Quaderni di Statistica
Vol. 10, 2008

Analysis of early life influences on cognitive development in childhood using multilevel ordinal models

Leah Li
MRC Centre of Epidemiology for Child Health/ Centre for Paediatric Epidemiology and Biostatistics, UCL Institute of Child Health
E-mail: l.li@ich.ucl.ac.uk

Summary: Studies of cognitive development in children are often based on tests designed for specific ages. Examination of the changes of these scores over time may not be meaningful. This paper investigates the influence of early life factors on cognitive development using maths and reading test scores at ages 7, 11, and 16 years in a British birth cohort born in 1958. The distributions of these test scores differ between ages, for example, 20% participants scored the top mark in the reading test at 7 and the distribution of reading score at 16 is heavily skewed. In this paper, we group participants into 5 ordered categories, approximately 20% in each category according to their test scores at each age. Multilevel models for a repeated ordinal outcome are applied to relate the ordinal scale of maths and reading ability to early life factors.

Keywords: Maths and reading scores, Repeated ordinal responses, Multilevel models.

1. Introduction

Cognitive development in childhood is associated with early life determinants (Lawlor *et al.* 2005) and adult health outcomes (Starr *et al.*2004; Hart *et al.*2003; Kuh *et al.*2004; Batty and Deary 2004). One of the possible explanations of the association is that individuals with higher scores in intelligence tests tend to have a higher social economic status and better health behaviour, which are inversely associated with adult disease risks. It is also possible, that cognitive development acts as an indicator of early life circumstances, which are known to have an effect on adult health. In order to establish how childhood cognition influences health in later life, it is important to understand how early life circumstances are associated with cognitive development in childhood.

Most studies of influences on cognition in childhood focus on cognitive ability measured at one particular age. Few have explored influences on cognitive trajectories through childhood and adolescence. A previous study of the 1958 British cohort examined the influences of birthweight and social class on maths test scores at multiple ages in childhood, and their reading scores at these ages were not fully explored because of the highly skewed distributions at some ages (Jefferis *et al.* 2002). In the current paper we group individuals into five ordered categories (i.e. quintiles) according to their ranks in maths and reading tests within the sample. We apply multilevel models for repeated ordinal outcomes to gain additional insights on the influences of a range of early life factors, including birthweight, maternal smoking during pregnancy, maternal age at child birth, maternal education, breastfeeding, social class, house tenure, and mother read regularly to the child, on trajectories of maths and reading ability during childhood, accounting for the within individual correlations of the test scores. The trajectories of cognitive development in childhood are characterised as maths and reading scores in early childhood (7 years) and progress during primary (7-11 years) and secondary school years (11-16 years). Specifically, we aim to investigate whether these early life influences on maths and reading ability strengthen with increasing age.

2. Data

We use the data from the 1958 British birth cohort, which includes all children born in one week in 1958 in England, Wales, and Scotland. Approximately 17,000 live births were followed up on 8 occasions from birth to 45 years (Power and Elliott 2006). At ages 7, 11, and 16 years, the participants took age appropriate tests at school for maths and reading ability. Tests of perceptual and motor ability skills were conducted at ages 7 and 11 years, and a general ability test approximated the conventional intelligence test, with verbal and nonverbal components (Douglas 1964) was conducted at age 11 years. In the current study we concentrate on maths and reading test scores. The analysis includes 12,903 individuals with maths or reading test scores at one or more ages and information on early life factors considered in this study.

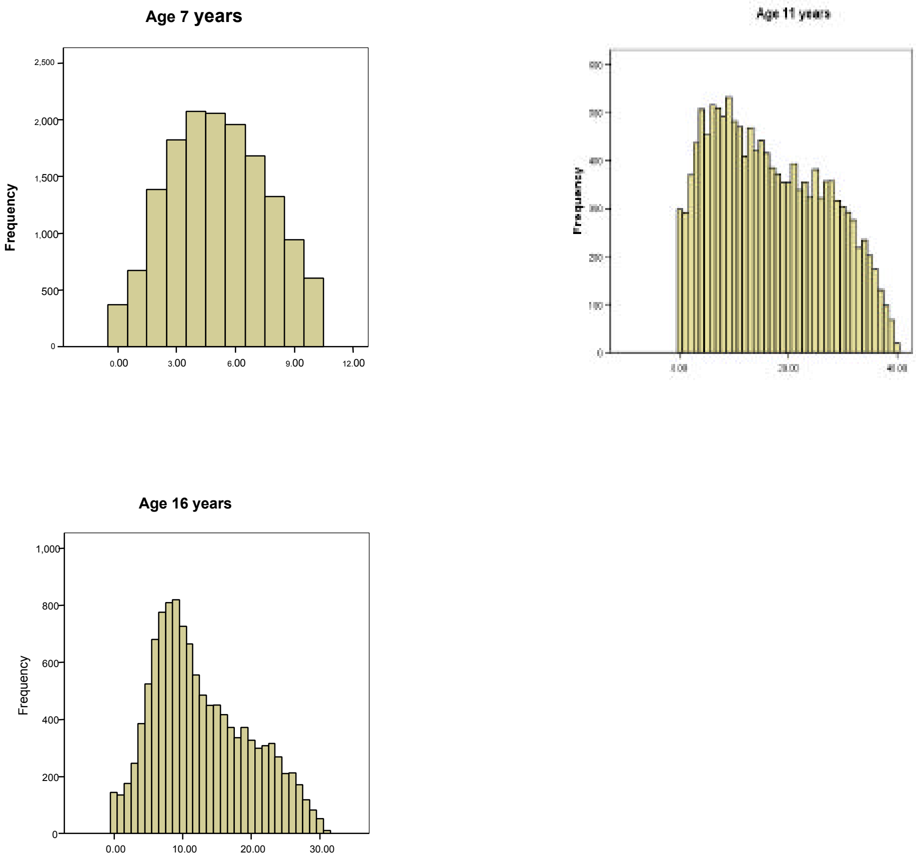

Figure 1. Distributions of maths test scores at ages 7, 11, and 16 years

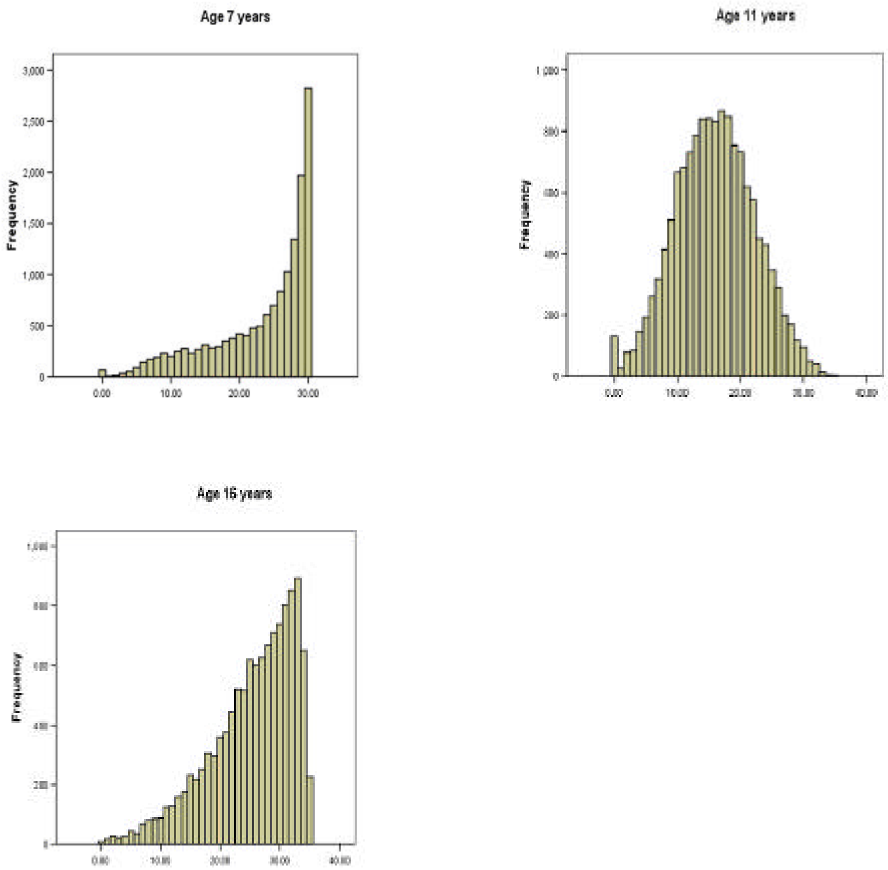

Figure 2. Distributions of reading test scores at ages 7, 11, and 16 years

Maths and reading tests in childhoods

The maths test at age 7 years comprised ten arithmetic problems with graded levels of difficulty (range 0-10). At age 11 the maths test (range 0-40) was constructed by the National Foundation for Educational Research in England and Wales. At age 16 a maths comprehension test (range 0-31) was constructed at Manchester University. The Southgate reading test (range 0-30) was used at age 7 to detect poor readers (Southgate 1962): children selected one word corresponding to a picture from a list of several words and teachers read out words that the children had to identify from a printed list. The reading comprehension tests used at ages 11 and 16 (range 0-35) were parallel to the Watts Vernon comprehension test.

These maths and reading tests were designed for specific ages and their scales differed. Thus the examination of changes in raw scores between ages is not meaningful. At ages 7 and 16 years, most children scored high marks on their reading tests, resulting in highly skewed distributions; at 7 years, 20% of children score the top mark of 30 (range 0-30, mean 23.4), and at 16 years 35% of children scored higher than 30 (range 0-35, mean 25.4) (Figure 2). Reading score at 11 years and maths scores at all ages discriminate a range of abilities (Figures 1 and 2).

At each survey, the tests were conducted over several months, thus the exact age is adjusted for so that the maths and reading scores are centred at ages 7, 11, and 16 years. We then group participants into quintiles, each has approximately 20% of the sample according to their maths or reading scores at each age. For each test, those scores in the ‘’top fifth’’ of the distribution are defined as ‘’highest scores’’ and the bottom fifth ‘’lowest scores’’. The categories at age 7 years and changes in categories between 7-16 years reflect the individual’s cognitive trajectories.

Explanatory variables

Prenatal factors considered include birthweight, both as a continuous variable (kg) and a categorical variable (<2.5 kg, 2.5-3 kg, 3-3.5 kg, 3.5-4 kg, >4 kg), gestational age (weeks), maternal smoking during pregnancy (yes vs no), maternal age at child birth (years), and maternal education (whether the mother stayed at school after the minimum school-leaving age). Early life factors considered are breastfeeding (yes vs no), social class (based on occupation of the male head, non-manual vs manual social class), housing tenure (social housing vs owner/private renter), and mother regularly (weekly) read to the child.

3. Statistical methods

The five ordered categories of responses (corresponding to quintiles for test scores) are denoted by integer labels k =1 to 5. We apply two-level ordinal models to maths and reading scores separately (Goldstein 1995). Let y_{ij} be the observed ordinal response of jth measurement (level-1) for ith individual (level-2), the cumulative probability P(y_{ij} £ $\boldsymbol{k}$) is the probability that a response variable will take on a value less than or equal to a particular value $\boldsymbol{k}$ (1, 2, 3, or 4). Because there are only three time points, we entered age as a categorical variable. We chose age 11 years as the baseline category, thus variables x_7 and x_{16} in the model contrast the ages 7 and 16 years with the age 11 years respectively. The independent variable z_i is an individual level covariate (i.e. an early life factor). Following the single level models defined by McCullagh and Nelder (1989), we apply a two-level model relating the ordinal scale of the test score to age terms and an early life factor (z) written as:

$$\mathrm{Log}\left[\frac{P(y_{ij} \le k)}{1-P(y_{ij} \le k)}\right] = a_{0k} + b_1 x_{7i} + b_2 x_{16i} + b_3 z_i + b_4 (x_{7i} z_i) + b_5 (x_{16i} z_i) + m_i$$

The left side of model specifies 4 cumulative logits, each contrasting the combined first k (1, 2, 3, 4) categories to the remaining (5 - k) categories, i.e. maths or reading scores £ 20% vs >20%, £ 40% vs >40%, £ 60% vs >60%, and £ 80% vs >80% in the distribution.

The individual level random effect μ_{i} is assumed to be N (0, s_m^2) distributed, thus s_m^2 is the level-2 between-individual variance. Individual responses (level-1) are assumed to follow a multinomial distribution.

In terms of regression coefficients, a_{0k} are the 4 intercept terms to model the marginal frequencies in the 5 ordered response categories. The regression coefficient b is the log odds ratio of a response measure that falls into a category or below for each unit change of the covariate. The model assumes the covariates (b s) are homogeneous across *k*-1 cumulative logits (proportional odds assumption). Thus parameters b s do not carry the *k* subscript, and a single effect is estimated for each covariate. As the highest category (top fifth) is the reference category, a negative value for *ß* implies a positive association between the response and the covariate, for example, an increased chance to have a higher response category (or higher test score) for a higher value of the covariate.

The effect of an early life factor on cognitive development over time is represented by the interaction terms between the fixed covariate and the dummy variable of age. As age 11 years is the baseline, the effect of a specific early life factor (z) on maths or reading score at age 7 years, progress between 7-11 years, and progress between 11-16 years (i.e. changes in relative positions in the distribution) are represented by $b_3 + b_4$, $-b_4$, b_5 respectively. A negative value of $b_3 + b_4$ indicates a higher score at age 7 years, and negative values of $-b_4$ and b_5 indicate improvement between 7-11 years and between 11-16 years respectively.

Table 1. Description of maths and reading scores at ages 7, 11, and 16 years

	N	Range	Mean (sd)	
Maths score at			Male	Female
7 years	14,898	0-10	5.22 (2.50)	5.00 (2.48)
11 years	14,129	0-40	16.81 (10.60)	16.44 (10.08)
16 years	11,921	0-31	13.39 (7.27)	12.09 (6.64)
Reading score at				
7 years	14,931	0-30	22.44 (7.43)	24.29 (6.70)
11 years	14,133	0-35	15.93 (6.55)	16.03 (6.02)
16 years	11,987	0-35	25.35 (7.34)	25.26 (6.83)

Table 1 displays the descriptive measures for maths and reading scores at each age. Tables 2 and 3 provide results from the multilevel models for repeated ordinal responses assuming homogeneous covariate effects. We first include gender, age, an early factor, and its interactions with the age terms in the model to assess the effect of the early life factor on maths or reading categories at 7 years, progress between 7-11 years and between 11-16 years. As birthweight shows a nonlinear effect, it is also considered as a categorical variable. Second, we extend the model by further including a range of individual characteristics such as birthweight, gestational age (for birthweight only), maternal age at child birth, maternal education, breastfeeding, social class, and housing tenure.

We fitted the two-level ordinal models using second order marginal quasi-likelihood (MQL) in the MLwiN (Goldstein 1995). The effects on score at 7 years ($b_3 + b_4$), progress between 7-11 years ($-b_4$), and 11-16 years (b_5) were estimated from the models and tested using contrast tests (Goldstein 1995).

4. Results

While a majority of individuals performed consistently over time, some individuals improved or declined relative to their peers within the study population. The correlations of test scores between 7-11 years (0.60 for maths and 0.55 for reading tests) are weaker than those between 11-16 years (0.79 and 0.74 respectively), suggesting that there was more mobility during primary school than secondary school years.

Gender differences Tables 2 and 3 show that there were gender differences in maths and reading scores, and the patterns differed between the two subjects. For maths scores, a positive coefficient for girls at 7 years indicates that girls were more likely to be in categories of lower score than boys. Although girls made more progress between 7-11 years, boys were more likely to move upwards relative to girls between 11-16 years (Table 2). Therefore at age 16 boys performed better on average than girls. For example, 23.3% of boys were in the top 20% of the distribution compared to 16.6% of girls. For reading scores, a negative coefficient for girls at age 7 years indicates that girls were more likely to be in categories of higher scores than boys. But throughout childhood boys made more progress relative to girls (Table 3). Thus there was little gender difference by age 16.

Most early life factors had a similar influence on maths and reading scores. Their influence on scores at 7 years became weaker (some non-significant) after the adjustment for birthweight, maternal age, maternal education, breastfeeding, social class, and housing tenure. But their influence on progress between 7-11 years persisted and was stronger than their influence on progress between 11-16 years.

Birthweight had a significant quadratic association with both maths and reading scores at 7 years. But its interactions with age terms were non-significant, suggesting that there was no association with changes in scores between 7-11 years and between 11-16 years. Thus the effect of birthweight on test scores was evident at age 7 years and remained throughout childhood. The non-linear association suggests that the effect weakened with increasing birthweight. This is confirmed by the analysis using birthweight as a categorical variable: the effect of birthweight was greater for those <3.5 kg.

Maternal smoking was associated with lower maths and reading scores at 7 years and poorer progress between 7-11 years, but not with progress during secondary school years. After the adjustment, the association with maths and reading scores at 7 years weakened substantially and was no longer significant for maths score, but the effect of maternal smoking on progress between 7-11 years remained.

Maternal age was positively associated with higher maths and reading scores at 7 years and progress between 7-11 years, but not with progress between 11-16 years. After the adjustment, its association with scores at 7 years was no longer significant, but its association with changes between 7-11 years persisted.

Table 2. Estimated effects (se) on maths score at 7 years, progress between 7-11 and between 11-16 years from multilevel ordinal models

Parameters	Effect on score at	Maths score Unadjusted†	Adjusted‡
Girls	7 years	0.400 (0.046)*	
	7-11 years	-0.348 (0.047)*	
	11-16 years	0.363 (0.051)*	
Birthweight (kg)#	Linear	-2.011 (0.267)*	-1.620 (0.261)*
	Quadratic	0.220 (0.039)*	0.175 (0.038)*
<2.5 kg		1.032 (0.078)*	0.838 (0.079)*
2.5-3 kg		0.567 (0.049)*	0.438 (0.048)*
3-3.5 kg		0.243 (0.040)*	0.186 (0.039)*
3.5-4 kg		-	-
>4 kg		-0.045 (0.061)	-0.069 (0.058)
Maternal smoking	7 years	0.303 (0.043)*	0.032 (0.041)
	7-11 years	0.302 (0.048)*	0.311 (0.050)*
	11-16 years	0.055 (0.045)	0.041 (0.055)
Maternal age	7 years	-0.011 (0.004)*	-0.003 (0.003)
	7-11 years	-0.014 (0.004)*	-0.012 (0.004)*
	11-16 years	-0.004 (0.005)	-0.005 (0.005)
Maternal education	7 years	-0.899 (0.049)*	-0.442 (0.047)*
	7-11 years	-0.585 (0.054)*	-0.587 (0.055)*
	11-16 years	-0.127 (0.059)*	-0.101 (0.059)
Breastfeeding	7 years	-0.212 (0.022)*	-0.070 (0.022)*
	7-11 years	-0.119 (0.027)*	-0.118 (0.027)*
	11-16 years	-0.064 (0.030)*	-0.045 (0.030)
Non-manual class	7 years	-0.771 (0.045)*	-0.264 (0.046)*
	7-11 years	-0.708 (0.053)*	-0.707 (0.053)*
	11-16 years	-0.120 (0.058)*	-0.092 (0.058)
Social housing	7 years	0.598 (0.041)*	0.186 (0.041)*
	7-11 years	0.505 (0.048)*	0.506 (0.048)*
	11-16 years	0.027 (0.052)	0.014 (0.052)
Mother read weekly	7 years	-0.098 (0.039)*	-0.138 (0.039)*
	7-11 years	-0.359 (0.047)*	-0.370 (0.047)*
	11-16 years	-0.115 (0.051)*	-0.086 (0.051)

† unadjusted models include each early life factor and gender.
‡ adjusted models include each early life factor and gender, also birthweight, gestational age (for birthweight only), breastfeeding, maternal age and education, social class, housing tenure.
#quadratic relationship between birthweight and maths and reading scores, no significant interaction terms with age.
* significant at level 0.05.

Table 3. Estimated effects (se) on reading score at 7 years, progress between 7-11 and between 11-16 years from multilevel ordinal models

Parameters	Effect on score at	Reading score Unadjusted†	Reading score Adjusted‡
Girls	7 years	-0.603(0.044)*	
	7-11 years	0.526(0.048)*	
	11-16 years	0.220(0.052)*	
Birthweight (kg)#	Linear	-2.331(0.273)*	-1.886(0.263)*
	Quadratic	0.269(0.041)*	0.220(0.039)*
<2.5 kg		1.016(0.088)*	0.836(0.083)*
2.5-3 kg		0.525(0.054)*	0.386(0.050)*
3-3.5 kg		0.210(0.043)*	0.137(0.041)*
3.5-4 kg		-	-
>4 kg		-0.038(0.065)	-0.041(0.061)
Maternal smoking	7 years	0.453(0.047)*	0.142(0.044)*
	7-11 years	0.143(0.053)*	0.156(0.052)*
	11-16 years	0.050(0.058)	0.042(0.057)
Maternal age	7 years	-0.010(0.005)*	0.001(0.004)
	7-11 years	-0.034(0.004)*	-0.034(0.004)*
	11-16 years	-0.001(0.001)	-0.001(0.001)
Maternal education	7 years	-1.115(0.048)*	-0.586(0.049)*
	7-11 years	-0.466(0.053)*	-0.456(0.055)*
	11-16 years	-0.111(0.051)*	-0.112(0.060)*
Breastfeeding	7 years	-0.245(0.025)*	-0.084(0.023)*
	7-11 years	-0.119(0.028)*	-0.123(0.027)*
	11-16 years	-0.067(0.030)*	-0.060(0.030)*
Non-manual class	7 years	-1.108(0.046)*	-0.535(0.048)*
	7-11 years	-0.443(0.053)*	-0.436(0.054)*
	11-16 years	-0.091(0.058)	-0.091(0.058)
Social housing	7 years	0.868(0.042)*	0.404(0.050)*
	7-11 years	0.304(0.048)*	0.315(0.048)*
	11-16 years	0.099(0.052)	0.088(0.052)
Mother read weekly	7 years	-0.301(0.043)*	-0.034(0.040)
	7-11 years	-0.342(0.048)*	-0.351(0.047)*
	11-16 years	-0.089(0.052)	-0.065(0.052)

† unadjusted models include each early life factor and gender.
‡ adjusted models include each early life factor and gender, also birthweight, gestational age (for birthweight only), breastfeeding, maternal age and education, social class, housing tenure.
#quadratic relationship between birthweight and maths and reading scores, no significant interaction terms with age.
* significant at level 0.05.

Breastfeeding had a positive association with maths and reading scores at 7 years, progress between 7-11 years, and a weaker association with progress between 11-16 years. After the adjustment, the association with scores at 7 years weakened substantially, but the association with progress between 7-16 years changed little.

Factors indicating socio-economic background such as *social class*, *housing tenure*, and *maternal education* were all strongly associated with maths and reading ability in childhood. Children from non-manual social class, owner occupied/privately rented accommodations, or mother stayed in school already performed better at 7 years, they improved considerably in primary school years compared to their peers. Although they did not make more progress between 11-16 years, they continued to perform better compared to their peers. The associations remain after adjustment for other factors.

Mother read to the child was associated with maths and reading scores at 7 years and progress in primary school, and in secondary school only for maths. After the adjustment, the association with maths score at 7 years remained significant, and its association with progress in maths and reading between 7-11 years remained unchanged, and its association with progress in maths between 11-16 years became non-significant.

5. Discussion

The evidence of early life influences on cognitive trajectories in childhood is limited as this requires repeated measurements on cognitive ability at different ages. This study has many strengths with respect to influences on cognitive development. The 1958 birth cohort has repeated maths and reading test scores from early childhood through to adolescence and information on early life circumstances. It is therefore an ideal sample to examine how prenatal and early childhood factors influence the cognitive trajectory in childhood. Such data allow the use of multilevel models, which take into account the strong correlation of repeated measures (correlation coefficients 0.55-0.80). The models using age as fixed occasions are practical for the number of repeated measurements available in our data. Importantly in longitudinal studies with sample attrition, multilevel models have the flexibility to include all subjects with one or more test scores. Thus the majority of study participants are included in the analysis.

Ordinal regression models are often used when data are collected on an ordered categorical scale (Lunn *et al.* 2001; Ribaudo *et al.* 1999). In studies of cognitive development in children, tests are often designed as age specific and are collected on a continuous scale as the maths and reading scores in the 1958 cohort. It is normally not desirable to categorise continuous variables as information from the measurements may be lost. In the current study however, repeated measures of response variables have different distributions. Moreover, there are inflated or skewed outcomes at some ages, and transformations do not sufficiently improve these distributions. Thus we group individuals into quintiles according to their test scores at each age and apply multilevel models for ordinal outcomes to examine their changes in relative positions of maths and reading ability within the study sample.

Models allowing heterogeneous effects for the early life factors (i.e. non-proportional odds assumption) are fitted to the data. There are no consistent patterns suggesting that the effects of these early life factors differ across the four cumulative logits. Therefore, only results from proportional odds models are presented here.

Prenatal and childhood factors have been found to be associated with childhood cognition at particular ages in many studies (Lawlor *et al.* 2005; Gomez-Sanchiz *et al.* 2003; Richards *et al.* 2002). Multilevel ordinal models allow us to examine the associations between early life factors and cognitive trajectories from early childhood to adolescence. Our results show that trajectories for maths and reading abilities differ between boys and girls. For early life factors, the broad pattern of associations is similar for maths and reading tests. Except for birthweight, the effect of early life factors strengthens with increasing age during primary school years (7-11 years), the effect remains but changes little during secondary school years (11-16 years). In general, poor maths and reading scores at 7 years and poor progress in primary school are associated with maternal smoking during pregnancy and social housing, while higher scores at 7 years and better progress in primary school are associated with increasing birthweight and maternal age, mother stayed in school after minimum age, breastfeeding, non-manual social class, and mother read to the child regularly.

Our findings suggest that cognitive development in childhood is influenced by early life circumstances, which are known to have a long term impact on adult health (Whalley and Deary 2001). Although cohort members born in 1958 would have experienced more socio-economic adversity compared to current British children, the findings of the inter-relationships between early life circumstances and cognitive trajectories in childhood are important as they are relevant to disease risks in current adult population.

Acknowledgments: Dr Leah Li is funded by the Medical Research Council Career Development Award in Biostatistics. This work was undertaken at the Great Ormond Street Hospital for Children/University College Londlon Institute of Child Health (London, UK) who received a proportion of funding from the Department of Health's NIHR Biomedical Research Centres funding scheme.

References

Batty, G. D. and I. J. Deary (2004), Early life intelligence and adult health, *British Medical Journal*, 329, 585-586.

Douglas, J.W.B. (1964), *The home and the school*, MacGibbon and Kee, London.

Goldstein, H. (1995), *Multilevel Statistical Models*, 2nd edition, John Wiley Inc., NewYork.

Gomez-Sanchiz, M., R. et al. (2003), Influence of breast-feeding on mental and psychomotor development, *Clinical .Pediatrics,* 42, 35-42.

Hart, C. L., et al. (2003), Childhood IQ, social class, deprivation, and their relationships with mortality and morbidity risk in later life: prospective observational study linking the Scottish Mental Survey 1932 and the Midspan studies, *Psychosomatic Medicine,* 65, 877-883.

Jefferis, B. J., C. Power, and C. Hertzman (2002), Birth weight, childhood socioeconomic environment, and cognitive development in the 1958 British birth cohort study; *British Medical Journal;* 325, 305-310.

Kuh, D., et al. (2004), Childhood cognitive ability and deaths up until middle age: a post-war birth cohort study, *International Journal of Epidemiology*, 33, 408-413.

Lawlor, D. A., et al. (2005), Early life predictors of childhood intelligence: evidence from the Aberdeen children of the 1950s study, *Journal of Epidemiology and Community Health*, 59, 656-663.

Lunn, D. J., J. Wakefield, and A. Racine-Poon (2001), Cumulative logit models for ordinal data: a case study involving allergic rhinitis severity scores, *Statistics in Medicine,* 20, 2261-2285.

McCullagh, B. and J. A. Nelder (1989), *Generalized Linear Models*, 2nd edition. Chapman&Hall, London.

Power, C. and J. Elliott (2006), Cohort profile: 1958 British birth cohort (National Child Development Study), *International Journal of Epidemiology*, 3534-41.

Ribaudo, H. J., M. Bacchi, J. Bernhard, and S. G. Thompson (1999), A multilevel analysis of longitudinal ordinal data: evaluation of the level of physical performance of women receiving adjuvant therapy for breast cancer, *Journal of the Royal Statistical, A*, 162, 349-360.

Richards, M., R. Hardy, D. Kuh, and M. E. Wadsworth (2002), Birthweight, postnatal growth and cognitive function in a national UK birth cohort, *International Journal of Epidemiology,* 31, 342-348.

Southgate, V. (1962), *Southgate group reading tests: manual of instructions,* University of London Press, London.

Starr, J. M., et al. (2004), Childhood mental ability and blood pressure at midlife: linking the Scottish Mental Survey 1932 and the Midspan studies, *Journal of Hypertension,* 22, 893-897.

Whalley, L. J. and I. J. Deary (2001), Longitudinal cohort study of childhood IQ and survival up to age 76; *British Medical Journal*, 322, 815-819.