

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		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	(%)
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														
<i>Female</i>	558	(24.9)	1,399	(54.5)	586	(20.6)	<0.001	679	(30.2)	1,326	(51.4)	534	(18.4)	0.259
<i>Male</i>	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														
<i>No</i>	724	(32.2)	1,328	(51.5)	481	(16.3)	0.715	668	(29.9)	1,286	(50.7)	569	(19.4)	0.198
<i>Yes</i>	748	(33.0)	1,311	(51.5)	434	(15.5)		733	(32.5)	1,243	(49.4)	513	(18.1)	
Smoked during pregnancy														
<i>Never</i>	909	(27.3)	1,913	(53.6)	738	(19.1)	<0.001	872	(26.5)	1,790	(50.6)	88	(22.9)	<0.001
<i>Gave up</i>	170	(29.9)	334	(55.1)	103	(15.0)		166	(29.3)	331	(53.5)	110	(17.2)	
<i>Kept smoking</i>	393	(49.0)	392	(43.6)	74	(7.4)		363	(45.3)	408	(46.5)	84	(8.2)	
Mod/heavy alcohol in preg.														
<i>No</i>	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
<i>Yes</i>	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														
<i>Higher (NVQ 4 and 5)</i>	341	(22.3)	773	(49.5)	448	(28.2)	<0.001	315	(20.2)	704	(45.5)	539	(34.3)	<0.001
<i>Medium (NVQ 3)</i>	106	(22.0)	261	(55.3)	109	(22.7)		92	(20.1)	266	(55.3)	117	(24.6)	
<i>Lower (NVQ 1 and 2)</i>	711	(33.7)	1,288	(55.2)	274	(11.1)		682	(32.2)	1,245	(54.0)	340	(13.8)	
<i>Other/None[§]</i>	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														
<i>Managerial/Professional</i>	555	(22.7)	1,330	(54.3)	612	(23.0)	<0.001	504	(20.5)	1,242	(50.6)	745	(28.9)	<0.001
<i>Intermediate</i>	424	(33.6)	758	(53.3)	192	(13.1)		398	(31.4)	761	(55.3)	211	(13.3)	
<i>Semi-routine/routine</i>	397	(45.8)	470	(46.7)	85	(7.5)		409	(46.2)	437	(45.1)	103	(8.7)	
<i>Not applicable</i>	96	(51.9)	81	(39.1)	26	(9.0)		90	(52.4)	89	(39.4)	23	(8.2)	
Mother working														
<i>No</i>	854	(39.5)	1,168	(46.7)	400	(13.8)	<0.001	808	(37.6)	1,153	(46.8)	455	(15.6)	<0.001
<i>Yes</i>	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														
<i>Married</i>	834	(27.1)	1,730	(52.9)	696	(20.0)	<0.001	796	(25.4)	1,636	(50.8)	824	(23.8)	<0.001
<i>Cohabitation</i>	398	(36.7)	654	(51.2)	169	(12.1)		370	(35.0)	642	(51.1)	201	(13.9)	
<i>Single mother</i>	240	(46.1)	255	(46.7)	50	(7.2)		235	(46.4)	251	(45.1)	57	(8.5)	
Maternal ethnicity														
<i>White</i>	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
<i>Other ethnicity</i>	213	(26.5)	424	(55.5)	142	(18.0)		232	(31.3)	381	(51.7)	165	(17.9)	
Language spoken at home														
<i>English only</i>	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
<i>English + other language</i>	170	(24.4)	345	(55.4)	131	(20.2)		183	(27.0)	296	(49.3)	167	(23.7)	
Index of Multiple Deprivation														
<i>Poorest quintile</i>	491	(44.6)	624	(46.0)	168	(9.4)	<0.001	474	(43.4)	617	(45.0)	190	(11.6)	<0.001
<i>Second quintile</i>	355	(35.6)	579	(50.7)	182	(13.7)		331	(34.1)	574	(50.8)	206	(15.1)	
<i>Middle quintile</i>	259	(30.8)	556	(53.3)	182	(15.9)		265	(30.4)	529	(53.1)	203	(16.5)	
<i>Fourth quintile</i>	199	(25.1)	455	(55.2)	179	(19.7)		180	(23.2)	419	(52.0)	228	(24.8)	
<i>Richest quintile</i>	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								
<i>Female</i>	865	(38.2)	1,674	(61.8)	<0.001	49.4	(18.7)	<0.001
<i>Male</i>	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								
<i>No</i>	970	(42.9)	1,553	(57.1)	0.347	47.5	(18.7)	0.197
<i>Yes</i>	1,012	(44.4)	1,477	(55.6)		46.7	(19.0)	
Smoked during pregnancy								
<i>Never</i>	1,230	(36.9)	2,320	(63.1)	<0.001	50.1	(19.6)	<0.001
<i>Gave up</i>	245	(43.5)	362	(56.5)		46.5	(17.3)	
<i>Kept smoking</i>	507	(62.7)	348	(37.3)		38.8	(15.0)	
Mod/heavy alcohol in preg.								
<i>No</i>	1,818	(43.2)	2,817	(56.8)	0.088	47.3	(18.8)	0.009
<i>Yes</i>	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								
<i>Higher (NVQ 4 and 5)</i>	438	(28.4)	1,120	(71.6)	<0.001	55.4	(21.4)	<0.001
<i>Medium (NVQ 3)</i>	151	(32.3)	324	(67.7)		51.8	(19.7)	
<i>Lower (NVQ 1 and 2)</i>	990	(46.4)	1,277	(53.6)		44.8	(16.3)	
<i>Other/None</i> [§]	403	(62.2)	309	(37.8)		39.3	(16.5)	
Highest parental social class								
<i>Managerial/Professional</i>	737	(30.5)	1,754	(69.5)	<0.001	53.1	(19.7)	<0.001
<i>Intermediate</i>	586	(45.8)	784	(54.2)		45.2	(16.9)	
<i>Semi-routine/routine</i>	535	(60.4)	414	(39.6)		40.0	(15.5)	
<i>Not applicable</i>	124	(67.7)	78	(32.3)		37.4	(18.1)	
Mother working								
<i>No</i>	1,116	(51.4)	1,300	(48.6)	<0.001	44.2	(18.9)	<0.001
<i>Yes</i>	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								
<i>Married</i>	1,128	(36.1)	2,128	(63.9)	<0.001	50.6	(19.8)	<0.001
<i>Cohabitation</i>	531	(49.2)	682	(50.8)		43.9	(16.9)	
<i>Single mother</i>	323	(62.4)	220	(37.6)		39.2	(14.9)	
Maternal ethnicity								
<i>White</i>	1,669	(44.1)	2,565	(55.9)	0.145	46.9	(18.4)	0.030
<i>Other ethnicity</i>	313	(39.7)	465	(60.3)		48.8	(22.0)	
Language spoken at home								
<i>English only</i>	1,737	(44.3)	2,629	(55.7)	0.001	46.8	(18.3)	<0.001
<i>English + other language</i>	245	(35.4)	401	(64.6)		50.8	(23.7)	
Index of Multiple Deprivation								
<i>Poorest quintile</i>	658	(58.8)	623	(41.2)	<0.001	41.0	(19.1)	<0.001
<i>Second quintile</i>	478	(47.8)	633	(52.2)		45.5	(18.9)	
<i>Middle quintile</i>	368	(42.3)	629	(57.7)		47.0	(17.4)	
<i>Fourth quintile</i>	268	(34.9)	559	(65.1)		51.4	(17.4)	
<i>Richest quintile</i>	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.