effect of schooling using instrumental variables estimators. We estimated mean differences using Wald estimators, and risk differences using additive structural mean models, for the continuous and binary outcomes, respectively. These models can be identified by making one of three assumptions. First, for the continuous outcomes, we could assume that staying in school has the same effect on the outcomes for all participants. This identifies the average effects of staying in school but is implausible for binary outcomes. Second, for the binary outcomes, we could assume a monotonic relationship between the reform and the participants’ likelihood of staying in school after 15 years of age. In the potential outcomes framework, that \( E[Y^1] - E[Y^0] = E[Y^1(y^0)] - E[Y^0(y^0)] \). This requires that there were no participants who were ‘defiers’, who would have remained in school if they were not affected by the reform, but would have left school if they were affected by the reform. Under monotonicity, the instrumental variable estimators estimate a local average treatment effect. This is the effects of treatment in the subgroup of participants whose decisions were affected by the reform; that is, the people in the year after the reform who would have chosen to leave school at 15 years of age had the reform not been introduced. Finally we could assume that the effects of education are not affected by the reform (no effect modification). This would identify the effects of education on participants who remained in school. We report the partial F-statistic of the association of \( E_{y_0} \) and \( D \). We also report the test for endogeneity (using a C-statistic, which is a heteroskedasticity robust Hausman test) of \( E_{y_0}(y^1) - E_{y_0}(y^0) \). This implicitly tests for differences between the linear regression and instrumental variable estimates. All estimates allow for clustered standard errors by year and month of birth, and include controls for sex and month of birth.

Difference-in-difference. We were concerned that differences between the two school years may occur because the participants affected by the reform were 1 year younger on average than participants unaffected by the reform. To investigate this, we estimated the year-on-year differences in each outcome for the five non-overlapping 2-year cohorts in the 10 years before and after the reform. Otherwise, we used an identical specification to the regression discontinuity analysis above. There were no changes to the school leaving ages between each of these years. Thus, any year-on-year differences observed in these ‘negative control cohorts’ must be due to other factors, such as age effects, and cannot be an effect of raising the school leaving age in 1972. We compared these estimates using forest plots, which are reported in the Supplementary Information. We pooled the year-on-year differences from the five negative control samples from before and after the reform using the Stata command metan. We calculated the difference between this pooled estimate and the difference between the years before and after the reform. We estimated the difference and the standard error of this difference using Bland–Altman tests.

Life Sciences Reporting Summary. Further information on experimental design is available in the Life Sciences Reporting Summary.

Code availability. All analyses were conducted in Stata MP 14.0 (ref. 5). The code used to generate these results can be found at https://github.com/nmdavies/UKbiobankRUSLA.

Data availability. The data used in this study can be accessed by contacting the UK Biobank (www.ukbiobank.ac.uk). This analysis was approved by the UK Biobank access committee as part of project 8786. The protocol for this study is available in the Supplementary Information. Consent was sought by the UK Biobank as part of the recruitment process.

Received: 22 November 2016; Accepted: 5 December 2017; Published online: 29 January 2018

References
1. Lager, A. C. J. & Torsosander, J. Causal effect of education on mortality in a quasi-experiment on 1.2 million Swedes. Proc. Natl Acad. Sci. USA 109, 8461–8466 (2012).
2. Clark, D. & Royer, H. The effect of education on adult mortality and health: evidence from Britain. Am. Econ. Rev. 103, 2087–2120 (2013).
3. Cutler, D. & Lleras-Muney, A. Education and Health: Evaluating Theories and Evidence (National Bureau of Economic Research, 2006).
4. Dickson, M. The causal effect of education on wages revisited: the causal effect of education. Q. J. Beall. Econ. 75, 477–498 (2013).
5. Fry, A. et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. Am. J. Epidemiol. 166, 1026–1034 (2017).
6. Harmon, C. & Walker, I. Estimates of the economic return to schooling for the United Kingdom. Am. Econ. Rev. 85, 1278–1286 (1995).
7. Bulik-Sullivan, B. K. et al. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat. Genet. 47, 291–295 (2015).
8. Connelly, G., Heckman, J. & Urzua, S. The education–health gradient. Am. Econ. Rev. 100, 234–238 (2010).
9. Naess, O., Hof, D. A., Lawlor, D. & Mortensen, L. H. Education and adult cause-specific mortality—examining the impact of family factors shared by 871 367 Norwegian siblings. Int. J. Epidemiol. 41, 1683–1691 (2012).
10. Meghie, C., Palme, M. & Simeonova, E. Education, Health and Mortality: Evidence from a Social Experiment (National Bureau of Economic Research, 2012).
11. Nordahl, H. et al. Education and cause-specific mortality: the mediating role of differential exposure and vulnerability to behavioral risk factors. Epidemiology 25, 389–396 (2014).
12. Mackenbach, J. P. et al. Variations in the relation between education and cause-specific mortality in 19 European populations: a test of the ‘fundamental causes’ theory of social inequalities in health. Soc. Sci. Med. 127, 51–62 (2015).
13. Strand, B. H. et al. Educational inequalities in mortality over four decades in Norway: prospective study of middle aged men and women followed for cause specific mortality, 1960–2000. BMJ 340, e654 (2010).
14. Baker, D. P., Leon, J., Smith Greenaway, E. G., Collins, J. & Movit, M. The education effect on population health: a reassessment. Popul. Dev. Rev. 37, 307–332 (2011).
15. Spearman, C. ‘General intelligence,’ objectively determined and measured. Am. J. Psychol. 15, 201–292 (1904).
16. Davey Smith, G. et al. Education and occupational social class: which is the more important indicator of mortality risk? J. Epidemiol. Community Health 52, 153–160 (1998).
17. Schafer, M. H., Wilkinson, L. R. & Ferraro, K. F. Childhood (mis)fortune, educational attainment, and adult health: contingent benefits of a college degree? Soc. Forces 91, 1007–1034 (2013).
18. Clouston, S. A. et al. Benefits of educational attainment on adult fluid cognition: international evidence from three birth cohorts. Int. J. Epidemiol. 41, 1729–1736 (2012).
19. Richards, M. & Sacker, A. Is education causal? Yes. Int. J. Epidemiol. 40, 516–518 (2011).
20. Davey Smith, G. & Ebrahim, S. Epidemiology—is it time to call it a day? Int. J. Epidemiol. 30, 1–11 (2001).
21. Learner, E. Let’s take the con out of ecometrics. Am. Econ. Rev. 73, 31–43 (1983).
22. Deary, I. J. & Johnson, W. Intelligence and education: causal perceptions drive analytic processes and therefore conclusions. Int. J. Epidemiol. 39, 1362–1369 (2010).
23. Angrist, J. D. & Krueger, A. B. Instrumental variables and the search for causal effects. J. Econ. Perspect. 15, 69–85 (2001).
24. Nguyen, T. T. et al. Instrumental variable approaches to identifying the causal effect of educational attainment on dementia risk. Am. J. Epidemiol. 26, 71–76.e3 (2016).
25. Layard, R. in The Youth Labor Market Problem: Its Nature, Causes, and Consequences (eds Freeman, R. B. & Wise, D. A.) 499–542 (Univ. Chicago Press, Chicago, IL, 1982).
26. McCulloch, G., Cowan, S. & Woodin, T. The British Conservative Government and the raising of the school leaving age, 1959–1964. J. Educ. Policy 27, 509–527 (2012).
27. Powdthavee, N. Does education reduce the risk of hypertension? Estimating the biomarker effect of compulsory schooling in England. J. Hum. Cap. 4, 173–202 (2010).
28. Jürges, H., Kruk, E. & Reinhold, S. The effect of compulsory schooling on health—evidence from biomarkers. J. Popul. Econ. 26, 645–672 (2013).
29. Collins, R. What makes UK Biobank special? Lancet 379, 1173–1174 (2012).
30. O’Keeffe, A. G. et al. Regression discontinuity designs: an approach to the evaluation of treatment efficacy in primary care using observational data. BMJ 349, g5293 (2014).
31. Okbay, A. et al. Genome-wide association study identifies 74 loci associated with educational attainment. Nature 533, 539–542 (2016).
32. Torsosander, J. From child to parent? The significance of children’s education for their parents’ longevity. Demography 50, 637–659 (2013).
33. Mori, H. The manipulation of the running variable in the regression discontinuity design: a density test. J. Econom. 142, 698–714 (2008).
34. Calonico, S., Cattaneo, M. D. & Titiunik, R. Robust nonparametric confidence intervals for regression-discontinuity designs: robust nonparametric confidence intervals. Econometrica 82, 2295–2326 (2014).
56. Keller, M. C. Gene × environment interaction studies have not properly controlled for potential confounders: the problem and the (simple) solution. **Biol. Psychiatry** 75, 18–24 (2014).

57. Lee, D. S. & Lemieux, T. Regression discontinuity designs in economics. **J. Econ. Lit.** 48, 281–355 (2010).

58. Hernán, M. A. & Robins, J. Instruments for causal inference: an epidemiologist’s dream? **Epidemiology** 17, 360–372 (2006).

59. Hayashi, F. **Econometrics** (Princeton Univ. Press, Princeton, NJ, 2000).

60. Hausman, J. A. Specification tests in econometrics. **Econometrica** 46, 1251–1271 (1978).

61. Altman, D. G. & Bland, J. M. Interaction revisited: the difference between two estimates. **BMJ** 326, 219 (2003).

62. Stata Statistical Software: Release 14 (StataCorp, 2015).

### Acknowledgements

We thank the Social Science Genetic Association Consortium for providing the coefficients from the educational attainment genome-wide association study, and G. Hemani, L. Paternoster, D. Carslake, J. Bowden, L. Zuccolo, E. Stergiakouli and E. Sanderson for helpful comments on an earlier draft. All mistakes remain our own. The Medical Research Council (MRC) and the University of Bristol fund the MRC Integrative Epidemiology Unit (MC_UU_12013/1 and MC_UU_12013/9). N.M.D. is supported by the Economics and Social Research Council (ESRC) via a Future Research Leaders grant (ES/N000757/1). The research described in this paper was specifically funded by a grant from the ESRC for Transformative Social Science. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. This work was carried out using the computational facilities of the Advanced Computing Research Centre (www.bris.ac.uk/acrc/) and the Research Data Storage Facility of the University of Bristol (www.bris.ac.uk/acrc/storage/). This research was conducted using the UK Biobank Resource.

### Author contributions

N.M.D. obtained funding for this study, analysed and cleaned the data, interpreted results, and wrote and revised the manuscript. M.D. interpreted the results and wrote and revised the manuscript. G.D.S. obtained funding for this study, interpreted results and wrote and revised the manuscript. F.W. obtained funding for this study, interpreted results and wrote and revised the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** is available for this paper at https://doi.org/10.1038/s41562-017-0279-y.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Correspondence and requests for materials** should be addressed to N.M.D.

**Publisher’s note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

### Experimental design

1. **Sample size**
   - Describe how sample size was determined.
   - The sample size consisted of all participants with valid data values for the outcome, education and date of birth.

2. **Data exclusions**
   - Describe any data exclusions.
   - Participants who were not born in England or Wales. The raising of the school leaving age may have had different effects in Scotland.

3. **Replication**
   - Describe whether the experimental findings were reliably reproduced.
   - NA

4. **Randomization**
   - Describe how samples/organisms/participants were allocated into experimental groups.
   - NA

5. **Blinding**
   - Describe whether the investigators were blinded to group allocation during data collection and/or analysis.
   - NA

6. **Statistical parameters**
   - For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).
   - **n/a** Confirmed
   - The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
   - A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
   - A statement indicating how many times each experiment was replicated
   - The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
   - A description of any assumptions or corrections, such as an adjustment for multiple comparisons
   - The test results (e.g. P values) given as exact values whenever possible and with confidence intervals noted
   - A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
   - Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.
Software

Describe the software used to analyze the data in this study.

We used Stata 14.0 MP for all analyses.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). Nature Methods guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

No unique methods were used. All data are available from UK Biobank.

No antibodies were used.

No eukaryotic cell lines were used.

No eukaryotic cell lines were used.

No eukaryotic cell lines were used.

No commonly misidentified cell lines were used.

Animals and human research participants

Provide details on animals and/or animal-derived materials used in the study.

No animals were used.

This information is provided in Table S1.