

Relative age in the school year and risk of mental health problems in childhood, adolescence and young adulthood

Thomas Broughton,^{1,2} Kate Langley,^{1,3} Kate Tilling,⁴ and Stephan Collishaw^{1,2}

¹Division of Psychological Medicine and Clinical Neurosciences, MRC Centre for Neuropsychiatric Genetics and Genomics, School of Medicine, Cardiff University, Cardiff, UK; ²Wolfson Centre for Young People's Mental Health, Cardiff University, Cardiff, UK; ³School of Psychology, Cardiff University, Cardiff, UK; ⁴MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK

Purpose: Relative age within the school year ('relative age') is associated with increased rates of symptoms and diagnoses of mental health disorders, including ADHD. We aimed to investigate how relative age influences mental health and behaviour before, during and after school (age range: 4–25 years). **Method:** We used a regression discontinuity design to examine the effect of relative age on risk of mental health problems using data from a large UK population-based cohort (Avon Longitudinal Study of Parents and Children (ALSPAC); $N = 14,643$). We compared risk of mental health problems between ages 4 and 25 years using the parent-rated Strengths and Difficulties Questionnaire (SDQ), and depression using self-rated and parent-rated Short Mood and Feelings Questionnaire (SMFQ) by relative age. **Results:** The youngest children in the school year have greater parent-rated risk of mental health problems, measured using parent-rated SDQ total difficulties scores. We found no evidence of differences before school entry [estimated standardised mean difference (SMD) between those born on 31 August and 1 September: .02 (–.05, .08)]. We found that estimates of effect size for a 1-year difference in relative age were greatest at 11 years [SMD: .22 (.15, .29)], but attenuated to the null at 25 years [SMD: –.02 (–.11, .07)]. We did not find consistent evidence of differences in self-rated and parent-rated depression by relative age. **Conclusions:** Younger relative age is associated with poorer parent-rated general mental health, but not symptoms of depression. **Keywords:** Relative age; Avon Longitudinal Study of Parents and Children (ALSPAC); depression; mental health.

Introduction

Up to one in eight children in the United Kingdom meet diagnostic criteria for mental health disorders (Sadler et al., 2018). Mental health disorders have become more prevalent in children over time and are associated with immediate and long-term physical and psychological impairment (Sadler et al., 2018; Vos et al., 2017). Policy and practice interventions targeting potentially modifiable risk and protective factors will only be effective if the risk factors are truly causal (Thapar & Rutter, 2019). In observational studies, it is hard to infer causality due to the possibility of unmeasured confounding (Thapar & Rutter, 2019). To overcome this, we can use 'natural experiments' which approximate the random assignment of risk or protective conditions. (Thapar & Rutter, 2019).

One potential risk factor is a child's age within the school year, henceforth, relative age. In England and Wales, this is determined by the cut-off date for school entry of the 1st of September. Children born in September are the oldest in the school year, and those born in August the youngest. This leads to an age difference of up to 12 months between children in the same class; at school entry, it is assumed that the oldest children will be approximately 20% more

mature than the youngest children in the school year (Holland & Sayal, 2019). Relative age in England and Wales is assigned quasi-randomly since one cannot directly influence a child's exact delivery date (assuming natural birth), even though children's dates of birth are unevenly distributed across the school year with a higher birth rate in late September (Borja & Martin, 2017). Potential confounders of mental health outcomes would be plausibly independent of birth being either side of the school entry cut-off within a narrow timeframe around that date, with this assumption becoming less plausible as the bandwidth increases. The assumption that being born either side of the cut-off is randomly allocated cannot be tested, but its association with observed confounders can be. Thus, observational studies of relative age allow for causal inference, using regression discontinuity methods where there is a continuous variable (day of year born) which has a cut-off (1st of September) at which the treatment (entry to school) is applied (Moscoe, Bor, & Bärnighausen, 2015; Thistlethwaite & Campbell, 1960).

The youngest children in the school year are over-represented in mental health disorder statistics (Root et al., 2019). Children born in the last quarter of an academic year were approximately 23% more likely to be diagnosed with depression than those born in the first quarter (Root et al., 2019). It is

Conflict of interest statement: No conflicts declared.

unclear whether this represents a real difference in disorder risk by relative age, or whether this is due to differences in help-seeking patterns or by referrals to mental health services.

Relative age effects on mental health are relatively unexplored from an epidemiological perspective. Three UK studies suggest that the youngest children in the school year have higher parent, teacher and self-rated mental health symptom scores (Crawford, Dearden, & Greaves, 2013; Norbury et al., 2016; Patalay et al., 2015). Crawford et al. (2013) compared teacher and parent-rated versions of the Strengths and Difficulties Questionnaire (SDQ, Goodman, 1997) in 7- to 13-year-old children in the Avon Longitudinal Study of Parents and Children (ALSPAC) cohort. Children born in August had greater total SDQ scores relative to their September-born peers at age 7. The researchers also found that relative age effects on mental health attenuate to the null as the gap between oldest and youngest children becomes proportionally smaller as chronological age increases (Crawford et al., 2013). Patalay et al. (2015) investigated self-rated SDQ scores in 11- to 13-year-old children in England, grouped into thirds by relative age. The youngest third of children had higher SDQ scores, primarily driven by emotional and peer problems. Lastly, Norbury et al. (2016) found that the youngest children in their first year of compulsory education were at higher risk of behaviour problems as evidenced by higher SDQ scores, as well as poorer language skills, relative to their peers (Norbury et al., 2016).

Cross-national comparisons of large representative population surveys of mental health disorders have shown that being born in the latest third of any academic year was associated with an increased risk of depression diagnosis, as well as increased self-rated, parent-rated and teacher-rated risk of mental health problems (Goodman, Gledhill, & Ford, 2003). Relative age effects were present across nations with different school entry cut-off dates, for example, in England (September cut-off), Scotland (March cut-off) and Northern Ireland (July cut-off). The same study found that relatively young children did not differ from older peers on age-standardised ability tests (Goodman et al., 2003). This evidence shows that relative age effects are independent of other birth date effects such as season of birth, and that younger children are more likely to have a greater risk of mental health problems regardless of rater, chronological age and a country's school entry cut-off.

There are several important knowledge gaps. First, it is not known whether differences in mental health by relative age occur prior to starting school. Typically, before school entry children are not grouped together by the school entry cut-off; if mental health differences by relative age emerge after, but not before, school entry, we can infer that these differences are caused by this grouping.

Second, it is not known whether relative age effects extend into adulthood. To get a clearer picture of relative age effects and potential interventions for these effects, it is not only important to consider what happens when relatively young children enter compulsory education, but also when they leave. Third, psychometric tests, such as the SDQ show different rates of mental health problem risk, and different patterns of associations with a range of risk factors, depending on the informant used – parent, young person or teacher (Collishaw, Goodman, Ford, Rabe-Hesketh, & Pickles, 2009). This emphasises the need for multi-informant sources of risk of mental health problems. Lastly, most previous studies have looked at general mental health symptom screens such as the SDQ. However, these studies highlight some specificity across different domains of mental health (Patalay et al., 2015). To better understand mental health domain-specific effects, more sensitive and specific measurements of particular mental health outcomes are needed, such as the Mood and Feelings questionnaire (MFQ; Angold et al., 1995) for depression.

This study aims to investigate how relative age influences mental health before, during and after school, using data from a large UK population-based cohort (ALSPAC) repeatedly assessed between the ages of 4 and 25 years. We hypothesise that the youngest children in the school year will have greater reported risk of mental health problems than their relatively older peers, that relative age differences will first emerge at school entry, and these differences will be strongest in the early school years. This is because the difference in chronological age (and therefore physiological and psychological maturity) is greatest between oldest and youngest in those years. We hypothesised that there would be no difference by relative age before school entry because we assumed that pre-school children are largely taught through play, and so are not usually subject to formal assessments and classroom streaming to the same degree as children who have entered the school system.

Method

Sample

The Avon Longitudinal Study of Parents and Children (ALSPAC) is an ongoing prospective longitudinal birth cohort of individuals born in the historical Avon administrative county in South-West England (Boyd et al., 2013; Fraser et al., 2013; Northstone et al., 2019).

Pregnant women resident in Avon, UK with expected dates of delivery 1/4/1991 to 31/12/1992 were invited to take part in the study. The initial number of pregnancies enrolled is 14,541 resulting in 14,062 live births and 13,988 children alive at 1 year of age (Boyd et al., 2013).

When the oldest children in the cohort were approximately 7 years of age, the initial sample was bolstered with 913 eligible children who did not initially join the study. The total sample size for analyses using any data collected after the age

of 7 is therefore 15,454 pregnancies, resulting in 15,589 foetuses. Of these 14,901 were alive at 1 year of age.

The ALSPAC study website contains details of all available data, available through a fully searchable data dictionary and variable search tool (<http://www.bristol.ac.uk/alspac/researchers/our-data/>). Study data gathered from participants at 22 years and onwards were collected and managed using REDCap electronic data capture tools hosted at the University of Bristol (Harris et al., 2009). REDCap (Research Electronic Data Capture) is a secure, web-based software platform designed to support data capture for research studies.

Individuals were included in the current study when data were available on week, month and year of birth, together with their school year at any given assessment age; participants were excluded if they were not alive at 1 year of age ($N = 688$), were not in the expected school year given their chronological age ($N = 12$), and if they were the younger of a twin pair ($N = 186$). After excluding these participants, 14,643 individuals were included in the analyses.

Ethical considerations

Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. Informed consent for the use of data collected via questionnaires and clinics was obtained from participating families following the recommendations of the ALSPAC Ethics and Law Committee at the time.

Relative age

Children's exact birth dates were not provided due to the risk of deanonymisation. Children's birth dates were grouped into one 8-day block (1–8 September) and 51 consecutive 7-day blocks ending on 31 August. Children were assigned a score reflecting week of birth relative to the academic year (range 0 (Oldest, 1 September)–51 (Youngest, up to 31 August)).

Child mental health

Two well-validated child and adolescent mental health screening questionnaires were administered at several timepoints between 4 and 25 years; the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) and the Short Mood and Feelings Questionnaire (SMFQ; Angold et al., 1995). We chose to use SDQ and SMFQ data continuously for statistical (increased power) and conceptual reasons. Previous research suggests that mental health and neurodevelopmental conditions such as depression and ADHD lie at the end of a continuous distribution of underlying symptom traits, with similar aetiological and outcome profiles (Thapar, 2018; Thapar, Collishaw, Pine, & Thapar, 2012). Figure 1 shows the timepoints for administration. We present further details of these questionnaires in the Supporting Information (Appendix S1).

Covariates

An advantage of regression discontinuity methods is that they do not rely on controlling for confounders as in usual observational studies (Oldenburg, Moscoe, & Bärnighausen, 2016), it is still important to test assumption violations such as unbalanced covariate distribution. We included various early pre-assignment covariates:

Maternal background: Age of mother at birth, mother's education (highest qualification), maternal depression at

18 weeks gestation (Edinburgh Postnatal Depression Scale; Cox, Holden, & Sagovsky, 1987).

Pregnancy and birth: Caesarean birth, birth size (i.e., single or multiple birth), birthweight, gestational age, maternal alcohol use during last 2 months of pregnancy and maternal smoking during pregnancy.

Child factors: household crowding, child ethnic background, home ownership, parity and Sex.

We controlled for age of child at questionnaire completion (in months, to nearest month). The questionnaires were sent out to parents to map on to certain timepoints throughout development. The age at which questionnaires was filled in had a wide spread across participants and was not different between August born and September born children (see Table S1). For further information relating to children's ages and questionnaire timepoints, we include a table in the Supporting Information (Table S1).

Design

We used a regression discontinuity design to compare the relative risk of mental health problems by relative age (Hilton Boon, Craig, Thomson, Campbell, & Moore, 2021; Moscoe et al., 2015; Oldenburg et al., 2016; Venkataramani, Bor, & Jena, 2016). The exposure is 'age at starting school' and the running variable is 'week of birth', with the discontinuity on 1 September. We assumed a 'sharp' RD design given the strict cut-off date for school year selection in England and Wales; we assumed (based on LEA rules at the time) that children started school at the same time and during the school year they turn 5 years old, and that schools and LEAs adhere to this cut-off. Regression discontinuity relies on the assumptions that:

1. The decision rule and cut-off value are known.
2. There are no confounders of age at starting school and outcome that are related to the cut-off.
3. The outcome (risk of mental health problems) would not show a discontinuity at the threshold in the absence of the exposure (being admitted to school) (Moscoe et al., 2015).

Statistical analysis

We checked whether regression discontinuity design assumptions were met by comparing covariate distributions across the school year to ensure that the cut-off is not associated with any other variables besides relative age. We checked distributions of maternal depression, maternal age, gestation, birthweight, birth size, pre-natal alcohol use, smoking, caesarean status, crowding, home ownership status, mother's education and parity across the cut-off. We also checked the distribution of births across the months of the school year. We carried out the analysis in three ways; first, one where the exposure is continuous throughout the school year and there is no selection window near the cut-off (henceforth, 'no bandwidth'), second, restricted to those born 4 weeks either side of the 1st September cut-off ('4 weeks') and lastly, restricted to those born 8 weeks either side of the cut-off ('8 weeks'). The 4-week and 8-week bandwidths only compare those with birthdates within those time windows around 1st September. We chose these bandwidths because in narrower bandwidths (e.g., 4 weeks) the assumption of no confounding is more plausible. We viewed the estimation of effects of relative age as local randomness near the cut-off, by limiting analysis to observations that lie within one bandwidth on either side of the 1st September cut-off (i.e., 4 weeks either side of the cut-off, in the case of a 4-week bandwidth). After bandwidth selection, we fit local linear regressions on observations within the bandwidth to estimate the effect of relative age.

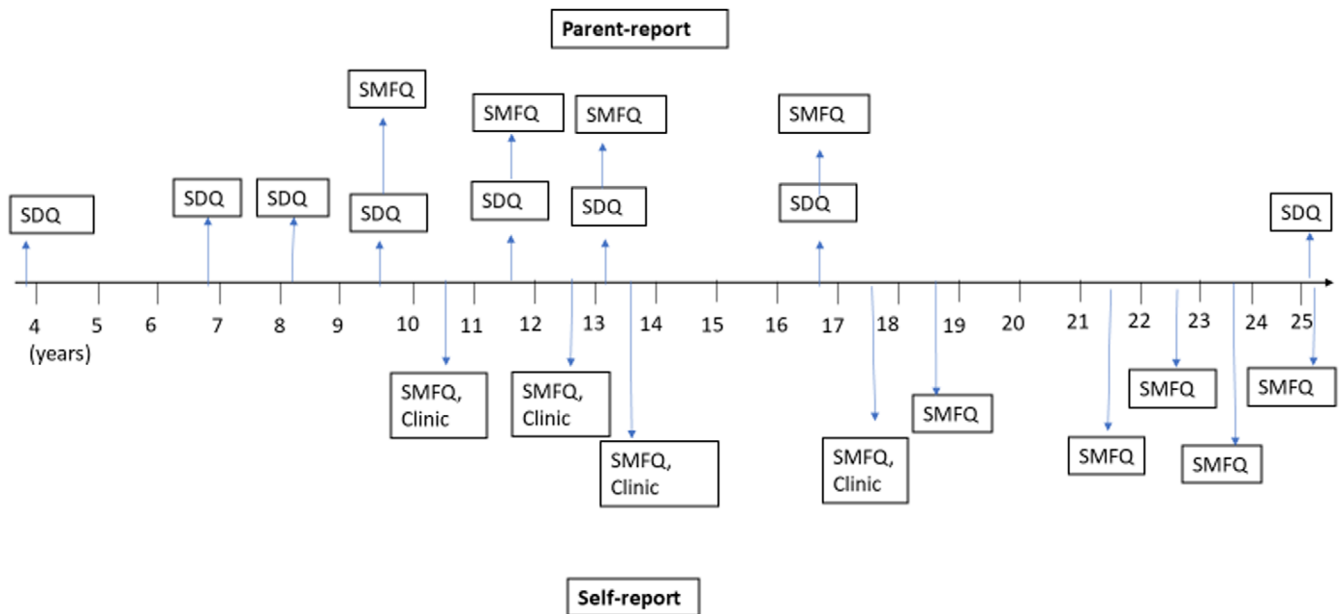


Figure 1 Timeline of ALSPAC mental health measures

Previous studies have used similar bandwidths (Crawford et al., 2014). A bandwidth of 8 weeks corresponds to approximately the length of a school half-term. A problem with using longitudinal cohort data is participant attrition over time; at age 4 years there were 9,312 participants, decreasing to 4,076 at 25 years. We compared complete cases with those who had incomplete data and then we used multiple imputation by chained equations (see Appendix S2 and Table S2 for further information) to account for selective attrition. We included maternal depression, maternal age, gestation, birthweight, birth size, alcohol use in last 2 months of pregnancy, smoking, caesarean status, crowding, home ownership status, mother's education and parity in the imputation models. We imputed to a maximum sample of those with at least one measure for each mental health outcome (N SDQ = 11,116; N Self SMFQ = 9,468; N Parent SMFQ = 9,146). We then used linear regressions for each outcome on age at starting school, then also adjusted for all abovementioned covariates. We checked all models for Monte Carlo error following guidelines (White, Royston, & Wood, 2011), further details in the Supporting Information (Appendix S2 and Table S2).

To test robustness of findings, we also report findings from a complete-case analysis at each outcome point.

We also used a generalised estimating equation (GEE) approach to model our outcomes (parent-rated SDQ, self-rated SMFQ and parent-rated SMFQ) by relative age, and other covariates, as a sensitivity analysis. The advantages of this approach are that GEE models have some robustness to attrition, but do not use imputation (i.e., only analyse the available data) (Liang & Zeger, 1986).

Secondary analyses tested SDQ subscales by relative age, and potential interactions by child sex. All analyses were conducted using Stata (v16.1 SE, StataCorp LLC, College Station, TX).

Results

Testing assumptions of the regression discontinuity design – covariate pattern by month of birth

Tables S3 and S4 display demographic characteristics. First, we found similar patterns of covariate distribution between children born August versus September, showing no discontinuous relationship between 'pre-

treatment' covariates and relative age. Second, we found a similar distribution of births in the study sample across the months of the year (Figure S1), so we found no evidence of a discontinuity in distribution of births around 1st September. Lastly, we found similar patterns of missing data (Tables S5–S10). Thus, we found no evidence that the regression discontinuity design assumptions, listed above, were violated.

Comparison of participants with complete versus incomplete data

Participants with complete records at all ages and assessments were more likely to be female, first-born, white, with parents who are older, non-smokers, higher educated and less depressed than those with incomplete data (Table S11). This suggests that a complete-case analysis approach would be biased (Lee et al., 2021).

Descriptive data on mental health outcome measures

Table S12 presents descriptive, unstandardised data on mental health outcome variables for the whole sample (with available data at a given time point), and for children born in August and September. Table S13 presents the same data for individuals included in the 4-week and 8-week bandwidths. For each measure, higher scores indicate greater reported SDQ scores, which are indicative of greater risk of mental health problems.

Relative age

As shown in Figure 2 (Parent-rated general mental health) we found no evidence of an effect of relative

age on parent-rated mental health before entry into school [age 4 Standardised Mean Difference (SMD): .02, 95% CI: (-.05, .08)]. At the earliest point after school entry (7 years) we found that a 1-year decrease in relative age in the school year was associated with a difference of approximately one-sixth of a standard deviation in SDQ total difficulties [SMD: .15, 95% CI: (.08, .22)]. Being relatively young in the school year was associated with higher SDQ total difficulty scores, indicative of poorer parent-rated child mental health. These differences persisted throughout the school years, with the strongest effect at 11 years [SMD: .22, 95% CI: (.15, .29)]. These differences attenuated to the null at 25 years [SMD: -.01, 95% CI: (-.1, .09)]. Results were materially unchanged after adjusting for covariates.

As shown in Figure 3 (self-rated depression symptoms) we found that younger children in the school year are more likely to report self-rated depression symptoms at 14 years [SMD: .12, 95% CI: (.04, .20)] and at 25 years of age [SMD: .14, 95% CI: (.04, .23)], but there was no evidence of difference by relative age at other ages.

Figure 4 (parent-rated depression) shows that relatively young children scored higher in parent-rated depression at 9 years [SMD: .12, 95% CI: (.05, .19)] and 11 years [SMD: .16, 95% CI: (.09, .23)]. We found no other evidence of differences by relative age on parent-rated SMFQ scores.

Regression tables of our findings are also available in the Supporting Information (Tables S14–S22).

Sensitivity analysis

We found that most of the identified relative age effects on parent-rated SDQ scores are materially unchanged when restricted to children born 4 weeks either side of the cut-off, except for the association at 16 years attenuating to the null. Additionally, restricting the bandwidths to those born in the months closest to the cut-off yielded wider 95% confidence intervals. For the self-rated SMFQ, we found no evidence of differences within the 4-week bandwidth, except at 25 years.

The complete-case results resembled the analysis using multiple imputation, with some differences.

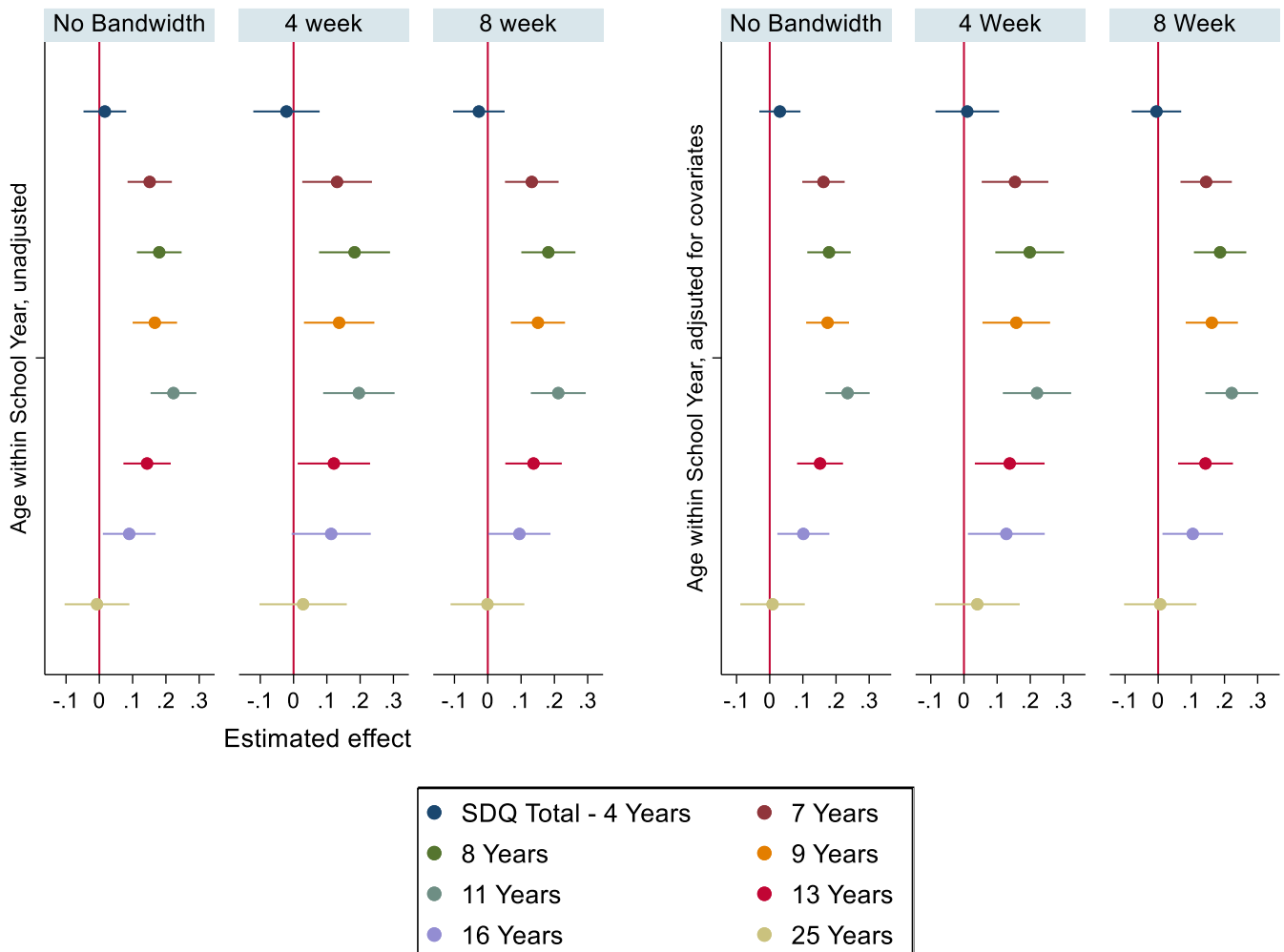


Figure 2 Parent-rated Mean Standardised SDQ total difficulties coefficient plots, unadjusted models (left), models adjusted for covariates (right), imputed data. 'Estimated effect' represents SMD by relative age. 'No Bandwidth' = All participants included; '4 Week' = Restricted to participants born up to 4 weeks either side of 1st September Cut-off; '8 week' = Restricted to participants born 8 weeks either side of the 1st September cut-off. $N = 11,116$

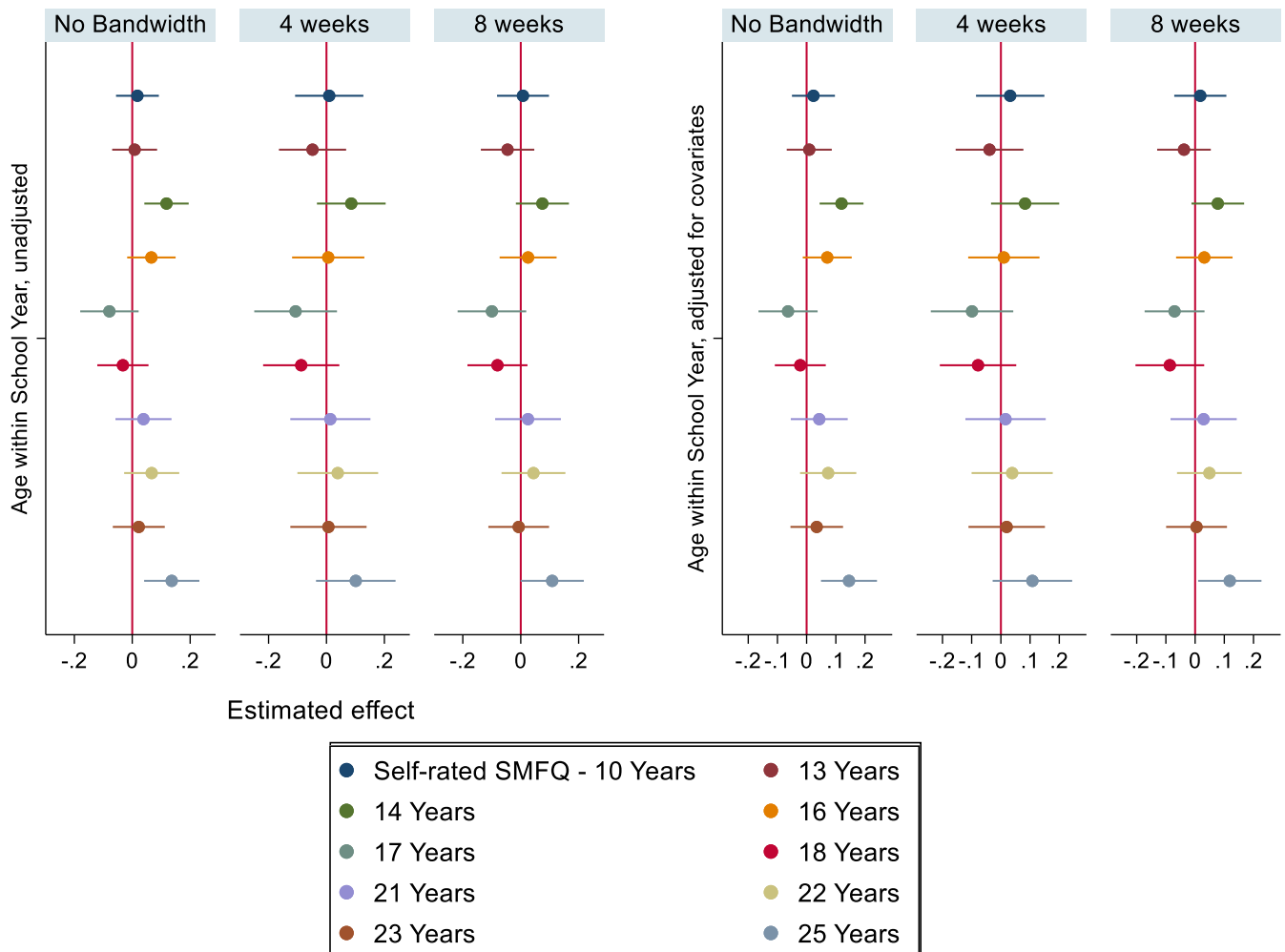


Figure 3 Self-rated SMFQ coefficient plots – unadjusted models (left), models adjusted for covariates (right). Imputed data. Estimated effect represents SMD by relative age. $N = 9,468$

We present these findings in Tables S23–S25. Complete-case analysis identified effects of relative age on parent-rated SMFQ scores at 16 years [SMD: .10 (.01, .20)]; these attenuated to the null in the multiple imputation analysis. The multiple imputation analysis showed an effect of relative age on self-rated SMFQ scores at 25 years [SMD: .16 (.06, .25)], in contrast to complete-case analyses [SMD: .09 (–.02, .20)], albeit with overlapping confidence intervals.

As an alternative approach to multiple imputation, we produced General Estimating Equations (GEE) models of relative age effects on all outcomes. We found that the GEE models strongly support the models using imputed data (Tables S26–S28).

Secondary analyses

SDQ subscales

We explored SDQ subscale scores by relative age. All subscale regression model information is provided in Table S14. We found that the hyperactivity subscale followed the same pattern as the total difficulties

scores, in addition to showing the largest SMDs by relative age of all subscales, up to .25 of a standard deviation at age 11 years [95% CI: .18, .32]. Parent-rated emotional problems and peer problems also showed effects by relative age in the same direction [SMD emotional at 7 years: .13; (95% CI: .06, .20); SMD peer at 11 years; .15; (95%CI: .08, .22)]. We did not find consistent relative age differences for conduct problems.

Interactions by sex

We found no consistent difference in relative age effect sizes for males and females, for the SDQ and both versions of the SMFQ (available on request).

Discussion

The present study aimed to investigate the effect of relative age in the school year on risk of mental health problems in a general population longitudinal cohort studied across childhood, adolescence and into young adulthood. Using the fact that relative age corresponds to a regression discontinuity design, we

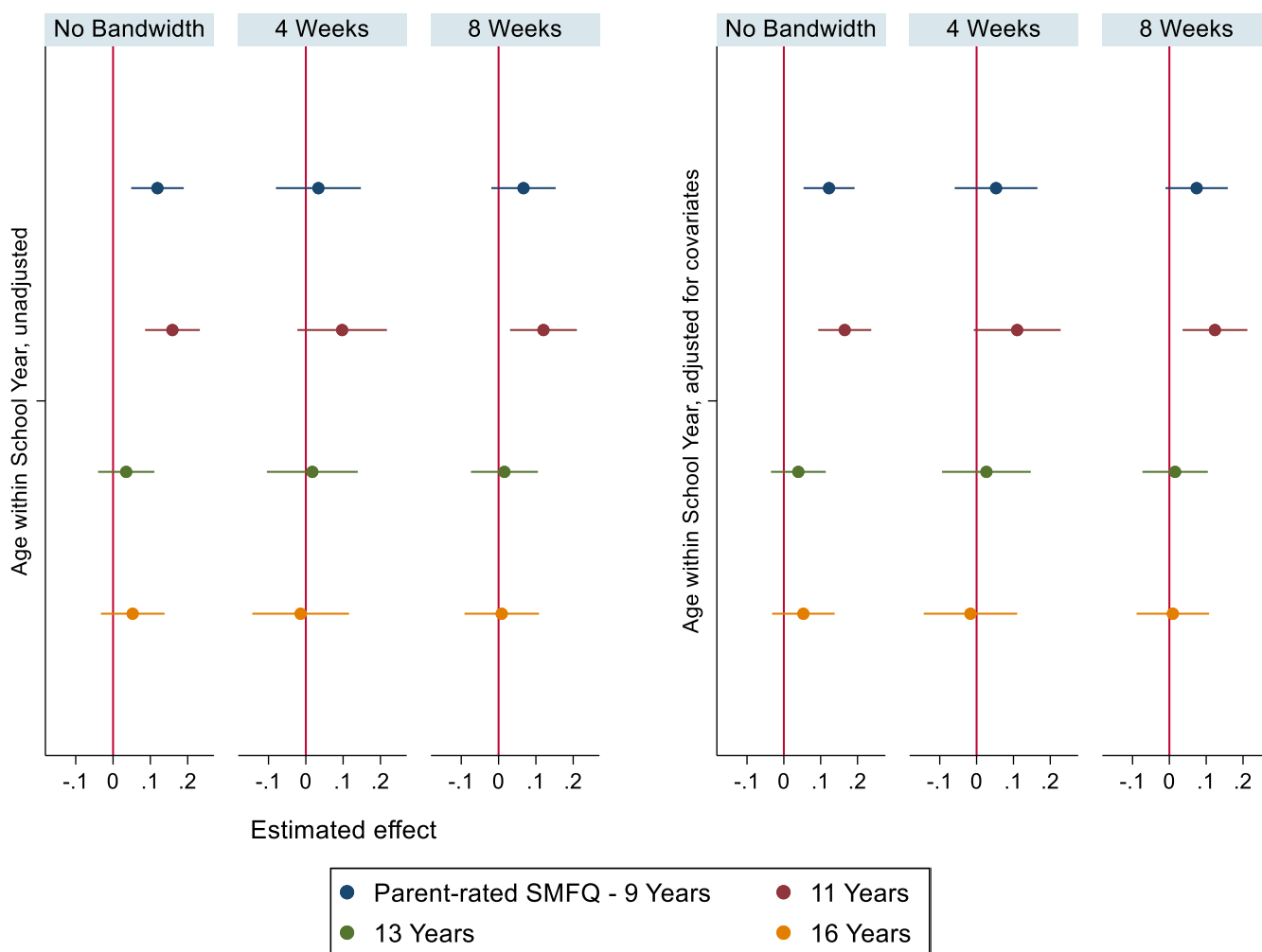


Figure 4 Parent-rated SMFQ coefficient plots – unadjusted models (left), models adjusted for covariates (right). Imputed data. Estimated effect represents SMD by relative age. $N = 9,146$

found that the youngest children in the academic year have greater parent-rated general risk of mental health problems, and that these findings were materially unchanged after adjusting for covariates. We found that the hyperactivity subscale accounts for the largest differences by relative age, with emotional problems and peer problems also contributing. Further, the findings support our prior hypothesis that relative age differences are only present after school entry.

We also found that relative age effects persist into secondary education, but wane after young people leave school. The largest differences by relative age were at 11 years. This differs from our prior hypothesis that relative age differences are strongest at the earliest measurement after school entry, and then subsequently weaken as children develop, on the basis that the difference of up to 1 year between children becomes increasingly less noticeable by chronological age. The transition to secondary school appears to be a particular period of vulnerability for relatively young children. We suggest that relative age effects attenuate to the null by adulthood as the gap between oldest and youngest in the year

becomes proportionately less as children age chronologically. The gap between the oldest and youngest is up to 1 year. At the beginning of school, ages 4–5 years, the oldest children will be approximately 20% more psychologically and physiologically mature than a child born in August. By age 16, this is reduced to 6.25%. We also expected to find relative age effects on specific measures of depression. However, while we did find some differences between the oldest and youngest children, we found no consistent pattern of effects across development; the majority of the self-rated and parent-rated SMFQ measurements showed no meaningful relationship between relative age and depression symptoms.

Investigating potential mechanisms through which relative age affects mental health was beyond the scope of the present study, but previous research shows that younger children in the school year have poorer education attainment and are more likely to be bullied (Crawford et al., 2013; Mühlenweg, 2010). Both are associated with increased risk of mental health problems in children and young people (Klomek, Sourander, & Elonheimo, 2015). Therefore, a further question for research to consider is whether

bullying influences the relationship between relative age and mental health. A second question for further research is whether pubertal changes influence this relationship (Copeland, Worthman, Shanahan, Costello, & Angold, 2019). A third suggestion is exploring whether relative age has differential effects on anxiety vs depression symptoms. We found relative age effects on the parent-rated SDQ emotional symptom subscale, which includes both anxiety and depression symptoms, but found no consistent effects for the SMFQ, which specifically measures depressive symptom traits. Fourth, research shows that the youngest children in the school year are perceived to have less developed language skills relative to older peers (Norbury et al., 2016). Further research is needed on whether language development mediates relative age effects on risk of mental health problems.

Lastly, identifying relative age effects in more vulnerable groups, such as children with neurodevelopmental disorders (Addicoat, Thapar, Riglin, Thapar, & Collishaw, 2019) and exploring the extent to which neurodevelopmental disorder traits moderate the relationship between relative age and mental health would be an interesting direction for future research. This is because the hyperactivity subscale, a measure of ADHD symptoms, showed the strongest relative age effects in the analysis. Relatively young children are more likely to be diagnosed with neurodevelopmental disorders such as ADHD (Root et al., 2019), but it is not known whether relative age effects are especially pronounced among children with neurodevelopmental disorders, and whether children with neurodevelopmental disorders who are relatively young constitute a particularly high-risk group that would warrant additional support as they start school. We used the same measure (the SDQ) across development, using the same rater (the parent). Evidence supports the reliability and validity of the SDQ as a measure of mental health in children and adolescents (Goodman, 1997, 2001) and emerging evidence supports its use in young adulthood (Riglin et al., 2021). However, further evidence is needed assessing measurement invariance of the SDQ and its subscales across the age range covered by this study.

Our findings support Patalay et al.'s (2015) findings that relative age influences emotional problems in children, and Crawford et al.'s findings that relatively young children show poorer social and emotional development as reflected by poorer outcomes on the SDQ. However, Crawford et al. (2013) found that parent-rated differences in SDQ scores by relative age are not present beyond the age of 9 years in the ALSPAC cohort, whereas in the same cohort we find that relative age effects persist up to the age of 16 years. However, in both studies, estimates were in the same direction. In contrast to Crawford et al., we used varying approaches to account for participant attrition (which became more

pronounced as children grew older) and a more precise measure of relative age (in weeks). The two studies also differ in the covariates included.

The present study adds to previous findings by investigating the relationship between relative age and mental health in the ALSPAC cohort both prior to and during school age, as well as extending to after young people they have left school. We also add to previous findings by testing relative age effects on a specific measure of depression symptom traits (SMFQ), and by using week of birth as a more precise measurement of age within school year. We used imputed data and longitudinal analyses to attempt to control for biases that may arise from missing data resulting from attrition in the ALSPAC cohort, and we used alternative selection bandwidths as sensitivity analyses. The strengths of the study are that we used rich data from a single longitudinal population cohort with data collected throughout development, and the study used consistent and widely implemented multi-informant measures of mental health.

We checked for violations of the RD design, there was no evidence that covariates act as confounders because there was no inequality in distribution across the school entry cut-off. Therefore, this is suggestive of a causal effect of young relative age on mental health in school age children. We are confident that our findings may represent causality given that regression discontinuity approaches can estimate causal effects of 'treatment' (i.e., being relatively old for the school year), when other experimental methods such as randomised controlled trials are not feasible (Moscoe et al., 2015; Venkataramani et al., 2016). A causal interpretation relies on the assumption that, apart from relative age, individuals born on either side of the 1st September cut-off rule, but close to it, are similar in other characteristics that are not affected by this rule. If this assumption is not violated, then we can infer that individuals close to the cut-off are effectively randomly selected for 'treatment'. We have demonstrated that there is no evidence that the assumption does not hold in the present study (see Table S3). Therefore, we can be confident that the increase in risk of mental health problems in the relatively youngest individuals was the result of being young in the school year.

This study has limitations. First, data missingness due to participant attrition may result in bias due to differences between participants who are retained versus drop out. We attempted to mitigate this by using a multiple imputation approach to missing data, however, if children at greater risk of mental health problems are more likely to drop out of the study, the use of multiple imputation will not remove all the bias (Lee et al., 2021). The pattern of findings was also the same when using GEE models – a different method for dealing with attrition. Second, there was a long gap between some measurements,

including a 9-year gap between the last two SDQ measurements, so we did not know precisely when effects attenuated to the null.

Third, we did not consider participants' childcare arrangements before school entry; some children may have experienced more formal pre-school settings, and we did not test if effects differed compared to children without these childcare arrangements. Similarly, we did not consider the type of schooling children entered, and whether classes were comprised of single or multiple academic year groups taught together. Lastly, we did not know exactly when the children started school, which would lead to a 'fuzzy' regression discontinuity design. In the UK, education authorities often offer some flexibility about when in the calendar year children start their first year of school (e.g., at the start of the autumn or spring terms) with a view to helping ensure school readiness including among summer born children. Nevertheless, summer born children typically remain the youngest in their school year, and further research is needed to assess whether there are impacts of delayed school entry on children's mental health.

Implications

The effect of relative age on risk of mental health problems is modest at an individual level but may be larger at a population level. It appears that the effects of relative age may specifically impact children of school age; therefore, any intervention should be applied at this period. Relative age effects may attenuate to the null by adulthood, but evidence suggests there are links with broader psychosocial outcomes, such as educational attainment and employment (Lopez-Lopez et al., 2019).

Previous authors have suggested school admissions and entry system changes, including delaying school entry (Dee & Sievertsen, 2018) or age-based assessment adjustments (Crawford et al., 2013). Changes to established school structures and admission systems may be difficult to implement in practice because of the likelihood of increased disruption for schools, teachers and families. Evidence is mixed on the benefits of this practice to children (Dhuey, Figlio, Karbownik, & Roth, 2019), and more evidence is required on whether age adjustments to grades and examinations counteract relative age effects. It is important to remember that in any classroom some children will be younger than others, even when combining more than 1-year group together. We are cautious about suggesting what the effect of relative age on clinical aspects of mental health problem risk looks like in practice, given the modest size of these effects, and that we did not consider risk of clinically diagnosed mental health problems. Being young for the school year may not affect individuals equally, and it is important to further investigate relative age effects on

mental health using measures that consider clinical impacts both in terms of presentation of symptoms and their impacts on functioning.

An alternative approach may involve organising the school register by age within school year to raise awareness of who is relatively young within the classroom (Norbury et al., 2016) which may facilitate differentiated instruction and assessment. We suggest that schools become more aware of relative age effects on education achievement and mental health on their pupils.

Conclusion

The present study provides a long-term longitudinal follow-up examining relative age effects across development using consistent measures of mental health. The study assumptions of the regression discontinuity design were examined for plausibility and adjusted for observed covariates, suggesting a causal relationship between relative age and mental health. We found that younger children in the school year have higher parent-rated general risk of mental health problems compared to their older peers. Young relative age may be a potentially modifiable causal factor for child and adolescent risk of mental health problems, but we did not consistently find the same effect for depression.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article:

Appendix S1. Questionnaires used to measure risk of mental health problems in the present study.

Appendix S2. Multiple imputation.

Table S1. Table summary of ages of ALSPAC participants at completion of questionnaires. Note IQR = Interquartile Range.

Table S2. Multiple Imputation variables, models used, and percentage of data missing from these variables/models.

Table S3. Demographic characteristics of the study sample.

Table S4. Demographic information for participants born 4 ("4 weeks") and 8 ("8 weeks") weeks either side of the September 1st Cut-off.

Table S5. Missing table patterns of SDQ total difficulties.

Table S6. Missing table patterns of SDQ total difficulties.

Table S7. Missing table patterns of self-rated SMFQ.

Table S8. Missing table patterns of self-rated SMFQ.

Table S9. Missing table patterns of parent rated SMFQ.

Table S10. Missing table patterns of parent rated SMFQ.

Table S11. Logistic regressions of covariates on being a complete-case (i.e., having all data at all mental health timepoints).

Table S12. Descriptive statistics of outcome measures in August and September born children (Parent-rated SDQ Total Difficulties, self-rated and parent-rated SMFQ, unstandardized, complete-case data).

Table S13. Descriptive statistics of outcome measures (Parent-rated SDQ Total Difficulties, self-rated and parent-rated SMFQ, unstandardized, complete-case data; restricted to 4 week bandwidth (left) and 8 week bandwidth (right)).

Table S14. Regression results for SDQ total difficulties and subscales by relative age, Imputed data ($N = 11,116$).

Table S15. Regression results for Self-rated SMFQ by relative age, Imputed data ($N = 9,468$).

Table S16. Regression results for parent-rated SMFQ by relative age, Imputed data ($N = 9,164$).

Table S17. Regression results for parent-rated SDQ total difficulties scores by relative age, restricted to 4 weeks either side of September 1st cut-off, Imputed data ($N = 2,035$).

Table S18. Regression results for Self-rated SMFQ by relative age, restricted to 4 weeks either side of September 1st cut-off, Imputed data ($N = 1,690$).

Table S19. Regression results for parent-rated SMFQ by relative age, restricted to 4 weeks either side of September 1st cut-off, Imputed data ($N = 1,642$).

Table S20. Regression results for parent-rated SDQ total difficulties scores by relative age, restricted to 8 weeks either side of September 1st cut-off, Imputed data ($N = 4,042$).

Table S21. Regression results for Self-rated SMFQ by relative age, restricted to 8 weeks either side of September 1st cut-off, Imputed data ($N = 3,383$).

Table S22. Regression results for parent-rated SMFQ by relative age, restricted to 8 weeks either side of September 1st cut-off, Imputed data ($N = 5,297$).

Table S23. Complete-case analysis. Outcome: Parent-Rated SDQ (4–25 years) Independent variable: Relative Age (week of birth).

Table S24. Complete-case analysis: Self-Rated SMFQ (10–25 years) Independent variable: Relative Age (week of birth), Coefficient represents difference in standardised self-report SMFQ score between children born between 1st September–31st August.

Table S25. Complete-case analysis– outcome: Parent-Rated SMFQ (9–16 years) Independent variable:

Relative Age (week of birth), Coefficient represents difference in standardised parent-report SMFQ score between children born between 1st September–31st August.

Table S26. Generalized estimating equation (GEE) results for parent-rated SDQ total difficulties ($N = 11,116$).

Table S27. Generalized estimating equation (GEE) results for self-rated SMFQ scores ($N = 9,468$).

Table S28. Generalized estimating equation (GEE) results for parent-rated SMFQ scores ($N = 9,164$).

Figure S1. Histogram of birthdates by month. Red lines denote a period of one complete school year (September 1991–August 1992).

Acknowledgements

The authors are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council and Wellcome (Grant ref: 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors and Thomas Broughton, Kate Langley, Kate Tilling and Stephan Collishaw will serve as guarantors for the contents of this paper. A comprehensive list of grants funding is available on the ALSPAC website (<http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf>). The primary outcome measures used in the paper were specifically funded by the Wellcome Trust (204895/Z/16/Z) for age 25 data. This work was undertaken as part of a PhD scholarship funded by Cardiff University Medical School. S.C. is grateful for support from the Wolfson Foundation. The authors have declared that they have no competing or potential conflicts of interest.

Correspondence

Thomas Broughton, Hadyn Ellis Building, Maindy Road, Cardiff CF24 4HQ, UK;
Email: broughtontr1@cardiff.ac.uk

Key points

- Younger relative age in the school year is associated with increased risk of mental health problems.
- School-specific effect of age within school year on risk of mental health problems, with the peak difference occurring at 11 years, attenuating to the null in adulthood.
- Implications for child and adolescent mental health services, educators and education policymakers.

Data Availability Statement

The informed consent obtained from ALSPAC participants does not allow for the data to be made freely available through any third party maintained public

repository. However, data used for this submission can be made available on request to the ALSPAC Executive. The ALSPAC data management plan describes in detail the policy regarding data sharing, which is through a system of managed open access. Full instructions for

applying for data access can be found here: <http://www.bristol.ac.uk/alspac/researchers/access/>. The ALSPAC study website contains details of all the data that are available (<http://www.bristol.ac.uk/alspac/researchers/our-data/>)

References

- Addicoat, A., Thapar, A.K., Riglin, L., Thapar, A., & Collishaw, S. (2019). Adult mood problems in children with neurodevelopmental problems: Evidence from a prospective birth cohort followed to age 50. *Social Psychiatry and Psychiatric Epidemiology*, 55, 351–358.
- Angold, A., Costello, E., Messer, S., Pickles, A., Winder, F., & Silver, D. (1995). The development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *International Journal of Methods in Psychiatric Research*, 5, 1–12.
- Borja, M.C., & Martin, P. (2017). What is the most popular birthday in England and Wales? *Significance*, 14, 6–7.
- Boyd, A., Golding, J., Macleod, J., Lawlor, D.A., Fraser, A., Henderson, J., ... & Davey Smith, G. (2013). Cohort profile: The 'children of the 90s' – The index offspring of the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology*, 42, 111–127.
- Collishaw, S., Goodman, R., Ford, T., Rabe-Hesketh, S., & Pickles, A. (2009). How far are associations between child, family and community factors and child psychopathology informant-specific and informant-general? *Journal of Child Psychology and Psychiatry*, 50, 571–580.
- Copeland, W.E., Worthman, C., Shanahan, L., Costello, E.J., & Angold, A. (2019). Early pubertal timing and testosterone associated with higher levels of adolescent depression in girls. *Journal of the American Academy of Child and Adolescent Psychiatry*, 58, 1197–1206.
- Cox, J.L., Holden, J.M., & Sagovsky, R. (1987). Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *The British Journal of Psychiatry*, 150, 782–786.
- Crawford, C., Dearden, L., & Greaves, E. (2013). *When you are born matters: Evidence for England*. Available from: <https://doi.org/10.1920/re.ifs.2013.0080>
- Crawford, C., Dearden, L., & Greaves, E. (2014). The drivers of month-of-birth differences in children's cognitive and non-cognitive skills. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 177(4), 829–860. Portico. <https://doi.org/10.1111/rssa.12071>
- Dee, T.S., & Sievertsen, H.H. (2018). The gift of time? School starting age and mental health. *Health Economics*, 27, 781–802.
- Dhuey, E., Figlio, D., Karbownik, K., & Roth, J. (2019). School starting age and cognitive development. *Journal of Policy Analysis and Management*, 38, 538–578.
- Fraser, A., Macdonald-Wallis, C., Tilling, K., Boyd, A., Golding, J., Davey Smith, G., ... & Lawlor, D.A. (2013). Cohort profile: The Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *International Journal of Epidemiology*, 42, 97–110.
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, 38, 581–586.
- Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40, 1337–1345.
- Goodman, R., Gledhill, J., & Ford, T. (2003). Child psychiatric disorder and relative age within school year: Cross sectional survey of large population sample. *BMJ*, 327, 472–470.
- Harris, P.A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J.G. (2009). Research electronic data capture (REDCap) – A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, 42, 377–381.
- Hilton Boon, M., Craig, P., Thomson, H., Campbell, M., & Moore, L. (2021). Regression discontinuity designs in health: A systematic review. *Epidemiology*, 32, 87–93.
- Holland, J., & Sayal, K. (2019). Relative age and ADHD symptoms, diagnosis and medication: A systematic review. *European Child & Adolescent Psychiatry*, 28, 1417–1429.
- Klomek, A.B., Sourander, A., & Elonheimo, H. (2015). Bullying by peers in childhood and effects on psychopathology, suicidality, and criminality in adulthood. *The Lancet Psychiatry*, 2, 930–941.
- Lee, K.J., Tilling, K.M., Cornish, R.P., Little, R.J.A., Bell, M.L., Goetghebeur, E., ... & Carpenter, J.R. (2021). Framework for the treatment and reporting of missing data in observational studies: The treatment and reporting of missing data in observational studies framework. *Journal of Clinical Epidemiology*, 134, 79–88.
- Liang, K.-Y., & Zeger, S.L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73, 13–22.
- Lopez-Lopez, J.A., Kwong, A.S.F., Washbrook, E., Pearson, R.M., Tilling, K., Fazel, M.S., ... & Hammerton, G. (2019). Trajectories of depressive symptoms and adult educational and employment outcomes. *BJPsych Open*, 6, e6.
- Moscoe, E., Bor, J., & Bärnighausen, T. (2015). Regression discontinuity designs are underutilized in medicine, epidemiology, and public health: A review of current and best practice. *Journal of Clinical Epidemiology*, 68, 122–133.
- Mühlenweg, A. (2010). Young and innocent: International evidence on age effects within grades on victimization in elementary school. *Economics Letters*, 109, 157–160.
- Norbury, C.F., Gooch, D., Baird, G., Charman, T., Simonoff, E., & Pickles, A. (2016). Younger children experience lower levels of language competence and academic progress in the first year of school: Evidence from a population study. *Journal of Child Psychology and Psychiatry*, 57, 65–73.
- Northstone, K., Lewcock, M., Groom, A., Boyd, A., Macleod, J., Timpson, N., & Wells, N. (2019). The Avon Longitudinal Study of Parents and Children (ALSPAC): An update on the enrolled sample of index children in 2019 [version 1; peer review: 2 approved]. *Wellcome Open Research*, 4, 51.
- Oldenburg, C.E., Moscoe, E., & Bärnighausen, T. (2016). Regression discontinuity for causal effect estimation in epidemiology. *Current Epidemiology Reports*, 3, 233–241.
- Patalay, P., Belsky, J., Fonagy, P., Vostanis, P., Humphrey, N., Deighton, J., & Wolpert, M. (2015). The extent and specificity of relative age effects on mental health and functioning in early adolescence. *The Journal of Adolescent Health*, 57, 475–481.
- Riglin, L., Agha, S. S., Eyre, O., Bevan Jones, R., Wootton, R. E., Thapar, A. K., Collishaw, S., Stergiakouli, E., Langley, K., & Thapar, A. (2021). Investigating the validity of the Strengths and Difficulties Questionnaire to assess ADHD in young adulthood. *Psychiatry Research*, 301, 113984. <https://doi.org/10.1016/j.psychres.2021.113984>
- Root, A., Brown, J.P., Forbes, H.J., Bhaskaran, K., Hayes, J., Smeeth, L., & Douglas, I.J. (2019). Association of relative age in the school year with diagnosis of intellectual disability, attention-deficit/hyperactivity disorder, and depression. *JAMA Pediatrics*, 173, 1068–1075.
- Sadler, K., Vizard, T., Ford, T., Goodman, A., Goodman, R., & McManus, S. (2018). *Mental health of children and young people in England, 2017: Trends and characteristics*. Leeds, UK: NHS Digital.
- Thapar, A. (2018). Discoveries on the genetics of ADHD in the 21st century: New findings and their implications. *The American Journal of Psychiatry*, 175, 943–950.
- Thapar, A., & Rutter, M. (2019). Do natural experiments have an important future in the study of mental disorders? *Psychological Medicine*, 49, 1079–1088.

- Thapar, A., Collishaw, S., Pine, D.S., & Thapar, A.K. (2012). Depression in adolescence. *The Lancet*, *379*, 1056–1067.
- Thistlethwaite, D., & Campbell, D. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology*, *51*, 309–317.
- Venkataramani, A.S., Bor, J., & Jena, A.B. (2016). Regression discontinuity designs in healthcare research. *BMJ*, *352*, i1216.
- Vos, T., Abajobir, A.A., Abate, K.H., Abbafati, C., Abbas, K.M., Abd-Allah, F., ... & Murray, C.J.L. (2017). Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *The Lancet*, *390*, 1211–1259.
- White, I.R., Royston, P., & Wood, A.M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine*, *30*, 377–399.

Accepted for publication: 1 June 2022