This is a post-peer-review, pre-copyedit version of an article published in *European child & adolescent psychiatry*. The final authenticated version is available online at: https://doi.org/10.1007/s00787-016-0872-z.

# Title Page

**Authors:** Frances M Cronin¹, Ricardo Segurado², Fionnuala M McAuliffe³, Cecily C Kelleher¹, Richard E Tremblay¹, ⁴.

**Affiliations:**

¹ School of Public Health, Physiotherapy and Sports Science, University College Dublin, Ireland

² UCD CSTAR, School of Public Health, Physiotherapy and Sports Science, University College Dublin, Ireland.

³ UCD Obstetrics & Gynaecology, School of Medicine and Medical Science, University College Dublin, National Maternity Hospital, Dublin, Ireland.

⁴ Departments of Pediatrics and Psychology, University of Montreal, Canada.

**Manuscript title:** Gestational age and chronic ‘body-mind’ health problems in childhood: Dose-response association and risk factors

**Corresponding Author:** Frances M Cronin. School of Public Health, Physiotherapy and Sports Science, University College Dublin, Ireland. T: +353 1 716 3478 E: fran.cronin@ucd.ie.

**Conflicts of Interest and Source of Funding:** This work was supported by the Health Research Board, Ireland [grant number HRA_PHS/2011/35] to RT. There are no conflicts of interest to disclose.
Abstract

Understanding the developmental course of all health issues associated with preterm birth is important from an individual, clinical and public health point-of-view. Both the number of preterm births and proportion of survivors have increased steadily in recent years. The UK Millennium Cohort Study (n=18 818) was used to examine the association of gestational age with maternal ratings of general health and behavior problems at ages 5 and 11 years using binary and multinomial logistic regression analyses. The association between mothers’ ratings of general health and behavior problems was relatively weak at each time point. Children rated as being in poor general health remained constant over time (4.0% at age 5, 3.8% at age 11), but children rated as having behavioral problems increased by almost 100% (5.6% at 5; 10.5% at 11). A gradient of increasing risk with decreasing gestational age was observed for a composite health measure (general health problems and/or behavior problems) at age 5, amplified at age 11 and was strongest for those with chronic problems (poor health at both age 5 and age 11). This association was found to be compounded by child sex, maternal characteristics at birth (education, employment, marital status) and duration of breast feeding. Integrated support to at-risk families initiated during, or soon after pregnancy, may prevent chronic problems and might potentially reduce long term health costs for both the individual and health services.

**Keywords** Preterm Birth, Behavior Problems, Health Problems, Sex Differences, Breast Feeding, Millennium Cohort Study.
Introduction

A recent worldwide study of the increase in premature births estimated that every year 15 million babies are born preterm (PT). Of these, over one million babies a year die from prematurity, and survivors are at high risk of long term health impairments [1]. However, while the World Health Organisation (WHO) defines health as a ‘state of complete physical, mental and social well-being’ [2], it is striking that, to the best of our knowledge, most studies of health issues associated with being born prematurely have focused either on physical health outcomes only or mental health outcomes only. Thus, the extent of the impact of PT birth on health and health services is difficult to ascertain as a distinction between mental and physical illness continues in both lay and medical circles, and persists in scientific publications [3].

This silo mentality regarding the consequences of PT birth is well illustrated in a recent paper published in a high impact medical journal [4]. The Boyle et al., (2012) paper used the Millennium Cohort Study (MCS) - a longitudinal study with a large population sample - and described an increase in health problems during early childhood as gestational age (GA) decreases. However, while numerous aspects of physical health (e.g. hospital admissions, longstanding illnesses, asthma etc.) were assessed, the study did not include any dimension of mental health, although data on the children’s behavior problems were available.

Unless physical and mental outcomes are extensively comorbid in most cases, the silo approach to understanding the consequences of prematurity is underestimating the number of premature babies that suffer long term adverse health outcomes. Furthermore, to plan appropriate preventive and treatment services it is important to understand the developmental courses of both physical and mental health problems related to PT births.

There appears to be a general consensus that PT children of school-going age - compared to their full term (FT) peers - are at increased risk of symptoms associated with behavior
problems [5, 6]. However, study findings for behavior problems outcomes have been inconsistent [7]: some report those born PT as being at a ‘significantly higher risk’ of later behavioral problems [5], and others report a relatively ‘modest’ risk of later behavioural problems (e.g. [8]). Clarity in this area is especially needed as behaviour problems in childhood have been identified as risk factors for later adult mental health problems [9-12].

An additional challenge to synthesising studies of children’s physical health and studies of behavior problem outcomes related to prematurity is the use of a variety of methodologies, instruments, sample recruitment criteria [13-16], and overlapping gestation categories [17]. For example, while most studies compare cohorts of children born FT (39-41 weeks) with cohorts of PT children, the selection of ‘preterm’ categories can differ between studies (e.g. ≤33 weeks gestation [8]; <30 weeks [6]; <26 weeks [18]; <25 weeks [19]). Others focus only on the higher-risk children (e.g. those born <22 weeks [20]), or compare specific ranges of GAs (e.g. 22–32 weeks [5], or ‘mature preterm’ 32-36 weeks gestation [4]) with FT births. Finally, possibly due to the financial and organizational challenges inherent in long-term, population-based follow-ups, most long-term studies of children born PT have assessed health outcomes at only one point in time [21]. Repeated assessments of outcomes are important in order to understand to what extent adverse health outcomes are of a chronic nature and how the different types of problems might be associated with each other (e.g. [22, 23]).

The main objective of the present study was to replicate, as far as possible the methodological approach of the Boyle et al., (2012) paper on physical health to investigate the association of PT birth with both physical health problems and behavioral problems. Their population-based longitudinal study compared children born at FT with PT children categorized by GA: namely 37-38 weeks, 34-36 weeks, 32-33 weeks, and 23-31 weeks [24]. They reported a ‘dose-response’ effect of prematurity on children’s health: as GA decreased, parents were
more likely to assess their child’s general health as ‘Poor’ or ‘Fair’, and report more physical health problems (i.e. hospital admissions, longstanding illnesses and asthma) up to age five years. Studies of the effect of GA commonly adjust for factors known to affect children’s health and development [4]. Multiple studies conducted over the past decade have demonstrated the significant impact of a variety of demographic and environmental factors on the health outcomes of children of school-going age. Factors include the child’s sex [25], ethnicity [26] and age at time of interview; maternal characteristics at time of the pregnancy including maternal education [27], marital status, occupation and age at delivery and maternal behavior including smoking and/or drinking during pregnancy [28] and duration of breastfeeding [29].

The MCS gathered a wealth of information on a wide range of factors that have been associated with childhood health, including familial and parental living arrangements, use and type of early childcare facilities used by the family, and parental psychosocial well-being, including self-esteem, life satisfaction, feeling of control, quality of their relationship and social support. However, these (and others available) factors were not included in the present study in order to facilitate comparison of sample descriptives, and results of modelling with the Boyle et al (2012) study on which the current study is based.

Thus, the present study used the same dataset, established the same sample, and matched adjustment variables and outcomes, in order to extend the ‘health outcome’ findings reported by Boyle et al (2012) to an assessment of behavior problem.

In summary, this research had three aims: Firstly, to examine the association between a general assessment of health and a general assessment of behavioral problems at the start and at the end of middle childhood (age 5 and age 11). Secondly, to describe the association between categories of GA and a composite measure of chronic general health problems and/or
behavioral problems at age 5 and at age 11. The third aim was to identify within this sample what perinatal risk factors were associated with chronic ‘composite’ health problems at ages 5 and 11.

**Materials and Methods**

**Sample**

The MCS is a well-documented, on-going, nationally representative prospective cohort study begun in 2002, and based in the United Kingdom (UK). Full details of the study have been reported previously [30]. The population-based cohort study (n=18818) comprised all children born in the UK who were living in the jurisdiction at the time of the first survey, and the UK register of social payment of child benefit was used as eligibility criterion. The present study used the data on behavior and general health collected at age 5 (first year of primary school) and at age 11 (final year of primary school).

In order to extend the outcome to include behavioral problems, and thereby facilitate comparison with Boyle et al., previously published findings [4] their baseline measures, categorical estimators, adjustment variables and specific selection strategies were replicated.

As this research involved secondary analysis of the MCS, ethical approval was not required.

**Gestational age**

GA categories replicate those used in Boyle et al. [4] and are consistent with commonly used categories stratified on established guidelines set out by the European Foundation for the Care of Newborn Infants (EFCNI, [24]: very preterm (VPT, <31 weeks), moderate preterm (MPT, 32-33 weeks), late preterm (LPT, 34-36 weeks), early term (ET, 37-38 weeks) and FT (39-41 weeks).
Outcome measures

General Health Status

General health at each time point was assessed by using a parent’s response (usually the mother) to the question ‘How would you rate your child’s health?’ using a five-point Likert scale: ‘Excellent’, ‘Very good’, ‘Good’, ‘Fair’ or ‘Poor’ [5]. Responses to this question have been shown to be strongly correlated with alternative or ‘hard’ measures of children’s health including number of episodes of hospitalization in the previous 12 months, chronic childhood conditions (e.g. asthma, epilepsy, heart condition) and number of bed days due to illness [31], and has been used previously to ascertain health status in a non-clinical sample [4, 5].

Behavioral Health Status

Behavior problems at both 5 and 11 years were assessed using a parent’s rating of behavior (usually the mother) on the Strengths and Difficulties Questionnaire (SDQ, Goodman [32]). The SDQ asks the respondent (primary care giver) to rate the child over 20 items, divided equally between four behavior problem categories: Emotional problems (e.g. ‘Many fears, easily scared’); Conduct problems (e.g. ‘Has temper tantrums or hot tempers’); Hyperactivity-inattention (e.g. ‘Restless, overactive, cannot stay still for long’); and Peer problems (e.g. ‘Picked on or bullied by other children’) [32, 33]. Scored on a three point scale: ‘not true’ (scoring 2), ‘somewhat true’ (scoring 1) and ‘certainly true’ (scoring 0) (and with items reverse scored as appropriate), the Total Difficulties (TD) score ranges from 0 to 40. A TD score of 17 (90th centile) or more is considered ‘abnormal’ and of use to screen for child psychiatric disorders in a community sample [34]. Furthermore, it is regarded as helpful in identifying the course and prediction of mental health problems in children and adolescents in the general population [35].
The internal consistency of the TD score for both age 5 (α coefficient = .79; [36]) and age 11 (.88) were in line with Goodman (.82) [32].

**Composite health**

This binary variable created two groups at each time point (age 5 and 11 years): a) those who were rated as having ‘Poor’ or ‘Fair’ health and/or given a rating of 17 or higher on the SDQ Total Difficulties score; b) those who were not rated as having ‘Poor’ or ‘Fair’ health and/or given a rating of 17 or higher on the SDQ Total Difficulties score.

**Longitudinal health**

To capture the changes in the ‘composite health’ assessment from age 5 to age 11, a second binary variable of ‘longitudinal health’ status was created. This variable identified children who had poor composite health at both time points and those who did not.

**Statistical methods**

The sample design of the MCS was such that a representation of the total population was possible, while at the same time sufficient numbers of key subgroups (areas of high child poverty, ethnic minorities and residents of smaller countries) were included for analysis.

It included geographical clustering (at UK electoral ward level) and disproportionate stratification (using nine different strata - each country in the UK (England, Wales, Scotland and Northern Ireland) having two strata: advantaged and disadvantaged and one additional strata for ethnic minorities for England). Multiple births (twins and triplets) were included for comparison purposes in the descriptive overview of the MCS sample, but were excluded from the regression analyses as multiple births are more likely to be PT - or induced PT [37].
Because of the MCS sampling design, methods employed for simple random sampling or independence of observations could not be used as this would lead to an underestimation of standard errors and subsequent invalid significance tests. All analyses were therefore completed using Complex Sampling Plans (CSPLAN; SPSS Statistics 20) commands (for further details see [38]) resulting in all reported numbers, percentages and means being appropriately weighted for non-response and attrition for each particular survey dataset(s). Furthermore, it was necessary to analyze the subpopulation of ‘singleton births only’ using a ‘subgroup’ function within the CSPLAN that adjusts the non-response and attrition weights appropriately. It should be noted that failure to employ a ‘subgroup’ function when interrogating any subset of these MCS datasets leads to incorrect variance estimation and misleading significance testing because of the misapplication of the pre-defined weighting scheme [38].

Baseline measures of interest in predicting child health at ages 5 and 11 were those identified and reported by Boyle [4]. These comprised established risk factors which were dichotomized. Adjustment variables included child gender, child race/ethnicity (White British/Other), mother’s marital status (Single mother/Other), maternal use of alcohol during pregnancy (≥3 units per week or ≥3 units per occasion), maternal smoking during pregnancy, breastfeeding (3-level variable ‘Never tried’, ‘Breastfed for less than 4 months’ (reference level) and ‘Breastfed for at least 4 months’). Adjustments were also made for mother’s education (Degree/No degree) and work status (3-level variable ‘Professional/Managerial’ (reference level), Long-term unemployed (Y/N), all others). Finally the mother’s age at time of birth was also included.

Poor physical health was identified by primary care-giver rating of the child’s general health as being ‘Poor’ or ‘Fair’ on the General Health measure [31]. Behavior problems at age 5 and age 11 were identified by using a dichotomized score of ≥17 on the TD score of the SDQ.
[32]. These two outcomes were used to create one binary variable showing the child’s ‘composite health status’ at each time point (age 5, age 11 and at both age 5 and 11). This group (with difficulty in at least one area of health during at least one time point) was compared to the cohort of children whose parents had rated them with both a TD score of <17 and also a general health rating as ‘Excellent’, ‘Very good’ or ‘Good’ at all time points.

Logistic regression with appropriate CSPLANS and subgroup analyses for ‘singleton births only’ was used to investigate the associations of adverse composite health outcomes (physical and/or mental health) with GA category at both age 5 and age 11. Unadjusted (results not shown) and fully-inclusive models were run for each exposure-outcome group in order to adjust for potential confounders. The FT group was used as the reference level for each GA category. For the longitudinal analysis, multinomial logistic regression was applied to avoid assumptions of proportional odds in the outcome groups.

**Results and Discussion**

**The Sample**

Subject-level exclusions were matched to those made by Boyle *et al.* [4] and included: children with missing GA, those born later than 42 weeks gestation, those whose birth weight and GA were deemed improbable (i.e. outside twice the interquartile range from the median birth weight centile based on Bonellie’s [39] centile charts for GA, sex and parity) and respondents other than the natural birth mother.

Table 1 presents the characteristics of the mothers of children who were included in the analysis and replicate the findings from Boyle *et al.* This confirmed the same baseline measurements from which to continue the current longitudinal investigation. Mothers of children born <32 weeks (VPT) were, on average, younger than mothers of children born
closer to term. They were least likely to be educated to degree level, and most likely to have no formal qualification. They were found to be most likely to have smoked during pregnancy and least likely to have breastfed for longer than four months.

**Table 1** Characteristics of mothers of children included in the analysis. Values are n(%) unless otherwise stated.

<table>
<thead>
<tr>
<th>Mother’s Characteristics&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Gestation at birth (weeks)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>P values comparing GA groups&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>VPT (&lt;32) (n=206)</td>
<td>MPT (32-33) (n=184)</td>
<td>LPT (34-36) (n=1058)</td>
<td>ET (37-38) (n=3607)</td>
<td>FT (39-41) (n=12565)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at time of child’s birth (years) Mean: (95% CI)</td>
<td>28.4 (27.3, 29.5)</td>
<td>29.7 (28.5, 30.9)</td>
<td>28.9 (28.4, 29.5)</td>
<td>29.2 (28.9, 29.5)</td>
<td>28.7 (28.5, 29.0)</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>Single mother</td>
<td>41 (18.5)</td>
<td>26 (9.4)</td>
<td>208 (16.4)</td>
<td>603 (14.0)</td>
<td>2104 (14.1)</td>
<td>0.065</td>
<td></td>
</tr>
<tr>
<td>Moderate/heavy alcohol use during pregnancy&lt;sup&gt;c&lt;/sup&gt;</td>
<td>12 (7.0)</td>
<td>12 (6.8)</td>
<td>61 (6.9)</td>
<td>237 (6.7)</td>
<td>885 (7.3)</td>
<td>0.917</td>
<td></td>
</tr>
<tr>
<td>Smoked during pregnancy</td>
<td>59 (29.4)</td>
<td>46 (17.9)</td>
<td>286 (25.0)</td>
<td>824 (21.3)</td>
<td>2734 (20.2)</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>Socioeconomic status:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Managerial/Professional</td>
<td>50 (24.6)</td>
<td>45 (27.9)</td>
<td>275 (30.6)</td>
<td>922 (29.0)</td>
<td>3356 (30.5)</td>
<td>0.414</td>
<td></td>
</tr>
<tr>
<td>Long term unemployed</td>
<td>21 (7.5)</td>
<td>21 (8.8)</td>
<td>118 (8.6)</td>
<td>417 (8.6)</td>
<td>1422 (8.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degree</td>
<td>59 (29.8)</td>
<td>50 (31.3)</td>
<td>297 (30.0)</td>
<td>994 (30.2)</td>
<td>3752 (33.7)</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>No qualifications</td>
<td>47 (21.6)</td>
<td>28 (11.3)</td>
<td>184 (14.1)</td>
<td>626 (14.4)</td>
<td>1948 (12.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration of breastfeeding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>62 (28.2)</td>
<td>52 (27.9)</td>
<td>415 (35.1)</td>
<td>1327 (33.4)</td>
<td>4220 (29.5)</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>≥ 4 months</td>
<td>39 (19.2)</td>
<td>36 (23.0)</td>
<td>222 (23.6)</td>
<td>951 (28.9)</td>
<td>3883 (34.6)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CI Confidence Interval; ET Early term; FT Full term; LPT Late preterm; MPT Moderate preterm; VPT Very preterm.

<sup>a</sup>Numbers, means and percentages are weighted with appropriate sampling and non-response weights for the dataset(s) used.

<sup>b</sup>χ² tests for categorical variables and F tests for continuous variables; calculated for weighted, clustered data.

<sup>c</sup>≥ 3 units/week or ≥ 3 units/occasion.

Characteristics of the children involved in this study are reported in Table 2 and match those reported by Boyle [4]. The average birth weight of the children increased as GA increased. A greater proportion of those born at <32 weeks were from ethnic groups other than ‘White British’. Being of multiple birth and/or first born increased the likelihood of being born <34 weeks.

**Table 2** Characteristics of children from the Millennium Cohort Study included in this analysis. Values are n(%) unless otherwise stated.

<table>
<thead>
<tr>
<th>Child’s Characteristics&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Gestation at birth (weeks)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>P values comparing GA groups&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>VPT (&lt;32) (n=206)</td>
<td>MPT (32-33) (n=184)</td>
<td>LPT (34-36) (n=1058)</td>
<td>ET (37-38) (n=3607)</td>
<td>FT (39-41) (n=12565)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

11
At age 5 and at age 11, the proportion of children who were rated as being in ‘Poor’ or ‘Fair’ health remained constant across the two time points (4.0% of the population at 5 (n=601), 3.8% at 11 (n=403)). However, across these same two time points there was almost 100% increase in the proportion of children rated as having behavior problems (5.6% at 5 (n=745); 10.4% at 11 (n=996)).

Interestingly, of the singletons with both health and behavior ratings (n=13051), the association between ratings of general health and behavior problems was relatively low at each time point. At age 5, 745 (5.6%) of the children were rated as having behavior problems. However, only 14.0% of these (n=103) were amongst the 601 (4.0%) children who had been rated as having ‘Poor’ or ‘Fair’ health. Similarly at age 11, 994 (10.5%) of the children were rated as having behavior problems. However, still only 12.6% of these (n=125) were amongst the 403 (3.8%) children who had been rated as having ‘Poor’ or ‘Fair’ health at this time point.
Association between gestational age categories and the composite physical/behavioral health outcome

Three logistic regression analyses were conducted – two binary logistic models with the cross-sectional datasets (data at age 5 and data at age 11) and one multinomial logistic model with the longitudinal dataset (combining data from age 5 and age 11) – using, where necessary, the subgroup function available in SPSS CSPLANs. A small dose-response of significant GA group effects at age 5 was observed, with a larger dose-response effect at age 11 (see Table 3a). Table 3b reports the same dose-response trend - albeit stronger - for those with health problems at both age 5 and age 11.

Early predictors of the composite physical/behavioral health problem outcomes

Table 3a shows the associations between the composite (physical and/or behavioral) health problem outcomes and the early risk factors included in the logistic regression analysis for the cross-sectional data at age 5 and age 11 and for the longitudinal analysis (to identify chronic cases).

The significant risk factors for poor outcome at age 5 were: being other than ‘White British’; mother not in managerial/professional role; mother smoking during pregnancy; mother being long-term unemployed; mother being single, mother not being educated to degree level; child being breastfed for less than 4 months and child being male. The significant risk factors at age 11 were: mother smoking during pregnancy; mother not being educated to degree level; mother being single; mother not in managerial/professional role; child being breastfed for less than 4 months and child being male.
Results from the longitudinal analysis (chronic problems i.e. occurring at both age 5 and age 11) indicate that while fewer risk factors were significant, effects were stronger than at a single time point (i.e. at age 5 only or age 11 only). Significant factors were mother being single; mother smoking during pregnancy; mother not being educated to degree level; mother’s age at time of birth; mother being long-term unemployed; child not being breastfed for at least 4 months and child being born male.

Associations between individual outcomes (‘Poor’/’Fair’ health or behavior problem) and the early risk factors for each, at each age (5 and 11 years) are reported in Supplementary Table 1.

Table 3

<table>
<thead>
<tr>
<th>Table 3</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a) Cross-Sectional Data</td>
<td>b) Longitudinal Data</td>
</tr>
<tr>
<td></td>
<td>Age 5a</td>
<td>Age 11a</td>
</tr>
<tr>
<td>Gestational Categoriesb</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very preterm (&lt;32 weeks)</td>
<td>2.0 (1.2, 3.2)c</td>
<td>3.9 (2.2, 7.1)c</td>
</tr>
<tr>
<td>Moderate preterm (32-33 weeks)</td>
<td>2.7 (1.5, 4.8)</td>
<td>2.6 (1.4, 4.7)</td>
</tr>
<tr>
<td>Late preterm (34-36 weeks)</td>
<td>1.4 (1.0, 1.8)</td>
<td>1.7 (1.3, 2.3)</td>
</tr>
<tr>
<td>Early term (37-38 weeks)</td>
<td>1.3 (1.0, 1.5)</td>
<td>1.1 (0.9, 1.3)</td>
</tr>
<tr>
<td>Additional Perinatal Risk Factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child sex (Male)</td>
<td>1.5 (1.3, 1.7)c</td>
<td>1.7 (1.4, 1.9)c</td>
</tr>
<tr>
<td>Ethnic group: not White British</td>
<td>1.8 (1.4, 2.2)d</td>
<td>1.1 (0.9, 1.4)</td>
</tr>
<tr>
<td>Not firstborn</td>
<td>0.9 (0.8, 1.1)</td>
<td>0.9 (0.8, 1.1)</td>
</tr>
<tr>
<td>Mother’s age at time of birth</td>
<td>1.0 (1.0, 1.0)</td>
<td>0.98 (0.96, 0.99)</td>
</tr>
<tr>
<td>Single mother</td>
<td>1.5 (1.3, 1.8)c</td>
<td>1.5 (1.2, 1.8)c</td>
</tr>
<tr>
<td>Not manager/professional role</td>
<td>1.7 (1.4, 2.1)c</td>
<td>1.3 (1.0, 1.6)d</td>
</tr>
<tr>
<td>Being long term unemployed</td>
<td>1.6 (1.3, 1.9)c</td>
<td>1.2 (1.0, 1.6)</td>
</tr>
<tr>
<td>Not degree level education</td>
<td>1.5 (1.2, 1.8)c</td>
<td>1.6 (1.4, 2.0)c</td>
</tr>
<tr>
<td>Alcohol use during pregnancyc</td>
<td>1.0 (0.8, 1.3)</td>
<td>1.2 (0.9, 1.6)</td>
</tr>
<tr>
<td>Smoked during pregnancy</td>
<td>1.7 (1.5, 5.0)c</td>
<td>1.6 (1.3, 1.8)c</td>
</tr>
<tr>
<td>Breastfed &lt;4 months</td>
<td>1.4 (1.2, 1.8)c</td>
<td>1.3 (1.0, 1.5)d</td>
</tr>
<tr>
<td>Child age at assessment (age 5)</td>
<td>0.8 (0.6, 1.1)</td>
<td>n/a</td>
</tr>
<tr>
<td>Child age at assessment (age 11)</td>
<td>n/a</td>
<td>0.8 (0.6, 1.0)</td>
</tr>
</tbody>
</table>

OR Odds Ratio; CI Confidence Interval.

aResults are weighted with the sampling and non-response weights appropriate for the dataset(s) used.
bFull term used as reference level.
cSig @ < 0.001.
dSig @ ≤ 0.05.
≥ 3 units/week or ≥ 3 units/occasion.
Discussion

This study used a very large population-based cohort to examine the association between gestational age (GA) and childhood health problems at age 5 and 11 years as well as changes in this association from age 5 to age 11 years. To the best of our knowledge this is the first large scale longitudinal study to assess both general health problems and behaviour problems in their association with GA.

Results showed that the comorbidity between parental assessments of general health problems and behaviour problems is relatively small - both at age 5 and 11 years - suggesting that parents (and most likely the general public) do not consider serious behaviour problems as ‘health’ problems. This suggests that studies which aim to assess health problems according to the WHO comprehensive definition [2] are generally missing a crucial health component which is estimated to affect more children than physical health problems. This tendency might hamper early identification of, and therefore delay early interventions for, behavior problems that have been identified as risk factors for later adult mental health problems (see [11, 40, 41]).

The associations between GA and poor general health and/or clinically relevant behavior problem ratings showed a dose-response trend of GA at age 5 and at age 11. The strongest dose-response of GA was observed for children with chronic general and/or behavioral health problems (i.e., problems at both age 5 and age 11). These results indicate that the negative associations of low GA with comprehensive health problems during the elementary school years is more clearly estimated by using a composite physical and mental health index over an extended period of time. This is an important finding with reference to understanding the long term impact of low GA. Previous studies which focus on either physical or mental health problems have clearly underestimated the amplitude of the negative consequences of low GA.
Further analyses - with this longitudinal study and others - are needed to identify the developmental course of the different health components within our composite health assessments as well as to identify their putative interactions. For example, a sub-group may have physical health problems at age 5 which lead to an onset of behavioral problems during the elementary school years (or vice versa), while another sub-group may have one or other forms of health problems during the elementary school years only. It is important to identify these different developmental courses and their predictors in order to put in place effective preventive and therapeutic measures.

The study also identified perinatal factors that are associated with chronic general and/or behavioral health problems during childhood: child sex, maternal education, maternal employment, single motherhood, maternal smoking during pregnancy and duration of breastfeeding. Except for child sex and single motherhood, these risk factors are also significantly associated with low GA (Table 1). Thus, our results suggest that preventive interventions targeting pregnant women with these risk factors might reduce not only the risk of low GA [42], but also the long term negative outcomes associated with low GA [43-46].

A major advantage of using the MCS to examine the association between GA, chronic behavior problems and general health problems during middle childhood is that the sample size allows a comparison of relatively small GA groups. However, despite being one of the largest, on-going, population-based longitudinal studies of children’s development from early childhood onwards, the number of children in the lowest GA categories is still relatively small. Furthermore, as in many other studies of this type, the reliance on the maternal ratings of child health might also be considered a limitation as maternal ratings alone have been shown to be influenced by factors including maternal characteristics and socio economic status [13, 47, 48]. Finally, further studies using this cohort might consider expanding crucial developmental variables (e.g. breastfeeding duration and/or socioeconomic status of the
family) beyond the binary versions preferred by Boyle et al. and possibly including other

crucial developmental variables that are available including parental habitation status,

childhood caregiving arrangements etc. However, considerations on extending the number of
adjustments must be cognizant of the small sample size in some of the GA groups, and the

possibility of ‘over-adjusting’ for factors which may be on the causal chain from early GA
(and its own causes) to health outcomes.

In summary, this study indicates that GA - or degree of prematurity - has a strong ‘dose-
response’ association with chronic health problems over the period of middle childhood.

While studies at one single time point and/or employing a single dimension (general health or
behavioral health) reflect this trend, longitudinal analyses of a composite measure of health
provides a clearer picture of the strength of association of poor health outcomes with
prematurity. This dose-response association is compounded by child sex (males of low GA
are at more risk than females), maternal characteristics at birth (education, employment,
marital status) and duration of breast feeding. Integrated support to the at-risk families
initiated during the perinatal period and appropriate corrective interventions during early
childhood may reduce the adverse outcomes, and in turn substantially reduce long term
physical/mental health costs to both the individual and health services. Further analyses of the
joint developmental trajectories of the different components of the composite health
assessment with the MCS and similar studies might also provide insight into the
developmental interactions between physical and mental health problems, thus providing
useful information to plan preventive and corrective interventions from early childhood to
pre-adolescence.
Supplementary Information

**Supplementary Table 1** Binary Logistic Regression for Cross-Sectional Data. Odds ratios (OR) and Confidence Intervals (95% CI) for A) Physical Health Problems only and B) Behavior Health Problems only at Age 5 and at Age 11 Adjusted for All Covariates.

<table>
<thead>
<tr>
<th>Odds ratios of parent rating (OR (95% CI))</th>
<th>A) Child’s health as ‘Poor’ or ‘Fair’ compared with ‘Good’, ‘Very good’ or ‘Excellent’</th>
<th>B) Child’s behavior with Total Difficulty score ≥ 17 compared with a Total Difficulty score of &lt;17</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Age 5 only</td>
<td>Age 11 only</td>
</tr>
<tr>
<td>Very preterm (&lt;32 weeks)</td>
<td>2.04 (1.05-3.95)</td>
<td>6.11 (2.77-13.50)</td>
</tr>
<tr>
<td>Moderate preterm (32-33 weeks)</td>
<td>3.64 (1.86-7.14)</td>
<td>3.32 (1.66-6.64)</td>
</tr>
<tr>
<td>Late preterm (34-36 weeks)</td>
<td>1.55 (1.05-2.30)</td>
<td>1.84 (1.24-2.72)</td>
</tr>
<tr>
<td>Early term (37-38 weeks)</td>
<td>1.44 (1.12-1.84)</td>
<td>1.40 (1.04-1.87)</td>
</tr>
<tr>
<td>Additional perinatal risk factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>0.81 (0.66-0.98)</td>
<td>0.83 (0.65 – 1.06)</td>
</tr>
<tr>
<td>Ethnicity: not White British</td>
<td>1.83 (1.39-2.42)</td>
<td>1.79 (1.27 – 2.52)</td>
</tr>
<tr>
<td>Not firstborn</td>
<td>0.91 (0.73-1.13)</td>
<td>0.87 (0.66 – 1.15)</td>
</tr>
<tr>
<td>Mother’s age</td>
<td>1.0 (0.98-1.02)</td>
<td>0.97 (0.95-0.99)</td>
</tr>
<tr>
<td>Single mother</td>
<td>1.30 (1.00-1.70)</td>
<td>1.21 (0.88 – 1.66)</td>
</tr>
<tr>
<td>Not managerial/professional</td>
<td>1.37 (1.02-1.85)</td>
<td>1.09 (0.73 – 1.61)</td>
</tr>
<tr>
<td>Long term unemployed</td>
<td>1.20 (0.90-1.60)</td>
<td>1.12 (0.75 – 1.69)</td>
</tr>
<tr>
<td>Not educated to degree level</td>
<td>1.45 (1.08-1.94)</td>
<td>1.98 (1.34 – 2.94)</td>
</tr>
<tr>
<td>Used alcohol during pregnancy e</td>
<td>0.88 (0.61-1.27)</td>
<td>0.97 (0.60 – 1.57)</td>
</tr>
<tr>
<td>Smoked during pregnancy</td>
<td>1.71 (1.35-2.16)</td>
<td>1.51 (1.13 – 2.02)</td>
</tr>
<tr>
<td>Breastfed for &lt; 4 months</td>
<td>1.34 (1.04-1.74)</td>
<td>1.22 (0.91 – 1.64)</td>
</tr>
<tr>
<td>Child age at age 11 interview</td>
<td>n/a</td>
<td>1.09 (0.79 – 1.50)</td>
</tr>
<tr>
<td>Child age at age 5 interview</td>
<td>0.86 (0.58-1.29)</td>
<td>n/a</td>
</tr>
</tbody>
</table>

*a Results are weighted with the sampling and non-response weights appropriate for the dataset used.

*b Full term was used as the reference level.

*c p < 0.001

d p ≤ 0.05

*e ≥ 3 units/week or ≥ 3 units/occasion
References


