Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: https://orca.cardiff.ac.uk/id/eprint/159239/

This is the author’s version of a work that was submitted to / accepted for publication.

Citation for final published version:


Publishers page: http://dx.doi.org/10.1136/bmjopen-2022-065587
<http://dx.doi.org/10.1136/bmjopen-2022-065587>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher’s version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See http://orca.cf.ac.uk/policies.html for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.
BMJ Open

Early childhood parent-reported speech problems in small and large for gestational age term-born and preterm-born infants: a cohort study

Gabrielle Jee 1,1, Sarah Joanne Kotecha,2, Mallinath Chakraborty 1,2, Sailesh Kotecha 1,2,3, David Odd 1,2,3,4

ABSTRACT

Objective  (1) To assess if preterm and term small for gestational age (SGA) or large for gestational age (LGA) infants have more parent-reported speech problems in early childhood compared with infants with birth weights appropriate for gestational age (AGA). (2) To assess if preterm and term SGA and LGA infants have more parent-reported learning, behavioural, hearing, movement and hand problems in early childhood compared with AGA infants.

Design Cohort study.

Setting Wales, UK.

Participants 7004 children with neurodevelopmental outcomes from the Respiratory and Neurological Outcomes of Children Born Preterm Study which enrolled 7129 children, born from 23 weeks of gestation onwards, to mothers aged 18–50 years of age were included in the analysis.

Outcome measures Parent-reported single-answer questionnaires were completed in 2013 to assess early childhood neurodevelopmental outcomes. The primary outcome was parent-reported speech problems in early childhood adjusted for clinical and demographic confounders in SGA and LGA infants compared with AGA infants. Secondary outcomes measured were parent-reported early childhood learning, behavioural, hearing, movement and hand problems.

Results Median age at the time of study was 5 years, range 2–10 years. Although the adjusted OR was 1.19 (0.92 to 1.55) for SGA infants and OR 1.11 (0.88 to 1.41) for LGA infants, this failed to reach statistical significance that these subgroups were more likely to have parent-reported speech problems in early childhood compared with AGA infants. This study also found parent-reported evidence suggestive of potential learning difficulties in early childhood (OR 1.51 (1.13 to 2.02)) and behavioural problems (OR 1.35 (1.01 to 1.79)) in SGA infants.

Conclusion This study of 7004 infants in Wales suggests that infants born SGA or LGA likely do not have higher risks of parent-reported speech problems in early childhood compared with infants born AGA. To further ascertain this finding, studies with wider population coverage and longer-term follow-up would be needed.

INTRODUCTION

Background

Birth weight is a complex product of intrinsic and extrinsic influences on the feto-maternal interaction and is an important predictor of perinatal morbidity and mortality.1–3 Population studies have shown trends that infants are being born heavier. In England and Wales, there was an average increase of 40 g over 26 years for all live births recorded between 1986 and 2012 with a 8%–10% increased risk of being born with a high birth weight during this period.4 Trends for heavier birth weights were also observed in other countries including Canada, USA and Sweden, and the cause for this shift is not understood.5 This drives a need to better understand predictors of birth weight and its association with longer-term outcomes.

Outcomes for certain subgroups such as prematurity and small for gestational age (SGA) have been well explored but remains unclear for large for gestational age (LGA) infants particularly mid-term to long-term outcomes. LGA is associated with birth complications and indirect effects persisting into adulthood.6 Around birth, larger term infants have increased risk of shoulder dystocia, meconium aspiration, lower 5-min
Apical scores and death.5,7 Postnatally, they are more likely to have polycythemia, hypoglycemia and respiratory distress syndrome; all conditions associated with poor long-term outcomes.5-8 In preterm births, the prognosis for being LGA is unclear, with some literature suggesting an advantage; reporting lower perinatal mortality in preterm LGA infants, but with higher risks of early-onset sepsis and intraventricular hemorrhage.9

Overall, data on the mid-term to long-term effects of being LGA across all gestational ages are lacking, and little is known about neurodevelopmental outcomes. A retrospective cohort study by Moore et al found an increased risk of autism in term infants born SGA between 23 and 31 weeks, whereas being born large may have conferred some protective effect.10 In view of speech being a dynamic product of higher function cognitive and sensorimotor feedback processes involving multiple cortical and subcortical areas for planning, selecting, sequencing and motor programming, it was selected as the primary outcome of interest. Due to the complex interaction between neurodevelopmental domains including speech, this study also evaluated learning, behavioural, hearing, movement and hand problems.11

Objectives
1. To assess if preterm and term SGA or LGA infants have more parent-reported speech problems in early childhood compared with infants with birth weights appropriate for gestational age (AGA).
2. To assess if preterm and term SGA and LGA infants have more parent-reported learning, behaviour, hearing, movement and hand problems in early childhood compared with AGA infants.

METHODS
Study design
This study was conducted using data collected from the Respiratory and Neurological Outcomes of Children Born Preterm study (RANOPS), a cross-sectional population study conducted in Wales in 2013. This study recruited equal numbers of preterm (n=13361) and term-born children (n=13361) in years 2003, 2005, 2007, 2009 and 2011 to complete questionnaires on respiratory and neurodevelopmental outcomes if their child has ever had any active or resolved problems (online supplemental files 1 and 2).

Term infants were selected to be comparable with the preterm infants, for date of birth, sex and locality. A total of 7129 responses were received including consent for data usage and access to health databases. Characteristics between those who enrolled and those that did not are shown in online supplemental file 3. Parents with preterm infants were more likely to respond to the questionnaire than those with term births (p<0.001) and responders were less likely (39.5% vs 53.8%) to live in the most deprived half of Wales. However, the proportions of males (p=0.52) and those with low, normal or high birth weights for their gestation were similar (p=0.61).

Parent-reported answers on neurodevelopmental outcomes were collected in 2013 for all ages. For example, ‘Does your child have any problems with their speech?’; followed by a yes, or no, option. In the event neither option was selected, the response was recorded as unsure. Baseline demographics, including birth and maternal characteristics, were collected from national health databases.11-13

The primary neurodevelopmental outcome was parent-reported early childhood speech problems.11 Secondary neurodevelopmental outcomes were parent-reported early childhood learning, behavioural, hearing, movement and hand problems. For this study, all ‘unsure’ responses were re-coded as ‘no’ and included in the primary analysis.

Study population
The eligible population were all children enrolled in RANOPS, born from 23 weeks of gestation to mothers from 18 to 50 years of age with available speech outcomes (n=7004). Exposure measures were gestational age at birth and birthweight centile by category. Birthweight centiles were calculated for each sex and gestation (in weeks) using the LMS Growth programme (Medical Research Council).14 SGA was defined as <10th centile on the UK-WHO growth charts and >90th centile for LGA. These are values generally accepted across England and Wales, with Scotland using the 5th and 95th centiles.15,16 Gestational age was categorised (as per WHO definitions) as extremely preterm (<28 weeks), very preterm (28–31 weeks) and moderate to late preterm (32–37 weeks).17 Post-maturity was defined as gestational age at birth greater than or equal to 42 weeks.18 SGA and LGA infants were compared with AGA infants across gestations.

Covariates
Covariates included neonatal and maternal influences known a priori to influence birth weight in-utero. Neonatal factors are singleton or multiple births, gestational age at birth and sex. Maternal factors accounted for, including smoking during pregnancy, the Welsh Index of Multiple Deprivation (WIMD) score (a measure of relative deprivation for small areas with scores of 1–1999, 1 being the most deprived) and their age.19 These overlap with potential influences on the primary outcome, parent-reported speech problems in early childhood. While complex and multifactorial in potentially influencing the primary outcome, maternal socioeconomic status was accounted for using WIMD scores.20 Predictors of low birth weight include low birth weight,21 multiple pregnancies22 and foetal sex.23 Mode of delivery was not assumed to impact birth weight and considered likely multifactorial due to perinatal clinical practice surrounding estimated birth weight or centile, birth complications and maternal preference.24,25 The determinants of high birth weight are, however, less clear.
Evidence suggests that active smoking is associated with increased risk of low birth weight and preterm birth, and prevalence of smoking is associated with socioeconomic status. There is also an increased risk of low birth weight with younger mothers, perhaps influenced by socioeconomic status. However, evidence surrounding birth weight and advanced maternal age is inconsistent, with some studies showing increased prevalence of infants born LGA, while others showing a U-shaped trend of low birth weight with increasing age.

Statistical analysis
Duplicates and infants with missing exposure or primary outcome data were removed. Initially, neonatal and maternal characteristics at birth for all infants were compared across their birthweight centile category. Comparisons were performed using the Kruskal-Wallis equality of populations ranks test for birthweight centile category, gestational age and WIMD score at birth. Sex, number of births, mode of delivery and maternal smoking during pregnancy were compared using the χ² test and maternal age at delivery using the analysis of variance test.

Next, the proportion of children with speech problems between birthweight centile categories was compared. Using a logistic regression model, the unadjusted and adjusted for potential confounders, ORs for speech problems, comparing SGA and LGA infants to those born AGA were derived. Stratifying by preterm and term infants and using a logistic regression model, the ORs for speech problems were also derived for SGA and LGA infants compared with AGA infants. In view of a single primary outcome, p values corresponding to each statistical test were not corrected.

Six sensitivity analyses were performed. First, we repeated the logistic regression using a random-effects model to cluster by week of gestational age, and second by the age of the child at time of the survey. We then repeated using the 5th and 95th centile cut-offs, and then again removing all responders where ‘unsure’ was coded for the primary outcome. The main analysis was also repeated with an ordinal logistic regression analysis looking at the odds of an increasing number of reported neurodevelopmental disorders. Finally, the analysis was repeated using a missing data technique (Multiple Imputation with Chain Equations (details in online supplemental file 4)) to assess the impact of missing outcome and covariate data on the association seen. Likelihood ratio tests were used to compare models. Analysis was performed using Stata SE V.17 (Statacorp LLC).

Patient and public involvement
Patients or the public were not involved in the design, conduct, reporting, or dissemination plans of this research.

Table 1 Neonatal and maternal birth characteristics for all including preterm-born and term-born infants

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Birthweight centile</th>
<th>n</th>
<th>SGA (&lt;10th)</th>
<th>AGA (10th–90th)</th>
<th>LGA (&gt;90th)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at survey (median years)</td>
<td>7003</td>
<td>4.08 (2.67–7.67)</td>
<td>5.17 (2.75–7.67)</td>
<td>4.37 (2.75–7.67)</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Neonatal factors</td>
<td>Male</td>
<td>7004</td>
<td>313 (48.68%)</td>
<td>2458 (45.50%)</td>
<td>450 (46.92%)</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>Singleton</td>
<td>7004</td>
<td>507 (78.85%)</td>
<td>4647 (86.02%)</td>
<td>904 (94.26%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Gestation (median weeks)</td>
<td>7004</td>
<td>35 (33–38)</td>
<td>36 (34–39)</td>
<td>36 (34–38)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Preterm (&lt;37 weeks)</td>
<td>7004</td>
<td>458 (71.23%)</td>
<td>3105 (57.48%)</td>
<td>623 (64.96%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Post-term (≥42 weeks)</td>
<td>7004</td>
<td>25 (3.89%)</td>
<td>99 (1.83%)</td>
<td>5 (0.52%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Mode of delivery</td>
<td>4091</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Unassisted vaginal delivery</td>
<td>2037</td>
<td>119 (32.16%)</td>
<td>1659 (52.33%)</td>
<td>259 (47.01%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Instrument</td>
<td>406</td>
<td>27 (7.30%)</td>
<td>334 (10.54%)</td>
<td>45 (8.17%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Elective caesarean section</td>
<td>511</td>
<td>66 (17.84%)</td>
<td>349 (11.01%)</td>
<td>96 (17.42%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Emergency caesarean section</td>
<td>1137</td>
<td>158 (42.70%)</td>
<td>828 (26.12%)</td>
<td>151 (27.40%)</td>
<td></td>
</tr>
<tr>
<td>Maternal factors</td>
<td>Age (mean years)*</td>
<td>6225</td>
<td>29.85 (6.00)</td>
<td>30.35 (5.66)</td>
<td>31.15 (5.37)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>WIMD score (median)†</td>
<td>6804</td>
<td>845 (428–1351)</td>
<td>985 (503–1447)</td>
<td>985 (524–1444)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>WIMD decile†</td>
<td>6804</td>
<td>5 (3–7)</td>
<td>6 (3–8)</td>
<td>6 (3–8)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Maternal smoking‡</td>
<td>6725</td>
<td>133 (21.38%)</td>
<td>655 (12.65%)</td>
<td>75 (8.11%)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Values are n (%), mean (SD) or median (IQR) as appropriate. Denominator between measures differs due to missing data.
*Maternal age at the time of delivery.
†Welsh Index of Multiple Deprivation/Deciles of WIMD (lower values reflecting more deprivation).
‡Maternal smoking during pregnancy.
AGA, appropriate for gestational age; LGA, large for gestational age; SGA, small for gestational age; WIMD, Welsh Index of Multiple Deprivation.

AGA, appropriate for gestational age; LGA, large for gestational age; SGA, small for gestational age; WIMD, Welsh Index of Multiple Deprivation.

AGA, appropriate for gestational age; LGA, large for gestational age; SGA, small for gestational age; WIMD, Welsh Index of Multiple Deprivation.

Evidence suggests that active smoking is associated with increased risk of low birth weight and preterm birth, and prevalence of smoking is associated with socioeconomic status. There is also an increased risk of low birth weight with younger mothers, perhaps influenced by socioeconomic status. However, evidence surrounding birth weight and advanced maternal age is inconsistent, with some studies showing increased prevalence of infants born LGA, while others showing a U-shaped trend of low birth weight with increasing age.

Statistical analysis
Duplicates and infants with missing exposure or primary outcome data were removed. Initially, neonatal and maternal characteristics at birth for all infants were compared across their birthweight centile category. Comparisons were performed using the Kruskal-Wallis equality of populations ranks test for birthweight centile category, gestational age and WIMD score at birth. Sex, number of births, mode of delivery and maternal smoking during pregnancy were compared using the χ² test and maternal age at delivery using the analysis of variance test.

Next, the proportion of children with speech problems between birthweight centile categories was compared. Using a logistic regression model, the unadjusted and adjusted for potential confounders, ORs for speech problems, comparing SGA and LGA infants to those born AGA were derived. Stratifying by preterm and term infants and using a logistic regression model, the ORs for speech problems were also derived for SGA and LGA infants compared with AGA infants. In view of a single primary outcome, p values corresponding to each statistical test were not corrected.

Six sensitivity analyses were performed. First, we repeated the logistic regression using a random-effects model to cluster by week of gestational age, and second by the age of the child at time of the survey. We then repeated using the 5th and 95th centile cut-offs, and then again removing all responders where ‘unsure’ was coded for the primary outcome. The main analysis was also repeated with an ordinal logistic regression analysis looking at the odds of an increasing number of reported neurodevelopmental disorders. Finally, the analysis was repeated using a missing data technique (Multiple Imputation with Chain Equations (details in online supplemental file 4)) to assess the impact of missing outcome and covariate data on the association seen. Likelihood ratio tests were used to compare models. Analysis was performed using Stata SE V.17 (Statacorp LLC).

Patient and public involvement
Patients or the public were not involved in the design, conduct, reporting, or dissemination plans of this research.
RESULTS
Four thousand two hundred and eighty-four preterm and 2865 term infants were enrolled in RANOPS (n=7149). Twenty duplicates were removed, leaving 7129 children. Those born to mothers younger than 18 or above 50 years of age were excluded (n=83), leaving 7046 eligible children. Eight infants had missing birthweight centiles, and a further 34 had missing data on the primary outcome, early childhood speech problems, leaving a study population of 7004 participants for the primary analysis (99.40% of the eligible responders).

Stratifying by birthweight centile categories of SGA, AGA and LGA infants, baseline neonatal and maternal characteristics of the study cohort are shown in table 1. The median age at time of the survey was 4.08–5.17 (p=0.69). Besides the distribution of sex across birthweight centile categories, there were significant differences in other baseline neonatal and maternal characteristics analysed.

Distribution of early childhood speech problems in preterm and term infants by birthweight centile is as shown in figure 1. Analysis of the distribution of the primary outcome, parent-reported speech problems (in early childhood) showed that SGA and LGA infants were not more likely than their AGA counterparts to have problems, OR 1.19 (0.92 to 1.55) and OR 1.11 (0.88 to 1.41), respectively. Analysis of parent-reported secondary outcomes found evidence that SGA infants have an increased risk of learning difficulties in early childhood compared with those born AGA, OR 1.51 (1.13 to 2.02). There was also some evidence that SGA infants were more likely to have early childhood hearing problems, OR 1.35 (1.01 to 1.79). This study did not find that infants born SGA or LGA were more likely to have early childhood motor problems, OR 1.35 (1.01 to 1.79). There was also no evidence of interaction between the exposures measured (Pinteraction =0.999) (table 4).

Analysis was repeated using a random-effects regression model, stratified by preterm-born or term-born, and the adjusted ORs did not demonstrate that preterm infants born SGA or LGA were more likely to have parent-reported speech-problems than infants born AGA. There was also no evidence of interaction between the exposures measured (Pinteraction =0.999) (table 4).

Multilevel regression model of SGA and LGA babies compared with those born AGA across each gestational age category (using child age as the clustering variable) also showed compatible results (SGA, OR 1.19 (0.92 to 1.55); LGA, OR 1.11 (0.87 to 1.41)) with the main analysis. When ‘unsure’ responses excluded from analysis, logistic regression repeated showed no increased likelihood in primary neurodevelopmental outcome, parent-reported
early childhood speech problems in babies born SGA, OR 1.20 (0.92 to 1.56) or LGA, OR 1.13 (0.90 to 1.44) compared with those born AGA. Repeat of the analysis using the 5th and 95th centiles as cut-offs for SGA showed that babies born SGA were not more likely than babies born AGA to have parent-reported early childhood speech problems, OR 1.35 (0.98 to 1.86).16 There was also no evidence to suggest more parent-reported early childhood speech problems in babies born LGA, OR 1.11 (0.84 to 1.48). Repeat of the main analysis examining the odds of an increasing number of reported developmental disorders found results compatible with the primary analysis (SGA, OR 1.17 (0.96 to 1.43); LGA, OR 1.11 (0.93 to 1.32)). Finally, repeating the main analysis

![Proportion of disorders (speech, hearing, behavioural, learning, movement, and hand problems) stratified by gestation age at birth, birthweight centile category, and WIMD category (quintile)](image)

**Figure 2** Proportion of neurodevelopmental disorders stratified by gestation age at birth, birthweight centile category and WIMD category (quintile). AGA, appropriate for gestational age; LGA, large for gestational age; SGA, small for gestational age; WIMD, Welsh Index of Multiple Deprivation.

| Table 2 Early childhood neurodevelopmental outcomes across all birthweight categories |
|-----------------------------------|-----------------|-----------------|-----------------|-----------------|
| **Outcome measure**              | **n**           | **SGA (<10th), n (%)** | **AGA (10th–90th), n (%)** | **LGA (>90th), n (%)** | **P value**   |
| **Primary**                      |                 | **AGA**          | **LGA**          | **P value**     |
| Speech problems                  | 7004            | 93 (14.46)       | 603 (11.16)      | 112 (11.68)     | 0.05         |
| Learning difficulties            | 6980            | 83 (12.95)       | 404 (7.50)       | 70 (7.34)       | <0.001       |
| Behavioural problems             | 6963            | 86 (13.46)       | 459 (8.54)       | 77 (8.11)       | <0.001       |
| Hearing problems                 | 6985            | 49 (7.66)        | 371 (6.89)       | 73 (7.62)       | 0.59         |
| Movement problems                | 6990            | 45 (7.01)        | 240 (4.45)       | 49 (5.12)       | 0.01         |
| Hand problems                    | 7000            | 40 (6.22)        | 206 (3.82)       | 36 (3.76)       | 0.01         |

AGA, appropriate for gestational age; LGA, large for gestational age; SGA, small for gestational age.
While the absolute numbers with standardised developmental objective measures in preterm-born and term-born infants, and are increasingly used in developmental screening for at risk infants. Example of such is the validated PARCA-R (Parent Report of Children’s Abilities) for very-preterm infants at 2 years. Although some studies have reported variation between parental estimates and objective measures of speech development, parent-reported speech outcomes especially speech intelligibility still remains widely used in clinical practice. While the absolute numbers reported here are large, they do only represent a relatively low proportion of those invited to enrol in the study; although they had similar low or high birthweights compared with those who did not enrol. However, they also appeared to come from less deprived areas than the wider population, and interpretation of our findings should consider this.

The production of speech is complex, relying on not only intact cognitive, motor and sensory functions but also complicated by hearing loss, particularly the child’s age at time of hearing loss. Reassuringly, in this work, using a multiple imputation model (n=7004) (SGA, OR 1.21 (0.95 to 1.54); LGA, OR 1.06 (0.85 to 1.32)) was also compatible, with no clear associations seen.

**DISCUSSION**

The literature supports that parents’ perceptions on their child’s development have shown significant consistency with standardised developmental objective measures in preterm-born and term-born infants, and are increasingly used in developmental screening for at risk infants. Example of such is the validated PARCA-R (Parent Report of Children’s Abilities) for very-preterm infants at 2 years. Although some studies have reported variation between parental estimates and objective measures of

---

**Table 3** Unadjusted and adjusted ORs for early childhood neurodevelopmental outcomes in SGA and LGA infants compared with AGA controls

<table>
<thead>
<tr>
<th>Outcome measures</th>
<th>Unadjusted</th>
<th>Adjusted for neonatal features*</th>
<th>Adjusted for neonatal* and maternal features†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Birthweight centile</td>
<td>Birthweight centile</td>
<td>Birthweight centile</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>SGA (&lt;10th)</td>
<td>LGA (≥90th)</td>
</tr>
<tr>
<td>Primary</td>
<td>n</td>
<td>SGA (&lt;10th)</td>
<td>LGA (≥90th)</td>
</tr>
<tr>
<td>Speech problems</td>
<td>7004</td>
<td>1.34 (1.06 to 1.70)</td>
<td>1.05 (0.85 to 1.30)</td>
</tr>
<tr>
<td>Secondary</td>
<td>6980</td>
<td>1.83 (1.42 to 2.36)</td>
<td>0.98 (0.75 to 1.27)</td>
</tr>
<tr>
<td>Learning difficulties</td>
<td>6936</td>
<td>1.67 (1.30 to 2.13)</td>
<td>0.95 (0.74 to 1.22)</td>
</tr>
<tr>
<td>Behavioural problems</td>
<td>6985</td>
<td>1.12 (0.82 to 1.53)</td>
<td>1.12 (0.86 to 1.45)</td>
</tr>
<tr>
<td>Hearing problems</td>
<td>6990</td>
<td>1.62 (1.16 to 2.25)</td>
<td>1.16 (0.84 to 1.59)</td>
</tr>
<tr>
<td>Movement problems</td>
<td>7000</td>
<td>1.67 (1.18 to 2.37)</td>
<td>0.98 (0.69 to 1.41)</td>
</tr>
<tr>
<td>Hand problems</td>
<td>7000</td>
<td>1.27 (0.98 to 1.66)</td>
<td>1.04 (0.81 to 1.33)</td>
</tr>
</tbody>
</table>

*Adjusted for foetal sex, singleton or multiple birth.
†Adjusted for maternal age and WIMD score at time of birth and maternal smoking status during pregnancy.

---

**Table 4** Unadjusted and adjusted ORs for primary neurodevelopmental outcome of early childhood speech problems stratified by preterm or term in SGA and LGA infants compared with AGA controls

<table>
<thead>
<tr>
<th>Gestation</th>
<th>Birthweight category</th>
<th>Unadjusted (n=7004)</th>
<th>Adjusted for neonatal features* (n=7004)</th>
<th>Adjusted for neonatal* and maternal features† (n=5935)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR (95% CI)</td>
<td>P interaction</td>
<td>OR (95% CI)</td>
<td>P interaction</td>
</tr>
<tr>
<td>Preterm (&lt;37 weeks)</td>
<td>SGA (&lt;10th)</td>
<td>1.27 (0.98 to 1.66)</td>
<td>0.864</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LGA (≥90th)</td>
<td>1.04 (0.81 to 1.33)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Term (≥37 weeks)</td>
<td>SGA (&lt;10th)</td>
<td>1.16 (0.89 to 1.96)</td>
<td>0.92 (0.60 to 1.42)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LGA (≥90th)</td>
<td>1.21 (0.71 to 2.05)</td>
<td>1.41 (0.99 to 1.94)</td>
<td></td>
</tr>
</tbody>
</table>

*Adjusted for foetal sex, singleton or multiple birth.
†Adjusted for maternal age and WIMD score at time of birth and maternal smoking status during pregnancy.

---

**Footnotes**

1. AGA, appropriate for gestational age; LGA, large for gestational age; SGA, small for gestational age; WIMD, Welsh Index of Multiple Deprivation.
2. *Adjusted for foetal sex, singleton or multiple birth.
3. †Adjusted for maternal age and WIMD score at time of birth and maternal smoking status during pregnancy.

---

**References**

we were unable to identify clear associations between SGA, or LGA, infants and adverse parent-reported speech outcomes in early childhood; in either the unadjusted or adjusted models, as well as in the analysis of hearing problems. Also reassuringly, this study also did not find an increased risk of adverse parent-reported behavioural and hand problems in SGA and LGA infants, movement problems in SGA infants and learning difficulties in LGA infants. However, CIs are relatively wide and important increases in morbidity cannot be excluded and more work with precise estimate may be warranted.

Evidence from this study suggest that SGA infants may have higher risks of parent-reported learning difficulties and behavioural problems in early childhood but not hearing, movement or hand problems. A single-centre cohort study of term infants in Australia born to women presenting antenatally between 1981 and 1984 showed that children born SGA had significantly more learning difficulties when followed up at 14 years of age using a parent-reported survey and academic achievement test. This study reported a comparable 19.8% and 20.5% incidence for those who completed psychometric testing and behavioural questionnaires, respectively. This study also demonstrated long-term attention difficulties in extremely SGA (3rd centile and below) term-born female adolescents. Similar findings of poorer attention, executive function and memory were reported in a small cohort study of SGA preterm-born and term-born young adults who underwent neuropsychological assessment. Similarly, a cross-sectional study of 5181 children’s behaviour, between ages 4 and 15 years in England, using a validated parent-reported questionnaire and stratifying for sociodemographic factors, suggested an association between birth weight and behavioural problems in children.

We also saw that the proportion of all parent-reported early childhood disorders showed a strong relationship with gestational age at birth (seen in figure 2). This appears consistent with the wider literature, and this work was designed to adjusted, in part, for this by using birthweight centiles. Alternatively, this may, in part, reflect the complexities and co-dependency of many of these outcomes as gestational age lowers, and further work is currently underway to look at the phenotypes and interactions of challenges these infants demonstrate.

Analysis of baseline characteristics noted a significant difference in maternal age, and demographics at birth between birthweight centile categories. While this study was not designed to explore the relationship between maternal age and birthweight centile categories, previous studies have reported significant findings of higher maternal age in LGA infants. In addition, the association between maternal smoking, deprivation and birthweight centile categories was sobering, with 20% of SGA infants exposed to in-utero smoking. In this work, more than 1 in 5 mothers smoked during pregnancy, identifying an important public health focus on a modifiable risk factor for low birth weight to address, with even first trimester cessation of smoking having substantial benefits.

There was also a significant difference in the mode of delivery across birthweight centile categories likely due to clinical practice managing high-risk pregnancies. This may include singleton or multiple pregnancies, gestational age and estimated birth weight or centile at time of birth. Sex was not associated with SGA or LGA, compatible with findings of a large cohort study in Netherlands.

CONCLUSIONS

Findings from this work, on a subset of 7004 infants in Wales suggest that infants born SGA or LGA may not have higher risks of speech problems in early childhood when compared with AGA infants. While enrolment was achieved on only a subset of less deprived infants, important differences may still exist, and we found that some infants being born SGA may have increased parent-reported learning difficulties and behavioural problems compared with AGA infants. Further longer-term studies on infants born SGA and LGA would be of value to better understand the association of birth weight on neurodevelopment.

Acknowledgements We would like to thank all participants in this study.

Contributors GJ conceived and designed this work, wrote statistical analysis plan, cleaned and analysed data, drafted the first manuscript, revised draft manuscript and edited and updated the final paper. DO conceived and designed this work, wrote statistical analysis plan, analysed data and drafted the first manuscript, revised draft manuscript and edited and updated the final paper. DO is also responsible for the overall content as guarantor. SJK and SK conceived, designed, performed original RANOPS including data collection, revised draft manuscript and edited and updated the final paper, MC revised the draft manuscript and edited and updated the final paper.

Funding No additional funding was received for this analysis. The original study RANOPS was funded by Medical Research Council Experimental Medicine Challenge Grant (reg: MR/M022552/1), Mason Medical Research Foundation grant and Children and Young Peoples Research Network Wales.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval Ethical approval was sought at initiation of RANOPS and approved by South East Wales Research Ethics Committee (Research Ethics Committee 12/ WA/0155 Project 91349). Parents provided written consent to participate.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. All data must be held securely at Cardiff University as per ethical approval given for this research. Anonymous data will be available from the Department of Child Health at Cardiff University to bonafide researchers provided that ethical approval is obtained from a research ethics committee in the UK for any suggested studies. Requests for data access should be sent to SK (Kotechas@cardiff.ac.uk).

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
REFERENCES
20. Open access This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) licence, which permits others to copy, redistribute, remix, and transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: https://creativecommons.org/licenses/by/4.0/.
21. ORCID iDs
Gabrielle Jee http://orcid.org/0000-0001-6647-1110
Mallinath Chakraborty http://orcid.org/0000-0002-1721-6532
Sailesh Kotecha http://orcid.org/0000-0003-3553-7627
David Odd http://orcid.org/0000-0002-6416-4966


1. **Background:**

   Babies born late preterm, between 33 and 36 weeks gestation (6-7% of all UK births), have a greater risk of developing breathing problems than previously appreciated. Our research group has recently shown that late preterm born 8-9 year olds have worse lung function than similarly aged children born at term but their lung function improves by 14-17 years of age (http://www.bbc.co.uk/news/uk-wales-south-east-wales-15071922). In 2009 in England and Wales there were 706,248 live-births and over 42,000 were born late preterm. This group of children born late preterm tends to be treated the same as children born full term however there may be a need to monitor this group of children more closely.

**Definition of terms:**

Gestational groups - ≤32 weeks gestation = extremely preterm
- 33-36 weeks gestation = late preterm
- ≥37 weeks gestation = term

**1.1 Research Question**

   We hypothesise that children born extremely preterm (23-32 weeks gestation) and late preterm (33 to 36 weeks gestation):
   (a) have increased respiratory symptoms and disease,
   (b) have increased neurodevelopmental problems,
   (c) have increased health care utilisation, mainly due to respiratory morbidity in infancy, in the pre-school and early school years, when compared to age-matched term-born children (control group).
1.2 Methods

Study Design: The morbidity and health utilisation of late preterm infants

In Wales, the average number of live-born deliveries over the last 5 years is 34,464 per annum.\(^2\) Approximately 2,000 each year are born prematurely with 500 born extremely preterm and 1,500 late preterm.\(^2\)

<table>
<thead>
<tr>
<th>AWPS data: Total number of survivors at 1 year of age per year of birth and gestational age in Wales</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year of birth</td>
</tr>
<tr>
<td>---------------</td>
</tr>
<tr>
<td>Gestational age</td>
</tr>
<tr>
<td>23-32 weeks</td>
</tr>
<tr>
<td>33-36 weeks</td>
</tr>
<tr>
<td>37-43 weeks</td>
</tr>
<tr>
<td>Other/unknown</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

For late preterm infants, we will address:
(a) respiratory morbidity,
(b) neurodevelopmental problems and
(c) health care utilisation, especially due to respiratory reasons, via a questionnaire study and also the well-linked health databases in Wales with the collaboration of local neonatologists.

The data will be compared to term born and extremely preterm-born groups. Our team, including Professor John Henderson as well as Professor Frank Dunstan, has great expertise in questionnaire studies assessing health of children and adults.\(^4,5\) A modified ISAAC questionnaire focusing on respiratory symptoms (rather than allergy) will be used to assess respiratory health including wheezing, drug usage and physical activity as well as visits to hospitals and general practitioners in school-aged children\(^6\) and a validated questionnaire by Powell et al will be used for preschool children.\(^7,8\) The questionnaires will be used to assess respiratory health and for any neurodevelopmental difficulties in a cross-sectional survey of children aged 1, 2, 3, 5, 7 and 9 years. The questionnaires will allow us to assess the impact of late preterm birth on respiratory outcomes, developmental delay and health service utilisation. The questionnaires will be age-appropriate (one for <5 years of age and another for 5-9 years of age). The two age-appropriate questionnaires are based on previous extensively validated questionnaires for both respiratory symptoms and neurodevelopmental outcomes in both preschool and school-aged children.\(^6-11\)

We shall send 2,000 questionnaires with an invitation letter and parent information sheet for each age-group for the preterm groups (including 1,500 late preterm- and 500 extremely preterm-born children) together with a similar number for age-matched term controls. The matching of the term controls will be based on sex, place of birth and chronological age (DOB), although in the analysis of the data, the corrected postnatal age (this is the age a premature baby would be if born at term) of the preterm infants will be taken into account as a potential confounding factor. The total of 24,000 questionnaire packs will be sent via the NHS Wales Informatics Service (NWIS), which have expertise for identifying subjects from the NHS administrative register and then providing DST, the mailing house with a database to send out the questionnaires. Furthermore, children who have moved from the area or have died are identified and excluded, by using data from the All Wales Perinatal Survey (AWPS) which is directed by Professor Sailesh Kotec ha with Professor Frank Dunstan a member of the steering committee.\(^2\) AWPS maintains data on children who die before their first birthday, and this data will be used along with the Welsh Demographics Service, which also holds information on children who have passed away.
Prior to any reminder packs being sent out to the families, NWIS will be updated of all families who have returned a questionnaire pack. NWIS will then update the contact database and also repeat the safe guard of who has died, and then send the new database of contact information to DST, who will send out the reminder packs (1st after 1 month from the initial mailing date and 2nd after 3-4 months).

Secondly, using linked databases (including the Patient Episode Database for Wales, PEDW and the National Community Child Health Database, NCCHD and the Welsh Demographics Service, WDS) for hospital episodes, we will identify the general health service usage of these children. For those children with a returned, signed (consented) questionnaire we will be able to link the hospital data to their completed questionnaire to allow for more accurate analysis. The information gathered from PEDW, NCCHD and WDS will also contain data on all children less than 10 years of age in Wales (anonymised) and this will allow us to determine the representativeness of the sample returning questionnaires. The social status of a family will be determined by location using the Welsh Index for Multiple Deprivation Child Index Score 2011 (taken from WDS).

The linked cross sectional data at 5 age points will be summarised in terms of rates of wheezing, rates of hospital admissions, etc. The exploration of the data will be hypothesis led to investigate if being born late preterm has an effect on the health outcome of children when compared to term born controls. The health outcomes will include hospital admissions, visits to GP, wheezing episodes, use of prescribed medication, doctor diagnosed conditions, long term disability and activity levels. The formal analysis will use general linear models to compare rates between different gestational age groups adjusting for confounders such as social class, maternal smoking during pregnancy, birth weight, gender etc.

Sample size: For a question with binary response (e.g. symptom present or absent), assuming a 50-60% response rate to the questionnaire, we will have 95% power for identifying a difference between the preterm- and term-born children if the true symptom rate is 15% in preterm and 10% in term children. We have estimated the response rates based on previous similar studies in Wales.4,12

2. NWIS collaboration
We have already liaised with NWIS and shall ask their team to identify children born between 1/01/2003 to 31/12/2011. NWIS will identify all preterm infants (<37 weeks of gestation at birth) for each age group (1, 2, 3, 5, 7 & 9 years of age) using the national database (NCCHD and WDS). NWIS will then select the term controls by identifying gender matched children born on the same day and in the same area (hospital or midwifery led unit) as each preterm infant. So there will be approximately 2,000 term controls for each age group. For each age (1, 2, 3, 5, 7 and 9 years of age) there will be three groups; children born at full term (37-43 weeks of gestation), children born late preterm (33-36 weeks of gestation) and children born extremely preterm (≤32 weeks of gestation). Each child will be assigned a study number based on their DOB/ gestational age/ a unique reference number.

The information obtained from the health databases (WDS, NCCHD and PEDW) for each child will include the child’s name, current address, gestational age, DOB, birth weight, breast feeding at birth and at 8 weeks of age, Welsh index deprivation score, hospital admissions of the last year and discharge diagnosis. The postcode for the current address will be used to identify the local health board (so that the appropriate local PI is identified on the invitation letter being sent to the family). This information will need to be checked to remove any children that have died via AWPS and the WDS database. This will involve AWPS sending data to NWIS with the name, DOB and address of each child that has passed away over the last 10 years (1/1/03 to 31/12/12). None of this data will be made available to the research team. NWIS will supply DST, the mailing house with a study
database split into two sections (> 5 years of age and > 5 years of age). The database will contain the study ID number, name, address and local health board of each study participant. Data will be transferred between NWIS and DST via File Transfer Protocol Secure (FTPS).

DST will divide each data set into 7 sections according to the local health board in which the child resides. DST will then print the individual questionnaires (including study ID numbers), invitation letters including the family address and signed with the appropriate LHB with details of the local PI, and information sheets and place these in envelopes with windows that allow the letter to show the family address, and these packs will then be sent by royal mail to the families. Stamped addressed envelopes (to our research department) will be included within the study pack. DST also provide a further data health check, informing us of any duplicates and of families who may have moved, thus avoiding unnecessary inappropriate mailing. No personal data will be released by NWIS to the research team at this stage.

NWIS will send study ID numbers with the child’s gestational age to the research team, so that when questionnaires are returned the correct gestational age for the child can be confirmed. Families will be given 4-8 weeks to return the questionnaires. A list of families who have returned questionnaires (by study ID numbers) will be sent to NWIS and they will send further information on each child that has a signed (consented) questionnaires, to include birth weight, antenatal details, hospital of birth, hospital admissions and discharge diagnosis over the last 12 months.

NWIS, following a re-check of the health database for any deaths, will create a list of any families that have not returned a questionnaire. This up-dated study database will be given to DST and the first reminder pack will be sent to these families at 1 month after the initial pack was sent out. If the response rate remains low then this process will be repeated with a second reminder pack being sent after 3-4 months, when the research team will give NWIS a list of all families who have returned questionnaires and NWIS will collate an up-dated list, having removed any families where a child has died in the interim period since last checking. This up-dated study database will be given to DST and the second reminder pack will be sent to those families who have not yet responded.

Questionnaires will be returned to the Department of Child Health at Cardiff University and processed by automated scanners and specialised software (Remark Office OMR 8). The questionnaires will be coded by the software, which also identifies any problems such as missing information. Any problems detected by the software will be reviewed by the research team. This will form part of the data quality control checks. Anonymised data will be available for the whole cohort to allow us to determine the representativeness of the responders. The returned questionnaires will be processed by Dr Martin Edwards and all data stored securely in the research office at the Department of Child Health based at the University Hospital of Wales, Heath Park, Cardiff. By returning the completed questionnaires it will be assumed that the family is consenting to take part in the study. If the questionnaire form is signed the family will be contacted with regards to clarifying any issues with the returned questionnaire; contacted in the future to take part in further research in this area; and for access to patient identifiable information in the health databases to link hospital admission and GP records information with that gathered from the questionnaires. It is clearly stated in the information leaflet what will happen with the data collected.

3. Ethical approval
The project has ethical and R&D approval from within Wales. The global governance check reference is IRAS91349 and the Ethics reference is 12/WA/0155. The sponsor for the project is Cardiff University (SPON1038-11).
4. Timescale
We aim to mail the questionnaires between January and February 2013. We shall allow the families 6 – 8 weeks to respond initially and send a second mailing of questionnaires between April and May 2013 and if necessary a third mailing of questionnaires in between July and August 2013, dependent upon the response rate. Our expectation is to reach a 50 – 60% response rate as has been achieved in many recent questionnaire studies.4-12-14

5. Resources
The study is currently funded by Departmental funds and research grant funding is being sought from several charities that support medical research – decision expected Jan/ Feb 2013.

6. Documents
i. Letter of invitation to parents – 1 side of A4 paper
ii. Questionnaire for preschool children (<5 years old) – 4 sides of A4 paper
iii. Questionnaire for school aged children (5-9 years old) – 4 sides of A4 paper
iv. Information sheet for families – 2 sides of A4 paper
v. Flow diagram of research protocol

7. References
Study No: 12351
Health questionnaire for preschool children (<5 years old)

Today’s date ________________ 2012

Person completing the questionnaire: Mothet  ○ 
Father  ○ 
Other  ○ 

Date of birth of child: _____/_____/_____

1. Has your child ever had wheezing or whistling in their chest at any time in their life?
   Yes ○  No ○  Unsure ○

   IF YOU HAVE ANSWERED “NO” PLEASE GO TO QUESTION 4.

2. In the last three months, during the day time (i.e. when awake) does your child:
   Wheeze: ○ a few days ○ some days ○ most days ○ every day ○
   Cough: ○ a few days ○ some days ○ most days ○ every day ○
   Short of breath: ○ a few days ○ some days ○ most days ○
   Rattly chest: ○ a few days ○ some days ○ most days ○
   Snuffly: ○ a few days ○ some days ○ most days ○

3. In the last three months, during the night time (i.e. when asleep) does your child:
   Wheeze: ○ a few nights ○ some nights ○ most nights ○ every night ○
   Cough: ○ a few nights ○ some nights ○ most nights ○
   Short of breath: ○ a few nights ○ some nights ○ most nights ○
   Rattly chest: ○ a few nights ○ some nights ○ most nights ○
   Snore: ○ a few nights ○ some nights ○ most nights ○

4. How many colds has your child had in the last three months?
   None ○ 1 to 3 ○ more than 3 ○ always has a cold ○

5. When my child has had a cold in the last three months, s/he has a:
   Wheeze: ○ a few colds ○ some colds ○ most colds ○ every cold ○
   Cough: ○ a few colds ○ some colds ○ most colds ○
   Short of breath: ○ a few colds ○ some colds ○ most colds ○
   Rattly chest: ○ a few colds ○ some colds ○ most colds ○

Pre-school age (<5) Questionnaire Version 1.0        Date: 01/05/12
Study No: 12351

e. Snuffy: ☐ ☐ ☐ ☐ ☐ ☐

6. When my child does NOT have a cold, in the last three months, s/he has a:

   | Not at all | a few days | some days | most days | every day |
---|------------|------------|-----------|-----------|-----------|
   | ☐          | ☐          | ☐         | ☐         | ☐         |

   a. Wheeze: ☐ ☐ ☐ ☐ ☐
   b. Cough: ☐ ☐ ☐ ☐ ☐
   c. Short of breath: ☐ ☐ ☐ ☐ ☐
   d. Rattly chest: ☐ ☐ ☐ ☐ ☐
   e. Snuffy: ☐ ☐ ☐ ☐ ☐

7. When my child has been more active (e.g. crawling, walking or when excited), in the last three months, s/he has a:

   | Not at all | a few days | some days | most days | every day |
---|------------|------------|-----------|-----------|-----------|
   | ☐          | ☐          | ☐         | ☐         | ☐         |

   a. Wheeze: ☐ ☐ ☐ ☐ ☐
   b. Cough: ☐ ☐ ☐ ☐ ☐
   c. Short of breath: ☐ ☐ ☐ ☐ ☐
   d. Rattly chest: ☐ ☐ ☐ ☐ ☐

8. Has your child ever been diagnosed with asthma by a doctor?
   Yes ☐ No ☐

9. Is there a family history of
   Yes ☐ No ☐

   a. Asthma ☐ ☐
   b. Eczema ☐ ☐
   c. Hayfever ☐ ☐
   d. Allergies ☐ ☐

10. In the last 12 months, has your child ever used any regular asthma inhalers (pumps) or medicines?
    Yes ☐ No ☐

    If yes, please provide the name (or colour of the pump) with details of how often used:

11. In the last 12 months, has your child had any chest infections?
    Yes ☐ No ☐

12. In the last 12 months, how many chest infections has your child had?

   Pre-school age (<5) Questionnaire Version 1.0
   Date: 01/05/12
Study No: 12351

13. In the last 12 months, how many courses of antibiotics has your child had?
   None  ○  1 to 3  ○  4 to 10  ○  more than 10  ○

14. In the last 12 months, how many admissions (overnight or longer) has your child had to hospital for breathing problems?
   None  ○  1 to 3  ○  4 to 10  ○  more than 10  ○

14a. Does the child's mother smoke cigarettes? Yes  ○  No  ○
   If yes, how many per day?  1 to 10  ○  11 to 20  ○  more than 20  ○
   If the child's mother smokes did she smoke during the pregnancy? Yes  ○  No  ○

14b. Does the child's father smoke cigarettes? Yes  ○  No  ○
   If yes, how many per day?  1 to 10  ○  11 to 20  ○  more than 20  ○

15. Do any other household members smoke cigarettes? Yes  ○  No  ○
   If yes, how many per day for the whole household?  1 to 10  ○
   11 to 20  ○
   more than 20  ○

(please add up all the cigarettes which are smoked by everyone living in the same household including the mother).

16. a. Has your child ever been diagnosed with any breathing problems (e.g. asthma, CF, TB etc.)? Yes  ○  No  ○
   If yes, please provide some details:

16. b. Has your child ever been diagnosed with any other conditions (e.g. diabetes, epilepsy etc.)? Yes  ○  No  ○
   If yes, please provide some details:

17. Is your child on any medication (please list all medicines your child is being given)?
   Yes  ○  No  ○
   If yes, please provide some details:

18. Does your child take part in any physical activity every week such as dancing, swimming, cycling?
   Yes  ○  No  ○
   If yes, please provide some details of how often and for how long:

Pre-school age (<5) Questionnaire Version 1.0

Date: 01/05/12
Study No: 12351

19. Does your child have any problems with moving?
   Yes ☐ No ☐
   If yes, please provide some details:

20. Does your child have any problems with using their hands?
   Yes ☐ No ☐
   If yes, please provide some details:

21. Does your child have any problems with their speech?
   Yes ☐ No ☐
   If yes, please provide some details:

22. Does your child have any problems with their vision?
   Yes ☐ No ☐
   If yes, please provide some details:

23. Does your child have any problems with their hearing?
   Yes ☐ No ☐
   If yes, please provide some details:

24. Does your child have any problems with their behaviour?
   Yes ☐ No ☐
   If yes, please provide some details:

25. Does your child have any learning difficulties?
   Yes ☐ No ☐
   If yes, please provide some details:

Thank you for filling in the form. The following section asks how you are happy for us to use the data or to contact you:

(A) If we need to clarify some of your answers, would you be willing to be contacted?
   Yes ☐ No ☐
   Please initial the box here ☐
   Address, Telephone number &/or email:

Pre-school age (<5) Questionnaire Version 1.0  Date: 01/05/12
(B) Most admissions and GP visits in Wales are stored in computer databases called NCCHD or PEDW. As part of this study, we would also like to study how children in Wales have used their GPs or have had admission to hospitals. Would you be happy for us to use your son’s or daughter’s records on these databases?

Yes ☐ No ☐ Please initial the box here ☐

(C) We may plan similar studies in the future, would you be willing to be contacted in the future?

Yes ☐ No ☐ Please initial the box here ☐

Name of Child __________________________ Name of Parent/Guardian __________________________ Date __________ Signature __________________________

Thank you very much for taking the time to fill in the form and for contributing to our research. Could you please send the form to:

Dr Martin Edwards,
Room UGT156
Department of Child Health
School of Medicine
Cardiff University
University Hospital of Wales
Heath Park
Cardiff CF14 4XN

In the enclosed self-addressed envelope.
Study No: [Redacted]

Health questionnaire for school age children (>5 years old)

Today's date ____________2012

Person completing the questionnaire:

Mother ○
Father ○
Other ○

Date of birth of child: _____/_____/_____

1. Has your child ever had wheezing or whistling in their chest at any time in the past?
   Yes ○ No ○ Unsure ○
   IF YOU HAVE ANSWERED “NO” PLEASE GO TO QUESTION 10.

2. Has your child had wheezing or whistling in the chest in the last 12 months?
   Yes ○ No ○ Unsure ○
   IF YOU HAVE ANSWERED “NO” PLEASE GO TO QUESTION 10.

3. How many attacks of wheezing has your child had in the last 12 months?
   None ○ 1 to 3 ○ 4 to 12 ○ more than 12 ○

4. In the last 12 months, how often, on average, has your child’s sleep been disturbed due to wheezing?
   Never woken with wheezing ○
   Less than one night per week ○
   One or more nights per week ○

5. In the last 12 months, has wheezing ever been severe enough to limit your child’s speech to only one or two words at a time between breaths?
   Yes ○ No ○

6. Has a doctor ever told you that your child has asthma?
   Yes ○ No ○

7. In the last 12 months, has your child's chest sounded wheezy during or after exercise?
   Yes ○ No ○

8. In the last 12 months, has your child ever used any regular asthma inhalers (pumps) or medicines?
   Yes ○ No ○

   If yes, please provide the name (or colour of the pump) with details of how often used:

   ________________________________________________________________
   ________________________________________________________________
   ________________________________________________________________

School age (>5) Questionnaire Version 1.0

Date: 01/05/12
Study No: [redacted]

9. In the last 12 months, has your child had a dry cough at night, apart from a cough associated with a cold or chest infection?
   Yes ☐ No ☐

10. A. Does the child’s mother smoke cigarettes?
    Yes ☐ No ☐
    If yes, how many per day? 1 to 10 ☐ 11 to 20 ☐ more than 20 ☐

    If the child’s mother smokes did she smoke during the pregnancy? Yes ☐ No ☐

10.B. Does the child’s father smoke cigarettes?
    Yes ☐ No ☐
    If yes, how many per day? 1 to 10 ☐ 11 to 20 ☐ more than 20 ☐

11. Do any other household members smoke cigarettes? Yes ☐ No ☐
    If yes, how many per day for the whole household? 1 to 10 ☐
    11 to 20 ☐ more than 20 ☐

    (please add up all the cigarettes which are smoked by everyone living in the same household including the mother).

12. In the last 12 months, has your child had any chest infections?
    Yes ☐ No ☐

13. In the last 12 months, how many chest infections has your child had?
    None ☐ 1 to 3 ☐ 4 to 10 ☐ more than 10 ☐

14. In the last 12 months, how many courses of antibiotics has your child had?
    None ☐ 1 to 3 ☐ 4 to 10 ☐ more than 10 ☐

15. In the last 12 months, how many admissions (overnight or longer) has your child had to hospital for breathing problems?
    None ☐ 1 to 3 ☐ 4 to 10 ☐ more than 10 ☐

16. Is there a family history of
    Yes ☐ No ☐
    a. Asthma ☐
    b. Eczema ☐
    c. Hayfever ☐
    d. Allergies ☐

17. Has your child ever been diagnosed with any breathing problems (e.g. asthma, CF, TB etc.)?
    Yes ☐ No ☐
    If yes, please provide some details:

_________________________________________________________________________

School age (>5) Questionnaire Version 1.0

Date: 01/05/12
Study No: [Barcode]

18. Has your child ever been diagnosed with any other conditions (e.g. diabetes, epilepsy etc.)?
   Yes  ☐  No  ☐
   If yes, please provide some details:

19. Is your child on any medication?
   Yes  ☐  No  ☐
   If yes, please provide some details:

20. Does your child take part in any physical activity such as dancing, cycling or swimming?
    Yes  ☐  No  ☐
    If yes, please provide some details of how often and for how long:

21. Does your child have any learning problems?
    Yes  ☐  No  ☐
    If yes, please provide some details:

22. Does your child have any problems with their behaviour?
    Yes  ☐  No  ☐
    If yes, please provide some details:

23. Does your child have an educational statement?
    Yes  ☐  No  ☐
    If yes, please provide some details:

24. Does your child have any problems with moving?
    Yes  ☐  No  ☐
    If yes, please provide some details:
Study No: 1234567890

25. Does your child have any problems with writing or using their hands?
   Yes ☐   No ☐
   If yes, please provide some details:

26. Does your child have any problems with speech?
   Yes ☐   No ☐
   If yes, please provide some details:

27. Does your child have any problems with their vision?
   Yes ☐   No ☐
   If yes, please provide some details:

28. Does your child have any problems with their hearing?
   Yes ☐   No ☐
   If yes, please provide some details:

29. Does your child have any problems with feeding?
   Yes ☐   No ☐
   If yes, please provide some details:

Thank you for filling in the form. The following section asks how you are happy for us to use the data or to contact you:

(A) If we need to clarify some of your answers, would you be willing to be contacted?
   Yes ☐   No ☐
   Please initial the box here ☐
Address, Telephone number &/or email:

School age (>5) Questionnaire Version 1.0

Date: 01/05/12
Study No: [Barcode]

(B) Most admissions and GP visits in Wales are stored in computer databases called NCCHD or PEDW. As part of this study, we would also like to study how children in Wales have used their GPs or have had admission to hospitals. Would you be happy for us to use your son’s or daughter’s records on these databases?

Yes ☐ No ☐ Please initial the box here ☐

(C) We may plan similar studies in the future, would you be willing to be contacted in the future?

Yes ☐ No ☐ Please initial the box here ☐

Name of Child __________ Name of Parent/Guardian __________ Date __________ Signature __________

Thank you very much for taking the time to fill in the form and for contributing to our research. Could you please send the form to:

Dr Martin Edwards,
Room UGT156
Department of Child Health
School of Medicine
Cardiff University
University Hospital of Wales
Heath Park
Cardiff CF14 4XN

In the enclosed self-addressed envelope.
Supplement 3. Comparison of responders and non-responders in the original study, RANOPs.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Responders</th>
<th>Non-Responders</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=7,129</td>
<td>n=19,593</td>
<td></td>
</tr>
<tr>
<td>Preterm</td>
<td>4264 (59.9%)</td>
<td>9080 (46.4%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Birthweight centile</td>
<td></td>
<td></td>
<td>0.606</td>
</tr>
<tr>
<td>&lt;10&lt;sup&gt;th&lt;/sup&gt;</td>
<td>1879 (9.6%)</td>
<td>666 (9.4%)</td>
<td></td>
</tr>
<tr>
<td>10&lt;sup&gt;th&lt;/sup&gt;-90&lt;sup&gt;th&lt;/sup&gt;</td>
<td>5478 (76.9%)</td>
<td>15,091 (77.1%)</td>
<td></td>
</tr>
<tr>
<td>&gt;90&lt;sup&gt;th&lt;/sup&gt;</td>
<td>976 (13.7%)</td>
<td>2603 (13.3%)</td>
<td></td>
</tr>
<tr>
<td>Male Sex</td>
<td>3846 (54.0%)</td>
<td>10,659 (54.5%)</td>
<td>0.523</td>
</tr>
<tr>
<td>Living in most deprived 50%</td>
<td>2736 (39.5%)</td>
<td>10,286 (53.8%)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Numbers are number (%) or mean (SD) as appropriate.
Supplement 4. Multiple imputation model

Imputation was based on the following distributions. 50 imputed sets of data were used and combined using Rubin’s rules.¹

<table>
<thead>
<tr>
<th>Variable</th>
<th>Imputation command</th>
<th>% missing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech Problems</td>
<td>-</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Learning Difficulties</td>
<td>logit</td>
<td>24 (0.3%)</td>
</tr>
<tr>
<td>Behaviour Problems</td>
<td>logit</td>
<td>41 (0.6%)</td>
</tr>
<tr>
<td>Movement Problems</td>
<td>logit</td>
<td>13 (0.2%)</td>
</tr>
<tr>
<td>Hand Problems</td>
<td>logit</td>
<td>4 (0.1%)</td>
</tr>
<tr>
<td>Maternal Age (years)</td>
<td>ologit</td>
<td>779 (11.1%)</td>
</tr>
<tr>
<td>Sex</td>
<td>-</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Single/Multiple pregnancy</td>
<td>-</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Mode of Delivery</td>
<td>mlogit</td>
<td>2913 (41.6%)</td>
</tr>
<tr>
<td>WIMD measure</td>
<td>regress</td>
<td>200 (2.9%)</td>
</tr>
<tr>
<td>Smoking Status</td>
<td>Logit</td>
<td>279 (4.0%)</td>
</tr>
<tr>
<td>Gestational age at birth</td>
<td>-</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Age at survey (years)</td>
<td>ologit</td>
<td>1 (0.01%)</td>
</tr>
<tr>
<td>Birthweight centile</td>
<td>-</td>
<td>0 (0.0%)</td>
</tr>
</tbody>
</table>