

## **Identifying potential effects of relative age in school year: an instrumental variable phenome-wide association study in the UK Biobank**

**Melanie A de Lange<sup>\*1,2</sup>, Neil M Davies<sup>1,3,4,5</sup>, Louise AC Millard<sup>#1,2</sup>, Kate  
Tilling<sup>#1,2</sup>**

<sup>1</sup>MRC Integrative Epidemiology Unit (IEU) at the University of Bristol, Bristol, United Kingdom.

<sup>2</sup>Population health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom.

<sup>3</sup>Division of Psychiatry and Department of Statistical Sciences, University College London, United Kingdom.

<sup>4</sup> Department of Statistical Sciences, University College London, London WC1E 6BT, UK

<sup>5</sup>K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, Norwegian University of Science and Technology, Norway.

\*Corresponding author. E-mail: [melanie.delange@bristol.ac.uk](mailto:melanie.delange@bristol.ac.uk)

# These authors contributed equally

Word count: 3707 (excluding abstract, key messages & references)

## **Abstract**

**Background:** Observational research shows that a child's relative age within their school year ('relative age') is associated with educational attainment and mental health. However, previous studies have only examined a small number of outcomes and evidence of the persistence of effects into adulthood is mixed. We conducted a hypothesis-free investigation of the effects of relative age.

**Method:** We used a regression discontinuity design and an instrumental variable (IV)-pheWAS in the UK Biobank (participants aged 40-69 years at baseline), using the PHEsANT software package. We created two IVs for relative age: being born in September vs. August ( $n=64\ 075$ ) and week of birth ( $n=383\ 309$ ). Outcomes passing the Bonferroni-corrected  $P$  value threshold for either instrument were plotted to identify those displaying a discontinuity at the school year transition.

**Results:** We found 21 traits associated with at least one of the instruments ( $P$  value below the Bonferroni threshold). Of these, 13 showed a discontinuity at the school year transition. These included previously identified effects including those with a younger relative age being less likely to have educational qualifications and more likely to have started smoking at an earlier age. We also identified a novel potential effect of a younger relative age in school year causing a better lung function as adults.

**Conclusion:** Educational policy should address educational inequality due to relative age. Further research should seek to replicate our identified effect on lung function in different populations, and investigate the mechanisms through which this effect may act.

**Key words:** Relative age, IV-pheWAS, summer-born, birth date, regression discontinuity

### **Key Messages**

- Children’s relative age within their school year has been associated with mental health in childhood and educational attainment.
- Our results supported previously identified effects, with those who were younger in their school year being less likely to have educational qualifications and more likely to report starting smoking at an earlier age.
- We also found a potential beneficial effect of a younger relative age in school year on lung function in adulthood.

## Introduction

Children's age within their school year, henceforth referred to as relative age, consistently relates to educational attainment. Children who are born later in the school year have lower average educational attainment, when it comes to tests taken on the same date, than those born at the beginning of the school year<sup>1-6</sup>. This means that in England, where entry for each school year runs from the 1st September to the 31st August, summer-born children on average perform worse than autumn-born children<sup>7</sup>.

Relative age differences in educational attainment occur across a range of education systems with different school year start dates, so are unlikely to be due to season of birth<sup>8</sup>. Although educational differences are greatest when comparing those born at the start and end of the school year (e.g. September vs. August in England), there is evidence of a linear change in educational outcomes over relative age<sup>7</sup>. These differences are largest when children start school and do diminish with age, but are still evident when children finish compulsory schooling<sup>7</sup>.

Children born later in the school year are also more likely to be diagnosed with special educational needs<sup>9-14</sup>, be bullied at school<sup>10, 15</sup>, engage in risky behaviour such as underage smoking<sup>7</sup>, and commit a crime in adolescence<sup>16</sup>. They also tend to have poorer mental health during childhood<sup>14, 17-19</sup> and higher suicide rates in adolescence and young adulthood (age 15-25)<sup>20</sup>.

A limited number of studies have examined whether the effects of relative age persist into adulthood. Some studies have found negative effects on whether adults are employed<sup>21</sup>, as well as their earnings<sup>22-24</sup>, self-confidence and risk taking<sup>25</sup>. Others have found little effect on whether people are in work and earnings<sup>7, 21, 26, 27</sup>, or health and mental wellbeing during adulthood<sup>7, 19, 21</sup>.

Many studies of relative age have used a regression discontinuity design (RDD). This is a quasi-experimental method that can be used to examine the effects of policies which assign people to a treatment or intervention based on a threshold on a continuous variable<sup>28, 29</sup>. Because a child's birth date cannot be controlled precisely (assuming natural birth), whether a child is assigned to the group born before or after the school year entry cut-off is quasi-random<sup>19</sup>. This means that, close to the threshold, covariates should be balanced between those born just before or just after the cut-off<sup>19</sup>. Thus, the RDD approach can give unbiased estimates of causal effects if the window around the threshold is small, and the underlying continuous variable (here the date of birth) cannot be manipulated<sup>28</sup>.

To date, research investigating relative age effects has focused on a small number of outcomes, mostly educational or related to mental health and wellbeing. In contrast to such hypothesis-driven approaches, hypothesis-free approaches test all measured factors, regardless of previous knowledge<sup>30</sup>. Phenome scans are a type of hypothesis-free analysis which can identify novel associations by testing the association of a trait of interest with a large number of other phenotypes ("the phenome")<sup>31</sup>. PHESANT is a software package for performing comprehensive phenome scans in UK Biobank<sup>31</sup>. Rather than being limited to a homogenous subset of phenotypes that can be analysed in the same way, PHESANT enables users to scan a heterogenous set of phenotypes (all UK Biobank continuous, integer and categorical fields)<sup>31</sup>. An instrumental variable (IV)-pheWAS is a type of pheWAS that uses an instrumental variable in order to investigate causality.

In this study we used PHESANT to conduct two IV-pheWAS, using a regression discontinuity, to identify phenotypes associated with relative age within ~383 000 UK Biobank participants born in England.

## Method

### Study population

The UK Biobank is a prospective cohort of around 500 000 adults aged mainly 40-69 years when recruited in in 2006-2010<sup>32</sup> (5.5% participation rate<sup>33</sup>). It contains extensive phenotypic data collected via lifestyle and health questionnaires; physical assessments; and blood, saliva and urine samples<sup>32</sup>. This research was conducted under UK Biobank application 16729 (dataset ID 48196).

Of the 502 448 UK Biobank participants, we removed 111 994 participants who were not born in England (as other UK countries have different school entry cut-off dates). We also removed 27 participants who withdrew their consent, giving a sample of 390 427 (see Figure 1).

### Instrumental variables (IVs) for relative age

Month of birth and week of birth were acquired from a central registry and updated by participants at their initial UK Biobank assessment centre visit. These variables were used to create two IVs for relative age.  $IV^{sep-aug}$  was a binary variable restricted to participants born in September or August, with September coded as one and August as zero.  $IV^{weeks}$  was week of birth as a continuous variable, with the first full week in September coded as zero and the last full week of August coded as 51. We excluded the week containing the August – September crossover in order to avoid misclassification (as we cannot know if these participants were born in August or September).

### Covariates

We included age at recruitment and sex as covariates in our models to reduce variation in our outcomes and increase precision. Age was derived from date of birth and date of attending initial assessment, and provided to us from UK Biobank as an integer. Sex was acquired from the NHS central registry at recruitment but was updated by the participant in some cases.

## Deriving outcomes with PHESANT

The UK Biobank Showcase (<https://biobank.ndph.ox.ac.uk/showcase/>) provides a searchable directory of variables available in the UK Biobank based on field type (integer, continuous, categorical (single) and categorical (multiple)). PHESANT was run with the 'save' option on to derive cleaned variables to be used as outcomes in the IV-pheWAS. This involved PHESANT categorising each outcome variable into one of four data types: continuous, ordered categorical, unordered categorical or binary, using its automated rule-based method described elsewhere<sup>31</sup>. The derived outcomes were stored as CSV files labelled according to data type.

## Statistical methods

### IV-pheWAS

We tested the association between our two IVs and our derived outcome variables. Regression analyses were adjusted for age and sex. We only tested outcomes with at least 500 participants, and at least 10 participants in each category in binary and unordered categorical variables. *P* values were generated for each outcome, using a likelihood ratio test comparing the model including the IV as a covariable versus the same model but not including the IV as a covariable. Results were ranked according to *P* value. In studies that test multiple hypotheses simultaneously, the family-wise error rate (the chances of getting type one errors (false positives)) rises as the number of tests increases<sup>34</sup>. To take multiple testing into account when evaluating the strength of the evidence of our results, we defined potential effects as any association of relative age with an outcome that passed the stringent Bonferroni-corrected significance threshold ( $0.05/\text{number of tests performed}$ ) in either IV-pheWAS. This controls the family-wise error rate assuming that tests are independent<sup>35</sup>.

We also ran a third IV-pheWAS, IV-months, where our IV was a set of 11 indicator variables describing the month of birth of each participant (excluding September) and compared each of the

11 months to September. However, we have not included this in our main results as it picked up many associations that were not specific to the discontinuity we were looking for (see Table S1 for results).

To explore whether our identified potential effects displayed a discontinuity at the transition between school years, and therefore were relative age effects rather than seasonal trends, estimates and confidence intervals for  $IV^{\text{months}}$  (month of birth relative to September) were plotted for each potential effect.

### **Sensitivity Analyses**

The results of  $IV^{\text{months}}$  were used to run a meta-regression analyses for each identified potential effect. For each potential effect we ran 11 meta-regressions, placing the origin of the linear function at the start of each month from October to August. We identified the model with the largest association. If the this model was for the time-origin of October, this was taken as evidence of a relative age in school year effect.

Analyses were performed in R version 4.1.0 and Stata version 17. Code is available at <https://github.com/MRCIEU/PHEASANT-IV-pheWAS-relative-school-age> (git tag v0.2 corresponds to the version presented here).

## **Results**

### **IV-pheWAS results**

Out of the 15 034 tests performed for  $IV^{\text{sep-aug}}$  and 23 164 tests performed for  $IV^{\text{weeks}}$ , we found 21 traits that had a  $P$  value below the Bonferroni threshold in at least one of these IV-pheWAS (Table 1; See Table S1 for full results for both IVs (note we do not have estimates for unordered categorical



variables)). The eight traits associated with  $IV^{\text{sep-aug}}$  were also associated with  $IV^{\text{weeks}}$ . Model convergence was not reached for a minority of outcomes and as a result these outcomes were not considered further (See Table S2 for a list of these variables).

Six of the eight traits that passed the Bonferroni-corrected threshold ( $P=3.33 \times 10^{-6}$ ) for  $IV^{\text{sep-aug}}$  were related to education (see Table 1). For example, participants born in September were more likely to have O levels/GCSEs ( $P=1.84 \times 10^{-14}$ ), A/AS levels ( $P=4.72 \times 10^{-08}$ ) or a college/university degree ( $P=4.06 \times 10^{-09}$ ) and less likely to have CSEs ( $P=1.22 \times 10^{-25}$ ) than those born in August. Those born in September were also more likely to complete full-time education at an older age than those born in August ( $P=7.24 \times 10^{-162}$ ). These findings are supported by the results of  $IV^{\text{weeks}}$  which found that participants born in a later week in the academic year had worse education outcomes (see Table 1).

Moving beyond education, participants born in September vs. August, and participants with an earlier week of birth in school year, were more likely to report being comparatively taller for their age at age 10 ( $P=1.25 \times 10^{-141}$  and  $P=8.55 \times 10^{-295}$ ). Participants born in September vs. August, and participants with an earlier week of birth in school year, who are former smokers also report starting smoking later ( $P=6.03 \times 10^{-09}$  and  $3.86 \times 10^{-07}$ ). However, little evidence was seen amongst current smokers ( $P=0.32$  and  $P=0.16$ ), although this may be due to the smaller sample size in this group hence lower statistical power.

**Table 1.** Traits associated with being born in September vs August ( $IV^{sep-aug}$ ) or week of birth ( $IV^{weeks}$ )

Outcome	Outcome type	Total $n$ ( $n$ in each category for binary outcomes)	$P$ value $IV^{sep-aug}$	$\beta$ (95% CI) $IV^{sep-aug}$	Total $n$ ( $n$ in each category for binary outcomes)	$P$ value $IV^{weeks}$	$\beta$ (95% CI) $IV^{weeks}$	Discontinuity at school year transition evident on forest plot	Best fitting discontinuity with origin in October (evidence of discontinuity)
Age completed full time education (3 categories: <16 years old, 16 years old, 17 years old or older)	Ordered categorical	43933	$7.24 \times 10^{-162}$	0.48 (0.45, 0.52)	261871	$4.81 \times 10^{-173}$	-0.007 (-0.008, -0.007)	✓	✓
Comparative height at age 10 (3 categories: shorter, about average, taller)	Ordered categorical	63039	$1.25 \times 10^{-141}$	0.39 (0.36, 0.42)	376987	$8.55 \times 10^{-295}$	-0.008 (-0.008, -0.007)	✓	✓
Year ended full time education	Continuous	16008	$1.96 \times 10^{-30}$	0.08 (0.06, 0.09)	96107	$1.03 \times 10^{-41}$	-0.001 (-0.001, -0.001)	✓	✓
Qualifications: CSEs or equivalent (versus not have CSEs or equivalent)	Binary	62941 (53530 vs 9411)	$1.22 \times 10^{-25}$	-0.25 (-0.30, -0.20)	376822 (320620 vs 56202)	$4.71 \times 10^{-60}$	0.005 (0.005, 0.006)	✓	✓
Qualifications: O levels/GCSEs or equivalent (versus not having O levels/GCSEs)	Binary	62941 (32791 vs 30150)	$1.84 \times 10^{-14}$	0.12 (0.09, 0.16)	376822 (195116 vs 181706)	$1.01 \times 10^{-17}$	-0.002 (-0.002, -0.002)	✓	✓
Qualifications: College or University degree (versus not having college of University degree)	Binary	62941 (43794 vs 19147)	$4.06 \times 10^{-09}$	0.10 (0.07, 0.14)	376822 (261214 vs 115608)	$3.03 \times 10^{-07}$	-0.001 (-0.002, -0.001)	✓	✓
Age started smoking in former smokers	Continuous	15647	$6.03 \times 10^{-09}$	0.09 (0.06, 0.12)	94093	$3.86 \times 10^{-07}$	-0.001 (-0.002, -0.001)	✓	✓
Qualifications: A/AS levels or equivalent (versus not having A/AS levels or equivalent)	Binary	62941 (46190 vs 16751)	$4.72 \times 10^{-08}$	0.10 (0.06, 0.14)	376822 (275200 vs 101622)	$3.93 \times 10^{-07}$	-0.001 (-0.002, -0.001)	✓	
Sensitivity/hurt feelings: Are your feelings easily hurt? (Yes versus no)	Binary	62277 (27659/34618)	0.27	-0.02 (-0.05, 0.01)	372349 (164547 vs 207802)	$2.54 \times 10^{-25}$	0.002 (0.002, 0.003)		
Year job started (calendar year)	Continuous	15892	$4.51 \times 10^{-06}$	0.05 (0.03, 0.07)	95445	$2.66 \times 10^{-10}$	-0.001 (-0.001, -0.001)	✓	
Ever had bowel cancer screening (versus never having had bowel cancer screening)	Binary	63077 (42549/20528)	0.75	-0.01 (-0.04, 0.03)	377262 (253638 vs 123624)	$4.93 \times 10^{-09}$	0.002 (0.001, 0.002)		
3mm asymmetry angle (right)	Continuous	14306	0.02	0.04 (0.01, 0.07)	85613	$4.44 \times 10^{-08}$	-0.001 (-0.002,		

Mean MO in tapetum on FA skeleton (right)	Continuous	5595	0.06	-0.05, (-0.10, 0.003)	33361	1.09x10 <sup>-07</sup>	-0.001	0.002 (0.001, 0.003)	✓	
Mean OD in tapetum on FA skeleton (right)	Continuous	5595	0.06	0.05 (-0.002, 0.10)	33359	2.75x10 <sup>-07</sup>	-0.002 (-0.003, -0.001)		✓	
Trouble concentrating in last 2 weeks (4 categories: Not at all, several days, more than half the days, nearly every day)	Ordered categorical	21750	0.01	-0.09 (-0.15, -0.02)	131316	3.83x10 <sup>-07</sup>	0.002 (0.001, 0.003)	✓		✓
Mood swings: Does your mood often go up and down? (Yes versus no)	Binary	62557 (34669/27888)	0.05	-0.03 (-0.06, 0.0004)	374169 (204225 vs 169944)	3.88x10 <sup>-07</sup>	0.001 (0.001, 0.002)			
Forced expiratory volume in 1-second (FEV1) Z-score (spirometry)	Continuous	50776	0.02	-0.02 (-0.04, -0.004)	303772	4.19x10 <sup>-07</sup>	0.001 (0.0003, 0.001)	✓		✓
Mean MO in inferior cerebellar peduncle on FA skeleton (left)	Continuous	5595	0.07	0.05 (-0.003, 0.10)	33361	4.43x10 <sup>-07</sup>	-0.002 (-0.003, -0.001)			✓
Neuroticism score (12 categories: 0 to 12)	Ordered categorical	52115	0.64	0.007 (-0.02, 0.04)	310355	1.17x10 <sup>-06</sup>	0.001 (0.001, 0.001)			
Job code: Typist (versus not having a job code of typist)	Binary	15684 (15141/543)	0.002	-0.27 (-0.45, -0.10)	94121 (90882 vs 3239)	1.58x10 <sup>-06</sup>	0.006 (0.004, 0.009)	✓		✓
Historical job code: Typist (versus not having a historical job code of typist)	Binary	15684 (15141/543)	0.002	-0.27 (-0.45, -0.10)	94121 (90882 vs 3239)	1.58x10 <sup>-06</sup>	0.006 (0.004, 0.009)	✓		✓

Continuous outcome: estimates are mean difference of standard deviation of inverse rank normal transformed outcome for September versus August born. Binary outcome: estimates are log odds of outcome September versus August born. Categorical ordered outcome: estimates are log odds of being in a higher outcome category for September versus August born. Estimates in grey are those that did not pass the Bonferroni correction for that IV-pheWAS.

A further 13 traits passed the Bonferroni-corrected threshold ( $P=2.16 \times 10^{-6}$ ) for  $IV^{\text{weeks}}$  alone. In terms of psychological traits,  $IV^{\text{weeks}}$  was positively associated with having your feelings hurt easily ( $P=2.54 \times 10^{-25}$ ), having mood swings ( $P=3.88 \times 10^{-07}$ ) and participants' neuroticism scores ( $P=1.17 \times 10^{-06}$ ), meaning those with a younger relative age (later week of birth) were more likely to have these traits. Week of birth was also found to be associated with a number of physical measurements, including being positively associated with forced expiratory volume in one second z-score ( $4.19 \times 10^{-07}$ ).  $IV^{\text{weeks}}$  was also negatively associated with 3mm asymmetry angle (right eye) ( $P=4.44 \times 10^{-08}$ ), mean OD in tapetum on FA skeleton (right) ( $P=2.75 \times 10^{-07}$ ) and mean MO in inferior cerebellar peduncle on FA skeleton (left) ( $P=4.43 \times 10^{-07}$ ), and positively associated with mean MO in tapetum on FA skeleton (right) ( $1.09 \times 10^{-07}$ ). However, the fact that brain and skeleton MRI, and eye measurement phenotypes passed the Bonferroni threshold may be the result of the sheer number of these variables in the UK Biobank, rather than true effects.

$IV^{\text{weeks}}$  was positively associated with ever being screened for bowel cancer ( $P=4.93 \times 10^{-09}$ ) but little evidence of an association with people reporting having had bowel cancer itself (small bowel cancer  $P=0.89$ ; large bowel cancer  $P=0.19$ ), possibly due to inadequate statistical power.  $IV^{\text{weeks}}$  was positively associated with having a job code or historical job code of typist ( $P=1.58 \times 10^{-06}$  for both). It was also positively associated with answering yes to the question "Over the last 2 weeks, how often have you been bothered by trouble concentrating on things, such as reading the newspaper or watching television?" ( $P=3.83 \times 10^{-07}$ ). Another related phenotype, "Over the last 2 weeks, how often have you been bothered by any of the following problems? [depressive symptoms] Trouble concentrating on things, such as reading the newspaper or watching television", was associated in a direction consistent with our finding ( $P=0.03$ ), although it did not pass the Bonferroni threshold.

### **Meta-analysis forest plots**

Of the 21 potential effects identified in the IV-pheWAS, 20 had an output from PHESANT for  $IV^{\text{months}}$  and were plotted by month of birth. A discontinuity at the transition point between school years was visually identified for 13 of these (see Table 1 and Figures S1-S20) – educational qualifications (CSEs, O level/GCSEs, A/AS Levels, College/University degree), the year and age full time education ended, the year a person’s job started, comparative height at age 10, the age former smokers started smoking, having trouble concentrating in the last two weeks, forced expiratory volume in one second (FEV1) z-score and having a job code or historical job code of a typist.

### **Sensitivity analyses**

Our meta-regression to identify the location of the best-fitting discontinuity (results shown in Table 1 and Table S3). This showed that the outcomes with the largest association with the origin in October were: age and year ended full time education, having O levels/GCSEs, having CSEs, having a college/university degree, comparative height at age 10, age starting smoking in former smokers, trouble concentrating in last two weeks, forced expiratory volume in one second (FEV1) Z-score, job code of typist, historical job code of typist, mean MO in tapetum on FA skeleton (right), mean OD in tapetum on FA skeleton (right) and mean MO in inferior cerebellar peduncle on FA skeleton (left).

### **Discussion**

In this study we conducted hypothesis-free IV-pheWAS to explore the effects of relative age in school year in around 383 000 UK Biobank participants born in England. We found 21 phenotypes that were associated with relative age, 13 of which displayed a discontinuity at the transition between school years. A number of our results are somewhat trivial, being directly due to the difference in participants age in their school year. This includes associations of younger relative age with being more likely to finish full-time education at an earlier age. It also includes an association between those with a younger relative age being more likely to report being shorter than their peers

at age 10 (where participants likely think of their school year as their peers). Consequently, researchers using this UK Biobank variable in their analyses need to bear in mind that comparative height at age 10 is a somewhat biased measure of height in childhood. Furthermore, those with a younger relative age are more likely to start smoking at an earlier age (amongst former smokers). This is also somewhat expected given that peers are likely to start smoking at the same point in time, so those younger in the year will be younger when they start. Novel (non-trivial) results include those with a younger relative age being more likely to have better lung function in adulthood and having had trouble concentrating in the previous two weeks.

We found that younger relative age is associated with a lower likelihood of having educational qualifications, which is consistent with previous studies that report an association between younger relative age and lower educational attainment<sup>1-7</sup>. Several strategies have been proposed to address this inequality. Crawford, Dearden & Greaves (2013) found that children being different ages when they take national standardised tests is the main driver of relative age differences in educational attainment. They subsequently demonstrated that age-adjusting national achievement test scores (changing the number of points needed to achieve a certain grade depending on the age of the child, whilst keeping an absolute measure of children's scores) effectively removes relative age differences, including the percentage of students reaching the government's expected level<sup>7</sup>.

Looking at the persistence of relative age effects into adulthood, in contrast to some previous studies<sup>22-24</sup>, but in line with others<sup>7, 21, 26, 27</sup>, we found little evidence of long term effects on employment prospects or earnings, with our only potential effect being having a job code or historical job code of a typist. However, this may be due to the extent and nature of employment variables in the UK Biobank. For example, analyses using individual occupation types will lack power. In addition, we did not find evidence of an effect of relative age on current employment status, but this may be due to the fact that around a third of UK Biobank participants are retired, so this variable

is not necessarily a measure of occupational success. Nevertheless, a higher average total household income was strongly associated with being born in September vs August and week of birth, although it did not pass the Bonferroni threshold of either IV-pheWAS. Furthermore, in accordance with existing research<sup>7, 19, 21</sup>, we also found little evidence of long term effects on adult mental health, with mental health phenotypes that passed the Bonferroni threshold, such as sensitivity and mood swings, not displaying a discontinuity when plotted.

The little body of research previously conducted into the effect of relative age on adult physical health found no substantial differences in self-reported health<sup>19</sup>. In line with this, our IV-pheWAS only found one potential effect related to physical health. We found that adults with a younger relative age performed better in forced expiratory lung function tests. This result is supported by the fact that when we searched our results we also found that forced vital capacity z-score (another measure of lung function) was also positively associated with week of birth ( $P=5.97 \times 10^{-05}$ ), meaning those with a younger relative age have a greater capacity. However, this did not pass the Bonferroni threshold. Our finding could possibly be the effect of children with a younger relative age being pushed in childhood to keep up with their bigger, older peers, with this process in effect training their lungs to give them better lung function as an adult. Alternative explanations include random chance or selection bias. This therefore warrants further exploration.

Our other potential effect of relative age on adult outcomes was that those with a younger relative age were also more likely to have had trouble concentrating in the last two weeks. Again, this could be a spurious finding, due to chance or selection. However, the fact that we also found having had “recent trouble concentrating on things” to be associated with relative age (albeit beyond the Bonferroni threshold) and that relative age has been associated with childhood diagnosis of attention-deficit/hyperactivity disorder<sup>11-14</sup> suggests the effects of relative age on concentration in adulthood could warrant further investigation.

A key strength of our hypothesis-free IV-PheWAS approach is that it enables potentially novel effects to be uncovered. This possibility is amplified by the fact that the UK Biobank includes an extensive number and range of phenotypes, and that the PHESANT software allowed us to perform comprehensive phenome scans, rather than our set of outcomes being restricted to a homogenous subset<sup>31</sup>. We used the Bonferroni corrected  $P$  value threshold to reduce the likelihood of identifying effects that are due to chance. However, this method may be overly strict and therefore could have potentially resulted in our analyses missing some outcomes where relative age does have a true small effect. Hence, we have also reported the FDR threshold and provide the full set of results of our IV-pheWAS, so other researchers are able to follow-up other outcomes in future studies (see Table S1).

Some limitations of this work include the fact that the UK Biobank is not representative of the general UK population, with participants tending to be older, healthier, female and from less socioeconomically deprived areas<sup>33</sup>. If relative age in school year also affects participation in UK Biobank (which is likely given its suggested effect on educational attainment, which in turn is strongly associated with UK Biobank participation), then this could cause selection bias in estimating causal effects of relative age.

Further limitations include that PHESANT uses an automatic rule-based method to decide how each outcome is tested and it is possible that this may treat some variables inappropriately. In addition, models did not converge for a small number of traits so their relationship with relative age could not be reported. Lastly, because pheWAS aim to generate hypotheses, associations discovered should be replicated in an independent sample and using other study designs which have contrasting assumptions and biases<sup>35</sup>.



In summary, this study adds to the body of evidence that suggests that children born later in the academic year have lower educational attainment than those born earlier in the school year. Our results have implications for educational policy, as age-adjustment of national achievement test scores could potentially help to reduce these inequalities if presented alongside original scores. This could be particularly important for standardised national tests that determine children's access to different types of schooling (Year Six SATs or Eleven Plus) or access to further education (GCSEs and A Levels). Future research should seek to replicate our identified effect on lung function in different populations, and investigate the mechanisms through which this effect may act. It could also further investigate the medium-term consequences of relative age on career prospects, and physical and mental health in early adulthood.

## **Ethics approval**

UK Biobank is approved by the National Health Service National Research Ethics Service (ref. 11/NW/0382; UK Biobank application number 16729).

## **Data availability**

The UK Biobank dataset used to conduct the research in this paper is available via application directly to the UK Biobank. Applications are assessed for meeting the required criteria for access, including legal and ethics standards. More information regarding data access can be found at [<https://www.ukbiobank.ac.uk/enable-your-research>].

## **Supplementary data**

Supplementary data are available at *IJE* online.

## Author contributions

K.T. initiated the study concept. M.A.dL., L.A.C.M. and N.M.D performed the statistical analysis.

M.A.dL drafted the first version of the manuscript. All authors were involved in the interpretation of the data, critically reviewed and revised the manuscript, and approved the final version as submitted.

## Funding

This research was funded in whole, or in part, by the Wellcome Trust [grant number 226909/Z/23/Z].

For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

M.A.dL is funded by the Wellcome Trust [grant number 226909/Z/23/Z]. M.A.dL, K.T. and L.M. work in a unit that receives support from the University of Bristol and the UK Medical Research Council [grant numbers MC\_UU\_00011/1, MC/UU/00011/3]. N.M.D is supported by the Norwegian Research Council [grant number 295989].

## Acknowledgements

This research has been conducted using the UK Biobank Resource under Application Number 16729.

We thank Dr Beate Leppert for contributing to the early stages of this work. This work was carried out using the computational facilities of the Advanced Computing Research Centre, University of Bristol - <http://www.bristol.ac.uk/acrc/>.

## Conflict of interest

None declared.

## References

1. Russell R, Startup M. Month of birth and academic achievement. *Pers Individ Dif*. 1986;**7**:839-46.
2. Bell J, Daniels S. Are summer-born children disadvantaged? The birthdate effect in education. *Oxf Rev of Educ*. 1990;**16**:67-80.
3. Sharp C, Hutchison D, Whetton C. How do season of birth and length of schooling affect children's attainment at key stage 1? *Educ Res*. 1994;**36**:107-21.
4. Alton A, Massey A. Date of birth and achievement in GCSE and GCE A-level. *Educ Res*. 1998;**40**:105-9.
5. Urruticoechea A, Oliveri A, Vernazza E, Giménez-Dasí M, Martínez-Arias R, Martín-Babarro J. The relative age effects in educational development: a systematic review. *Int J Environ Res Public Health*. 2021;**18**:8966.
6. Givord P. How a student's month of birth is linked to performance at school: new evidence from PISA. OECD Education Working Paper 221. 2020. <https://www.oecd-ilibrary.org/docserver/822ea6ce-en.pdf?expires=1677840185&id=id&accname=guest&checksum=B7A3A5368B8E947FB7F5BCB4E415AC6C> (03 March 2023, date last accessed)
7. Crawford C, Dearden L, Greaves E. When you are born matters : evidence for England. Institute for Fiscal Studies Report R80. 2013. <https://ifs.org.uk/comms/r80.pdf> (03 March 2023, date last accessed)
8. Bedard K, Dhuey E. The persistence of early childhood maturity: international evidence of long-run age effects. *Q J Econ*. 2006;**121**:1437-72.
9. Dhuey E, Lipscomb S. Disabled or young? relative age and special education diagnoses in schools. *Econ Educ Rev*. 2010;**29**:857-72.
10. Department for Education. Month of birth and education. DfE Research Report RR017. 2010. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/182664/DFE-RR017.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/182664/DFE-RR017.pdf) (03 March 2023, date last accessed)
11. Chen K, Fortin N, Phipps S. Young in class: implications for inattentive/hyperactive behaviour of Canadian boys and girls. *Can J Econ*. 2015;**48**:1601-34.
12. Folgar M, Boubeta A, Lamas M, Mociño L. Evaluation of impact of relative age effect on school performance, bullying, self-esteem, ADHD diagnosis and consumption of tobacco in the transition from primary education to high school. *Rev Iberoam Diagn Ev* 2017;**44**:92-104.
13. Morrow RL, Garland EJ, Wright JM, Maclure M, Taylor S, Dormuth CR. Influence of relative age on diagnosis and treatment of attention-deficit/hyperactivity disorder in children. *Can Med Assoc J*. 2012;**184**:755-62.
14. Root A, Brown JP, Forbes HJ et al. Association of relative age in the school year with diagnosis of intellectual disability, attention-deficit/hyperactivity disorder, and depression. *JAMA Pediatr*. 2019;**173**:1068.
15. Mühlenweg A. Young and innocent: international evidence on age effects within grades on victimization in elementary school. *Econ Lett*. 2010;**109**:157-60.
16. Landersø R, Nielsen H, Simonsen M. School starting age and crime. Aarhus University, Economics Working Papers. 2013. [https://econ.au.dk/fileadmin/site\\_files/filer\\_oekonomi/Working\\_Papers/Economics/2013/wp13\\_03.pdf](https://econ.au.dk/fileadmin/site_files/filer_oekonomi/Working_Papers/Economics/2013/wp13_03.pdf) (03 March 2023, date last accessed)
17. Goodman R. Child psychiatric disorder and relative age within school year: cross sectional survey of large population sample. *BMJ*. 2003;**327**:472-0.
18. Ando S, Usami S, Matsubayashi T et al. Age relative to school class peers and emotional well-being in 10-year-olds. *PLoS One*. 2019;**14**:e0214359.
19. Broughton T, Langley K, Tilling K, Collishaw S. Relative age in the school year and risk of mental health problems in childhood, adolescence and young adulthood. *J Child Psychol Psychiatry*. 2023;**64**:185-96.

20. Matsubayashi T, Ueda M. Relative age in school and suicide among young individuals in Japan: a regression discontinuity approach. *PLoS One*. 2015;**10**:e0135349.
21. Crawford C, Dearden L, Greaves E. The impact of age within academic year on adult outcomes. Institute for Fiscal Studies Working Paper W13/07. 2013. [https://ifs.org.uk/sites/default/files/output\\_url\\_files/wp201307.pdf](https://ifs.org.uk/sites/default/files/output_url_files/wp201307.pdf) (03 March 2023, date last accessed)
22. Kawaguchi D. Actual age at school entry, educational outcomes, and earnings. *J Jpn Int Econ*. 2011;**25**:64-80.
23. Fredriksson P, Ockert B. Is early learning really more productive? the effect of school starting age on school and labour market performance. Institute for the Study of Labor (IZA), Discussion Paper 1659. 2005. <https://docs.iza.org/dp1659.pdf> (03 March 2023, date last accessed)
24. Peña P. Creating winners and losers: date of birth, relative age in school, and outcomes in childhood and adulthood. *Econ Educ Rev*. 2017;**56**:152-76.
25. Page L, Sarkar D, Silva-Goncalves J. Long-lasting effects of relative age at school. *J Econ Behav Organ*. 2019;**168**:166-95.
26. Balestra S, Eugster B, Liebert H. Summer-born struggle: The effect of school starting age on health, education, and work. *Health Econ*. 2020;**29**:591-607.
27. Pehkonen J, Viinikainen J, Böckerman P, Pulkki-Råback L, Keltikangas-Järvinen L, Raitakari O. Relative age at school entry, school performance and long-term labour market outcomes. *Appl Econ Lett*. 2015;**22**:1345-8.
28. Moscoe E, Bor J, Bärnighausen T. Regression discontinuity designs are underutilized in medicine, epidemiology, and public health: a review of current and best practice. *J Clin Epidemiol*. 2015;**68**:132-43.
29. Bor J, Moscoe E, Mutevedzi P, Newell M-L, Bärnighausen T. Regression discontinuity designs in epidemiology. *Epidemiology*. 2014;**25**:729-37.
30. Escobar-Morreale HF, Insenser M, Cortón M, Millán JLS, Peral B. Proteomics and genomics: a hypothesis-free approach to the study of the role of visceral adiposity in the pathogenesis of the polycystic ovary syndrome. *Proteomics Clin Appl*. 2008;**2**:444-55.
31. Millard LAC, Davies NM, Gaunt TR, Davey Smith G, Tilling K. Software application profile: PHESANT: a tool for performing automated phenome scans in UK Biobank *Int J Epidemiol*. 2018;**47**:29-35.
32. Sudlow C, Gallacher J, Allen N et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*. 2015;**12**:e1001779.
33. Fry A, Littlejohns TJ, Sudlow C et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am J Epidemiol*. 2017;**186**:1026-34.
34. Nicholson KJ, Sherman M, Divi SN, Bowles DR, Vaccaro AR. The role of family-wise error rate in determining statistical significance. *Clin Spine Surg*. 2022;**35**:222-3.
35. Millard LAC, Davies NM, Tilling K, Gaunt TR, Davey Smith G. Searching for the causal effects of body mass index in over 300 000 participants in UK Biobank, using Mendelian randomization. *PLoS Genet*. 2019;**15**:e1007951.

**Figure 1.** Participant flow diagram

