# Education and myopia: a Mendelian randomisation study

Dr Edward Mountjoy[1,2], *postgraduate researcher*, Dr Neil M Davies[1,2], *research fellow*, Denis Plotnikov[3], *postgraduate researcher*, Prof George Davey Smith[1,2], *professor*, Dr Santiago Rodriguez[1,2], *senior lecturer*, Ms Cathy E Williams[2], *reader*, Prof Jeremy A Guggenheim[3†*], *professor*, Dr Denize Atan[4†*], c*onsultant senior lecturer*

**Affiliations:**

[1]MRC Integrative Epidemiology Unit, Bristol Medical School, University of Bristol, Bristol BS8 2BN

[2]Population Health Sciences, Oakfield House, Bristol Medical School, University of Bristol, Bristol BS8 2BN

[3]School of Optometry and Vision Sciences, Cardiff University, Cardiff CF24 4HQ

[4]Translational Health Sciences, Biomedical Sciences Building, Bristol Medical School, University of Bristol, Bristol BS8 1TD

***Correspondence to:**

Dr Denize Atan, Denize.Atan@bristol.ac.uk, @DenizeAtan

**Abstract**

**Objective** To determine whether more years spent in education is a causal risk factor for myopia, or myopia for more years in education.

**Design** Bidirectional, two-sample Mendelian randomisation study, using genetic variants as proxies for years spent in education and myopia to minimise bias from confounding and reverse causation. Publically available genetic data from two consortia were applied to a large, independent population cohort.

**Setting** Genetic proxies for myopia and years of education were derived from two large genome wide association studies, from 23andMe and the Social Science Genetic Association Consortium (SSGAC), respectively. Standard regression analyses of the association between myopic refractive error and years of education in UK Biobank participants were performed and compared with the results of bidirectional Mendelian randomisation analyses to determine: (1) the causal effect of education on myopia; and (2) the causal effect of myopia on education. Finally, the results were analysed for evidence of confounding.

**Participants** Adult men and women from England, Scotland and Wales in the UK Biobank cohort with available information for years of completed education and refractive error (N=67,798).

**Main exposure and outcome measures** Mendelian randomisation analyses were performed in two directions. In the first, the exposure was the genetic predisposition to myopia, measured with 44 genetic variants strongly associated with myopia in 23andMe, and the outcome was years in education. In the second, the exposure was

the genetic predisposition to higher levels of education, measured with 69 genetic variants from SSGAC, and the outcome was refractive error.

**Results** Conventional regression analyses of the observational data suggested that every additional year of education was associated with a more myopic refractive error of -0.18 Dioptres (D) per year (95% confidence intervals (CI): -0.19 to -0.17; p<2e-16). Mendelian randomisation analyses suggested the true causal effect was even stronger: -0.27 D/year (95% CI: -0.37 to -0.17; p=4e-8). By contrast, there was little evidence to suggest myopia affected education (years in education per Dioptre of refractive error = -0.008 years/D, 95% CI: -0.041 to 0.025, p=0.6). Sensitivity analyses showed minimal evidence for genetic confounding that could have biased the causal effect estimates.

**Conclusions** This study shows that exposure to more years in education contributes to the rising prevalence of myopia. Increasing the length of education may inadvertently increase myopia prevalence and potential future visual disability.

**What this study adds**

**Section 1: What is already known**

Myopia, or short-sight, is one of leading causes of visual disability in the World. The global prevalence is rising rapidly and has reached epidemic levels in the developed countries of East and Southeast Asia. For more than a century, numerous observational studies have reported strong associations between educational outcomes and myopia, but whether increasing exposure to education causes myopia, children with myopia are more studious, or indeed, an association with socioeconomic position leads to both myopia and higher levels of education was not

known with any certainty, since randomising children to different levels of education would be unethical.

**Section 2: What this study adds**

This study shows that more time spent in education is a causal risk factor for a greater level of myopia. Though an increased level of education has numerous benefits to population health and economics, this must be tempered by the rise in myopia prevalence caused by exposure to more years in education. This study highlights a need for further research and discussion about how educational practices might be improved to achieve better outcomes without adversely affecting the population's vision.

**Summary statistic:**

**For every additional year in education, there is an increase in myopic refractive error of -0.27 Dioptres/year (95% CI: -0.37 to -0.17; p=4e-8).** Thus, the cumulative effect of more years in education on refractive error means that a University graduate from the UK with 17 years of education would, on average, be at least -1D more myopic than an individual who left school at 16 (with 12 years of education). Myopia of this magnitude would be sufficient to necessitate the use of glasses for driving.

**Print abstract**

**Study question** Do more years spent in education cause increasing levels of myopia (short-sight)?

**Methods** Participants were adults from the UK Biobank cohort who had visual assessments to measure refractive error (short- or long-sight or no refractive error) and had provided information about their education in health questionnaires (N=67,798). Genetic variants that could be used as proxies for myopia and years spent in education were derived from two large genome wide association studies, 23andMe and the Social Science Genetic Association Consortium respectively. The genetic proxies were used in bidirectional Mendelian randomisation analyses to determine whether more years spent in education was a causal risk factor for myopia, or vice versa.

**Study answer and limitations** More time spent in education was a causal risk factor for myopia, whereas myopia was not a causal risk factor for more time in education. For every year in education, there was an increase in myopic refractive error of -0.27 Dioptres/year (95% CI: -0.37 to -0.17; p=4e-8). This study did not investigate how exposure to more years in education causes myopia. Instead, the results highlight the need for further research and discussion about how educational practices may be improved to achieve better outcomes without adversely affecting the population's vision.

**What this study adds** This study shows that more time spent in education is a causal risk factor for myopia (short-sight).

**Funding, competing interests, data sharing** This study was funded by the Medical Research Council, Global Education Program of the Russian Federation

116 government, National Eye Research Centre, National Institute for Health Research
117 (NIHR) and Economics and Social Research Council (ESRC). There were no
118 competing interests. Code implementing the statistical methods to analyse the data
119 are available online.

120

## Introduction

Myopia, or short-sight, is one of the leading causes of visual disability in the World, and the prevalence is rising rapidly [1 2]. Myopia is a refractive defect of the eye causing light to focus in front of, instead of on, the retina, usually because the axial length of the eye is too long. The result is that distant objects appear blurred while close objects appear clear (short-sight). The symptoms of myopia can be alleviated with spectacles, contact lenses or refractive surgery, but irrespective of visual correction, the risk of complications from potentially blinding conditions like retinal detachment, glaucoma and myopic maculopathy, increases with the longer axial lengths associated with high myopia [3-5]. Currently, 30-50% of adults in the United States and Europe are myopic, with epidemic levels of 80-90% reported in school leavers aged 17-18 in Singapore, South Korea, China and other high-income East and Southeast Asian countries [1 2 5-8], where myopic maculopathy has become one the most frequent causes of untreatable blindness [7]. Based on existing trends, the number of individuals affected by myopia worldwide is expected to rise from 1.4 billion currently, to 5 billion by 2050, affecting ~50% of the world population [7]. Almost 10% of these individuals (~9 million people) will have high myopia [7].

For more than a century, myopia has been associated with higher levels of educational attainment [9 10], but despite evidence from observational studies for an association between myopia and years of schooling or educational attainment, causal evidence for a role of education on myopia is lacking [6 11-13]. Both myopia and educational attainment have a heritable component [14-20], however genetics cannot explain the rapid rise in myopia prevalence over 1-2 generations. The current epidemic in myopia prevalence, particularly pathological and high levels of myopia, appears to be linked to an increasingly earlier age of onset and higher rate of myopia

progression in childhood[21 22] since myopia tends to remain relatively stable during adulthood (until myopic shifts occur secondary to the development of cataracts) [23]. Randomised controlled trials have demonstrated convincingly that time spent outdoors in childhood partially protects against myopia development [24 25], but the association between myopia and time spent by children doing near work activities, such as reading, is less consistent across studies [11 26]. Furthermore, the time children spend outdoors is typically independent of their near work activities, as measures of the two are generally uncorrelated [27-29]. Consequently, it is not known with any certainty whether exposure to more years in education causes myopia, children with myopia spend more time on near work leading to better educational outcomes, children with myopia are more intelligent, or indeed, an association with another confounding factor, such as socioeconomic position, leads to more years in education and myopia [6 11-13 30] since randomised trials limiting education in children would be unethical.

Mendelian randomisation (MR) is a type of instrumental variable (IV) analysis [31] that uses genetic variants associated with a risk factor, e.g. education, as proxies for an environmental exposure to make causal inferences about the impact of the exposure on the outcome of interest, e.g. myopia. It is an approach designed to reduce bias from confounding and reverse causation, to which observational epidemiology studies are susceptible. It exploits the fact that genotypes are randomly assigned at conception. Hence, Mendelian randomisation has been likened to a randomised trial by genotype, since genetic variants are not modifiable and largely free from confounding [32 33]. With the recent availability of large-scale genome-wide association study (GWAS) data for educational attainment [34] (N=293,723) and myopia [15] (N=191,843) together with the genotypes of approximately 488,000 participants in

the UK Biobank, an investigation of the causal relationship between years in education and myopia by bidirectional MR analyses [35] became possible with unprecedented statistical power. Hence, this study was able to address the question "Is more time spent in education a causal risk factor for myopia?"

## Methods

### Study cohorts

#### *1. 23andMe*

Pickrell *et al*.[15] reported the results of a GWAS for self-reported myopia in a sample of N=191,843 individuals of European descent (106,086 cases, 85,757 controls) carried out by the personal genomics company 23andMe. Myopia was ascertained by the questionnaire item, "Have you ever been diagnosed by a doctor with nearsightedness (near objects are clear, far objects are blurry)?".

#### *2. SSGAC*

Okbay *et al*.[34] reported the results of a large meta-analysis of GWAS for educational attainment in individuals of European descent (N=293,723). Educational attainment was defined as whether the participant attained a given level of schooling and was based on the International Standard Classification of Education (ISCED 1997) scale [36].

#### *3. UK Biobank*

Cross-sectional data from the baseline assessment of the UK Biobank project was collected between 2006 to 2010[37]. UK Biobank recruited 502,664 participants aged 40 to 69 years old through 22 assessment centres across the UK. Participants were genotyped using one of two platforms: the Affymetrix UK BiLEVE Axiom array or the Affymetrix UK Biobank Axiom array. All participants completed sociodemographic

questionnaires, which included questions on past educational and professional qualifications. In the latter stages of recruitment, an ophthalmic assessment was introduced, which approximately 23% of participants completed.

**(i) Definition of education**

*Time spent in education* was determined by questionnaire as defined by the question *age completed full-time education* (variable 845) in UK Biobank (N=336,826 participants completed the questionnaire at the baseline visit). The question was ascertained only for participants that did not have a college or University degree (variable 6138, answer 1). To harmonise the educational outcome measure in UK Biobank (time spent in education) with the *number of years spent in schooling* (*EduYears*) variable in the SSGAC study [34], participants with a college or University degree were coded as having left full-time education at the age of 21. Similarly, participants who reported their *age completed full-time education* was less than 15 years were assigned a value of 15 years. As schooling systems differ between countries, only participants born in England, Scotland or Wales were included in the analyses (variables 1647, 20115).

**(ii) Definition of refractive error**

Measures of visual function were not performed from the start of recruitment for UK Biobank. Consequently, only a subset of participants underwent measurements of refractive error (N=127,412). Refractive error was measured by non-cycloplegic autorefraction (Tomey RC5000 autorefractor) after removal of habitual spectacles or contact lenses. Although cycloplegic eye drops were not used (meaning that the effect of accommodation on measurements of refractive error was not controlled), only adult subjects were recruited to UK Biobank in whom the effects of accommodation would be minimal [38]. Up to 10 measurements were taken.

Measurements were excluded if the autorefractor reading was flagged as unreliable (variables 5090/5091). Spherical power (variables 5085/5084) and cylindrical power (variables 5086/5087) were averaged over repeat measurements. Mean spherical equivalent (MSE) refractive error for each eye was calculated ($Spherical\ power + 0.5 * Cylindrical\ power$). The mean of the left and right MSE (aveMSE) was taken as the participant's refractive error in Dioptres (D) and used in subsequent analyses (N=127,412). For participants with repeat measurements from separate visits (baseline visit and subsequent visits), only the baseline measurement was used. Individuals with pre-existing eye conditions that could affect refractive error were excluded from the analyses, namely: cataracts (variables 6148, 5324, 5441), refractive laser eye surgery (variable 5325), injury or trauma resulting in vision loss (variable 5419), or corneal graft surgery (variable 5328). For example, cataracts are associated with a myopic shift in refractive error. A total of 10,984 individuals with pre-existing eye conditions were excluded.

In total, 69,798 participants had valid education, refractive error and genetic data available (Figure 1).

**(iii) Genotype data**

The genetic data in UK Biobank underwent rigorous quality control procedures and was phased and imputed against a reference panel of Haplotype Reference Consortium (HRC), UK10K and 1000 Genomes Phase 3 haplotypes[39 40]. Due to an issue with the imputation of UK10K and 1000 Genomes variants, analyses were restricted to HRC variants only. Samples were excluded based on the following genotype-based criteria: non-European ancestry, relatedness, mismatch between

genetic sex and self-reported gender, putative aneuploidy (variable 22019), outlying heterozygosity, and excessive missingness (variable 22027)[39].

## Statistical analyses

### 1. Ordinary Least Squares (OLS) observational analyses

Observational associations between refractive error and years spent in education were assessed using linear regression adjusted for *sex* and *age* in UK Biobank. The regression was then repeated with adjustment for additional potentially confounding variables (for example, breastfeeding has been reported to be associated with both refractive error[41] and education[42]): *Townsend deprivation index (TDI), birth weight, whether breastfed*, and geographic coordinates of place of birth rounded to the nearest kilometre (*northing* and *easting* coordinates).

### 2. Generation of instrumental variables for Mendelian randomisation

Pickrell *et al*.[15] reported the 50 variants most strongly associated with myopia in 23andMe. Six variants (rs5022942, rs10887265, rs71041628, rs34016308, rs11658305 and rs201140091) were not in the HRC panel, leaving 44 for use as genetic instrumental variables in the MR analysis (Supplementary Data Table).

Okbay *et al*.[34] used UK Biobank as a replication cohort. Therefore, only genetic variants and summary statistics from their discovery analysis were used in this study (available at: http://ssgac.org/documents/EduYears_Discovery_5000.txt [accessed 30/03/2017]). The authors identified 74 variants associated with *educational attainment* in SSGAC. Five variants (rs9320913, rs148734725, rs544990728, rs114598875, rs8005528) were not in the HRC panel, leaving 69 variants for use as instrumental variables (Supplementary Data Table).

Multiple genetic variants were combined into a single weighted allele score for each trait. An allele score, compared to individual variants, has been shown to improve the coverage properties and reduce the bias of instrumental variable estimates[43]. Effect size estimates from the original GWAS publications were used to weight variants when constructing allele scores. Variants were harmonised with UK Biobank to ensure correct coding of the effect allele. Genotype probabilities were converted to effect allele (*a*) and non-effect allele (*A*) dosages. Allele scores were calculated by summing the product of the weights and dosages across all *n* variants:

$$Dosage = Prob(Aa) + 2 * Prob(aa)$$

$$Allele\ score = \sum_{i=1}^{n} weight_i * dosage_i$$

The proportion of variance in the phenotype variable explained by the allele score instrumental variable was calculated by regressing the phenotype on its respective allele score.

**3. Implementation of MR**

MR was implemented using the two-stage least squares method in the R package *ivpack* [44]. *Age* and *sex* were included as covariates. The strength of association in the first stage regressions between allele score and exposure were assessed with F-tests, to assess the risk of weak instrument bias [45]. Statistical power was assessed using the mRnd online calculator[46] for a Type I error level $\alpha$=0.05 (available at: http://cnsgenomics.com/shiny/mRnd/).

**4. Sensitivity analyses**

**(i) Confounding**

Confounding bias plots [47 48] were used to assess relative bias in the instrumental variable estimate compared to standard multivariable regression. Such analyses are designed to quantify the bias present in an MR analysis in a manner analogous to examining the effect of adjusting or not adjusting for a potential confounder in a standard regression analysis. Additionally, suspected confounding factors were included as covariates in supplementary analyses (Supplementary Table 4). The confounding variables considered[42 49 50] were the first 10 genetic *principal components (PC)*, *Townsend Deprivation Index (TDI)*, *birth weight*, *breastfeeding* as an infant, and place of birth (*northing* and *easting* coordinates).

**(ii) Horizontal (genetic) pleiotropy**

Two sensitivity analyses (MR-Egger and weighted median MR) were used to investigate the degree of bias in the initial MR causal estimates due to pleiotropic effects. MR-Egger is not valid for studies in which the instrumental variable-exposure and instrumental variable-outcome associations are calculated in the same sample (as was done for the main analyses in this study). Therefore, MR-Egger was run as a split sample analysis, by randomly splitting the sample in half (groups A & B). The associations of the variants and time spent in education and refractive error for each group are given in the Supplementary Data Table. MR-Egger and weighted median methods were implemented using the R package TwoSampleMR [51] (available at: github.com/MRCIEU/TwoSampleMR).

**(iii) Measurement error**

To ensure the association between time spent in education and myopia was not an artefact of the non-normal distribution of the *age completed full-time education* variable, time spent in education was recoded using two alternative methods: (1)

dichotomisation into age >16 years when completed education and age ≤16 years when completed education; and (2) excluding individuals who attended college or University. The results were compared with the original analyses using the continuous variable *age completed full-time education*.

The Durbin-Wu-Hausman (DWH) test is a method to check for the presence of endogenous variables in a regression model; the presence of such variables leads to biased effect estimates [52 53]. Effect estimates from the observational analysis and second-stage instrumental analysis were tested for endogeneity using the DWH test.

## Patient Involvement

Patients were not involved in the design or conduct of this study.

## Results

### Observational analyses: higher levels of education are associated with myopia

In agreement with previous studies [6 13 30], participants in UK Biobank who had spent longer in full-time education were more myopic; in other words, they had increasingly negative refractive errors (Table 1). The relationship was linear for those leaving full-time education between the ages of 15 to18 years, meaning that every additional year in education was associated with a higher myopic refractive error by -0.25D per year. For those leaving full-time education after the age of 18, the rate slowed to -0.10D/year (Figure 2). On average, every additional year spent in education was associated with a more myopic refractive error of -0.18D/year (95% Confidence Intervals [CI]: -0.19 to -0.17, p<2e-16). The association was largely unaffected by adjustment for measured potential confounders, including socioeconomic position

(*Townsend Deprivation Index*), *birth weight*, *breastfeeding* as an infant, and place of birth (*northing* and *easting* coordinates) (Table 1).

**Mendelian randomisation analyses: more time spent in education causes myopia**

Bidirectional MR was used to assess the causality and direction of the association between time spent in education and refractive error. Bidirectional MR analyses consist of two separate MR calculations - one in each direction. Firstly, the causal effect of education on myopia was calculated using a weighted education allele score as the instrumental variable. Secondly, the causal effect of myopia on time spent in education was calculated using a weighted myopia allele score as the instrumental variable. The allele score for time spent in education was derived from genetic variants identified by Okbay *et al.* [34] in a large meta-analysis of GWAS of individuals of European descent (N=293,723). Likewise, the allele score for myopia was derived from genetic variants reported by Pickrell *et al*.[15] in a GWAS of self-reported myopia (N=191,843).

The myopia allele score explained 4.32% (F=3155) of the variance in average mean spherical equivalent refractive error of participants in UK Biobank. The education allele score explained 0.71% (F=464) of the variance in time spent in education of participants in UK Biobank. These genetic variants were selected to use as instrumental variables because of their robust association with time spent in education and myopia, allowing us to construct strong aggregate instrumental variables for making MR inferences. The large F-statistics suggested that these analyses would not be affected by weak instrument bias.

Thus, using the allele score for time spent in education as the instrumental variable, MR analysis showed that every additional year spent in education resulted in a more myopic refractive error of -0.27 D/year (95%CI, -0.37 to -0.17, p=4e-8) (Table 2; Figure 3). The MR effect estimate was even greater in magnitude than the observational estimate (-0.27 vs. -0.18 D) suggesting that unmeasured confounders may have attenuated the latter relationship. Conversely, using the myopia allele score as the instrumental variable in MR analyses provided little evidence that refractive error affected time spent in education ($\beta_{IV}$ = -0.008 yr/D, 95% CI -0.041 to 0.025, p=0.6) (Table 2; Figure 3). With a sample size of N=69,798, there was 80% power to detect an effect of time spent in education on refractive error $\geq$0.14D/yr. In the reciprocal direction, there was 80% power to detect an effect $\geq$0.048yr/D (Supplementary Figure 1), suggesting that this study had sufficient power to detect an effect of myopia on education, if present.

**Sensitivity analyses: the results of Mendelian randomisation are robust to potential bias**

***(i) Confounding***

MR analyses are based on two pertinent assumptions: (i) the genetic instrumental variables are not associated with any confounders of the exposure-outcome relationship; and (ii) the genetic instrumental variables are only associated with the outcome via the exposure.

In tests of the association between the allele scores for time spent in education and myopia with potential confounders, there was evidence that the geographical co-ordinate, *northing* (measured northward distance in UK) was negatively associated with time spent in education ($\beta$ = -1.6e-6, 95% CI, -1.8e-6 to -1.5e-6) and positively

with refractive error ($\beta$ = 1.2e-6, 95% CI 9.8e-7 to 1.3e-6). *Northing* was also associated with the time spent in education (p=7e-5) and myopia (p=6e-3) allele scores (Supplementary Table 2). Compared to standard regression, the confounding bias plot suggested that inclusion of the *northing* variable in the instrumental variable analysis would result in a greater degree of bias for the education allele score (Figure 4A) but not for the myopia allele score (Figure 4B).

In contrast, the geographical *easting* coordinate was positively associated with time spent in education ($\beta$ = 8.9e-7, 95% CI, 6.8e-7 to 1.1e-6) and negatively associated with refractive error ($\beta$ = -1.0e-6, 95% CI, -1.3e-6 to -8.1e-6). It was weakly associated with the myopia allele score (p=0.01). However, there was little evidence to suggest a greater degree of bias in the instrumental variable analysis compared to a standard regression with the inclusion of the *easting* variable (Figure 4B). Sensitivity analyses suggested that confounding bias from the geographical coordinates had negligible impact on the MR results (Supplementary Table 4).

One further confounding variable, *population stratification principal component 9* (*PC9*), incurred a greater degree of bias in the instrumental variable regression compared to observational least squares regression. Additional analyses showed that *PC9* was associated with a self-reported place-of-birth in Wales (Supplementary Figure S2) and also with a -0.17 D (95% CI, -0.05 to -0.28) more myopic refractive error, on average (P=4e-3) compared to those who reported being born in England. An MR sensitivity analysis that adjusted for *PC1-10* provided very similar results to those prior to adjustment (Supplementary Table 4), suggesting that confounding due to *PC9* did not lead to appreciable bias.

While education legislation has not been different in England and Wales while the UK Biobank participants were in education, Scottish schools normally finish one year earlier and University degrees are correspondingly one year longer. This difference would impact on the years spent in education for Scottish individuals moving to England to attend University, and vice versa. However, the results of an MR sensitivity analysis restricted to participants born in England were essentially unchanged (Supplementary Table 4), providing evidence that imprecision in quantifying years spent in education due to differences in school leaving age did not adversely affect the results.

***(ii) Horizontal (genetic) pleiotropy***

Under the second assumption of MR, genetic variants with pleiotropic effects are invalid instrumental variables. This can be problematic when genetic variants are used without regard for the biological mechanisms through which they affect the exposure, e.g. if the genetic variants associated with more years in education also caused myopia independently of the education phenotype. MR-Egger, Weighted Mode and Weighted Median methods are alternative methods of integrating instrumental variable estimates across individual SNPs. These methods allow some of the assumptions of MR to be relaxed providing valid tests for causality despite the presence of invalid instrumental variables, e.g. due to genetic pleiotropy [54]. If the results across different MR methods are divergent, this may indicate that genetic pleiotropy is creating bias. However, all methods yielded similar causal estimates in magnitude and direction, such that increasing time spent in education led to a more myopic refractive error (by -0.17 to -0.40 D/year), while there was little evidence that a more myopic refractive error led to more time spent in education (Supplementary Table 3). With MR-Egger, a deviation of the intercept estimate from zero suggests

the existence of genetic pleiotropy, i.e. where certain genetic variants affect the outcome via a different biological pathway from the exposure under investigation. In practice, there was little evidence that the Egger intercept deviated from zero either for more time in education causing refractive error (intercept=0.007, SE=0.006, p=0.2) or refractive error causing more time in education (intercept=-0.002, SE=0.007, p=0.8), indicating that there was little evidence for directional genetic pleiotropy.

***(iii) Measurement error***

Encoding time spent in education as a dichotomous trait (>16 years vs $\leq$16 years of age when completed full-time education) produced the same pattern of causality as the continuous variable, *age completed full-time education*; i.e. more time spent in education had an effect on refractive error ($\beta_{IV}$ = -0.35 D/LOD(education) where LOD is the logarithm of odds for having spent >16 vs $\leq$16 years in education, 95% CI -0.48 to -0.22) while refractive error did not have an effect on time spent in education ($\beta_{IV}$ = -0.0004 LOD(education)/D, 95% CI -0.03 to 0.03) (Supplementary Table 4).

When individuals who had attended University or college were excluded from the analyses, there was a similar point estimate of the effect of time spent in education on refractive error ($\beta_{IV}$ = -0.23 D/yr, 95% CI -0.48 to 0.02, p=0.07) with larger standard errors. This was attributable, in part, to the reduced sample size (N=45,535). Again, there was little evidence that refractive error had an effect on time spent in education ($\beta_{IV}$ = -0.004 yr/D, 95% CI -0.04 to 0.03, p=0.8) (Supplementary Table 4).

Using the DWH test for endogeneity, there was weak evidence that the instrumental variable estimate using the time spent in education allele score differed from the

observational point estimate (DWH-p=0.06), with the instrumental variable estimate suggesting a larger negative association (Table 2). There was strong evidence that the instrumental variable estimate using the myopia allele score was a departure from the observational point estimate (DWH-p<2e-16) (Table 2).

## Discussion

### Principal findings

In this study, there was strong evidence that more time spent in education is a causal risk factor for myopia. More specifically, every additional year in education caused an increase in myopic refractive error of -0.27 Dioptres/year (95% CI: -0.37 to -0.17; p=4e-8). Thus, the cumulative effect of more years in education on refractive error means that a person attending University would be likely to have at least -1D more myopia compared to an individual who left school at age 16. A difference of this magnitude would blur vision on a Snellen visual acuity chart to 6/18 and affect the ability to drive without glasses. Individuals with myopia, by definition, have better near vision than distance vision and require less accommodative effort for near work and study, and so myopia has been proposed as an educational advantage [55]. Despite the general perception that individuals with myopia are more studious than those without myopia, there was little evidence that myopia caused people to remain in education for longer.

### Strengths and limitations of study

MR is a particularly powerful approach for testing causal hypotheses in epidemiology [56 57]. The large sample size and robustly-associated genetic instrumental variables used here meant that causal effects could be estimated with high precision. Consistent with other studies, the allele score for myopia explained only a small

fraction (4.32%) of the variance in refractive error of participants in UK Biobank [20]. Likewise, the education allele score explained only a small fraction of the variance (0.71%) in time spent in education of participants in UK Biobank [34]. However, power calculations confirmed these effects were sufficient to draw solid inferences from the MR analysis results presented here (Supplementary Figure S1). Given the ubiquity of exposure to education in populations with available genotype data, it is not possible to assess individuals who were completely free of the outcome, specifically education. Nor is it ethical to randomise children to different levels or years in education to assess the impact on refractive error. The advantage of MR is that participants are grouped based on their genotype - randomly allocated at conception and so analogous to a randomised controlled trial in which genetic variants are used as proxies for an environmental exposure to make causal inferences about the impact of the exposure on the outcome of interest. However, it Is not possible to determine exactly which components of educational practices in the last 5-7 decades have led to increases in myopic refraction using MR. Though more robust to confounding than standard observational studies, MR is not entirely immune. There was some evidence of confounding by the variables *northing* and *PC9,* of which the latter identified individuals from Wales. Although there are some differences between the education system in England & Wales compared with Scotland, the results held true when analyses were restricted to individuals from England. Another limitation of this study was selection bias. UK Biobank participants have been shown to be more highly educated, have healthier lifestyles and have fewer self-reported adverse health outcomes than expected in comparison to the general UK population [58]. This selection bias could have led to bias in both the observational and MR effect estimates [59].

When using MR, it is not necessary to know how the genetic variants used in the analysis cause the exposure. Yet, without knowing the function of the genetic variants and how they influence the traits described here, it is possible that some SNPs may influence the outcome through a pathway that does not involve the exposure, i.e. through horizontal pleiotropy. For example, educational outcomes and intelligence are highly correlated, and if intelligence caused myopia through a pathway that did not involve exposure to education, this could cause bias in the MR causal effect estimate (Supplementary Figure 3a). In contrast, vertical pleiotropy refers to SNPs that influence the outcome via an intermediate phenotype, e.g. if some SNPs affect exposure to education through their influence on intelligence. Vertical pleiotropy acting through intelligence would not bias the MR causal estimate obtained here (Supplementary Figure 3b). Sensitivity analyses (MR-Egger and mode-based MR; Supplementary Table S3) suggested little evidence of unbalanced horizontal pleiotropy in the relationship between education and myopia, though such bias cannot be ruled out unequivocally.

**Comparison with other studies**

In agreement with a substantial number of epidemiological studies dating back more than 100 years [6 13 30], the observational analyses in this study showed that more highly educated participants in UK Biobank were more myopic. The results of bidirectional MR analyses demonstrated that this association arises from exposure to factors related to education on myopia. The current epidemic of myopia in developed East and Southeast Asian countries over the last 1-2 generations appears to coincide with widening exposure to primary and secondary education, whereas educational outcomes, e.g. in scientific, reading and mathematical literacy, are less clearly associated with myopia, since many Western countries achieve top

527 international rankings in student assessments without the same high prevalence
528 rates of myopia [60]. Moreover, there are countries with poorly developed education
529 systems in which myopia prevalence is low [61-64], and hence any causal relationship
530 between intelligence and myopia is unlikely. There are other well-established
531 associations between myopia and urbanization, reduced light exposure,
532 socioeconomic position, near work and prenatal factors [21 65-67], and several of these
533 factors either confound the relationship between education and myopia or may work
534 synergistically to exacerbate the effect, e.g. in countries where myopia prevalence is
535 particularly high. Despite the robust associations between exposure to education and
536 myopia reported by many of these previous studies, they have not demonstrated
537 causality. Only one study has addressed the causal relationship between education
538 and myopia: in an MR analysis of 3 European-ancestry cohorts (combined N=5,649),
539 Cuellar-Partida *et al.*[68] reported that each year of education led to a more myopic
540 refractive error of -0.46 D/year (p=1e-3). However, the study was under-powered
541 and the authors did not investigate the possibility of horizontal genetic pleiotropy or
542 reverse causation[68]. Moreover, their methodology risked violating the key
543 assumptions of MR because they used several thousand SNPs (N=17,749) to
544 construct a polygenic risk score as an instrumental variable for their measure of
545 education. The number of SNPs used in this previous study means it was more likely
546 to include: (1) pleiotropic variants with direct effects on both exposure to education
547 and refractive error; and (2) SNPs that are in linkage disequilibrium (LD) with
548 refractive error variants. The much larger sample size in this study permitted the use
549 of a small number of strongly associated variants as instrumental variables for
550 exposure to education and refractive error. Thus, the risk of LD between the major
551 risk variants for the two traits explaining the underlying associations between

552 education and myopia was mitigated. Crucially, the analyses in this study provided
553 strong evidence that the relationship arose from a *causal* effect of exposure to
554 education on refractive error, and not via reverse causation or confounding by
555 influences such as socioeconomic position.

556 Exactly how increasing levels of education cause myopia cannot be inferred from
557 MR analyses, although the known environmental risk factors for myopia provide
558 intriguing clues. Children from developed East and Southeast Asian countries
559 consistently report that they spend less time outdoors than children from Australia or
560 the US [25 27 28 69-73] and randomised controlled trials have demonstrated that time
561 spent outdoors during childhood protects against myopia development [24 25].
562 Therefore, lack of time outdoors is a plausible mediator in the causal pathway linking
563 more time spent in education and myopia. Furthermore, engaging in higher levels of
564 near work activities, such as reading, is associated with the incidence and
565 progression of myopia, albeit less consistently than time spent outdoors [11 26 73-75].
566 Yet, measures of time spent on near work activities and time spent outdoors are
567 generally uncorrelated [27-29]. Thus, lack of time outdoors and excessive near work
568 may not be the only routes mediating the effects of exposure to education on
569 myopia. Children with myopia tend to engage in less physical activity, such as
570 sports, but physical activity per se does not appear to be protective [29 76]. Others have
571 correlated higher light exposure with lower myopia risk [66 77], and it is possible that
572 individuals who spend more years in education have less exposure to natural light.
573 The progression of myopia is faster in winter months, thus supporting the hypothesis
574 that exposure to natural light is important [78 79]. This hypothesis has been one of the
575 main drivers for recent investment in “Bright Light Classrooms” to protect against
576 myopia in South-East Asia [80]. Whether these classrooms provide any protection

against myopia that replicates the effects of increasing time spent outdoors is not currently known until the impact of this intervention has been measured. The best recommendation, based on the highest quality available evidence at the moment, is for children to spend more time outside (https://www.nhs.uk/conditions/short-sightedness/).

## Conclusions and policy implications

In summary, this study provides strong evidence that more time spent in education is a causal risk factor for myopia. With the rapid rise in the global prevalence of myopia and the economic burden of myopia and its vision-threatening complications, the findings of this study have important implications for educational practices. Axial eye growth occurs predominantly during the school years [81] and since levels of myopia tend to stabilise in adulthood [23], any interventions to halt or prevent myopia need to be delivered in childhood. Policy makers should be aware that the educational practices used to educate children and to promote personal and economic health may have the unintended consequence of causing increasing levels of myopia and later visual disability as a result.

## Competing interests

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 .

**Contributors:**

DA and JG conceived the project. EM, DP, and ND cleaned and analysed the data with input from GDS. EM wrote the first draft of the manuscript with DA and JG. DA and JG revised the draft. All authors contributed to data interpretation, critical revisions and final approval.

**Funding**

EM is funded by the Medical Research Council (MRC) and Bristol Centre for Systems Biomedicine; DP by the Global Education Program of the Russian Federation government; JAG & CW by the National Eye Research Centre (grant SAC015) and CEW by the NIHR (Senior Research Fellowship SRF-2015-08-005); NMD by the Economics and Social Research Council (ESRC) (Future Research Leaders grant [ES/N000757/1]). GDS works in the MRC Integrative Epidemiology Unit at the University of Bristol [MC_UU_12013/1, MC_UU_12013/8, MC_UU_12013/9]. The funders had no role in the study design, data collection, analysis, interpretation, writing, nor the decision to submit the article for publication.

**Data sharing**

Participants consented to data sharing as described in the population cohorts and are not identifiable in these analyses.

The UK Biobank Resource and Access Committee approved this research (application #8786). Anonymised phenotype and genetic data are available from UK Biobank on application. Additional ethical approval was not required.

Code implementing the statistical methods to analyse the data are available from https://github.com/edm1/myopia-education-MR.

**Transparency**

The corresponding author (DA) is guarantor of the paper and affirms that the manuscript is an honest, accurate and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

## References


1. Vitale S, Sperduto RD, Ferris FL. Increased prevalence of myopia in the United States between 1971-1972 and 1999-2004. *Archives of Ophthalmology* 2009;127(12):1632-39. doi: 10.1001/archophthalmol.2009.303

2. Williams KM, Verhoeven VJ, Cumberland P, et al. Prevalence of refractive error in Europe: the European Eye Epidemiology (E3) Consortium. *European Journal of Epidemiology* 2015;30:305–15. doi: 10.1007/s10654-015-0010-0

3. Verhoeven VJ, Wong KT, Buitendijk GH, et al. Visual consequences of refractive errors in the general population. *Ophthalmology* 2014;122(1):101–09. doi: 10.1016/j.ophtha.2014.07.030 [published Online First: 2014/09/12]

4. Flitcroft DI. The complex interactions of retinal, optical and environmental factors in myopia aetiology. *Progress in Retinal and Eye Research* 2012;31(6):622-60. doi: 10.1016/j.preteyeres.2012.06.004

5. Morgan IG, Ohno-Matsui K, Saw SM. Myopia. *Lancet* 2012;379(9827):1739-48. doi: 10.1016/S0140-6736(12)60272-4

6. Jung SK, Lee JH, Kakizaki H, et al. Prevalence of myopia and its association with body stature and educational level in 19-year-old male conscripts in seoul, South Korea. *Investigative Ophthalmology and Visual Science* 2012;53(9):5579-83. doi: 10.1167/iovs.12-10106 [published Online First: 2012/07/28]

7. Holden BA, Fricke TR, Wilson DA, et al. Global Prevalence of Myopia and High Myopia and Temporal Trends from 2000 through 2050. *Ophthalmology* 2016;123(5):1036–42. doi: 10.1016/j.ophtha.2016.01.006

8. Vitale S, Ellwein L, Cotch MF, et al. Prevalence of refractive error in the United States, 1999-2004. *Archives of Ophthalmology* 2008;126(8):1111-19.

9. Harman NB. The Education of High Myopes. *Journal of the Royal Society of Medicine* 1913;6(Sect Ophthalmol):146-63. doi: 10.1177/003591571300601644

10. Holm E. Investigations of myopia in Danish secondary schools. *Acta Ophthalmologica* 1925;3(1-2):121-30. doi: 10.1111/j.1755-3768.1925.tb03235.x

11. Morgan I, Rose K. How genetic is school myopia? *Progress in Retinal and Eye Research* 2005;24(1):1-38. doi: 10.1016/j.preteyeres.2004.06.004

12. Zylbermann R, Landau D, Berson D. The influence of study habits on myopia in Jewish teenagers. *Journal of Pediatric Ophthalmology and Strabismus* 1993;30:319-22. doi: 10.3928/0191-3913-19930901-12

13. Au Eong KG, Tay TH, Lim MK. Education and myopia in 110,236 young Singaporean males. *Singapore Medical Journal* 1993;34:489-92.

14. Kiefer AK, Tung JY, Do CB, et al. Genome-wide analysis points to roles for extracellular matrix remodeling, the visual cycle, and neuronal development in myopia. *PLoS Genetics* 2013;9(2):e1003299. doi: 10.1371/journal.pgen.1003299

15. Pickrell JK, Berisa T, Liu JZ, et al. Detection and interpretation of shared genetic influences on 42 human traits. *Nature Genetics* 2016;48(7):709-17. doi: 10.1038/ng.3570

16. Verhoeven VJM, Hysi PG, Wojciechowski R, et al. Genome-wide meta-analyses of multiancestry cohorts identify multiple new susceptibility loci for refractive error and myopia. *Nature Genetics* 2013;45(3):314-18. doi: 10.1038/ng.2554

678 17. Hammond CJ, Snieder H, Gilbert CE, et al. Genes and environment in refractive
679 error: The twin eye study. *Investigative Ophthalmology and Visual Science*
680 2001;42(6):1232-36.
681 18. Lopes MC, Andrew T, Carbonaro F, et al. Estimating heritability and shared
682 environmental effects for refractive error in twin and family studies.
683 *Investigative Ophthalmology and Visual Science* 2009;50:126-31.
684 19. Deary IJ, Johnson W, Houlihan LM. Genetic foundations of human intelligence.
685 *Human Genetics* 2009;126(1):215-32. doi: 10.1007/s00439-009-0655-4
686 20. Williams KM, Hysi PG, Yonova-Doing E, et al. Phenotypic and genotypic
687 correlation between myopia and intelligence. *Scientific Reports* 2017;7:45977.
688 doi: 10.1038/srep45977
689 21. Morgan IG, French AN, Ashby RS, et al. The epidemics of myopia: Aetiology and
690 prevention. *Progress in Retinal and Eye Research* 2018;62:134-49. doi:
691 10.1016/j.preteyeres.2017.09.004
692 22. Morgan IG, He M, Rose KA. Epidemic of Pathologic Myopia: What Can
693 Laboratory Studies and Epidemiology Tell Us? *Retina* 2016;(in press) doi:
694 10.1097/IAE.0000000000001272
695 23. Weale RA. Epidemiology of refractive errors and presbyopia. *Survey of*
696 *Ophthalmology* 2003;48(5):515-43.
697 24. Wu PC, Tsai CL, Wu HL, et al. Outdoor activity during class recess reduces
698 myopia onset and progression in school children. *Ophthalmology*
699 2013;120(5):1080-85. doi: 10.1016/j.ophtha.2012.11.009
700 25. He M, Xiang F, Zeng Y, et al. Effect of Time Spent Outdoors at School on the
701 Development of Myopia Among Children in China: A Randomized Clinical
702 Trial. *JAMA* 2015;314(11):1142-48. doi: 10.1001/jama.2015.10803

703 26. Guggenheim JA, St Pourcain B, McMahon G, et al. Assumption-free estimation
704 of the genetic contribution to refractive error across childhood. *Molecular*
705 *Vision* 2015;21:621-32.
706 27. Jones LA, Sinnott LT, Mutti DO, et al. Parental history of myopia, sports and
707 outdoor activities, and future myopia. *Investigative Ophthalmology and Visual*
708 *Science* 2007;48(8):3524-32. doi: 10.1167/iovs.06-1118
709 28. Rose KA, Morgan IG, Ip J, et al. Outdoor activity reduces the prevalence of
710 myopia in children. *Ophthalmology* 2008;115(8):1279-85. doi:
711 10.1016/j.ophtha.2007.12.019
712 29. Guggenheim JA, Northstone K, McMahon G, et al. Time outdoors and physical
713 activity as predictors of incident myopia in childhood: A prospective cohort
714 study. *Investigative Ophthalmology and Visual Science* 2012;53:2856-65. doi:
715 10.1167/iovs.11-9091
716 30. Rosner M, Belkin M. Intelligence, education and myopia in males. *Archives of*
717 *Ophthalmology* 1987;105:1508-11. doi:
718 10.1001/archopht.1987.01060110054030
719 31. Palmer TM, Sterne JAC, Harbord RM, et al. Instrumental Variable estimation of
720 causal risk ratios and causal odds ratios in Mendelian randomization
721 analyses. *American Journal of Epidemiology* 2011;173(12):1392-403. doi:
722 10.1093/aje/kwr026
723 32. Davey Smith G, Ebrahim S. ‘Mendelian randomization’: can genetic
724 epidemiology contribute to understanding environmental determinants of
725 disease? *International Journal of Epidemiology* 2003;32(1):1-22. doi:
726 10.1093/ije/dyg070

727 33. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal
728 inference in epidemiological studies. *Human Molecular Genetics*
729 2014;23:R89-R98. doi: 10.1093/hmg/ddu328
730 34. Okbay A, Beauchamp JP, Fontana MA, et al. Genome-wide association study
731 identifies 74 loci associated with educational attainment. *Nature*
732 2016;533(7604):539–42. doi: 10.1038/nature17671
733 35. Timpson NJ, Nordestgaard BG, Harbord RM, et al. C-reactive protein levels and
734 body mass index: elucidating direction of causation through reciprocal
735 Mendelian randomization. *Int J Obes (Lond)* 2011;35(2):300-8. doi:
736 10.1038/ijo.2010.137 [published Online First: 2010/08/18]
737 36. UNESCO. International Standard Classification of Education ISCED 1997, 2006.
738 <http://www.unesco.org/education/information/nfsunesco/doc/isced_1997.htm
739 >.
740 37. Sudlow C, Gallacher J, Allen N, et al. UK Biobank: An Open Access Resource
741 for Identifying the Causes of a Wide Range of Complex Diseases of Middle
742 and Old Age. *PLoS Medicine* 2015;12(3):e1001779. doi:
743 10.1371/journal.pmed.1001779
744 38. Sanfilippo PG, Chu BS, Bigault O, et al. What is the appropriate age cut-off for
745 cycloplegia in refraction? *Acta Ophthalmol* 2014;92(6):e458–e62. doi:
746 10.1111/aos.12388 [published Online First: 2014/03/20]
747 39. UK Biobank. Genotype imputation and genetic association studies of UK
748 Biobank. May 2015. 2015:www.ukbiobank.ac.uk/wp-
749 content/uploads/2014/04/UKBiobank_genotyping_QC_documentation-
750 web.pdf.

751 40. Mitchell R, Hemani G, Dudding T, et al. UK Biobank Genetic Data: MRC-IEU
752 Quality Control, Version 1: University of Bristol, 2017.
753 41. Sham WK, Dirani M, Chong YS, et al. Breastfeeding and association with
754 refractive error in young Singapore Chinese children. *Eye* 2009
755 42. Victora CG, Horta BL, de Mola CL, et al. Association between breastfeeding and
756 intelligence, educational attainment, and income at 30 years of age: a
757 prospective birth cohort study from Brazil. *The Lancet Global Health*
758 2015;3(4):e199-e205. doi: https://doi.org/10.1016/S2214-109X(15)70002-1
759 43. Burgess S, Scott RA, Timpson NJ, et al. Using published data in Mendelian
760 randomization: a blueprint for efficient identification of causal risk factors.
761 *European Journal of Epidemiology* 2015;30(7):543-52. doi: 10.1007/s10654-
762 015-0011-z
763 44. Jiang Y, Small D. ivpack: Instrumental Variable Estimation. R package ver. 1.2.
764 URL: CRAN.R-project.org/package=ivpack. 2014
765 45. Davies NM, Scholder SvHK, Farbmacher H, et al. The many weak instruments
766 problem and Mendelian randomization. *Statistics in Medicine* 2015;34(3):454-
767 68. doi: 10.1002/sim.6358
768 46. Brion M-JA, Shakhbazov K, Visscher PM. Calculating statistical power in
769 Mendelian randomization studies. *International Journal of Epidemiology*
770 2013;42(5):1497-501. doi: 10.1093/ije/dyt179
771 47. Davies NM. An even clearer portrait of bias in observational studies?
772 *Epidemiology* 2015;26(4):505-08. doi: 10.1097/EDE.0000000000000302
773 48. Jackson JW, Swanson SA. Toward a Clearer Portrayal of Confounding Bias in
774 Instrumental Variable Applications. *Epidemiology* 2015;26(4):498-504. doi:
775 10.1097/EDE.0000000000000287

776 49. White KR. The relation between socioeconomic status and academic
777 achievement. *Psychological Bulletin* 1982;91(3):461-81.
778 50. Sirin SR. Socioeconomic Status and Academic Achievement: A Meta-Analytic
779 Review of Research. *Review of Educational Research* 2005;75(3):417-53.
780 doi: 10.3102/00346543075003417
781 51. Hemani G, Zheng J, Wade KH, et al. MR-Base: a platform for systematic causal
782 inference across the phenome using billions of genetic associations. *bioRxiv*
783 2016:078972 doi:10.1101/72. doi: 10.1101/078972
784 52. Nakamura A, Nakamura M. On the performance of tests by Wu and by Hausman
785 for detecting the ordinary least squares bias problem. *Journal of Econometrics*
786 1985;29(3):213-27. doi: 10.1016/0304-4076(85)90153-8
787 53. Staiger D, Stock JH. Instrumental Variables Regression with Weak Instruments.
788 *Econometrica* 1997;65:557-86 doi: 10.3386/t0151
789 54. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid
790 instruments: effect estimation and bias detection through Egger regression.
791 *International Journal of Epidemiology* 2015;44(2):512-25. doi:
792 10.1093/ije/dyv080
793 55. Hirsch MJ. The relationship between refractive state of the eye and intelligence
794 test scores. *Optometry and Vision Science* 1959;36(1):12-21.
795 56. Davey Smith G, Ebrahim S. What can mendelian randomisation tell us about
796 modifiable behavioural and environmental exposures? *BMJ*
797 2005;330(7499):1076-9. doi: 10.1136/bmj.330.7499.1076
798 57. VanderWeele TJ, Tchetgen EJT, Cornelis M, et al. Methodological Challenges in
799 Mendelian Randomization. *Epidemiology* 2014;25(3):427-35. doi:
800 10.1097/ede.0000000000000081

801 58. Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of Sociodemographic and
802 Health-Related Characteristics of UK Biobank Participants With Those of the
803 General Population. *Am J Epidemiol* 2017;186(9):1026-34. doi:
804 10.1093/aje/kwx246 [published Online First: 2017/06/24]
805 59. Munafo M, Davey Smith G. Biased Estimates in Mendelian Randomization
806 Studies Conducted in Unrepresentative Samples. *JAMA Cardiol* 2017 doi:
807 10.1001/jamacardio.2017.4279 [published Online First: 2017/12/15]
808 60. Morgan IG, Rose KA. Myopia and international educational performance.
809 *Ophthalmic and Physiological Optics* 2013;33(3):329-38. doi:
810 10.1111/opo.12040 [published Online First: 2013/05/15]
811 61. Gao Z, Meng N, Muecke J, et al. Refractive error in school children in an urban
812 and rural setting in Cambodia. *Ophthalmic Epidemiology* 2012;19(1):16-22.
813 62. Garner LF, Kinnear RF, Klinger JD, et al. Prevalence of myopia in school
814 children in Vanuatu. *Acta Ophthalmologica* 1985;63:323-26. doi:
815 10.1111/j.1755-3768.1985.tb06814.x
816 63. Garner LF, Owens H, Kinnear RF, et al. Prevalence of myopia in Sherpa and
817 Tibetan children in Nepal. *Optometry and Vision Science* 1999;76(5):282-85.
818 64. Casson RJ, Kahawita S, Kong A, et al. Exceptionally Low Prevalence of
819 Refractive Error and Visual Impairment in Schoolchildren from Lao People's
820 Democratic Republic. *Ophthalmology* 2012;119(10):2021-7. doi:
821 10.1016/j.ophtha.2012.03.049 [published Online First: 2012/06/16]
822 65. Williams KM, Bentham GC, Young IS, et al. Association Between Myopia,
823 Ultraviolet B Radiation Exposure, Serum Vitamin D Concentrations, and
824 Genetic Polymorphisms in Vitamin D Metabolic Pathways in a Multicountry

825 European Study. *JAMA Ophthalmology* 2016 doi:
826 10.1001/jamaophthalmol.2016.4752
827 66. Hua W-J, Jin J-X, Wu X-Y, et al. Elevated light levels in schools have a
828 protective effect on myopia. *Ophthalmic and Physiological Optics*
829 2015;35(3):252-62. doi: 10.1111/opo.12207
830 67. Rahi JS, Cumberland PM, Peckham CS. Myopia over the lifecourse: Prevalence
831 and early life influences in the 1958 British Birth Cohort. *Ophthalmology*
832 2011;118(5):797-804.
833 68. Cuellar-Partida G, Lu Y, Kho PF, et al. Assessing the Genetic Predisposition of
834 Education on Myopia: A Mendelian Randomization Study. *Genetic*
835 *Epidemiology* 2016;40:66-72. doi: 10.1002/gepi.21936
836 69. French AN, Ashby RS, Morgan IG, et al. Time outdoors and the prevention of
837 myopia. *Experimental Eye Research* 2013;114:58–68. doi:
838 10.1016/j.exer.2013.04.018 [published Online First: 2013/05/07]
839 70. French AN, Morgan IG, Burlutsky G, et al. Prevalence and 5- to 6-year incidence
840 and progression of myopia and hyperopia in Australian schoolchildren.
841 *Ophthalmology* 2013;120(7):1482–91. doi: 10.1016/j.ophtha.2012.12.018
842 [published Online First: 2013/03/26]
843 71. French AN, Morgan IG, Mitchell P, et al. Risk factors for incident myopia in
844 Australian schoolchildren: The Sydney Adolescent Vascular and Eye Study.
845 *Ophthalmology* 2013;120(10):2100–08. doi: 10.1016/j.ophtha.2013.02.035
846 [published Online First: 2013/05/16]
847 72. Rose KA, Morgan IG, Smith W, et al. Myopia, lifestyle, and schooling in students
848 of Chinese ethnicity in Singapore and Sydney. *Archives of Ophthalmology*
849 2008;126(4):527-30.

850 73. Jones-Jordan LA, Sinnott LT, Cotter SA, et al. Time outdoors, visual activity, and
851 myopia progression in juvenile-onset myopes. *Investigative Ophthalmology*
852 *and Visual Science* 2012;53(11):7169-75. doi: 10.1167/iovs.11-8336
853 [published Online First: 2012/09/15]
854 74. Saw SM, Tan SB, Fung D, et al. IQ and the association with myopia in children.
855 *Investigative Ophthalmology and Visual Science* 2004;45(9):2943-48. doi:
856 10.1167/iovs.03-1296
857 75. Mutti DO, Zadnik K. Has Near Work's Star Fallen? *Optometry and Vision*
858 *Science* 2009;86(2):76-78.
859 76. Deere K, Williams C, Leary S, et al. Myopia and later physical activity in
860 adolescence: a prospective study. *British Journal of Sports Medicine*
861 2009;43(7):542-24. doi: 10.1136/bjsm.2008.049288
862 77. Williams KM, Bentham GC, Young IS, et al. Association Between Myopia,
863 Ultraviolet B Radiation Exposure, Serum Vitamin D Concentrations, and
864 Genetic Polymorphisms in Vitamin D Metabolic Pathways in a Multicountry
865 European Study. *JAMA Ophthalmology* 2017;135(1):47-53. doi:
866 10.1001/jamaophthalmol.2016.4752
867 78. Cui D, Trier K, Munk Ribel-Madsen S. Effect of day length on eye growth, myopia
868 progression, and change of corneal power in myopic children. *Ophthalmology*
869 2013;120(5):1074-79. doi: 10.1016/j.ophtha.2012.10.022 [published Online
870 First: 2013/02/06]
871 79. Gwiazda JE, Deng L, Manny RE, et al. Seasonal variations in the progression of
872 myopia in children enrolled in the Correction of Myopia Evaluation Trial.
873 *Investigative Ophthalmology and Visual Science* 2014;55:752-58. doi:
874 10.1167/iovs.13-13029 [published Online First: 2014/01/11]

80. Zhou Z, Chen T, Wang M, et al. Pilot study of a novel classroom designed to prevent myopia by increasing children's exposure to outdoor light. *PloS ONE* 2017;12(7):e0181772. doi: 10.1371/journal.pone.0181772

81. Wallman J, Turkel J, Trachtman J. Extreme myopia produced by modest changes in early visual experience. *Science* 1978;201:1249-51. doi: 10.1126/science.694514

## Figures

**Figure 1.** Numbers of participants in UK Biobank that passed validation for the MR study. MSE = Mean Spherical Error.

(figure_1.pdf)

**Figure 2. Observational association between age completed full-time education and refractive error.**

Graph of refractive error (y-axis) in Dioptres (D) against age completed full-time education (x-axis) for 69,798 individuals in UK Biobank with 95% Confidence Intervals (CI). On average, more educated individuals had higher levels of myopia (more negative refractive error).

(figure_2.png)

**Figure 3. Results of bidirectional Mendelian randomisation show that higher levels of education cause higher levels of myopia (more negative refractive error) but myopia does not cause higher levels of education in UK Biobank.**

**(A)** 69 variants associated with educational attainment in Okbay *et al.* (2016) were linked to higher levels of myopia (more negative Mean Spherical Error [MSE]) in UK Biobank (UKB) with 95% confidence intervals (w 95%CI); **(B)** 44 variants associated with myopia (more negative MSE) in Pickrell *et al.* (2016) were not linked with more time spent in education in UK Biobank (UKB) with 95% confidence intervals (w 95%CI). Regression line and standard errors fitted using robust linear regression.

(figure_3.png)

**Figure 4. Confounding bias plots.**

Plots showing the relative bias in the instrumental variable estimate (blue) and standard multivariable regression estimate (red) from potential confounders including: place of birth (*northing* and *easting* coordinates), *Townsend Deprivation Index*, *age, sex, breastfeeding, birth weight*, and the first 10 genetic *principal components (PC)*, when **(A)** estimating the effect of time spent in education on refractive error; and **(B)** estimating the effect of refractive error on time spent in education. *Townsend deprivation index* (TDI) was natural log-transformed.



(figure_4.pdf)

## Tables

**Table 1. Observational association between time spent in education and refractive error.**

*Time spent in education* was defined as the age full time education was completed in years (yrs) and *refractive error* was defined as the average measured mean spherical equivalent refractive error in Dioptres (D). Model A included *sex* and *age* as covariates. Model B included *age, sex, Townsend Deprivation Index (TDI), birth weight, whether breastfed, northing* and *easting* coordinates.

| | | **Model A** | | | **Model B** | | |
|---|---|---|---|---|---|---|---|
| **Exposure** | **Outcome** | **N** | **Effect size** | **p-value** | **N** | **Effect size** | **p-value** |
| Time spent in education (years) | Refractive error (Dioptres) | 69,798 | -0.178 (-0.185 to -0.170) D/yr | <2e-16 | 37,734 | -0.165 (-0.179 to -0.154) D/yr | <2e-16 |
| Refractive error (Dioptres) | Time spent in education (years) | 69,798 | -0.154 (-0.161 to -0.147) yr/D | <2e-16 | 37,734 | -0.136 (-0.145 to -0.128) yr/D | <2e-16 |

**Table 2. Causal association between time spent in education and refractive error.**

Results of conventional multivariable linear regression and bidirectional MR. All regressions included *age* and *sex* as covariates.

Abbreviation: DWH =Durbin-Wu-Hausman test for endogeneity.

| **Exposure** | **Outcome** | **N** | **Observational estimate (OLS)** | **MR regression** | | | |
|---|---|---|---|---|---|---|---|
| | | | **Effect size** | **Partial $R^2$** | **p-value (DWH)** | **Effect size** | **p-value** |
| **Time spent in education (years [yrs])** | **Refractive error (Dioptres [D])** | 69,798 | -0.178 (-0.185 to -0.170) D/yr | 0.71% | 0.06 | -0.270 (-0.368 to -0.173) D/yr | 4e-8 |
| **Refractive error (Dioptres [D])** | **Time spent in education (years [yrs])** | 69,798 | -0.154 (-0.161 to -0.147) yr/D | 4.32% | <2e-16 | -0.008 (-0.041 to 0.025) yr/D | 0.6 |

## Supplementary Figures and Tables

**Supplementary Figure 1. Mendelian randomisation power calculation for the available sample size in UK Biobank (N=69,798).** Causal association of: **(A)** time spent in education (years [yrs]) on refractive error (Dioptres [D]) using allele score with $R^2$=0.0073 (80% power = 0.14 D/yr); and **(B)** refractive error on time spent in education using allele score with $R^2$=0.0442 (80% power = 0.048 yr/D).

(figure_S1.png)

**Supplementary Figure 2. Genetic principal component 9 (PC9) is associated with self-reported place-of-birth in Wales and with refractive error. (A)** Scatter plot of geographic coordinates for place-of-birth vs. PC9 level. **(B)** Refractive error distribution in Dioptres (D) by PC9 level category. White rectangle shows inter-quartile range. Black circle shows median.

(figure_S2.png)

**Supplementary Figure 3. Schematic representations of horizontal and vertical pleiotropy in the relationship between the instrumental variables, education and IQ. (A)** Horizontal pleiotropy could lead to bias in the MR causal estimate, but only if there is a causal path between IQ and myopia. In communities with no formal education system there is a very low prevalence of myopia, making such a causal path between IQ and myopia unlikely. **(B)** Vertical pleiotropy would not lead to bias in the MR causal estimate.

(figure_S3.pdf)

**Supplementary Table 1. Baseline characteristics of UK Biobank cohort for refractive error, age completed education and potential confounder variables.**

Refractive errors (in Dioptres) were defined as: (i) myopia ≤ 0.75; (ii) -0.75 < emmetropia < 0.75; (iii) hypermetropia ≥ 0.75.

*Townsend deprivation index* (TDI) was natural log-transformed.

| **Variable** | **All** (N=69,798) | | **Myopic** (N=21,055) | | **Emmetropic** (N=25,209) | | **Hypermetropic** (N=23,534) | |
|---|---|---|---|---|---|---|---|---|
| | **Mean or %** | **SD** | **Mean or %** | **SD** | **Mean or %** | **SD** | **Mean or %** | **SD** |
| Refractive error (D) | -0.271 (-0.291 to -0.251) | 2.679 | -3.329 (-3.361 to -3.297) | 2.385 | 0.071 (0.066 to 0.076) | 0.400 | 2.099 (2.081 to 2.117) | 1.385 |
| Age completed education (years) | 18.15 (18.13 to 18.17) | 2.50 | 18.80 (18.76 to 18.83) | 2.39 | 18.08 (18.05 to 18.11) | 2.46 | 17.66 (17.62 to 17.69) | 2.51 |
| Age (years) | 57.09 (57.03 to 57.14) | 7.81 | 55.97 (55.86 to 56.07) | 7.70 | 55.14 (55.04 to 55.24) | 8.10 | 60.18 (60.09 to 60.26) | 6.56 |
| Sex | | | | | | | | |
| - Male | 47.2% | - | 47.4% | - | 48.6% | - | 45.5% | - |
| - Female | 52.8% | - | 52.6% | - | 51.4% | - | 54.5% | - |
| Birth weight | 3.325 (3.320 to 3.330) | 0.637 | 3.310 (3.302 to 3.318) | 0.622 | 3.330 (3.322 to 3.338) | 0.628 | 3.332 (3.324 to 3.340) | 0.662 |
| Breastfed | | | | | | | | |
| - Yes | 78.6% | - | 78.3% | - | 87.9% | - | 81.6% | - |
| - No | 20.4% | - | 20.7% | - | 22.1% | - | 18.4% | - |
| Townsend deprivation index | -0.012 (-0.016 to -0.008) | 0.495 | -0.012 (-0.019 to -0.005) | 0.493 | -0.005 (-0.011 to 0.001) | 0.494 | -0.019 (-0.025 to -0.013) | 0.499 |
| Northing | 3.016e+5 (3.007e+5 to 3.024e+5) | 1.135e+5 | 2.955e+5 (2.030e+5 to 2.971e+5) | 1.162e+5 | 3.015e+5 (3.001e+5 to 3.029e+5) | 1.124e+5 | 3.071e+5 (3.057e+5 to 3.085e+5) | 1.121e+5 |
| Easting | 4.279e+5 (4.272e+5 to 4.286e+5) | 8.911e+4 | 4.318e+5 (4.305e+5 to 4.330e+5) | 9.089e+4 | 4.318e+5 (4.305e+5 to 4.330e+5) | 9.089e+4 | 4.224e+5 (4.233e+5 to 4.256e+5) | 8.869e+4 |

**Supplementary Table 2. Observational (OLS) associations of confounding variables with education and myopia allele scores and outcomes.**

Only associations that are concordant with the observational direction of effect between time spent in education and refractive error are shown. These include *northing* and *easting* geographical coordinates, and genetic population stratification principal components 8 and 9.

| Exposure | Outcome | N | Beta | SE | P-value |
|---|---|---|---|---|---|
| Easting | Refractive error | 69797 | -1.0e-6 (-1.3e-6 to -8.1e-7) | 1.1e-7 | <2e-16 |
| Easting | Education | 69797 | 8.9e-7 (6.8e-7 to 1.1e-6) | 1.1e-7 | <2e-16 |
| Easting | Myopia allele score | 69797 | 3.8e-8 (8.6e-9 to 6.7e-8) | 1.5e-8 | 1e-2 |
| Easting | Education allele score | 69797 | 6.1e-9 (-2.2e-9 to 1.4e-8) | 4.2e-9 | 2e-1 |
| Northing | Education | 69797 | -1.6e-6 (-1.8e-6 to -1.5e-6) | 8.3e-8 | <2e-16 |
| Northing | Refractive error | 69797 | 1.2e-6 (9.8e-7 to 1.3e-6) | 8.9e-8 | <2e-16 |
| Northing | Education allele score | 69797 | -1.3e-8 (-2.0e-8 to -6.7e-9) | 3.3e-9 | 7e-5 |
| Northing | Myopia allele score | 69797 | -3.2e-8 (-5.5e-8 to -9.2e-9) | 1.2e-8 | 6e-3 |
| PC8 | Education | 69798 | 2.7e-2 (1.7e-2 to 3.7e-2) | 4.9e-3 | 3e-8 |
| PC8 | Refractive error | 69798 | -1.2e-2 (-2.3e-2 to -2.2e-3) | 5.2e-3 | 2e-2 |
| PC8 | Education allele score | 69798 | 4.2e-4 (4.2e-5 to 8.1e-4) | 1.9e-4 | 3e-2 |
| PC8 | Myopia allele score | 69798 | 1.1e-3 (-2.9e-4 to 2.4e-3) | 6.9e-4 | 1e-1 |
| PC9 | Education | 69798 | -2.1e-2 (-2.5e-2 to -1.7e-2) | 2.3e-3 | <2e-16 |
| PC9 | Education allele score | 69798 | -3.1e-4 (-4.8e-4 to -1.3e-4) | 9.1e-5 | 8e-4 |
| PC9 | Myopia allele score | 69798 | -7.8e-4 (-1.4e-3 to -1.5e-4) | 3.2e-4 | 2e-2 |
| PC9 | Refractive error | 69798 | 5.6e-3 (8.0e-4 to 1.0e-2) | 2.4e-3 | 2e-2 |

**Supplementary Table 3. Causal estimates of time spent in education on refractive error and refractive error on time spent in education using methods implemented in MR-Base and a split sample in UK Biobank.**

| *Exposure* | *Outcome* | *Method* | *N SNPs* | *Beta* | *SE* | *P-value* |
|---|---|---|---|---|---|---|
| Time spent in education | Refractive error | MR Egger | 67 | -0.399 | 0.125 | 2e-3 |
| | | Weighted median | 67 | -0.206 | 0.099 | 0.04 |
| | | Inverse variance weighted | 67 | -0.271 | 0.070 | 1e-4 |
| | | Simple mode | 67 | -0.169 | 0.547 | 0.8 |
| | | Weighted mode | 67 | -0.231 | 0.529 | 0.7 |
| Refractive error | Time spent in education | MR Egger | 43 | -0.002 | 0.057 | 1 |
| | | Weighted median | 43 | 0.003 | 0.037 | 0.9 |
| | | Inverse variance weighted | 43 | -0.014 | 0.026 | 0.6 |
| | | Simple mode | 43 | -0.091 | 0.071 | 0.2 |
| | | Weighted mode | 43 | 0.030 | 0.049 | 0.6 |

**Supplementary Table 4. Table of MR results from sensitivity analyses using alternative encodings of educational exposure and outcome, additional covariates and England-only restriction.**

| Instrument | Exposure | Outcome | Covariates | N | TSLS Beta | TSLS SE | TSLS P | Weak instr P | DWH P |
|---|---|---|---|---|---|---|---|---|---|
| Education | Education | Refractive error | sex + age | 69798 | -0.270 (-0.368 to -0.173) | 0.049 | 4e-8 | <2e-16 | 6e-2 |
| Education | Education (England only) | Refractive error | sex + age | 68152 | -0.269 (-0.369 to -0.171) | 0.050 | 8e-8 | <2e-16 | 7e-2 |
| Education | Education | Refractive error | sex + age + genechip | 69798 | -0.270 (-0.368 to -0.173) | 0.049 | 4e-8 | <2e-16 | 6e-2 |
| Education | Education | Refractive error | sex + age + PC1-10 | 69798 | -0.277 (-0.376 to -0.179) | 0.050 | 3e-8 | <2e-16 | 4e-2 |
| Education | Education | Refractive error | sex + age + birthweight | 42511 | -0.272 (-0.399 to -0.146) | 0.064 | 2e-5 | <2e-16 | 1e-1 |
| Education | Education | Refractive error | sex + age + breastfed | 53601 | -0.307 (-0.417 to -0.200) | 0.055 | 2e-8 | <2e-16 | 1e-2 |
| Education | Education | Refractive error | sex + age + northing | 69797 | -0.264 (-0.363 to -0.166) | 0.050 | 1e-7 | <2e-16 | 7e-2 |
| Education | Education | Refractive error | sex + age + easting | 69797 | -0.268 (-0.366 to -0.171) | 0.049 | 5e-8 | <2e-16 | 6e-2 |
| Education | Education | Refractive error | sex + age + TDI | 69737 | -0.271 (-0.371 to -0.173) | 0.050 | 6e-8 | <2e-16 | 6e-2 |
| Education | Education (Dichotomised) | Refractive error | sex + age | 69798 | -0.347 (-0.482 to -0.220) | 0.064 | - | - | - |
| Education | Education (No college) | Refractive error | sex + age | 45535 | -0.228 (-0.479 to 0.018) | 0.124 | 7e-2 | <2e-16 | 5e-1 |
| Myopia | Refractive error | Education | sex + age | 69798 | -0.008 (-0.041 to 0.025) | 0.017 | 6e-1 | <2e-16 | <2e-16 |
| Myopia | Refractive error | Education (England only) | sex + age | 68152 | -0.009 (-0.042 to 0.025) | 0.017 | 6e-1 | <2e-16 | <2e-16 |
| Myopia | Refractive error | Education (No college) | sex + age | 45535 | -0.004 (-0.035 to 0.027) | 0.016 | 8e-1 | <2e-16 | 7e-06 |
| Myopia | Refractive error | Education (Dichotomised) | sex + age | 69798 | 0.000 (-0.028 to 0.028) | 0.014 | - | - | - |
| Myopia | Refractive error | Education | sex + age + genechip | 69798 | -0.009 (-0.041 to 0.025) | 0.017 | 6e-1 | <2e-16 | <2e-16 |
| Myopia | Refractive error | Education | sex + age + PC1-10 | 69798 | -0.011 (-0.044 to 0.022) | 0.017 | 5e-1 | <2e-16 | <2e-16 |
| Myopia | Refractive error | Education | sex + age + birthweight | 42511 | 0.016 (-0.026 to 0.057) | 0.021 | 5e-1 | <2e-16 | 6e-15 |
| Myopia | Refractive error | Education | sex + age + breastfed | 53601 | 0.007 (-0.030 to 0.044) | 0.019 | 7e-1 | <2e-16 | <2e-16 |
| Myopia | Refractive error | Education | sex + age + northing | 69797 | -0.011 (-0.044 to 0.022) | 0.017 | 5e-1 | <2e-16 | <2e-16 |
| Myopia | Refractive error | Education | sex + age + easting | 69797 | -0.009 (-0.042 to 0.024) | 0.017 | 6e-1 | <2e-16 | <2e-16 |
| Myopia | Refractive error | Education | sex + age + TDI | 69737 | -0.008 (-0.040 to 0.026) | 0.017 | 7e-1 | <2e-16 | <2e-16 |

**Supplementary Data Table. Genetic variants used as instrumental variables for time spent in education and myopia in this study. Summary statistics are shown for associations with (1) time spent in education/myopia in the original study, (2) time spent in education and refractive error in the full UK Biobank sample, (3) time spent in education and refractive error in the two UK Biobank split sample groups.**

(supplementary_data_table_online_only_18-03-27.docx)