



OPEN ACCESS

Association between breastfeeding duration and educational achievement in England: results from the Millennium Cohort Study

Reneé Pereyra-Elías , Claire Carson , Maria A Quigley

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/archdischild-2022-325148>).

National Perinatal Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK

Correspondence to

Dr Reneé Pereyra-Elías, National Perinatal Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, OX1 2JD, UK; renee.pereyraelias@balliol.ox.ac.uk

Received 16 November 2022
Accepted 29 March 2023

ABSTRACT

Objective To evaluate the association between breastfeeding duration and educational outcomes at the end of secondary education among children from the Millennium Cohort Study.

Design Cohort study comparing school results at age 16 according to breastfeeding duration.

Setting England.

Participants Children born in 2000–2002 (nationally representative sample).

Exposure Self-reported breastfeeding duration (categorised).

Main outcome measures Standardised school assessments taken at the end of secondary education (General Certificate of Secondary Education (GCSEs), marked 9–1) in English and Mathematics, categorised as: 'fail, marks <4', 'low pass, marks 4–6' and 'high pass, marks ≥7 (equivalent to A–A*)'. Additionally, overall achievement was measured using the 'attainment 8' score (adding the marks of eight GCSEs, English and Mathematics double weighted; 0–90).

Results Approximately 5000 children were included. Longer breastfeeding was associated with better educational outcomes. For example, after full adjustment for socioeconomic markers and maternal cognitive ability, in comparison with children who were never breastfed, those who were breastfed for longer were more likely to have a high pass in their English and Mathematics GCSEs, and less likely to fail the English GCSE (but not the Mathematics GCSE). Additionally, compared with those never breastfed, those breastfed for at least 4 months had, on average, a 2–3 point higher attainment 8 score (coefficients: 2.10, 95% CI 0.06 to 4.14 at 4–6 months; 2.56, 95% CI 0.65 to 4.47 at 6–12 months and 3.09, 95% CI 0.84 to 5.35 at ≥12 months).

Conclusions A longer breastfeeding duration was associated with modest improvements in educational outcomes at age 16, after controlling for important confounders.

INTRODUCTION

There is some evidence that breastfeeding (BF) duration is associated with improved child's cognitive development, even after accounting for important confounders, such as socioeconomic position (SEP) and maternal intelligence.¹ However, not all studies observe significant differences between breastfeeding groups after adjustment for these confounders,^{1,2} and among those that do, the effect sizes tend to be modest.¹ Consequently, the importance of breastfeeding for improving cognitive

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Studies evaluating the association between breastfeeding duration and educational outcomes are relatively scarce and most do not conduct a comprehensive adjustment for potential confounders, primarily socioeconomic position (SEP) and maternal cognitive ability.

WHAT THIS STUDY ADDS

⇒ This study analysed a large British birth cohort study and found that longer breastfeeding durations were associated with modest improvements in educational outcomes at age 16, even after controlling for SEP and maternal intelligence.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Breastfeeding should continue to be encouraged when possible, as the potential improvements in academic achievement seen in this study constitute only one of its potential benefits.

outcomes continues to be debated. Educational attainment is related to cognitive ability, and is a strong predictor of life trajectories. Researchers have hypothesised that breastfed children would have better school results than non-breastfed children,³ although the empirical evidence is inconsistent.^{3–19}

In developed economies, including the UK, mothers from a higher SEP are more likely to breast feed,²⁰ and to have children who perform better in school. A study analysing 16 British cohorts found that a lower socioeconomic background was associated with poorer academic performance.²¹ Maternal intelligence is also an important confounder in the association of interest. While maternal intelligence is partly explained by SEP, strong evidence supports its heritability.²² Hence, its standalone confounding role has been recognised.^{1,2} However, the majority of published studies did not adjust for this variable.^{3–16 18 19}

Studies in the different British birth cohorts have found a positive association between breastfeeding and schooling results.^{5–10} For those born in 1946, breastfed babies scored higher in English/Mathematics tests at age 15,⁵ and were more likely to have higher educational qualifications at 26.⁶ For those born in 1991, BF was positively associated with national examinations' scores at ages 7, 11 and 14.⁷



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Pereyra-Elías R, Carson C, Quigley MA. *Arch Dis Child* Epub ahead of print: [please include Day Month Year]. doi:10.1136/archdischild-2022-325148

Similarly, studies analysing those born in the early 2000s (Millennium Cohort Study, MCS) found breastfeeding to be positively associated with school readiness at age 3^{8 9} and a statutory test at age 5.¹⁰ However, none of these studies adjusted for maternal cognitive ability, and to our knowledge, studies exploring the association in the MCS have only analysed outcomes at the time of entry to the education system.^{8–10} Therefore, the present study evaluated the association between breastfeeding duration and educational outcomes at age 16 among children from the MCS in England.

METHODS

Study population, design and setting

The MCS enrolled 18 818 children born in 2000–2002 and living in the UK, who were followed-up at ages 3, 5, 7, 11, 14, 17 and 22.²³ The present study analysed participants from England due to the differences in educational assessment between UK countries, and up to age 14 for survey data and age 16 for educational data. The MCS sampled a nationally representative cohort, and oversampled children from disadvantaged and ethnic minority backgrounds.²⁴ Singleton, term babies were included, for whom an English-speaking mother was the main respondent. Participants were included if they were present at follow-up at age 7 (consent for educational data requested), and age 14 (maternal cognitive performance measured).

Exposure

The exposure was the duration of any breastfeeding (maternal report), categorised as: never breast fed, <2, 2 to <4, 4 to <6, 6 to <12 and ≥12 months, to examine potential patterns in the association, such as dose–response relationships, and to explore what might be the minimal BF duration associated with improved academic achievement.

Outcomes

The MCS has been linked with the National Pupil Dataset (NPD), which stores longitudinal academic data of students enrolled in English state schools.²⁵ The English national curriculum provides a framework to be followed by all state-funded schools, and encompasses five key stages of learning (KS1–5).^{26 27} This study evaluated educational attainment in KS4 (ages 14–16, assessed at age 16).

KS4 ends with the assessment of the child's educational achievement through standardised tests in different subjects (General Certificate of Secondary Education, GCSE).^{27 28} The MCS children were examined in the transition from letter grading (A*–E) to numeric grading (9–1). GCSEs in English and Mathematics were marked in a 9–1 scale (9: highest performance; ≥4: pass). The rest of the subjects were marked in an A*–E scale (A*: highest; C: pass).²⁸ We categorised the GCSEs in English and Mathematics as: fail (<4), low pass (4–6) and high pass (7–9; equivalent to A–A*).²⁹

We also evaluated general measures of achievement: a binary measure (yes/no) if the child passed at least five GCSEs with marks ≥5 (including English and Mathematics), and the 'attainment 8' score. The latter is a measure that combines the best eight GCSE results, including English and Mathematics (both double weighted) and other subjects (eg, science, foreign languages, history). Exams marked in the A*–E scale were converted into 9–1 marks within the NPD. The final score adds the marks (range: 0–90).^{27 28 30} All participants with a score of zero were excluded.

Confounding

This study attempts to estimate the overall effect of breastfeeding on schooling outcomes; therefore, potential confounders were identified through a literature review, and the hypothesised underlying relationships (mediators not shown) were described using a directed acyclic graph (figure 1). The variables considered (categories in table 1) included:

1. A priori confounders: SEP markers, indicated by the highest parental social class and maternal education (National Vocational Qualifications; gestational age at birth, maternal ethnicity, language spoken at home, child's sex and age).
2. Other potential confounders: older siblings in the household, mother smoked during pregnancy, maternal age, employment, relationship status, Index of Multiple Deprivation.
3. Maternal verbal cognitive performance: proxy for cognitive ability evaluated when the participants were aged 14 years. The mothers' understanding of the meaning of 20 words was assessed through a vocabulary test (standardised: $\bar{x}=0$, $SD=1$).³¹

Except for age and maternal cognitive performance, all variables were measured at baseline.

Missing data

The most common reason for missing data was loss to follow-up (unit non-response). Differential non-response and attrition were corrected using longitudinal survey weights (age 14), thus minimising the risk of selection bias.³² Among those followed up to age 14, <5% had missing data for confounders. Considering this proportion was low, we followed a complete-case approach.

Statistical analysis

The association between BF duration and the schooling results was first evaluated using X^2 (categorical outcomes) and the Wald test/F-statistic (attainment 8, continuous). To estimate the association between breastfeeding duration and English/Maths GCSEs, we fitted multinomial models to calculate relative risk ratios (RRR), comparing 'fail' and 'high pass', with 'low pass' (reference), comparing the different categories of breastfeeding duration with non-BF children, while adjusting for confounders. To evaluate the association between breastfeeding duration and five GCSE passes (binary) and attainment 8 (continuous), modified Poisson regression was used to generate risk ratios (RR) and linear regression to estimate coefficients, respectively.

Confounder adjustment followed a sequential approach, adding the variables in blocks as outlined above. Model one was adjusted for the confounders identified a priori, then additional variables were included if they remained associated with the outcome ($p<0.10$), after full adjustment. Different models were fitted for English, Mathematics and general outcomes. The final and fully adjusted model incorporated maternal cognitive scores (model 2). All analyses were conducted in Stata 17.0 and considered the complex sampling design, and attrition weights.²⁴

RESULTS

Descriptive results

Of the 11 695 children enrolled in the MCS, 7645 were successfully linked to education data. A further 2402 children did not fulfil the eligibility criteria, and of these, 1292 were excluded because they were not followed up at age 14. 4940 children with complete data were included in the analyses (figure 2, online supplemental material 1).

The mean age at baseline was 9.2 months; 32.8% were never breast fed, and 9.5% were breast fed for ≥12 months. Half

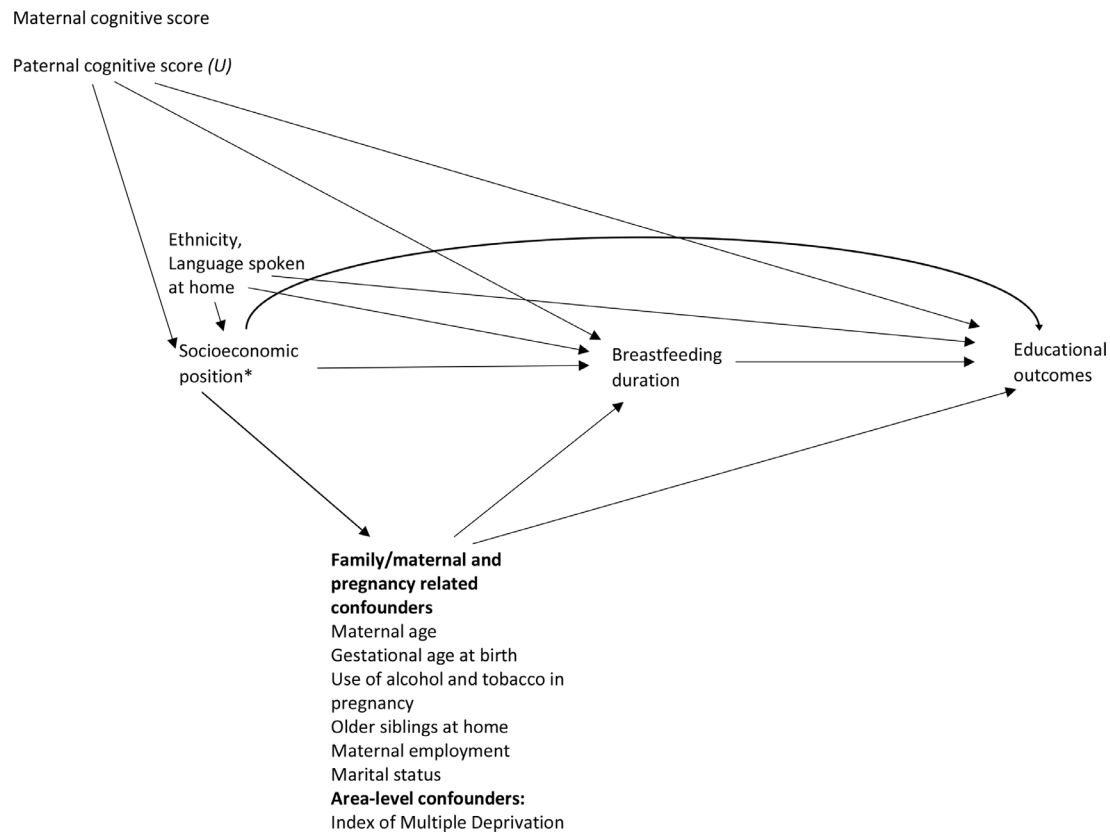


Figure 1 Directed acyclic graph used for the evaluation of the association between breastfeeding duration and educational outcomes. *Maternal education, parental social class. U, unobserved variable.

(49.2%) of the mothers had only secondary education, and 22.6% were in the semiroutine/routine social class. Approximately, one-third of participants failed their English (32.6%) and Mathematics (31.2%) GCSEs; 15.9% and 18.8% achieved a high pass (marks ≥ 7) in English and Mathematics, respectively; 56.4% passed five GCSEs with marks ≥ 5 , and the average attainment 8 score was 47.1 (SD 18.9) (table 1).

Children breast fed for longer were more likely to have older and more educated mothers, of higher social class. They were also less likely to have mothers that identified as White and spoke only English at home (table 1). The associations between covariates and outcomes are shown in online supplemental materials 2 and 3.

Association between breastfeeding duration and academic results

English GCSE

There is an apparent dose-response association between longer BF durations and being less likely to fail and more likely to have a high pass. For example, 19.2% of those BF for ≥ 12 months failed English compared with 41.7% of those never BF; 28.5% of those BF for ≥ 12 months achieved a high pass compared with 9.6% among non-breastfed participants (table 1). In crude models, compared with those not BF, the relative risk reductions (RRRs) for a high pass increased from 1.32 (95% CI 1.02 to 1.70) for BF < 2 months to 2.77 (95% CI 2.08 to 3.71) for BF ≥ 12 months. Adjustment explained most of the observed associations (RRR: 0.98; 95% CI 0.83 to 1.27 for BF < 2 months, and RRR: 1.38; 95% CI 1.00 to 1.90 for BF ≥ 12 months) (figure 3). There was a similar dose-response association for failing compared with achieving a low pass. Again, after adjustment,

only a borderline association remained in those who were BF for ≥ 12 months (RRR: 0.75; 95% CI 0.55 to 1.00).

Mathematics GCSE

For Mathematics, there was not a clear gradient between BF duration and failing. After adjustment, those breast fed for 4–6 months were less likely to fail in comparison with those never BF (0.72; 95% CI 0.52 to 0.99) (figure 3). However, this was not the case for those BF ≥ 12 months (RRR: 1.00, 95% CI 0.74 to 1.35).

There was a strong dose-response association between BF duration and having a high pass. After adjustment, the association weakened, although those breast fed for ≥ 12 months (vs non-breast fed) were 39% more likely to have a high pass than a lower pass (RRR: 1.39; 95% CI 1.01 to 1.90).

Five GCSEs with marks ≥ 5

In the crude analysis, all BF infants were more likely to pass five GCSEs than those never BF. The effect was similar for all those BF for ≥ 4 months: they were 1.6 times more likely to achieve this than those never BF. Adjustment attenuated the effect, but the association persisted for durations ≥ 4 months (RR: 1.13; 95% CI 1.02 to 1.23) (figure 4).

Attainment 8

There was a dose-response relationship between BF and the attainment 8 score. After adjustment, compared with those non-BF, those BF for ≥ 4 months had, on average, a 2–3 point higher attainment 8 score (coefficients: 2.10, 95% CI 0.06 to 4.14 for 4–6 months; 2.56, 95% CI 0.65 to 4.47 for 6–12 months; and

Table 1 Characteristics of the study participants according to their breastfeeding duration (n=4940)* among children from the UK Millennium Cohort Study (England)

Characteristics	Descriptive	Duration of any breastfeeding						P value†
	Whole sample	Never BF (n=1200) (32.8%)	<2 months (n=1253) (33.9%)	2 to <4 months (n=564) (10.6%)	4 to <6 months (n=543) (9.3%)	6 to <12 months (n=811) (13.9%)	≥12 months (n=569) (9.5%)	
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	
Pregnancy and child-related								
Gestational age at birth‡	39.8 (1.3)	39.8 (1.1)	39.8 (1.4)	39.8 (1.3)	39.9 (1.3)	39.9 (1.4)	39.8 (1.3)	0.067
Female	2517 (49.9)	641 (52.3)	606 (46.9)	291 (50.9)	273 (48.0)	401 (47.3)	305 (54.1)	0.015
Age at key stage 4‡	16.2 (0.3)	16.2 (0.3)	16.2 (0.3)	16.2 (0.3)	16.2 (0.3)	16.2 (0.3)	16.2 (0.3)	0.296
No older siblings	2492 (51.2)	546 (48.0)	717 (58.4)	326 (56.6)	291 (53.0)	378 (47.2)	234 (42.6)	<0.001
Smoked during pregnancy								
Never	3515 (64.4)	703 (51.8)	815 (58.9)	390 (64.8)	426 (74.9)	679 (81.6)	502 (85.6)	<0.001
Gave up	601 (13.4)	148 (12.8)	182 (16.1)	87 (17.6)	68 (13.0)	70 (9.6)	46 (9.9)	
Kept smoking	824 (22.2)	349 (35.4)	256 (25.0)	87 (17.6)	49 (12.1)	62 (8.8)	21 (4.6)	
Moderate/heavy alcohol in pregnancy§	370 (7.7)	94 (8.0)	89 (7.4)	31 (6.2)	41 (7.2)	70 (8.4)	45 (8.3)	0.745
Sociodemographic, age 9 m.								
Maternal age (years)‡	29.8 (6.0)	27.7 (5.2)	29.2 (5.9)	30.2 (6.0)	31.9 (5.5)	32.2 (5.6)	32.7 (5.7)	<0.001
Maternal education¶								
Higher (NVQ 4 and 5)	1551 (25.0)	131 (7.7)	316 (22.3)	173 (26.7)	260 (42.5)	392 (44.5)	279 (44.3)	<0.001
Medium (NVQ 3)	471 (8.5)	59 (4.0)	123 (8.7)	71 (12.1)	55 (10.3)	93 (11.3)	70 (13.5)	
Lower (NVQ 1 and 2)	2238 (49.2)	729 (59.9)	653 (53.6)	262 (51.0)	180 (37.2)	256 (35.8)	158 (30.3)	
Other/overseas	112 (2.0)	25 (1.4)	16 (1.4)	16 (3.2)	14 (2.9)	22 (2.2)	19 (3.9)	
None	568 (15.3)	256 (27.0)	145 (14.0)	42 (7.0)	34 (7.1)	48 (6.2)	43 (8.0)	
Highest parental social class								
Managerial/professional	2477 (44.2)	332 (23.2)	563 (42.6)	300 (51.7)	357 (62.9)	552 (66.1)	373 (62.1)	<0.001
Intermediate	1348 (28.2)	399 (30.6)	389 (32.7)	160 (28.0)	124 (25.0)	157 (21.2)	119 (22.7)	
Semiroutine/routine	921 (22.6)	386 (37.5)	256 (20.9)	87 (17.1)	51 (10.7)	81 (10.2)	60 (11.7)	
Not applicable	194 (5.0)	83 (8.7)	45 (3.8)	17 (3.2)	11 (1.4)	21 (2.6)	17 (3.6)	
Mother working	2571 (49.4)	506 (38.7)	665 (52.0)	329 (59.3)	343 (63.1)	458 (56.1)	270 (45.5)	<0.001
Maternal relationship status								
Married	3220 (57.9)	601 (41.4)	787 (58.2)	377 (63.7)	389 (68.3)	625 (74.5)	441 (73.4)	<0.001
Cohabitation	1196 (27.9)	372 (34.6)	328 (29.3)	133 (25.9)	113 (23.0)	154 (20.8)	96 (18.1)	
Single mother	524 (14.2)	227 (24.0)	138 (12.5)	54 (10.4)	41 (8.7)	32 (4.7)	32 (8.5)	
Maternal ethnicity: White	4174 (89.6)	1082 (93.9)	1069 (90.9)	459 (85.8)	439 (85.6)	677 (87.7)	448 (82.5)	<0.001
Language spoken at home								
English only	4305 (92.0)	1100 (95.1)	1105 (93.2)	483 (90.3)	452 (88.4)	705 (91.4)	460 (84.8)	<0.001
English+other language	635 (8.0)	100 (4.9)	148 (6.8)	81 (9.7)	91 (11.6)	106 (8.6)	109 (15.2)	
Index of Multiple Deprivation								
Poorest quintile	1248 (22.8)	437 (34.8)	335 (21.3)	119 (16.3)	94 (13.5)	156 (14.7)	107 (13.2)	<0.001
Second quintile	1089 (20.3)	309 (24.8)	269 (19.0)	132 (21.5)	120 (18.9)	139 (14.1)	120 (17.0)	
Middle quintile	989 (21.7)	224 (19.9)	264 (23.4)	104 (21.4)	125 (25.0)	168 (22.0)	104 (20.7)	
Fourth quintile	823 (18.4)	138 (12.1)	187 (18.3)	102 (20.5)	103 (22.2)	173 (24.8)	120 (24.5)	
Richest quintile	791 (16.8)	92 (8.4)	198 (18.0)	107 (20.3)	101 (20.4)	175 (24.4)	118 (24.6)	
Maternal Cognitive Score‡	0.0 (1.0)	-0.42 (0.72)	-0.05 (0.96)	0.06 (0.95)	0.42 (1.06)	0.43 (1.10)	0.52 (1.23)	<0.001
Schooling outcomes								
Key stage 4 (age 16)								
English GCSE								
Failed (marks <4)	1472 (32.6)	472 (41.7)	406 (34.0)	176 (32.0)	130 (24.4)	180 (23.2)	108 (19.2)	<0.001
Low pass (marks ≥4 and <7)	2639 (51.5)	627 (48.7)	686 (52.5)	309 (52.9)	282 (52.6)	441 (54.1)	294 (52.3)	
High pass (marks ≥7)	915 (15.9)	130 (9.6)	192 (13.5)	90 (14.9)	135 (23.0)	200 (22.7)	168 (28.5)	
Mathematics GCSE								
Failed (marks <4)	1401 (31.2)	467 (41.9)	358 (29.5)	169 (30.0)	111 (20.8)	173 (21.3)	123 (23.7)	<0.001
Low pass (marks ≥4 and <7)	2529 (50.0)	605 (47.1)	702 (54.8)	288 (50.7)	296 (54.8)	386 (48.8)	252 (44.9)	
High pass (marks ≥7)	1082 (18.8)	154 (11.0)	219 (15.7)	117 (19.3)	138 (24.4)	260 (29.9)	194 (31.4)	
Five GCSE passes (marks ≥5)**	3030 (56.4)	590 (43.8)	740 (56.0)	329 (55.4)	380 (69.2)	575 (69.5)	416 (71.0)	<0.001
Attainment 8 score††	47.1 (18.9)	41.4 (15.3)	46.5 (18.3)	47.6 (19.3)	52.3 (20.1)	52.9 (20.5)	54.2 (21.2)	<0.001

*Parameters were estimated considering the complex design of the Millennium Cohort Study sample. The counts reflect the absolute number of participants included.

†X² or mean difference, as appropriate.

‡Mean (SD).

§Consumption of ≥3 units per week or ≥3 units per occasion during pregnancy.

¶National Vocational Qualifications (NVQ): NVQ 4 and 5: higher/university education, NVQ3: A-levels, qualifications obtained at age 18, NVQ 1 and 2: qualifications obtained at age 16, at the end of secondary education.

**Five GCSEs passed with marks ≥5, including English and Mathematics.

††Attainment 8 is the additive score of the marks of a student in their best eight GCSEs, including English and Mathematics (double weighted) and can range between 0 and 90.

BF, breastfeeding; GCSE, General Certificate of Secondary Education.

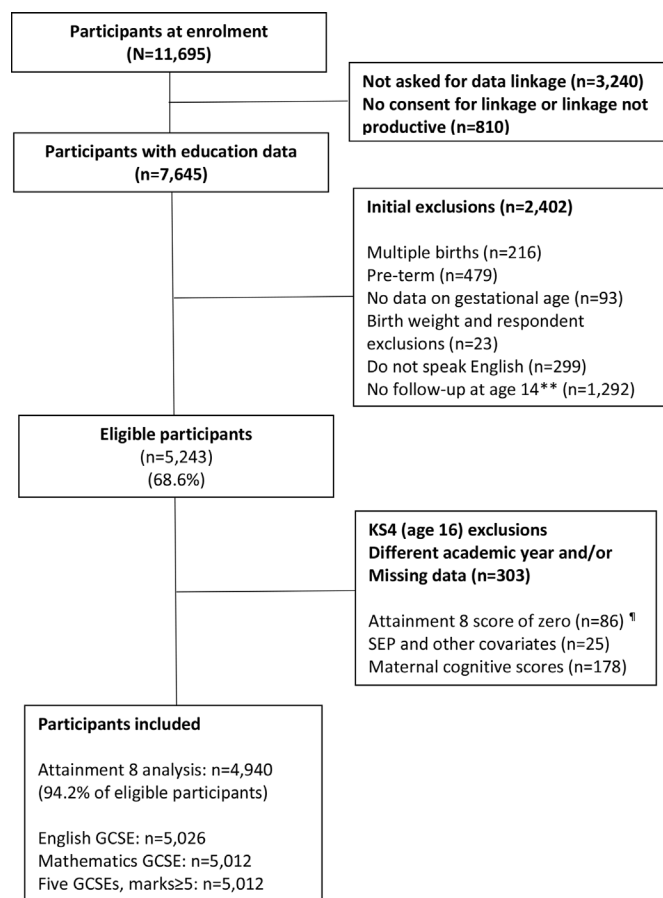


Figure 2 Flow chart of study participants, UK Millennium Cohort Study (England). *Reasons including no data for birth weight, birthweight exclusions, and if the respondent was not the natural mother. Grouped because individual numbers were low. **Excluded if not followed up at age, as maternal cognitive ability was assessed then. †Children with an attainment 8 score of 0 were excluded from the analysis of attainment 8 only. KS4, key stage 4; SEP, socioeconomic position.

3.09 to 95% CI 0.84 to 5.35 for those BF for ≥ 12 months) (figure 4).

DISCUSSION

Summary of key findings

This study evaluated the association between breastfeeding duration and school results at age 16 in England. We observed modest dose–response relationships, where children breast fed for longer were more likely to achieve high marks and less likely to fail their examinations. Additionally, children breast fed for ≥ 4 months were more likely to pass five GCSEs than non-breastfed children, and also scored higher in the consolidated (additive) mark for their GCSEs.

The association between breastfeeding and schooling results

Longer BF was associated with modest gains in academic outcomes. The strength of the association equates to a 38–39% increase in the probability of achieving high marks in both GCSEs and a 25% reduction in the probability of failing (English). The gains in the probability of passing five or more GCSEs were smaller (13% increase). The difference of approximately 3 points in attainment 8 scores (out of 90) between extreme categories of breastfeeding duration is also modest.

Previous studies have found similar associations, with comparable or stronger effect sizes.^{3–17} The effect seems to be present in studies that evaluated attainment at the beginning of the school system,^{8–10 15} in primary school,^{7 11–13 16 17} near the end of secondary school^{3 5 7 13 18 19} and later in life.^{4 6 14} However, the different nature of the outcomes hinders comparability: early-life school readiness assessments are not directly comparable to standardised tests taken in secondary school, or with the number of years of schooling completed by age 30. However, most studies seem to suggest a positive association between breastfeeding and schooling results.^{3–17} Comparability may also be limited by heterogeneity in the ‘non-breastfed’ group, who could receive cow’s milk, other liquids and/or formula, which are time and region-dependent.

Several mechanisms may explain the association between breastfeeding and educational outcomes. The main mechanism proposed is through improved cognitive development.^{3 4 6 13 19} Breast milk contains polyunsaturated fatty acids and micronutrients, which enhance neurodevelopment.³³ Adjustment for cognitive abilities at age 15 explained a considerable portion of association in the 1946 British birth cohort, which may indicate mediation.⁶ However, education may also be a marker of cognitive ability. Similarly, breastfeeding was associated with intergenerational upward social mobility (change in occupational social class) in the 1958 and 1970 British birth cohorts, and this effect was mediated by cognitive performance.³⁴ It has also been proposed that breastfeeding may enable mother–child bonding,^{1 13 15–17 33} which would favour both cognitive and academic performance. Another hypothesis includes mediation by improvements in self-regulation, which can predict academic achievement.¹⁵

Confounding

Some propose that the observed association is mostly/solely due to confounding,² predominantly by SEP and maternal intelligence.¹² In the present study, adjustment for SEP had the greatest explanatory effect. Further adjustment for maternal cognitive abilities explained a smaller portion of the association. A recent MCS study found that adjustment for maternal cognitive abilities explained a seemingly higher fraction of the association with cognitive outcomes³⁵—compared with its smaller explanatory power with educational outcomes in the present study. This suggests that socioeconomic circumstances may play a bigger role than maternal cognitive ability in predicting school results in the UK. In cohort studies from low- and middle-income countries, where SEP is negatively (or not) associated with breastfeeding duration, breastfeeding did not predict schooling outcomes.^{18 19} Additionally, in the 15-year time gap between exposure and outcome, SEP could have created other conditions that influence the schooling results, and/or accentuated the differences between breastfed and non-breastfed groups, thus influencing the observed associations. These evolving conditions are difficult to capture and deserve deeper exploration.

While SEP seems to be the strongest confounder, adjustment for maternal cognitive abilities also explained some of the association. This is important to consider when interpreting the results of most previous studies, as control for this variable is rare.^{3–16 19}

Strengths and limitations

Our findings are nationally representative for children enrolled in state schools in England. The large sample size allowed us to detect outcome differences between several breastfeeding duration groups. Additionally, the outcomes

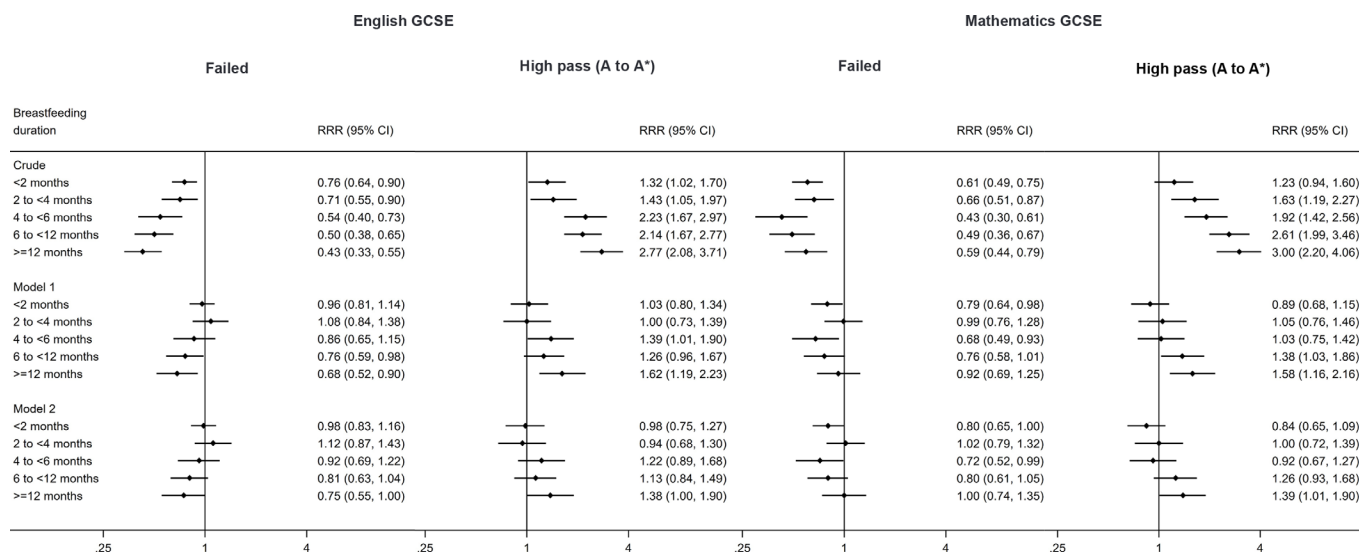


Figure 3 Association between breastfeeding duration and key stage 4 (KS4) results at age 16 in English (n=5026) and Mathematics (n=5012) GCSEs*, Millennium Cohort Study (MCS). *Parameters were estimated considering the complex design of the MCS sample, using multinomial regression, where 'obtaining low passing marks (excluding A to A*)' was the reference outcome category. Model 1: adjusted for sex, age at KS1, maternal ethnicity and language spoken in the household, socioeconomic position (maternal education and highest social class in the household), Index of Multiple Deprivation, older siblings in the household, mother working outside the home, partnership status, smoking during pregnancy (Mathematics :model 3 also includes maternal age and maternal alcohol use during pregnancy). Model 2: adjusted for model 1+Maternal Cognitive Score.

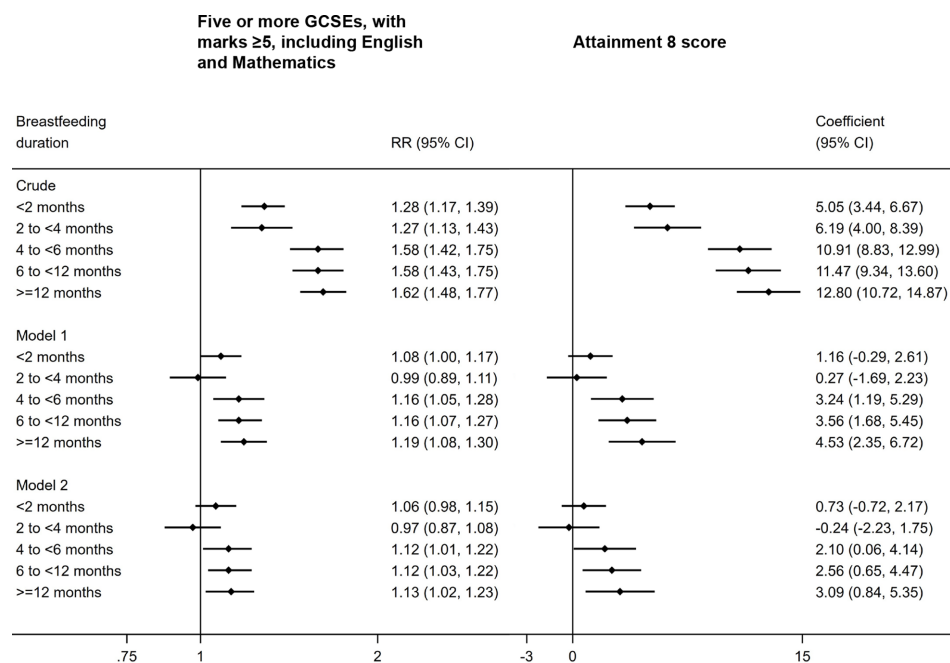


Figure 4 Association between breastfeeding duration and key stage 4 (KS4) results at age 16: Passing five GCSEs with marks ≥5, including English and Mathematics (n=5012) and attainment 8 score (n=4940)*[†], Millennium Cohort Study (MCS). *Parameters were estimated considering the complex design of the MCS sample. Model 1: adjusted for sex, age at KS4, maternal ethnicity and language spoken in the household, socioeconomic position (maternal education and highest social class in the household), Index of Multiple Deprivation, older siblings in the household, maternal age, mother working outside the home, partnership status, smoking and alcohol use during pregnancy. Model 2: adjusted for model 1+Maternal Cognitive Score. [†]Attainment 8 is the additive score of the marks of a student in their best eight GCSEs, including English and Mathematics (double weighted), and can range between 0 and 90.

were both subject-specific and general measures of achievement, assessed by the English education system, and have direct relevance to future opportunities. We have also

comprehensively controlled for the confounding effect of several markers of family-level and area-level SEP, and maternal cognitive abilities.

Some caveats should be considered when interpreting our results. First, linkage with the NPD was not possible for approximately 4000 children because they were lost to follow-up/did not consent. A further 1292 children were not followed up to age 14, when maternal cognitive ability was measured. However, the use of survey weights corrects for potential selection bias and estimates representative population parameters, even after exclusions and attrition (online supplemental material 2).^{23 24 31} Second, residual confounding may persist. The maternal cognitive score only captures a verbal dimension of cognitive ability. However, this measure correlates with general intelligence and has been previously used as its proxy.^{1 35} Third, adjustment for these maternal cognitive scores further explains the associations of interest even after accounting for SEP markers, which suggests that it represents more than just socioeconomic circumstances. Lastly, our results do not take into account potential mediators (eg, diet, parenting and early education) or the fact that confounding by SEP may change or be accentuated after baseline.

CONCLUSION

Breastfeeding duration was associated with improved educational outcomes at age 16 among children living in England, after controlling for important confounders. However, the effect sizes were modest and may be susceptible to residual confounding. breastfeeding should continue to be encouraged when possible, as potential improvements in academic achievement constitute only one of its potential benefits. Future studies should adjust for both socioeconomic circumstances (comprehensively) and maternal general intelligence.

Twitter Reneé Pereyra-Elías @ReneePereyra

Acknowledgements The authors would like to thank the participants of the Millennium Cohort Study, and the UK Data Service for providing the datasets. We would also like to thank Sarah Chamberlain for her help in producing the figures.

Contributors All authors conceptualised the study. RP-E carried out the data analysis and drafted a first version of the paper. CC and MAQ provided supervision during this process and critically reviewed the manuscript. All authors approved the final version of this manuscript. All authors are guarantors of the article.

Funding This work was supported by Nuffield Department of Population Health at the University of Oxford, as part of a DPhil scholarship held by Reneé Pereyra-Elías (MSD1920_1296925).

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by the Millennium Cohort Study. Protocol was approved by National Health Services (NHS) Research Ethics Committee (REC) of the South West. Further data collection also received ethical approval by the NHS REC system. The datasets contain no personal identifiers and are publicly available upon request (UK Data Service). No formal ethical approval was required for this secondary analysis of the pseudonymised data. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. Data may be obtained from a third party on request. The datasets used for this study are publicly available in the UK Data Service: <https://beta.ukdataservice.ac.uk/datacatalogue/series/series?id=2000031> Access to educational data requires an additional request and analyses should be conducted within a data safe haven. Data citation: University of London, Institute of Education, Centre for Longitudinal Studies. (2021). Millennium Cohort Study: First Survey, 2001-2003. [data collection]. 13th Edition. UK Data Service. SN: 4683, DOI: 10.5255/UKDA-SN-4683-5 University of London, Institute of Education, Centre for Longitudinal Studies. (2021). Millennium Cohort Study: Second Survey, 2003-2005. [data collection]. 10th Edition. UK Data Service. SN: 5350, DOI: 10.5255/UKDA-SN-5350-5 University of London, Institute of Education, Centre for Longitudinal Studies. (2021). Millennium Cohort Study: Third Survey, 2006. [data collection]. 8th Edition. UK Data Service. SN: 5795, <http://doi.org/10.5255/UKDA-SN-5795-5> University of London, Institute of Education, Centre for Longitudinal Studies. (2021). Millennium Cohort Study: Fourth

Survey, 2008. [data collection]. 8th Edition. UK Data Service. SN: 6411, <http://doi.org/10.5255/UKDA-SN-6411-8> University of London, Institute of Education, Centre for Longitudinal Studies. (2021). Millennium Cohort Study: Fifth Survey, 2012. [data collection]. 5th Edition. UK Data Service. SN: 7464, <http://doi.org/10.5255/UKDA-SN-7464-5> University of London, Institute of Education, Centre for Longitudinal Studies. (2020). Millennium Cohort Study: Sixth Survey, 2015. [data collection]. 7th Edition. UK Data Service. SN: 8156, <http://doi.org/10.5255/UKDA-SN-8156-8> University of London, Institute of Education, Centre for Longitudinal Studies. (2020). Millennium Cohort Study: Longitudinal Family File, 2001-2018. [data collection]. 4th Edition. UK Data Service. SN: 8172, <http://doi.org/10.5255/UKDA-SN-8172-4> University College London, UCL Institute of Education, Centre for Longitudinal Studies, Department for Education. (2021). Millennium Cohort Study: Linked Education Administrative Datasets (National Pupil Database), England: Secure Access. [data collection]. 2nd Edition. UK Data Service. SN: 8481, <http://doi.org/10.5255/UKDA-SN-8481-2>.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Reneé Pereyra-Elías <http://orcid.org/0000-0001-6398-4287>

Claire Carson <http://orcid.org/0000-0001-5490-699X>

Maria A Quigley <http://orcid.org/0000-0002-8058-6181>

REFERENCES

- Horta BL, Loret de Mola C, Victora CG. Breastfeeding and intelligence: a systematic review and meta-analysis. *Acta Paediatr* 2015;104:14–9.
- Der G, Batty GD, Deary IJ. Effect of breast feeding on intelligence in children: prospective study, sibling pairs analysis, and meta-analysis. *BMJ* 2006;333:945.
- Victora CG, Barros FC, Horta BL, et al. Breastfeeding and school achievement in Brazilian adolescents. *Acta Paediatr* 2005;94:1656–60.
- Victora CG, Horta BL, Loret de Mola C, et al. Association between breastfeeding and intelligence, educational attainment, and income at 30 years of age: a prospective birth cohort study from Brazil. *Lancet Glob Health* 2015;3:e199–205.
- Rodgers B. Feeding in infancy and later ability and attainment: a longitudinal study. *Dev Med Child Neurol* 1978;20:421–6.
- Richards M, Hardy R, Wadsworth MEJ. Long-term effects of breast-feeding in a national birth cohort: educational attainment and midlife cognitive function. *Public Health Nutr* 2002;5:631–5.
- Borra C, Iacovou M, Sevilla Sanz A. The effect of breastfeeding on children's cognitive and noncognitive development. *SSRN Journal* 2012.
- Del Bono E, Rabe E. Breastfeeding and child cognitive outcomes: evidence from a hospital-based breastfeeding support policy. Essex ISER; 2012. Available: <https://www.iser.essex.ac.uk/wp-content/uploads/files/working-papers/iser/2012-29.pdf>
- Fitzsimons E, Vera-Hernández M. Breastfeeding and child development. London IFS; 2014. Available: <https://ifs.org.uk/publications/breastfeeding-and-child-development>
- Heikkilä K, Kelly Y, Renfrew MJ, et al. Breastfeeding and educational achievement at age 5. *Matern Child Nutr* 2014;10:92–101.
- McCorry C, Layte R. The effect of breastfeeding on children's educational test scores at nine years of age: results of an Irish cohort study. *Soc Sci Med* 2011;72:1515–21.
- Oddy WH, Li J, Whitehouse AJO, et al. Breastfeeding duration and academic achievement at 10 years. *Pediatrics* 2011;127:e137–45.
- Horwood LJ, Fergusson DM. Breastfeeding and later cognitive and academic outcomes. *Pediatrics* 1998;101:E9.
- Lutter R, Ruhm C, Lin D, et al. Breastfeeding, schooling, and income: insights from the Indonesian Family Life Survey. *Matern Child Nutr* 2019;15:e12651.
- Tua Karing J, Tracy A, Gonzales CR, et al. Breastfeeding, early self-regulation, and academic achievement in kindergarten among disadvantaged children. *Matern Child Health J* 2021;25:1766–75.
- Huang J, Vaughn MG, Kremer KP. Breastfeeding and child development outcomes: an investigation of the nurturing hypothesis. *Matern Child Nutr* 2016;12:757–67.
- Kim JJ, Kim B-N, Kim J-W, et al. Breastfeeding is associated with enhanced learning abilities in school-aged children. *Child Adolesc Psychiatry Ment Health* 2017;11:36.
- Horta BL, Bas A, Bhargava SK, et al. Infant feeding and school attainment in five cohorts from low- and middle-income countries. *PLoS One* 2013;8:e71548.

- 19 Nandi A, Lutter R, Laxminarayan R. Breastfeeding duration and adolescent educational outcomes: longitudinal evidence from India. *Food Nutr Bull* 2017;38:528–41.
- 20 Simpson DA, Quigley MA, Kurinczuk JJ, *et al*. Twenty-five-year trends in breastfeeding initiation: the effects of sociodemographic changes in Great Britain, 1985–2010. *PLoS One* 2019;14:e0210838.
- 21 von Stumm S, Cave SN, Wakeling P. Persistent association between family socioeconomic status and primary school performance in Britain over 95 years. *NPJ Sci Learn* 2022;7.
- 22 Plomin R, von Stumm S. The new genetics of intelligence. *Nat Rev Genet* 2018;19:148–59.
- 23 Connelly R, Platt L. Cohort profile: UK Millennium Cohort Study (MCS). *Int J Epidemiol* 2014;43:1719–25.
- 24 Plewis I, Calderwood L, Hawkes D, *et al*. *The millennium cohort study: technical report on sampling (fourth edition)*. London: Centre for Longitudinal Studies (IoE, UCL), 2007. Available: https://doc.ukdataservice.ac.uk/doc/4683/mrdoc/pdf/mcs_technical_report_on_sampling_4th_edition.pdf
- 25 Jay MA, McGrath-Lone L, Gilbert R. Data resource: the National Pupil Database (NPD). *Int J Popul Data Sci* 2019;4:1101.
- 26 Department for Education. *The national curriculum in England. Key stages 1 and 2 framework document*. London, 2013. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/425601/PRIMARY_national_curriculum.pdf
- 27 Department for Education. *The national curriculum in England. Key stages 3 and 4 framework document*. London, 2014. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/840002/Secondary_national_curriculum_corrected_PDF.pdf
- 28 Department for Education. *Revised GCSE and equivalent results in England, 2016 to 2017*. London, 2018. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/676596/SFR01_2018.pdf
- 29 Jadhav C. GCSE 9 to 1 grades: a brief guide for parents. London Ofqual; 2018. Available: <https://ofqual.blog.gov.uk/2018/03/02/gcse-9-to-1-grades-a-brief-guide-for-parents/>
- 30 Department for Education. *Progress 8. How progress 8 and attainment 8 measures are calculated*. London, 2018. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/561021/Progress_8_and_Attainment_8_how_measures_are_calculated.pdf
- 31 Fitzsimons E. *Millennium cohort study sixth survey 2015–2016 user guide (first edition)*. London: Centre for Longitudinal Studies (IoE, UCL), 2017. Available: https://cls.ucl.ac.uk/wp-content/uploads/2018/10/mcs6_user_guide_28march2017.pdf
- 32 Seaman SR, White IR. Review of inverse probability weighting for dealing with missing data. *Stat Methods Med Res* 2013;22:278–95.
- 33 Chiurazzi M, Cozzolino M, Reinelt T, *et al*. Human milk and brain development in infants. *Reprod Med* 2021;2:107–17.
- 34 Sacker A, Kelly Y, Iacovou M, *et al*. Breast feeding and intergenerational social mobility: what are the mechanisms? *Arch Dis Child* 2013;98:666–71.
- 35 Pereyra-Elías R, Quigley MA, Carson C. To what extent does confounding explain the association between breastfeeding duration and cognitive development up to age 14? Findings from the UK Millennium Cohort Study. *PLoS One* 2022;17:e0267326.

Supplementary material 1. Characteristics (baseline) of participants included at each stage of the study, excluding twins*[†].

Characteristics	Stage 1		Stage 2		Stage 3		Stage 4	
	England		Linked		Eligible		Final sample	
	(n=11,372)		(n=7,429)		(n=5,290)		(n=4,940)	
	n	(%)	n	(%)	n	(%)	n	(%)
Pregnancy and child-related								
Gestational age at birth**	39.4	(2.0)	39.4	(1.9)	39.8	(1.3)	39.8	(1.3)
Female	5,549	(48.9)	3,690	(49.1)	2,686	(49.5)	2,517	(49.9)
Age at Key Stage 4**	16.2	(0.3)	16.2	(0.3)	16.2	(0.3)	16.2	(0.3)
No older siblings	5,747	(51.4)	3,709	(51.2)	2,641	(50.8)	2,492	(51.2)
Smoked during pregnancy								
<i>Never</i>	4,660	(65.9)	5,129	(65.1)	3,745	(63.8)	3,515	(64.4)
<i>Gave up</i>	1,384	(13.3)	899	(13.3)	635	(13.2)	601	(13.4)
<i>Kept smoking</i>	2,313	(20.8)	1,397	(21.6)	910	(23.0)	824	(22.2)
Mod/heavy alcohol in preg.	775	(7.5)	522	(7.4)	391	(7.7)	370	(7.7)
Sociodemographic								
Maternal age (years)**	30.1	(6.0)	29.8	(5.9)	29.7	(6.1)	29.8	(6.0)
Maternal education								
<i>Higher</i>	2,678	(27.0)	1,926	(24.8)	1,584	(23.6)	1,551	(25.0)
<i>Medium</i>	929	(8.8)	652	(8.5)	486	(8.1)	471	(8.5)
<i>Lower</i>	5,022	(46.3)	3,366	(47.6)	2,371	(48.7)	2,238	(49.2)
<i>Other</i>	424	(2.5)	230	(2.5)	137	(2.3)	112	(2.0)
<i>None</i>	2,279	(15.4)	1,243	(16.6)	706	(17.3)	568	(15.3)
Highest parental social class								
<i>Managerial/Professional</i>	4,392	(44.9)	3,208	(43.4)	2,549	(42.1)	2,477	(44.2)
<i>Intermediate</i>	3,200	(28.1)	2,101	(28.4)	1,445	(27.9)	1,348	(28.2)
<i>Semi-routine/routine</i>	2,951	(22.0)	1,720	(23.0)	1,056	(24.3)	921	(22.6)
<i>Not applicable</i>	829	(5.0)	400	(5.2)	240	(5.7)	194	(4.9)
Mother working at age 9 m.	4,758	(47.0)	3,470	(47.1)	2,669	(47.9)	2,571	(49.4)
Maternal relationship status								
<i>Married</i>	6,779	(60.1)	4,639	(58.7)	3,415	(56.5)	3,220	(57.9)
<i>Cohabitation</i>	2,751	(25.9)	1,821	(26.9)	1,277	(27.8)	1,196	(27.9)
<i>Single mother</i>	1,817	(13.9)	960	(14.4)	598	(15.7)	524	(14.2)
Maternal ethnicity: White	8,543	(86.4)	5,859	(86.1)	4,363	(88.1)	4,174	(89.6)
Language spoken at home								
<i>English only</i>	8,984	(90.9)	6,075	(90.9)	4,525	(90.7)	4,305	(92.0)
<i>English + other language</i>	1,770	(9.1)	1,001	(9.1)	765	(9.3)	635	(8.0)

Supplementary material 1 (cont.). Characteristics (baseline) of participants included at each stage of the study, excluding twins*[†].

Characteristics	Stage 1 England (n=11,372)		Stage 2 Linked (n=7,429)		Stage 3 Eligible (n=5,290)		Stage 4 Final sample (n=4,940)	
	n	(%)	n	(%)	n	(%)	n	(%)
Sociodemographic (cont.)								
Index of Multiple Deprivation								
<i>Poorest quintile</i>	4,041	(23.3)	2,325	(24.4)	1,430	(24.4)	1,248	(22.8)
<i>Second quintile</i>	2,469	(19.3)	1,617	(19.8)	1,168	(20.6)	1,089	(20.3)
<i>Middle quintile</i>	1,947	(20.5)	1,369	(20.9)	1,035	(21.3)	989	(21.8)
<i>Fourth quintile</i>	1,533	(18.9)	1,080	(17.8)	849	(17.7)	823	(18.4)
<i>Richest quintile</i>	1,381	(18.0)	1,038	(17.2)	808	(16.0)	791	(16.8)
Maternal cognitive score**	-	-	-	-	-	-	0.0	(0.99)
Schooling outcomes								
Key Stage 3 (age 16)								
English GCSE								
<i><Expected (Level <2)</i>	-	-	2,416	(34.3)	1,575	(33.4)	1,394	(31.3)
<i>Expected (Level ≥2 & <3)</i>	-	-	3,813	(50.4)	2,769	(51.0)	2,632	(52.5)
<i>>Expected (Level ≥3)</i>	-	-	1,195	(15.3)	944	(15.5)	914	(16.2)
Math GCSE								
<i><Expected (Level <2)</i>	-	-	2,354	(33.3)	1,517	(32.3)	1,329	(29.9)
<i>Expected (Level ≥2 & <3)</i>	-	-	3,638	(48.5)	2,658	(49.5)	2,529	(51.0)
<i>>Expected (Level ≥3)</i>	-	-	1,432	(18.2)	1,113	(18.2)	1,082	(19.1)
General outcomes								
Five GCSEs	-	-	4,186	(54.6)	3,152	(55.1)	3,030	(57.4)
Attainment 8	-	-	45.4	(19.7)	45.7	(20.0)	47.1	(18.9)

* Numbers do not exactly match those of the flowchart, as all these descriptive statistics have been produced excluding twins.

[†] All percentages consider the complex survey sampling of MCS and percentages from Stages 2-4 additionally consider longitudinal attrition.

** Mean (standard deviation)

Stage 1 - England: Initial MCS participants from England, considering baseline weights.

Stage 2 - Linkage: MCS participants from England with productive educational data linkage. Weights consider longitudinal attrition at age 7 (when educational data linkage was carried out).

Stage 3 - Eligible: Participants that, additionally, meet the eligibility criteria (excluding pre-term births, no data on gestational age, birth weight exclusions, not speaking English, and not being present at the follow-up visit at age 14). Weights consider longitudinal attrition at age 14.

Stage 4 - Final sample: Final sample of participants included in the analysis, after excluding those that were examined in a different academic year, had missing data for the educational outcomes, had an Attainment 8 score of zero, or had missing data for the confounders, including maternal cognitive ability. Weights consider longitudinal attrition at age 14.

Supplementary material 2. Univariate association between participants' characteristics and KS4 results at age 16 in English (n=5,026) and Mathematics (n=5,012)*, Millennium Cohort Study.

Characteristics	English						p	Math						p [¶]
	Failed		Low pass		High pass (A-A*)			Failed		Low pass		High pass (A-A*)		
	n	(%)	n	(%)	n	(%)		n	(%)	n	(%)	n	(%)	
Pregnancy and child-related														
Gestational age at birth**	39.7	(1.2)	39.8	(1.3)	39.9	(1.3)	0.062	39.7	(1.2)	39.8	(1.3)	39.9	(1.4)	0.069
Sex														
Female	558	(24.9)	1,399	(54.5)	586	(20.6)	<0.001	679	(30.2)	1,326	(51.4)	534	(18.4)	0.259
Male	914	(40.1)	1,240	(48.6)	329	(11.3)		722	(32.1)	1,203	(48.8)	548	(19.1)	
Age at Key Stage 4**	16.2	(0.3)	16.2	(0.3)	16.2	(0.3)	0.207	16.2	(0.3)	16.2	(0.3)	16.2	(0.3)	0.513
Older siblings														
No	724	(32.2)	1,328	(51.5)	481	(16.3)	0.715	668	(29.9)	1,286	(50.7)	569	(19.4)	0.198
Yes	748	(33.0)	1,311	(51.5)	434	(15.5)		733	(32.5)	1,243	(49.4)	513	(18.1)	
Smoked during pregnancy														
Never	909	(27.3)	1,913	(53.6)	738	(19.1)	<0.001	872	(26.5)	1,790	(50.6)	88	(22.9)	<0.001
Gave up	170	(29.9)	334	(55.1)	103	(15.0)		166	(29.3)	331	(53.5)	110	(17.2)	
Kept smoking	393	(49.0)	392	(43.6)	74	(7.4)		363	(45.3)	408	(46.5)	84	(8.2)	
Mod/heavy alcohol in preg.														
No	1,337	(32.0)	2,448	(51.8)	850	(16.2)	0.043	1,288	(30.9)	2,333	(50.0)	1,014	(19.1)	0.126
Yes	129	(39.0)	184	(47.6)	64	(13.4)		113	(33.9)	196	(51.4)	68	(14.7)	
Sociodemographic, age 9 m.														
Maternal age (years)**	28.5	(5.9)	30.0	(5.9)	31.6	(5.7)	<0.001	28.7	(6.0)	29.9	(5.9)	31.5	(5.5)	<0.001
Maternal education														
Higher (NVQ 4 and 5)	341	(22.3)	773	(49.5)	448	(28.2)	<0.001	315	(20.2)	704	(45.5)	539	(34.3)	<0.001
Medium (NVQ 3)	106	(22.0)	261	(55.3)	109	(22.7)		92	(20.1)	266	(55.3)	117	(24.6)	
Lower (NVQ 1 and 2)	711	(33.7)	1,288	(55.2)	274	(11.1)		682	(32.2)	1,245	(54.0)	340	(13.8)	
Other/None [§]	314	(48.7)	317	(42.2)	84	(9.1)	<0.001	312	(48.5)	314	(43.2)	86	(8.3)	<0.001
Highest parental social class														
Managerial/Professional	555	(22.7)	1,330	(54.3)	612	(23.0)	<0.001	504	(20.5)	1,242	(50.6)	745	(28.9)	<0.001
Intermediate	424	(33.6)	758	(53.3)	192	(13.1)		398	(31.4)	761	(55.3)	211	(13.3)	
Semi-routine/routine	397	(45.8)	470	(46.7)	85	(7.5)		409	(46.2)	437	(45.1)	103	(8.7)	
Not applicable	96	(51.9)	81	(39.1)	26	(9.0)		90	(52.4)	89	(39.4)	23	(8.2)	
Mother working														
No	854	(39.5)	1,168	(46.7)	400	(13.8)	<0.001	808	(37.6)	1,153	(46.8)	455	(15.6)	<0.001
Yes	618	(25.4)	1,471	(56.5)	515	(18.1)		593	(24.4)	1,376	(53.5)	627	(22.1)	

Supplementary material 2 (cont.). Univariate association between participants’ characteristics and KS4 results at age 16 in English (n=5,026) and Mathematics (n=5,012) GCSEs*, Millennium Cohort Study.

Characteristics	English						p	Math						p [¶]
	Failed		Low pass		High pass (A-A*)			Failed		Low pass		High pass (A-A*)		
	n	(%)	n	(%)	n	(%)		n	(%)	n	(%)	n	(%)	
Sociodemographic (cont.)														
Maternal relationship status														
Married	834	(27.1)	1,730	(52.9)	696	(20.0)	<0.001	796	(25.4)	1,636	(50.8)	824	(23.8)	<0.001
Cohabitation	398	(36.7)	654	(51.2)	169	(12.1)		370	(35.0)	642	(51.1)	201	(13.9)	
Single mother	240	(46.1)	255	(46.7)	50	(7.2)		235	(46.4)	251	(45.1)	57	(8.5)	
Maternal ethnicity														
White	1,259	(33.3)	2,215	(51.0)	773	(15.7)	0.047	1,169	(30.3)	2,148	(49.9)	917	(18.9)	0.734
Other ethnicity	213	(26.5)	424	(55.5)	142	(18.0)		232	(31.3)	381	(51.7)	165	(17.9)	
Language spoken at home														
English only	1,302	(33.3)	2,294	(51.2)	784	(15.5)	0.001	1,218	(31.5)	2,233	(50.1)	915	(18.4)	0.011
English + other language	170	(24.4)	345	(55.4)	131	(20.2)		183	(27.0)	296	(49.3)	167	(23.7)	
Index of Multiple Deprivation														
Poorest quintile	491	(44.6)	624	(46.0)	168	(9.4)	<0.001	474	(43.4)	617	(45.0)	190	(11.6)	<0.001
Second quintile	355	(35.6)	579	(50.7)	182	(13.7)		331	(34.1)	574	(50.8)	206	(15.1)	
Middle quintile	259	(30.8)	556	(53.3)	182	(15.9)		265	(30.4)	529	(53.1)	203	(16.5)	
Fourth quintile	199	(25.1)	455	(55.2)	179	(19.7)		180	(23.2)	419	(52.0)	228	(24.8)	
Richest quintile	168	(22.9)	425	(53.7)	204	(23.4)		151	(20.3)	390	(20.1)	255	(29.6)	
Maternal cognitive score**	-0.26	(0.91)	0.02	(0.96)	0.47	(1.11)	<0.001	-0.29	(0.91)	0.00	(0.95)	0.48	(1.08)	<0.001

*Parameters were estimated considering the complex design of the MCS sample. The counts reflect the absolute number of participants included.

**Mean (S.D.)

¶Chi-square or mean difference, as appropriate.

§No education and Other/overseas educational qualifications were displayed together as a safeguarding procedure, as some counts were <10.

Supplementary material 3. Univariate association between participants' characteristics and KS4 results at age 16: Passing five GCSEs with marks ≥ 5 , including English and Mathematics (n=5,012) and Attainment 8 score (n=4,940)*, Millennium Cohort Study.

Characteristics	Five GCSEs ≥ 5 marks				p	Attainment 8		p [¶]
	No		Yes			Mean	S.D.	
	n	(%)	n	(%)				
Pregnancy and child-related								
Gestational age at birth**	39.7	(1.2)	39.8	(1.3)	0.019	0.03		0.017
Sex								
Female	865	(38.2)	1,674	(61.8)	<0.001	49.4	(18.7)	<0.001
Male	1,117	(48.9)	1,356	(51.1)		44.9	(18.7)	
Age at Key Stage 4**	16.2	(0.3)	16.2	(0.3)	0.120	0.02		0.117
Older siblings								
No	970	(42.9)	1,553	(57.1)	0.347	47.5	(18.7)	0.197
Yes	1,012	(44.4)	1,477	(55.6)		46.7	(19.0)	
Smoked during pregnancy								
Never	1,230	(36.9)	2,320	(63.1)	<0.001	50.1	(19.6)	<0.001
Gave up	245	(43.5)	362	(56.5)		46.5	(17.3)	
Kept smoking	507	(62.7)	348	(37.3)		38.8	(15.0)	
Mod/heavy alcohol in preg.								
No	1,818	(43.2)	2,817	(56.8)	0.088	47.3	(18.8)	0.009
Yes	164	(48.4)	213	(51.6)		44.4	(19.5)	
Sociodemographic, age 9 m.								
Maternal age (years)**	28.6	(5.9)	30.7	(5.8)	<0.001	0.18		<0.001
Maternal education								
Higher (NVQ 4 and 5)	438	(28.4)	1,120	(71.6)	<0.001	55.4	(21.4)	<0.001
Medium (NVQ 3)	151	(32.3)	324	(67.7)		51.8	(19.7)	
Lower (NVQ 1 and 2)	990	(46.4)	1,277	(53.6)		44.8	(16.3)	
Other/None [§]	403	(62.2)	309	(37.8)		39.3	(16.5)	
Highest parental social class								
Managerial/Professional	737	(30.5)	1,754	(69.5)	<0.001	53.1	(19.7)	<0.001
Intermediate	586	(45.8)	784	(54.2)		45.2	(16.9)	
Semi-routine/routine	535	(60.4)	414	(39.6)		40.0	(15.5)	
Not applicable	124	(67.7)	78	(32.3)		37.4	(18.1)	
Mother working								
No	1,116	(51.4)	1,300	(48.6)	<0.001	44.2	(18.9)	<0.001
Yes	866	(35.6)	1,730	(64.4)		50.1	(18.3)	

Supplementary material 3 (cont.). Univariate association between participants’ characteristics and KS4 results at age 16: Passing five GCSEs with marks ≥5, including. English and Mathematics (n=5,012) and Attainment 8 score (n=4,940)*, Millennium Cohort Study.

Characteristics	Five GCSEs ≥5 marks				p	Attainment 8		p
	Yes		No			Mean	S.D.	
	n	(%)	n	(%)				
Sociodemographic (cont.)								
Maternal relationship status								
Married	1,128	(36.1)	2,128	(63.9)	<0.001	50.6	(19.8)	<0.001
Cohabitation	531	(49.2)	682	(50.8)		43.9	(16.9)	
Single mother	323	(62.4)	220	(37.6)		39.2	(14.9)	
Maternal ethnicity								
White	1,669	(44.1)	2,565	(55.9)	0.145	46.9	(18.4)	0.030
Other ethnicity	313	(39.7)	465	(60.3)		48.8	(22.0)	
Language spoken at home								
English only	1,737	(44.3)	2,629	(55.7)	0.001	46.8	(18.3)	<0.001
English + other language	245	(35.4)	401	(64.6)		50.8	(23.7)	
Index of Multiple Deprivation								
Poorest quintile	658	(58.8)	623	(41.2)	<0.001	41.0	(19.1)	<0.001
Second quintile	478	(47.8)	633	(52.2)		45.5	(18.9)	
Middle quintile	368	(42.3)	629	(57.7)		47.0	(17.4)	
Fourth quintile	268	(34.9)	559	(65.1)		51.4	(17.4)	
Richest quintile	210	(28.8)	586	(71.2)		52.7	(18.8)	
Maternal cognitive score**	-0.28	(0.92)	0.21	(1.01)	<0.001	0.28		<0.001

*Parameters were estimated considering the complex design of the MCS sample. The counts reflect the absolute number of participants included.

**Mean (S.D.). Correlation coefficient between continuous variables and Attainment 8 score.

[¶]Chi-square, mean difference or correlation, as appropriate.

[§]No education and Other/overseas educational qualifications were displayed together as a safeguarding procedure, as some counts were <10.