# Education, intelligence and Alzheimer's disease: Evidence from a multivariable two-sample Mendelian randomization study

**Emma L Anderson, research fellow[1,2], Laura D Howe, reader[1,2], Kaitlin H Wade, senior research associate [1,2], Yoav Ben-Shlomo, professor[2], W. David Hill, statistical geneticist[3], Ian J Deary, professor[3], Eleanor C Sanderson, senior research associate[1,2], Jie Zheng, senior research associate[1,2], Roxanna Korologou-Linden, research assistant[1,2], Evie Stergiakouli, lecturer[1,2,4], George Davey Smith, professor[1,2], Neil M Davies, senior research fellow[1,2]*, Gibran Hemani, senior research fellow[,2]*.**

***Joint last authors with equal contributions**

[1] Medical Research Council Integrative Epidemiology Unit, Bristol Medical School, University of Bristol, UK.

[2] Population Health Sciences, Bristol Medical School, University of Bristol, UK.

[3] Centre for Cognitive Ageing and Cognitive Epidemiology, Department of Psychology, University of Edinburgh, Edinburgh, UK.

[4] School of Oral and Dental Sciences, University of Bristol, Bristol, UK.

**Corresponding Author:**

Dr Emma L Anderson

MRC Integrative Epidemiology Unit

Oakfield House, Oakfield Grove, Clifton

Bristol, UK

BS8 2BN

Tel: 01173 313414

Email: emma.louise.anderson@bristol.ac.uk

**Word Count:** 4323

## ABSTRACT

**Objectives:** To examine whether educational attainment and intelligence have causal effects on risk of Alzheimer's disease (AD), independently of each other.

**Design:** Two-sample univariable and multivariable Mendelian Randomization (MR) to estimate the causal effects of education on intelligence and vice versa, and the total and independent causal effects of both education and intelligence on risk of AD.

**Participants:** 17,008 AD cases and 37,154 controls from the International Genomics of Alzheimer's Project (IGAP) consortium

**Main outcome measure:** Odds ratio of AD per standardised deviation increase in years of schooling and intelligence

**Results:** There was strong evidence of a causal, bidirectional relationship between intelligence and educational attainment, with the magnitude of effect being similar in both directions. Similar overall effects were observed for both educational attainment and intelligence on AD risk in the univariable MR analysis; with each SD increase in years of schooling and intelligence, odds of AD were, on average, 37% (95% CI: 23% to 49%) and 35% (95% CI: 25% to 43%) lower, respectively. There was little evidence from the multivariable MR analysis that educational attainment affected AD risk once intelligence was taken into account, but intelligence affected AD risk independently of educational attainment to a similar magnitude observed in the univariate analysis.

**Conclusions:** There is robust evidence for an independent, causal effect of intelligence in lowering AD risk, potentially supporting a role for cognitive training interventions to improve aspects of intelligence. However, given the observed causal effect of educational attainment on intelligence, there may also be support for policies aimed at increasing length of schooling to lower incidence of AD.

## INTRODUCTION

Alzheimer's disease (AD) is the leading cause of death in England and Wales[1]. Existing treatments are currently unable to reverse or delay progression of the disease. Thus, strategies for reducing the incidence of the disease by intervening on modifiable risk factors are important. Higher educational attainment is associated with a lower risk of dementia[2-5]. However, the mechanisms underlying the associations of educational attainment with AD risk are uncertain and this has implications for intervention design. In particular, what is the role of intelligence? The degree to which education affects intelligence, versus intelligence being largely fixed in early life and acting as a determinant of educational attainment, has been debated for decades[6-10] and studies have provided evidence of an effect in both directions.[8 11] If the principal direction of causality is intelligence to educational attainment, intelligence would induce confounding bias in the association between educational attainment and AD. In this case, interventions aiming to increase educational attainment (e.g. raising the school leaving age to increase years of schooling) are unlikely to affect risk of AD, but alternative prevention strategies such as cognitive training may prove effective. In contrast, if the principal direction of causality is such that greater educational attainment increases intelligence (i.e. intelligence lies on the causal pathway from educational attainment to AD risk), then interventions designed to prolong the duration of education may reduce AD risk, either directly or indirectly through subsequently increasing intelligence.

Determining the relative contributions of education and intelligence to AD risk is of clear importance for designing appropriate policy interventions to reduce AD risk. Using observational methods to unpick these associations is challenging due to bias from measurement error, confounding and reverse causation. More recently, studies have attempted to estimate causal effects of educational attainment on AD risk using methods such as univariable Mendelian randomization (MR). MR is a form of instrumental variable analysis, in which genetic variants are used as proxies for a single environmental exposure[12]. Due to their random allocation at conception, genetic variants associated

with a particular risk factor are largely independent of potential confounders, that may otherwise bias the association of interest when using observational methods. Genetic variants also cannot be modified by subsequent disease, thereby eliminating potential bias by reverse causation. Thus, MR can be a useful tool for helping to establish whether the association between an exposure and an outcome is likely to be causal. However, these methods can be problematic with traits that are highly genetically and phenotypically correlated (such as educational attainment and intelligence)[13] [14]. Figure 1 illustrates possible models underlying the observed associations of educational attainment and intelligence with AD risk. In all models shown, causal effects for both exposures on AD risk would be implied from univariable MR analyses. However, depending on the underlying model, intervention targets will differ. Multivariable MR is an extension of univariable MR in which multiple exposures are included within the same model. It can estimate causal effects of one trait, independently of another related trait. Thus, extending MR analyses from the univariable to the multivariable setting may be a useful tool for further disentangling these relationships and establishing the respective roles of both education and intelligence in AD risk[13]. In this study, we estimated (i) the effect of educational attainment on intelligence and vice versa, (ii) the overall effects of educational attainment and intelligence on risk of AD and (iii) the independent effects of both education and intelligence on risk of AD (i.e. the effects of educational attainment and intelligence on AD risk that are independent of the other trait).

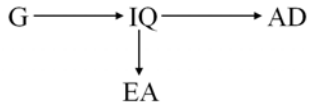
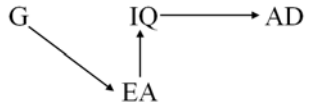
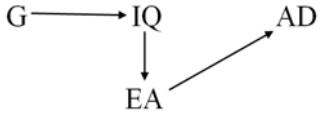
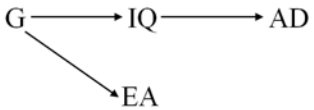
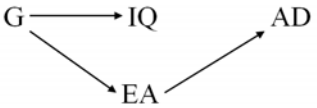
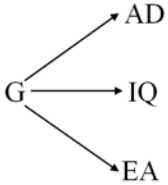
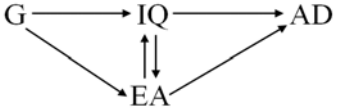
| | Causal effect implied by univariable analysis | Causal effects identified in multivariable analyses | Intervention Target |
|---|---|---|---|
| **(a) EA has no causal effect on AD but confounding by IQ causes spurious association**   | IQ EA | IQ | IQ |
| **(b) IQ has no causal effect on AD but confounding by EA causes spurious association**   | IQ EA | EA | EA |
| **(c) Causal effect of EA on AD is mediated through IQ**   | IQ EA | IQ | IQ EA |
| **(d) Causal effect of IQ on AD is mediated through EA**   | IQ EA | EA | IQ EA |
| **(e) No causal effect of EA on AD but spurious association induced by horizontal pleiotropic pathway through IQ**   | IQ EA | IQ | IQ |
| **(f) No causal effect of IQ on AD but spurious association induced by horizontal pleiotropic pathway through EA**   | IQ EA | EA | EA |
| **(g) No causal effect of EA or IQ on AD but G has an independent causal effect on all three traits**   | IQ EA | - | - |
| **(h) Joint independent causal effects of both EA and IQ on AD**   | IQ EA | IQ EA | IQ EA |

**Figure 1. A non-exhaustive list of possible models underlying the observed causal effects of educational attainment, intelligence and risk of Alzheimer's disease.** Please note that these are not intended to be directed acyclic graphs. IQ denotes intelligence. EA denotes educational attainment and AD denotes Alzheimer's Disease. G denotes a set of instruments, which are drawn as a single node for visual simplicity. Panel (a) illustrates a model in which G is identified in a genome wide association study of EA, because it is

associated with EA indirectly through IQ. IQ has an independent effect on AD but EA does not. A spurious association between EA and AD is induced due to confounding by IQ. Accounting for IQ in multivariable analysis would reveal no independent effect of EA on AD risk and the intervention target should be IQ. Panel (b) illustrates a model in which G is identified in a genome wide association study of IQ because it is associated with IQ indirectly through EA. EA has an independent effect on AD but IQ does not. A spurious association between IQ and AD is induced due to confounding by EA. Accounting for EA in multivariable analysis would reveal no independent effect of IQ on AD risk and the intervention target should be EA. Panel (c) illustrates a model in which the effect of EA on AD risk is entirely mediated by IQ (i.e. IQ lies on the causal pathway between EA and AD). Multivariable analyses would reveal an independent effect of IQ on AD risk, but no independent effect of EA. The intervention target could be either IQ or EA. Panel (d) illustrates a model in which the effect of IQ on AD risk is entirely mediated by EA (i.e. EA lies on the causal pathway between IQ and AD). Multivariable analyses would reveal an independent effect of EA on AD risk, but no independent effect of IQ. The intervention target could be either EA or IQ. Panel (e) illustrates a model in which there is full horizontal pleiotropy through IQ. Horizontal pleiotropy occurs when G has a causal effect on disease independently of its effect on the exposure. In this case, multivariate analyses would reveal an independent effect of IQ on AD risk, but no independent effect of EA and the intervention target should be IQ. Panel (f) illustrates a model in which there is full horizontal pleiotropy through EA. Multivariate analyses would reveal an independent effect of EA on AD risk, but no independent effect of IQ and the intervention target should be EA. Panel G illustrates a model in which G independently effects all three traits, but the three traits have no causal effect on each other. Multivariable analysis would show no independent effects of EA or IQ on AD risk. Panel (h) illustrates a model in which there are joint independent effects of both EA and IQ on AD risk. Multivariate analysis would show independent effects of both IQ and EA and the intervention target could be either IQ or EA. Here, the bi-directional relationship between IQ and EA does not affect the qualitative interpretation.

## METHODS

### Mendelian Randomization

MR is a form of instrumental variable analysis that uses genetic variants to proxy for environmental exposures. Two-sample MR[15] is an extension in which the effects of the genetic instrument on the exposure and on the outcome are obtained from separate genome-wide association studies (GWAS). This method is particularly useful for trying to identify early life risk factors for later life diseases like AD, because unlike in observational studies, rich longitudinal data across the whole life course (which are scarce) are not needed. MR is based on three key assumptions: (i) genetic variants must be robustly associated with the exposure of interest, (ii) genetic variants must not be associated with potential confounders of the association between the exposure and the outcome and (iii) there must be no effects of the genetic variants on the outcome, that do not go via the exposure (i.e. no horizontal pleiotropy)[16]. To-date, MR studies have typically been univariable (i.e. examining the effect of one exposure on an outcome), thereby estimating the total effect of the exposure on the outcome through all possible pathways. More recently, multivariable MR methods have been proposed to investigate the independent effects of multiple traits on an outcome. Methods for conducting a multivariable MR analysis have been published elsewhere[13 17 18].

### Data

For educational attainment, we used the GWAS (discovery and replication meta-analysis, n=293,723)[19] which identified 162 approximately independent genome-wide significant ($p<5x10^{-8}$) single nucleotide polymorphisms (SNPs) associated with years of schooling. SNP coefficients were per standard deviation (SD) units of years of schooling (SD=3.6 years). For intelligence, we used the largest (n= 248,482) and most recent iteration of the Multi-Trait Analysis of Genome-wide association studies[20], which identified 194 approximately independent ($r^2$ threshold <0.01 within a 10mb window using 1000 genomes reference panel[21]) genome-wide significant SNPs. SNP coefficients were per one SD increase in the intelligence test scores. F statistics provide an indication of instrument strength[22] and are a function of $R^2$ (how much variance in the trait is explained by the

set of genetic instruments being used), the number of instruments being used and the sample size. The F statistics for the educational attainment and intelligence instruments are 43.5 and 50.45, respectively (F>10 indicates the analysis is unlikely to suffer from weak instrument bias)[23]. For the outcome (AD) we used the large-scale GWAS of AD conducted by the International Genomics of Alzheimer's Project (IGAP, n=17,008 AD cases and 37,154 controls)[24]. SNP coefficients were log odds ratios of AD. Ethical approval was granted for each of the original GWAS studies and details can be found in the respective publications.

**Estimating the bidirectional association between intelligence and educational attainment**

After (i) excluding non-independent SNPs (ii) excluding SNPs that overlapped between the two GWAS and (iii) harmonization across both GWAS, there were 148 genome-wide significant SNPs for educational attainment and 180 for intelligence available for these analyses. Full details of the harmonization procedure are provided in the online supplement. Univariable MR was used to estimate the total effect of intelligence on educational attainment, and educational attainment on intelligence. This was done using inverse-variance-weighted (IVW) regression analysis[25]. Briefly, IVW regression is where causal effect estimates for each genetic variant are averaged using an inverse-variance weighted formula (taken from the meta-analysis literature) to provide an overall causal estimate of the exposure on the outcome[26]. In this regression, the intercept is constrained to zero, which makes the assumption of no horizontal pleiotropy. Results are presented in SD units to enable a comparison of the magnitude of effect across both exposures.

**Estimating the total and independent effects of education and intelligence on Alzheimer's disease**

There were 142 genome-wide significant SNPs for educational attainment and 185 for intelligence available for these analyses, after excluding non-independent SNPs and harmonization across both GWAS (full details of harmonization in online supplement). Univariable MR was used to estimate the total effects of both intelligence and educational attainment (separately) on risk of AD, through all possible pathways, using in an inverse-variance-weighted (IVW) regression analysis (described

above)[25]. As mentioned previously, this univariable method has been shown to yield biased effect estimates if the genetic instruments being used are non-specific for the hypothesised exposure.[13 14] Thus, to demonstrate these effects as they would be observed in a typical univariable analyses, we did not exclude the 9 SNPs that overlapped across education and intelligence GWAS. We then used multivariable MR to estimate the independent effects of both educational attainment and intelligence on risk of AD, by including both exposures within the same model[13]. After clumping the full list of SNPs from both the education and intelligence GWAS (to ensure only independent SNPs are included) and restricting to those SNPs (or proxies) found in the AD GWAS, a total of 231 SNPS were available for the multivariable MR analyses (84 for education and 156 for intelligence, 9 of which overlap between both GWAS).

**Sensitivity analyses**

Firstly, in the bidirectional analysis between educational attainment and intelligence, we endeavoured to rule out the possibility that the genetic instruments used to proxy for educational attainment are actually instruments for intelligence and vice versa (i.e. we wanted to test that the hypothesised causal direction was correct for each SNP used). To do this we performed Steiger filtering[27] for each SNP to examine whether it explains more variance in the exposure than it does in the outcome (which should be true if the hypothesised causal direction from exposure to outcome is correct). We then re-ran analyses excluding those SNPs for which there was evidence that it explained more variance in the outcome than the exposure. Secondly, to check that the SNPs do not exert a direct effect on the outcome apart from through the exposure (which would violate a key MR assumption of no horizontal pleiotropy[12]), we compared results from all univariable (both the bidirectional education on intelligence analyses and the analysis of education and intelligence on AD risk) and multivariable IVW regressions to those obtained with MR-Egger regression. In MR-Egger regression, the intercept is not constrained to zero, thus, the assumption of no horizontal pleiotropy is relaxed.[16 26 28] The estimated value of the intercept in MR-Egger regression can be interpreted as an estimate of the average pleiotropic effect across the genetic variants. An intercept term that

differs from zero is therefore indicative of horizontal pleiotropy, and the causal effect estimate obtained from an MR-Egger regression is adjusted for the degree of pleiotropy detected.[16] Full details of the MR-Egger regression analyses are provided in the online supplement. Thirdly, we conducted a leave-one-out analysis for the univariable models in which we systematically removed one SNP at a time to assess the influence of potentially pleiotropic SNPs on the causal estimates[29]. If any single SNP was invalid, there would likely be distortion in the distribution of the causal effects estimates. Fourth, in all univariable analysis, we assessed whether causal estimates from different genetic variants were comparable (i.e. heterogeneity) using Cochran's Q statistic[16]. Considerable heterogeneity would imply that the MR assumptions may not be valid for all the variants included in the analysis. Finally, funnel plots were generated to enable the visual assessment of the extent to which pleiotropy is balanced across the set of instruments used in each analysis. Symmetry in these plots provides evidence against directional pleiotropy.

## RESULTS

### Bidirectional effects of intelligence on educational attainment, and their influences on AD risk

Using 180 and 148 genetic instruments for intelligence and educational attainment, respectively (and no overlapping SNPs), we found strong evidence of causal effects both of intelligence on educational attainment, and of educational attainment on intelligence (Table 1). However, the magnitude of the effect was over two-fold greater for educational attainment on intelligence compared with intelligence on educational attainment.

197 **Table 1: Bidirectional effect of intelligence on years of schooling**

| | Causal effect estimates | | |
|---|---|---|---|
| **Total effects** | **N SNPs** | **Standardised β (95% CI)** | **P** |
| Intelligence on years of schooling | 180 | 0.51 (0.49, 0.54) | 1.77e-95 |
| Years of schooling on intelligence | 148 | 1.04 (0.99, 1.10) | 9.36e-80 |

198 SNP – single nucleotide polymorphism. β – beta coefficient. CI – confidence interval. Results are
199 interpreted per one standard deviation increase years of schooling and intelligence test scores.

200

The main IVW regression using all SNPs from the educational attainment GWAS showed that, with each SD more years of schooling (i.e. ~3.6 years), the odds of AD were, on average, 37% lower (95% CI: 23% to 49%). Per one SD higher intelligence test score, the odds of AD were, on average, 35% lower (95% CI: 25% to 43%, Figure 2 and Table C of the online supplement).

**Multivariable analysis of education and intelligence on AD**

When both intelligence and educational attainment were included within a single multivariable model, there was little evidence of an effect of educational attainment on AD risk, independent of intelligence (Figure 2 and Table C of the online supplement). There was, however, evidence that higher intelligence lowers risk of AD, independently of educational attainment. On average, after accounting for educational attainment, odds of AD were 38% lower (95% CI: 12% to 56%) per one SD higher intelligence test score Figure 2 and Table C of the online supplement).

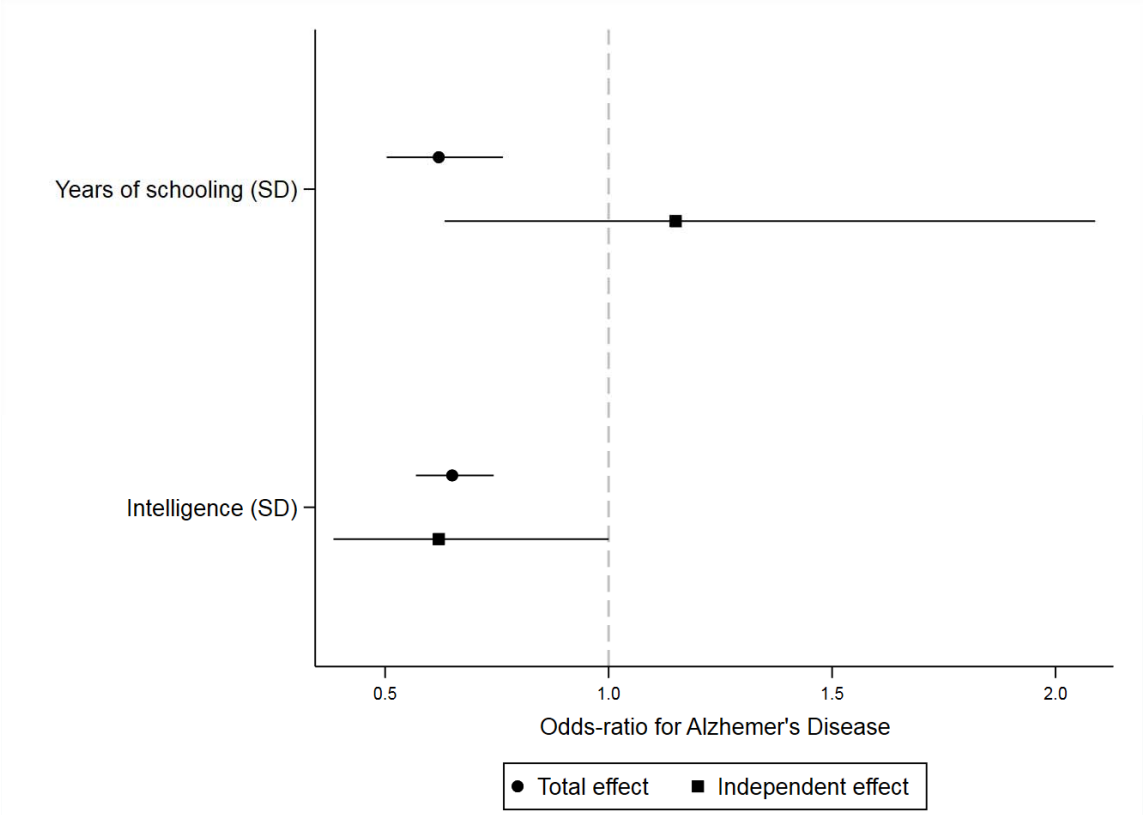



**Figure 2: Forest plot showing (i) total effect estimates for years of schooling (in standard deviations) and intelligence (in standard deviations) on odds of AD and (ii) independent effect estimates for both years of schooling and intelligence on odds of AD, when each exposure is adjusted for the other.**

**Sensitivity analyses**

The Steiger filtering provided evidence that all intelligence SNPs explained more variance in intelligence than educational attainment, suggesting they were all in the correct causal direction (i.e. from intelligence to education). However, there was evidence that 125 (85%) of the 148 education SNPs explained more variance in intelligence than educational attainment, suggesting the hypothesised causal direction is incorrect and is more likely to go from intelligence to education. This left 23 education SNPs. When using only these 23 education SNPs, there was still strong evidence of a causal effect of educational attainment on intelligence (standardised $\beta$ =0.57, 95% CI: 0.48 to 0.66, Table A of the online supplement), but the magnitude attenuated so that it was comparable to the effect of intelligence on educational attainment (as opposed to the main analysis which showed over 2-fold greater magnitude of effect for education on intelligence than vice versa). There was some evidence of horizontal pleiotropy only in the estimate of the total effect of intelligence on AD risk (Tables B and C of the online supplement). However, for all univariable and multivariable analyses (including the bidirectional effects of intelligence on educational attainment), MR-Egger effect estimates adjusting for pleiotropy were consistently comparable to those from the IWV regressions (Tables B and C of the online supplement). As expected the standard errors were much larger for MR-Egger estimates, because MR-Egger regression provides estimates of two parameters (i.e. both an intercept and a slope) compared to the single parameter in the IVW regressions (i.e. only the slope). The MR-Egger estimate for the total effect of intelligence on risk of AD went in the opposite direction to the IVW estimate (i.e. greater rather than lower odds of AD per SD increase in the intelligence score); however, the confidence intervals were very wide, and the effect estimate could plausibly go in either direction (OR: 1.36, 95% confidence interval: 0.75, 2.48). There was no distortion in the leave-one-out plots for univariable analyses (Figures A to D), suggesting that no single SNP was driving the observed effect from any analysis. There was evidence of heterogeneity in the causal effect estimates from all univariable analyses (P values for all analyses <0.02, Tables B and C of the online supplement). However, provided the pleiotropic effects of genetic variants are

242 equally likely to be positive or negative (i.e. no directional pleiotropy), the overall causal estimate
243 based on all genetic variants is likely to be unbiased and the funnel plots showed little evidence of
244 departure from symmetry (Supplemental figures E to H).

245 **DISCUSSION**

246 **Bidirectional causal effects in the relationship between of educational attainment and intelligence**

247 In this study we examined the bidirectional effects of intelligence on educational attainment. We
248 found that the relationship between intelligence and educational attainment is indeed likely to be
249 bidirectional in nature (i.e. there is evidence of an effect in both directions), with the magnitude of
250 effect being similar in both directions after filtering SNPs to check they are instrumenting the correct
251 exposure. A recent meta-analysis of quasi-experimental studies of educational effects on intelligence
252 provides evidence that support our MR findings. Across 142 effect sizes from 42 data sets involving
253 over 600,000 participants, the authors reported consistent evidence for beneficial effects of
254 education on cognitive abilities of approximately one to five IQ points (contingent on study design,
255 inclusion of moderators, and publication-bias correction) for an additional year of education[11]. These
256 findings are similar to ours in respect to magnitude of effect. Assuming a SD of 15 for IQ (as
257 described in the meta-analysis[11]), intelligence was, on average, up to one-third of a SD higher per
258 year of schooling. In our study we show an average of 0.57 SD higher in intelligence per SD (or. 3.6
259 years) increase in years of schooling, which equates to 0.16 SD higher intelligence per one additional
260 year of schooling. It is worth nothing that in the quasi-experimental policy reform studies, levels of
261 prior intelligence (or underlying general cognitive ability) will be similar among individuals who left
262 school before and after the policy reforms, making confounding by prior intelligence unlikely.
263 Similarly, in the MR analyses, we endeavoured to exclude any SNPs for education for which there
264 was evidence that they explained more variance in intelligence than education, making it unlikely
265 that our findings for the effect of education on intelligence are a result of all genetic instruments
266 being associated with intelligence and not educational attainment. Thus, both genetic and non-

267 genetic instruments (which contain different sources of bias) provide consistent evidence that
268 educational attainment affects later intelligence. The underlying mechanisms by which educational
269 attainment improves intelligence are uncertain, but several hypotheses have been proposed
270 including the teaching of material directly relevant to the intelligence tests, the training of thinking
271 styles such as abstract reasoning, and the instilling of concentration and self-control[30]. It is also
272 established that learning increases the strength of synaptic connections between neurons in grey
273 matter[31 32], and human brain imaging has revealed structural changes in white matter after learning
274 complex tasks[33 34].

275

276 Longitudinal observational studies have previously reported associations between early-life
277 intelligence and educational attainment[8]. However, we are unaware of any longitudinal studies that
278 have compared the magnitude of effect for baseline intelligence on educational attainment, with
279 educational attainment on subsequent intelligence in the same sample. One previous study has
280 examined the association between education and lifetime cognitive change after controlling for
281 childhood IQ. The authors reported that (after controlling for childhood IQ score) education was
282 positively associated with IQ at ages 70 and 79 (with the two outcome ages being in different
283 samples), and more strongly for participants with lower initial IQ scores. Education, however,
284 showed no significant association with processing speed, measured at ages 70 and 83 (again, with
285 the two ages being in different samples)[35]. Another study examined associations between father's
286 occupation, childhood cognition, educational attainment, own occupation in the 3rd decade, and
287 self-reported literacy and numeracy problems in the 4th decade in the 1946 and 1958 Birth
288 Cohorts[36]. The authors report inverse associations between childhood cognition, educational
289 attainment and adult literacy and numeracy problems. Some studies have looked at genetic overlap
290 between the two traits[20 37] and reported correlations of up to 0.7[20 38] but to date, none have
291 explicitly tried to examine the direction of the association using genetic variants that are associated

292 with each of them. As mentioned previously, the largest and most robust evidence to date comes
293 from a recent meta-analysis of quasi-experimental studies of educational effects on intelligence.[11]

294 **Effects of educational attainment and intelligence on AD risk**

295 In addition to assessing the bidirectional causal effects in the relationship between educational
296 attainment and intelligence, we also examined the total and independent effects of these traits on
297 risk of AD. Our findings imply that the existing associations reported in the literature between
298 greater educational attainment and lower AD risk are likely to be largely driven by intelligence,
299 rather than there being an independent protective effect of staying in school for longer. This
300 provides evidence against the underlying models illustrated panels (b), (d), (f) and (h) in Figure 1 (i.e.
301 models in which there is an independent effect of educational attainment on AD risk). There are
302 then four main possible explanations for our finding. The first is that prior intelligence is a
303 confounder and induces a spurious association between education and AD risk (i.e. panel (a) in
304 Figure 1). However, given the evidence supporting an effect of education on later intelligence from
305 instrumental variable analyses using policy reforms to increase the school leaving age (in which prior
306 intelligence is randomly distributed among instrument arms and thereby cannot confound), the
307 model in panel (a) is unlikely. The second and third explanations relate to horizontal pleiotropy
308 (either a pathway through IQ as in panel (e) or G independently effecting all traits as in panel (g)).
309 Given our causal effect estimates were comparable when using methods to quantify and adjust for
310 horizontal pleiotropy, these models are also unlikely to fully explain our findings. The fourth
311 explanation is that there is an effect of educational attainment on AD risk, but it is largely mediated
312 by its effects on later intelligence (i.e. panel (c)). Given the existing evidence supporting an effect of
313 education on later intelligence from quasi-experimental studies[11], and from our own MR analyses,
314 this explanation seems most plausible.
315
316 Together, these findings suggest that increasing education attainment (for example, by increasing
317 years of schooling) may have beneficial consequences for future AD incidence. As such, they offer

318 support to the most recent change in school policy in the United Kingdom (in 2013), which now
319 requires young people to remain in at least part-time education until age 18 years (as opposed to 16
320 years). Our findings also suggest that there may potentially be other ways of reducing risk of AD by
321 improving various aspects of intelligence (e.g. with cognitive training), which may be particularly
322 effective in those with lower educational attainment or in populations where increasing years of
323 schooling is not feasible (e.g. older populations). However, it is worth nothing that it is not clear
324 what type of training (if any) would be beneficial (i.e. memory tasks, abductive reasoning tasks,
325 creative tasks) or when in the life course (and indeed disease course) such training would confer
326 protection (e.g. completing training earlier in life, versus much later but prior to onset of preclinical
327 disease, versus throughout early disease stages).
328
329 Our findings are consistent with the 'brain reserve' and the 'cognitive reserve' hypotheses. Brain
330 reserve refers to structural differences in the brain itself that may increase tolerance of pathology.
331 Cognitive reserve refers to differences in the ability to tolerate and compensate for the effects of
332 brain atrophy, using pre-existing cognitive-processing approaches or compensatory mechanisms[39]. In
333 support of this, higher levels of education have been shown to be associated with whole brain and
334 ventricular volume as well as cortical thickness[40-42]. However, it is important to note that these
335 studies often do not consider the potential confounding effects of prior intelligence. One previous
336 study that examined associations between education and brain structure at 73 years found that that
337 the majority of associations observed between education and brain structure (cortical thickness in
338 bilateral temporal, medial-frontal, parietal, sensory and motor cortices) attenuated to the null after
339 accounting for childhood intelligence at age 11, and that neither education nor age 11 IQ was
340 associated with total brain atrophy or tract-averaged fractional anisotropy[43]. A post-mortem study of
341 130 elderly patients who had undergone cognitive assessment approximately 8 months before death
342 also showed that, at any given level of brain pathology, higher education was associated with better
343 cognitive function[44]. Higher educational attainment may lead to extrinsic compensation through

adaptations. Hence, more educated people will usually have occupations that are more intellectually demanding or have greater resources to partake in intellectual activities, resulting in greater cognitive stimulation and consistent with the "use it or lose it" hypothesis[45]. These compensatory mechanisms may confer protection against advancing AD pathology by increasing the time it takes for an individual to reach the threshold of cognitive impairment, whereby daily living is adversely affected, and a clinical AD diagnosis is made. In addition to compensatory mechanisms, higher education is also associated with avoidance of other potential downstream risk factors such as smoking and excessive alcohol consumption, as well as better engagement with health care systems surrounding primary and secondary prevention (e.g. uptake of and adherence to statin or anti-hypertensive medications).

**Limitations**

There are a number of limitations to our study. Firstly, in two-sample MR, "winner's curse" (i.e. where the effect sizes of variants identified within a single sample are likely to be larger than in the overall population, even if they are truly associated with the exposure) can bias causal estimates towards the null. However, we used SNPs identified in the meta-analysis of the discovery and replication samples of the educational attainment GWAS[19] making it unlikely that the estimate of the independent effect of education is biased to the null. Secondly, in the presence of weak instruments (i.e. SNPs that are not associated with the exposure at the genome-wide significance level), sample overlap in two-sample MR can bias estimates towards the confounded observational estimate[46]. There were no overlapping samples in the analysis of educational attainment and intelligence on AD risk, but there was considerable overlap in the samples used for the bidirectional educational attainment on intelligence analysis. Given that all instruments used in the analysis were strong (associated with the exposure at $p < 5 \times 10^{-08}$), any bias should be minimal. Thirdly, it is currently not possible to estimate the F statistic (a measure of instrument strength) for multivariable MR in a two- or three-sample setting. Thus, we are unable to assess the conditional strength of our instruments for each exposure, once the SNP effect on the other exposure is taken into account[13]. Fourth, the

370 estimated effect of an exposure on an outcome, that are both associated with mortality, may be
371 susceptible to survival bias.[47] For example, if individuals with lower educational attainment are more
372 likely to die before the age of onset of AD, bias may occur because those individuals with a genetic
373 predisposition for higher educational attainment are likely to live longer, thus having greater risk of
374 being diagnosed with AD. This may induce a non-zero causal effect estimate even if no true
375 biological association exists. In a previous study, we performed simulations to investigate whether
376 our estimates of the effect of educational attainment on AD risk may be biased by survival and found
377 no evidence to suggest this was the case[5]. Fifth, the phenotype used in the GWAS of intelligence was
378 typically brief (a 2-minute, 13-item test) and heterogeneous. Thus, results may be different if a
379 better phenotype of intelligence was available for GWAS studies. Finally, the educational attainment
380 GWAS only assessed years of full-time academic training from primary education through to
381 advanced qualifications (e.g. degree). Therefore, it remains unclear whether the same genetic
382 variants would be associated with other aspects of education, such as completing vocational courses
383 or completing part-time as opposed to full-time courses. It's also not clear whether education needs
384 to be completed in a formal setting (such as school or college), or whether any form of learning (e.g.
385 learning new skills 'on the job' such as in an apprenticeship during adolescence, or through career
386 development and training courses as an adult in existing full-time employment) would confer the
387 same degree of cognitive protection. This likely depends on the mechanism driving the association
388 between education and AD, thus further studies to unpick the mechanisms may help to shed light on
389 which forms of learning may confer cognitive benefits later in life and in turn, reduce AD risk.

390 **Conclusions**

391 Our findings imply that there is a bidirectional effect of intelligence on educational attainment and
392 that the magnitude of effect is likely to be similar in both directions. There is robust evidence for an
393 independent, causal effect of intelligence in reducing AD risk. The implications of this are uncertain,
394 but it potentially increases support for a role of cognitive training interventions to improve various
395 aspects of fluid intelligence. However, given that greater educational attainment also increases

396 intelligence, there is potentially also support for policies aimed at increasing length of schooling in

397 order to lower incidence of AD.

## References

1. Statistics OfN. Deaths registered in England and Wales (Series DR): 2015, 2015.
2. Meng X, D'Arcy C. Education and Dementia in the Context of the Cognitive Reserve Hypothesis: A Systematic Review with Meta-Analyses and Qualitative Analyses. *PLOS ONE* 2012;7(6):e38268. doi: 10.1371/journal.pone.0038268
3. Sharp ES, Gatz M. Relationship between education and dementia: an updated systematic review. *Alzheimer Dis Assoc Disord* 2011;25(4):289-304. doi: 10.1097/WAD.0b013e318211c83c
4. Larsson SC, Traylor M, Malik R, et al. Modifiable pathways in Alzheimer's disease: Mendelian randomisation analysis. *BMJ* 2017;359:j5375. doi: 10.1136/bmj.j5375 [published Online First: 2017/12/08]
5. Anderson EL, Wade KH, Hemani G, et al. The Causal Effect Of Educational Attainment On Alzheimer's Disease: A Two-Sample Mendelian Randomization Study. *bioRxiv* 2017
6. Sewell WH, Shah VP. Socioeconomic Status, Intelligence, and the Attainment of Higher Education. *Sociology of Education* 1967;40(1):1-23. doi: 10.2307/2112184
7. Deary IJ, Johnson W. Intelligence and education: causal perceptions drive analytic processes and therefore conclusions. *Int J Epidemiol* 2010;39(5):1362-9. doi: 10.1093/ije/dyq072 [published Online First: 2010/05/28]
8. Deary IJ, Strand S, Smith P, et al. Intelligence and educational achievement. *Intelligence* 2007;35(1):13-21. doi: https://doi.org/10.1016/j.intell.2006.02.001
9. Richards M, Sacker A. Is education causal? Yes. *International Journal of Epidemiology* 2011;40(2):516-18. doi: 10.1093/ije/dyq166
10. Singh-Manoux A. Commentary: Is it time to redefine cognitive epidemiology? *International Journal of Epidemiology* 2010;39(5):1369-71. doi: 10.1093/ije/dyq123
11. Ritchie SJ, Tucker-Drob EM. How Much Does Education Improve Intelligence? A Meta-Analysis. *Psychol Sci* 2018:956797618774253. doi: 10.1177/0956797618774253 [published Online First: 2018/06/19]
12. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet* 2014;23(R1):R89-98. doi: 10.1093/hmg/ddu328
13. Sanderson E, Davey Smith G, Windmeijer F, et al. An examination of multivariable Mendelian randomization in the single sample and two-sample summary data settings. *bioRxiv* 2018
14. Hemani G, Bowden J, Davey Smith G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum Mol Genet* 2018 doi: 10.1093/hmg/ddy163 [published Online First: 2018/05/18]
15. Lawlor DA. Commentary: Two-sample Mendelian randomization: opportunities and challenges. *Int J Epidemiol* 2016;45(3):908-15. doi: 10.1093/ije/dyw127 [published Online First: 2016/07/19]
16. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* 2015;44(2):512-25. doi: 10.1093/ije/dyv080
17. Burgess S, Thompson SG. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *Am J Epidemiol* 2015;181(4):251-60. doi: 10.1093/aje/kwu283 [published Online First: 2015/01/30]
18. Burgess S, Dudbridge F, Thompson SG. Re: "Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects". *Am J Epidemiol* 2015;181(4):290-1. doi: 10.1093/aje/kwv017 [published Online First: 2015/02/11]
19. Okbay A, Beauchamp JP, Fontana MA, et al. Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* 2016;533(7604):539-42. doi: 10.1038/nature17671
20. Hill WD, Marioni RE, Maghzian O, et al. A combined analysis of genetically correlated traits identifies 187 loci and a role for neurogenesis and myelination in intelligence. *Mol Psychiatry* 2018 doi: 10.1038/s41380-017-0001-5 [published Online First: 2018/01/13]

21. Genomes Project C, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature* 2015;526(7571):68-74. doi: 10.1038/nature15393 [published Online First: 2015/10/04]
22. Burgess S, Thompson SG, Collaboration CCG. Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol* 2011;40(3):755-64. doi: 10.1093/ije/dyr036 [published Online First: 2011/03/19]
23. Staiger D, Stock JH. Instrumental Variables Regression with Weak Instruments. *Econometrica* 1997;65(3):557-86. doi: 10.2307/2171753
24. Lambert JC, Ibrahim-Verbaas CA, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* 2013;45(12):1452-8. doi: 10.1038/ng.2802
25. Bowden J, Davey Smith G, Haycock PC, et al. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol* 2016;40(4):304-14. doi: 10.1002/gepi.21965
26. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol* 2017;32(5):377-89. doi: 10.1007/s10654-017-0255-x [published Online First: 2017/05/21]
27. Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet* 2017;13(11):e1007081. doi: 10.1371/journal.pgen.1007081 [published Online First: 2017/11/18]
28. Rees JMB, Wood AM, Burgess S. Extending the MR-Egger method for multivariable Mendelian randomization to correct for both measured and unmeasured pleiotropy. *Stat Med* 2017;36(29):4705-18. doi: 10.1002/sim.7492 [published Online First: 2017/09/30]
29. Stone M. Cross-validatory choice and assessment of statistical predictions. *J R Stat Soc B* 1974;36:111-47.
30. J. Ceci S. How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence1991.
31. Driemeyer J, Boyke J, Gaser C, et al. Changes in gray matter induced by learning--revisited. *PLoS One* 2008;3(7):e2669. doi: 10.1371/journal.pone.0002669 [published Online First: 2008/07/24]
32. Koch K, Reess TJ, Rus OG, et al. Extensive learning is associated with gray matter changes in the right hippocampus. *Neuroimage* 2016;125:627-32. doi: 10.1016/j.neuroimage.2015.10.056 [published Online First: 2015/11/01]
33. Scholz J, Klein MC, Behrens TE, et al. Training induces changes in white-matter architecture. *Nat Neurosci* 2009;12(11):1370-1. doi: 10.1038/nn.2412 [published Online First: 2009/10/13]
34. Carreiras M, Seghier ML, Baquero S, et al. An anatomical signature for literacy. *Nature* 2009;461(7266):983-6. doi: 10.1038/nature08461 [published Online First: 2009/10/16]
35. Ritchie SJ, Bates TC, Der G, et al. Education is associated with higher later life IQ scores, but not with faster cognitive processing speed. *Psychol Aging* 2013;28(2):515-21. doi: 10.1037/a0030820 [published Online First: 2013/01/02]
36. Richards M, Power C, Sacker A. Paths to literacy and numeracy problems: evidence from two British birth cohorts. *J Epidemiol Community Health* 2009;63(3):239-44. doi: 10.1136/jech.2007.064923 [published Online First: 2008/08/23]
37. Deary IJ, Spinath FM, Bates TC. Genetics of intelligence. *Eur J Hum Genet* 2006;14(6):690-700. doi: 10.1038/sj.ejhg.5201588 [published Online First: 2006/05/25]
38. Johnson W, Deary IJ, Silventoinen K, et al. Family Background Buys an Education in Minnesota but Not in Sweden. *Psychological Science* 2010;21(9):1266-73. doi: 10.1177/0956797610379233
39. Stern Y. Cognitive reserve in ageing and Alzheimer's disease. *Lancet Neurol* 2012;11(11):1006-12. doi: 10.1016/S1474-4422(12)70191-6
40. Liu Y, Julkunen V, Paajanen T, et al. Education increases reserve against Alzheimer's disease--evidence from structural MRI analysis. *Neuroradiology* 2012;54(9):929-38. doi: 10.1007/s00234-012-1005-0

41. Coffey CE, Saxton JA, Ratcliff G, et al. Relation of education to brain size in normal aging: implications for the reserve hypothesis. *Neurology* 1999;53(1):189-96.
42. Sole-Padulles C, Bartres-Faz D, Junque C, et al. Brain structure and function related to cognitive reserve variables in normal aging, mild cognitive impairment and Alzheimer's disease. *Neurobiol Aging* 2009;30(7):1114-24. doi: 10.1016/j.neurobiolaging.2007.10.008
43. Cox SR, Dickie DA, Ritchie SJ, et al. Associations between education and brain structure at age 73 years, adjusted for age 11 IQ. *Neurology* 2016;87(17):1820-26. doi: 10.1212/WNL.0000000000003247 [published Online First: 2016/10/26]
44. Bennett DA, Wilson RS, Schneider JA, et al. Education modifies the relation of AD pathology to level of cognitive function in older persons. *Neurology* 2003;60(12):1909-15.
45. Whalley LJ, Dick FD, McNeill G. A life-course approach to the aetiology of late-onset dementias. *Lancet Neurol* 2006;5(1):87-96. doi: 10.1016/S1474-4422(05)70286-6
46. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. *Genet Epidemiol* 2016;40(7):597-608. doi: 10.1002/gepi.21998
47. Hernan MA, Alonso A, Logroscino G. Cigarette smoking and dementia: potential selection bias in the elderly. *Epidemiology* 2008;19(3):448-50. doi: 10.1097/EDE.0b013e31816bbe14

**Competing interests:** The authors have no competing interests to disclose. All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

**Funding statement**

This work was supported by a grant from the UK Economic and Social Research Council (ES/M010317/1) and a grant from the BRACE Alzheimer's charity (BR16/028). Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health under Award No. R01AG048835. LDH and ELA are supported by fellowships from the UK Medical Research Council (MR/M020894/1 and MR/P014437/1, respectively). The Economics and Social Research Council (ESRC) support NMD via a Future Research Leaders grant [ES/N000757/1].GH is supported by the Wellcome Trust and the Royal Society [208806/Z/17/Z]. ELA, KHW, RKL, GH, LDH, JB, GDS and ES work in a unit that receives funding from the University of Bristol and the UK Medical Research Council (MC_UU_00011/1). KHW is funded by the Wellcome Trust Investigator Award (202802/Z/16/Z, Principal Investigator: Professor Nicholas J Timpson). The content is solely the responsibility of the authors and does not necessarily represent the official views of any of the funders.

**Author contribution statement:** ELA, NMD and GH conceptualised the study. ELA completed all statistical analyses with guidance from GH, NMD, LDH and KHW. ELA drafted the first version of the manuscript. All authors provided critical comments on the manuscript. ELA is the guarantor and accepts full responsibility for the work and/or the conduct of the study, had access to the data, and controlled the decision to publish. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

**Keywords:** Education; Dementia; Alzheimer's disease; Mendelian Randomization