

Published in final edited form as:

NPJ Prim Care Respir Med. ; 25: 14112. doi:10.1038/npjpcrm.2014.112.

Birth data accessibility via primary care health records to classify health status in a multi-ethnic population of children: an observational study

Rachel Bonner^{1,*}, Vassiliki Bountziouka^{1,*}, Janet Stocks¹, Seeromanie Harding², Angela Wade³, Chris Griffiths⁴, David Sears^{1,5}, Helen Fothergill^{1,6}, Hannah Slevin^{1,7}, and Sooky Lum¹

¹Respiratory, Critical Care & Anaesthesia Section (Portex Unit), UCL, Institute of Child Health, 30 Guildford Street, London WC1N 1EH, UK

²MRC/CSO, Social and Public Health Sciences Unit, University of Glasgow, 4 Lilybank Gardens, Glasgow G12 8RZ, UK

³Clinical Epidemiology, Nutrition and Biostatistics Section, UCL, Institute of Child Health, 30 Guildford Street, WC1N 1EH, London, UK

⁴Centre for Primary Care and Public Health, Blizard Institute, Queen Mary University of London, 4 Newark Street, London E1 2AT, UK

⁵Lung Function Unit, Royal Brompton Hospital, Sydney St, London SW3 6NP, UK

⁶Torbay Hospital, South Devon NHS Trust, Newton Rd, Torquay, Devon TQ2 7AA, UK

⁷Faculty of Medicine, University of Southampton, Southampton SO17 1BJ, UK

Abstract

Background/Objective—Access to reliable birth-data (birthweight(BW) and gestational-age(GA)) is essential for identifying individuals at subsequent health risk. This study explored the feasibility of retrospectively collecting birth-data for school-children from parental questionnaires(PQ) and General Practitioners(GPs) in primary care clinics, in inner city neighbourhoods with high density of ethnic minority and disadvantaged populations.

Methods—Attempts were made to obtain birth-data from parents and GPs for 2171 London primary school-children (34% White, 29% Black-African origin, 25% South-Asians, 12% Other) as part of a larger study of respiratory health.

Corresponding author: Rachel Bonner (r.bonner@ucl.ac.uk), Respiratory, Critical Care & Anaesthesia Section (Portex Unit), UCL, Institute of Child Health, 30 Guilford Street, London WC1N 1EH, Telephone: 020 7905 2382.

*Joint first authors

Contributions:

SL and JS designed the study; SH and AW contributed to the analytical framework of the study; SL, CG, DS, HF and HS contacted PCTs and GPs for approval to participate in study; RB, DS, HF and HS undertook data retrieval from GP surgeries; VB analysed the data and together with RB had joint responsibility for the first draft of the manuscript. All authors contributed to the subsequent drafts of the manuscript and approved the final draft. SL and JS are guarantors.

Competing Financial Interests: S Harding reports personal fees as an editor of the Journal of Ethnicity and Health, outside the submitted work; RB, VB, JS, AW, CG, DS, HF, HS and SL declare that they have no conflict of interest in relation to this article.

Supplementary information is available at the Primary Care Respiratory Medicine journal's website

Results—Information on BW and/or GA were obtained from parents for 2052 (95%) children. Almost all parents (2045) gave consent to access their children's health records held by GPs. Based on parental information, GPs of 1785 children were successfully contacted and GPs of 1202 children responded. Birth-data were retrieved for only 482 children (22% of 2052). Missing birth-data from GPs were associated with non-White ethnicity, non UK-born, English not the dominant language in the home, or socio-economic disadvantage. Paired data were available in 376 children for BW and 407 for GA. No significant difference in BW or GA was observed between PQ and GP data, with <5% difference between sources regardless of normal or low birth weight, or term or pre-term status.

Conclusion—Parental recall of birth-data for primary school children yields high quality and rapid return of data and should be considered as a viable alternative where there is limited access to birth records. It provides the potential to include children with increased risk of health problems within epidemiological studies.

Keywords

Child; questionnaires; general practitioners; preterm birth; intra-uterine growth restriction

Introduction

Despite increasing evidence that pre- and early post-natal insults to the developing lung may impact later respiratory health,¹⁻⁴ remarkably little emphasis has been placed on the need for rapid access to reliable birth data such as birthweight (BW) and gestational age (GA).⁵ Without access to such data, it is difficult to identify individuals at subsequent risk, design intervention studies to prevent or minimize the impact of such insults, or select a population free of such risks when designing epidemiological studies. Population-based studies of lung function in children often exclude those born preterm (i.e. GA<37weeks) or low birthweight (LBW, BW<2.5kg) because of the known long-term influence on lung growth.^{6, 7} The most accurate means of obtaining BW and GA data should be via the child's medical record.⁸ The administrative process to obtain these data from hospitals or primary care centres can, however, be lengthy and complicated. The collection of birth data for many community-based longitudinal epidemiologic studies in the UK is via parental recall.^{9, 10} Evidence regarding the precision and reliability of parental recall of BW and GA is discrepant. While some studies have shown maternal recall to be reliable¹⁰⁻¹³ others suggest a bias, with poorer recall from mothers with more than one child or who are not of White European origin.^{14, 15} This study provided a unique opportunity to examine the feasibility of collecting essential information relating to birth status from both parents and GPs using data collected from the SLIC study, which is the largest study of lung function undertaken in a multi-ethnic population of London primary school children to date.

The aims of this study were to (1) determine the feasibility of collecting BW and GA from parents and General Practitioners (GP) in primary care surgeries, where all children are registered for health care, (2) assess the agreement of BW and GA data between GPs and parental recall and (3) estimate the extent to which reliance on parental data may bias identification of full term (i.e. 37 weeks GA) and appropriately grown (i.e. 2.5kg BW)

children for epidemiological studies, based on data collected as part of the Size and Lung function In Children (SLIC) study.¹⁶

Methods

As part of the SLIC study¹⁶(www.ucl.ac.uk/slic), an epidemiological study designed to explore ethnic differences in lung function and body size in a multi-ethnic population of London children, anthropometry and spirometry were undertaken in primary school children between December 2010 and July 2013. Primary schools in the London area with a high ethnic mix of pupils were identified and ranked by education performance within boroughs. The sampling was undertaken from each stratum of rankings to ensure a wide range of social-economic circumstances. In this study, an all-inclusive strategy was adopted to ensure no child would feel excluded from a study that was being undertaken in the school. Thus all children with written parental consent (n=2291) were eligible to participate in the study. The study was approved by the London-Hampstead Research Ethics Committee (REC: 10/H0720/53). Parents were requested to complete a study questionnaire that was sent home with the children. Information requested included relevant health information such as birth data and respiratory and medical history, ethnicity and socio-economic circumstances. A member of the study team was available to assist in person or over the phone where assistance in completing the questionnaire was required. Birth weight was reported in kilograms and grams or pounds and ounces (the latter being subsequently converted into kilograms and grams).¹⁶ Children were classified into 4 main ethnic groups: White (European descent), Black African origin (Black African or Black Caribbean descent), South-Asian (Indian sub-continent) and Other/mixed ethnicities, based on child's ethnicity information from parental questionnaires (PQ).

Socio-economic circumstances were assessed at area level using the English Index of Multiple Deprivation (IMD)¹⁷⁻¹⁹ and at individual level using the Family Affluence Scale (FAS).^{19, 20} Each child was assigned an IMD using area postcodes for both their registered GP surgery and their home address to examine potential associations of any bias between GP or parental data according to locality. The FAS, commonly used for collecting socio-economic data from children, included information such as number of cars and computers owned by the family, whether the child shared a bedroom,^{19, 21} and the dominant language spoken within the household.

Where parental consent was obtained for access to the child's and maternal GP records, GP surgeries were requested either to extract the relevant birth data from medical records, or to permit a designated researcher to extract such data. Approval from the relevant Primary Care Trusts was obtained to access GP records, with supplementary funding for service support costs being provided by the Local Comprehensive Research Networks to enable remuneration to be offered to GP surgeries.

Statistical analysis

For the purposes of this study, GP data were used as baseline, with discrepancies between PQ and GP exceeding 0.10kg for BW or two weeks for GA being considered to be of potential clinical or physiological significance.^{10, 11} These thresholds were used to estimate

the degree of potential under- and over-estimation from the PQ, if the GP report was assumed to be correct. Children were also classified as being of LBW (<2.5kg) and/or preterm (<37 weeks GA) according to both GPs and parental report. The Mann-Whitney U-test and binary logistic regression models were used to assess whether the child's test age or socio-economic circumstances distribution varied between GPs who a) did or did not respond to requests for data and b) could or could not provide the relevant birth data. Agreement between BW and GA reported by parents and GPs was assessed using the Bland and Altman method.²² Agreement between LBW and preterm classification according to PQ and GP was measured using the Kappa statistic. Multinomial logistic regression models were used to evaluate the extent and nature of any apparent parental misreport of birth information. Significance level was set at 0.05 and SPSS and R program were used for analyses.^{23, 24} Data were stored in a dedicated research database (Re-Base™ software, Re-Base Ltd).

Results

Out of the 2291 children with parental consent, 2171 children (median age 8.1 [range 5.2-12.0] years; 47% boys) participated in the SLIC study.¹⁶ Of these, parental reports for BW and/or GA were available for 2052 (95%) children with 2045 (94%) parents giving consent to access GP records. Of those with parental consent, 260 (13%) contact details for GPs were missing, and therefore GPs for 1785 children were approached. Although some GP information regarding past medical history was available for 1202 children (67% of requested), birth data were only available for 482 children (27% of those with parental consent and GP details, and 22% of the total study sample). Paired data (parent and GP) were available from 376 records for BW and 407 for GA, representing only 18% and 20% of those with parental consent to access GP data. Data availability from both sources is summarised in Figure 1.

The sex and age distributions were similar for children with or without GP data (median age (95% CI): 8.1(8.0; 8.2) vs. 8.3 (8.1; 8.6) years respectively). GP data were more likely to be missing for children of ethnicities other than White, who were not born in UK, where English was not the dominant language at home or who lived in the most deprived areas or were in low FAS households (Table 1). The distribution of area level or individual level SEC characteristics in terms of data availability was similar for all indices (Supplementary Table S1).

Among the 40% of GPs who responded, details regarding BW or GA were more frequently missing for older children (8.2 (95% CI: 8.1; 8.4) vs. 7.8 (7.7; 8.0) years, $p<0.0001$) and those from low FAS households (mean (95% CI)% missing data: 78 (69; 87)% for low FAS vs 42 (36; 48)% for high FAS, $p<0.0001$; Table 2). The proportion of missing data was independently associated with age, country of birth, dominant language, IMD and FAS, with the adjusted odds for not obtaining data being higher for older children and those not born in UK, without English as the dominant home language or who were living in more deprived areas. Low FAS was related to increased odds of missing GP birth data after accounting for the variables mentioned above (Table 2).

There was no significant bias between PQ and GP reports of either birth weight (bias (95% CI): -0.04 (-0.07 ; -0.01) kg) or gestational age (0.17 (0.04 ; 0.30) weeks), but the relatively wide limits of agreement (95% LoA (95% CI): [-0.63 (-0.68 ; -0.58); 0.55 (0.50 ; 0.60) kg] for birth weight; [-2.4 (-2.6 ; -2.2); 2.8 (2.5 ; 3.0) weeks] for gestational age) indicate that individual differences may exist (Figure 2). Although there was a trend for parents to underestimate BW or GA for heavier and full-term children when compared with GP data (-0.4 r -0.2 , $p = 0.1$ for all cases), the agreement regarding BW or GA was consistent across ethnicities, indicating that no ethnic bias was observed when estimating either BW or GA from PQ as compared with GP data (Figure 3). Differences in BW or GA were also found to be constant across the age range, although somewhat larger for children born LBW as compared with those of normal BW. Nevertheless, no trend towards over- or under-reporting data from PQ was evident ($\rho < 0.2$, $p > 0.1$ for both BW and GA). When the analysis was repeated after including the five extreme data-points, results remained very similar albeit the limits of agreement were slightly wider.

Parental ‘under-estimation’ of BW by at least 0.1kg occurred in 19% (95% CI: 15; 23%) of children, while ‘overestimation’ occurred in 12% (9; 16%). By contrast, parents under- or over-reported GA by at least 2 weeks on 4% (3; 7%) and 3% (1; 4%) respectively. The odds of parents under-estimating BW were ~ 2.5 (95% CI: 1.1; 5.5) times higher in Black-African origin children when compared with White children. In addition, lower FAS and increasing IMD quintile were both associated with increased misclassification of birth weight status (Table 3). No significant associations were observed between socio-economic circumstances and the likelihood of parents mis-estimating GA (Supplementary Table S2).

Nine percent (95% CI: 8; 11%) of children were classified as LBW and 6% (5; 8%) as preterm by parents compared with 6% (4; 9%) and 9% (7; 12%) respectively when classified by GPs. Among children with paired data, there was good agreement with respect to whether or not the child was of LBW (95.5%) or born preterm (97%) (Supplementary Table S3). When repeating the analysis including the five extreme data-points, no difference was seen in the proportion of misclassification and the agreement remained in the same level as the initial analysis. Significant association was found between parental mis-estimation of birthweight and socio-economic circumstances as indicated by IMD and FAS, with those from most deprived areas or lower FAS having higher odds to mis-estimate the child’s birthweight. No significant associations were found with GA and both of these results are in line with original data in the main manuscript.

Discussion

Main findings

These results demonstrate that it is currently not feasible to obtain essential birth data from GP records. Parental recall is an appropriate alternative, especially for birthweight. Comparison of GP data with PQ showed reasonable agreement on average. Although birth data from reliable health records would be undeniably preferable, our findings suggest that parental reports have the potential of yielding high quality data and quicker access to the data, for both gestational age and birth weight.

Feasibility of data retrieval

In contrast to the relative ease with which parental data was collected, obtaining birth data from GPs was difficult and information was less likely to be available if the child was older, born outside the UK or where English was not the dominant language at home. This raises a number of issues. First, owing to the overall low response from GPs, currently, this does not appear to be a feasible method for obtaining data for epidemiological studies. Further, the response rate was especially low for children from more deprived areas, thereby risking collection bias toward those with higher SEC. Second, the lower GP response rate for children not born in the UK or without English as their dominant home language may impact on provision of health care for migrant children. Increasing awareness of the potential long-term influence of early life events, including preterm birth and intra-uterine growth restriction highlights the need for GPs to try to obtain this information as accurately as possible at time of registration. Reassuringly, the barriers to obtaining information from the GP did not arise from parents. Not only was it feasible to collect birth data from virtually all parents via the questionnaire, but the vast majority gave consent for access to GP records, regardless of ethnicity or socio-economic circumstance.

Interpretation of findings in relation to previously published work

Comparison of paired PQ and GP data, where available, showed good agreement on average across all ethnic groups and socio-economic circumstances. Previous studies found contrasting results. In the Millennium Cohort Study, although there was 94% agreement in the reporting of GA within one week between parent and medical record, disagreement was associated with low SEC.²⁵ Likewise, significant underestimation of GA and less accurate BW reporting was found from mothers of non-white children.¹¹ However, in both these studies and the current study, disagreement only resulted in minimal misclassification of birth status, suggesting that parental reporting of BW and GA is accurate enough to provide appropriate classification of birth status, especially within the context of large-scale epidemiological studies. It should be noted that while GP data were used as the baseline for the purpose of analysis, with parental ‘under’ or ‘over-estimation’ based on the assumption that GP data should be the most accurate, this assumption was not necessarily correct, as shown by several physiologically impossible birth data or “outliers” provided by GPs (see Figure 2).

Strengths and limitations

The major limitation of this study was the relatively low sample size in certain sub-categories, mainly due to the high proportion of GP-missing data for children from more deprived households, meaning that potentially important differences pertaining to deprivation could not be discounted. In addition, although the best approach for collecting birth data would be via hospital birth records this would only have been feasible for children born in London. Given the nature of the multi-ethnic SLIC study, many of the children had been born outside London or indeed the UK, thereby precluding this approach. The use of GP data was as a baseline reference for comparisons with PQ, rather than a gold standard method. Nevertheless, the SLIC study is the largest study of lung function undertaken in a multi-ethnic population of London primary school children to date, providing a unique

opportunity to examine the feasibility of collecting essential information relating to birth status from both parents and GPs.

Implications for future research, policy and practice

A major current focus of the National Health Service in the UK is the development of administrative and informatics networks to develop linked electronic health records (EHR) for both clinical (<http://www.england.nhs.uk/ourwork/tsd/sst/>) and public health research (<http://www.cprd.com/intro.asp>) purposes. The intention is that data held in electronic health records be available to appropriate individuals via remote access, thus facilitating record/data access. However, contrary to time projections at the inception of the SLIC study, the software enabling this functionality has yet to be rolled out to the majority of GP surgeries in London. It was therefore impossible to access health records without the cooperation of the GP surgeries. Unfortunately, among the GPs contacted, ~1/3 failed to respond at all to requests for children's birth data. This was despite providing written parental consent for access to specific medical records, monetary incentives for participation and the option for a researcher to extract the necessary information. The reason for such a high rate of non-response is unclear, but may reflect either lack of resources or failure to appreciate the relevance of requests for such data. Inner city surgeries in deprived areas may find it difficult to engage with activities perceived to be research-based and quality of data collected by GP practice nurses may be compromised both by competing clinical priorities or lack of research-active practitioners.²⁶ There is a shift to EHR in the UK but even when fully functional, the issue of missing data for those not born in the UK will still remain.

In conclusion, detailed BW and GA data were difficult to retrieve from GP records. The proportion of missing birth data from GP records emphasises the need for more accurate and systematic recording of these data. Furthermore, it is essential that electronic health records are established within healthcare systems and the information made readily available, through the primary care practitioners to the appropriate personnel. Parental report of birth data, at least for primary school children is an appropriate alternative to health records for use in obtaining high quality data for epidemiological studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank S Lee and P Cottam for their help with data management. We would particularly like to thank the Head teachers and staff of participating schools for facilitating the recruitment and school assessments and in particular the children and families who participated in this study.

Funding: This work was supported by the Wellcome Trust [WT094129MA] and Asthma UK [10/013]. The SLIC study team acknowledges the support of the National Institute for Health Research, through the Comprehensive Clinical Research Network.

Funding source: The Wellcome Trust; Asthma UK and the National Institute for Health Research, through the Comprehensive Clinical Research Network.

References

1. Krauss-Etschmann S, Bush A, Bellusci S, Brusselle GG, Dahlén SE, Dehmel S, et al. Of flies, mice and men: a systematic approach to understanding the early life origins of chronic lung disease. *Thorax*. 2013; 68(4):380–4. [PubMed: 22781122]
2. Narang I, Bush A. Early origins of chronic obstructive pulmonary disease. *Semin Fetal Neonatal Med*. 2012; 17(2):112–8. [PubMed: 22265926]
3. Pike KC, Hanson MA, Godfrey KM. Developmental mismatch: consequences for later cardiorespiratory health. *BJOG*. 2008; 115(2):149–57. [PubMed: 18081597]
4. Stocks J, Hislop A, Sonnappa S. Early lung development: lifelong effect on respiratory health and disease. *Lancet Respir Med*. 2013; 1(9):728–42. [PubMed: 24429276]
5. Bolton CE, Bush A, Hurst JR, Kotecha S, McGarvey L, Stocks J, et al. Are early life factors considered when managing respiratory disease? A British Thoracic Society survey of current practice. *Thorax*. 2012; 67(12):1110. [PubMed: 22993167]
6. Nguyen TT, Thia LP, Hoo AF, Bush A, Aurora P, Wade A, et al. Evolution of lung function during the first year of life in newborn screened cystic fibrosis infants. *Thorax*. 2013
7. Borrego LM, Stocks J, Almeida I, Stanojevic S, Antunes J, Leiria-Pinto P, et al. Bronchodilator responsiveness using spirometry in healthy and asthmatic preschool children. *Arch Dis Child*. 2013; 98(2):112–7. [PubMed: 23292523]
8. McNamara BJ, Gubhaju L, Chamberlain C, Stanley F, Eades SJ. Early life influences on cardio-metabolic disease risk in aboriginal populations--what is the evidence? A systematic review of longitudinal and case-control studies. *Int J Epidemiol*. 2012; 41(6):1661–82. [PubMed: 23211415]
9. Boyle EM, Poulsen G, Field DJ, Kurinczuk JJ, Wolke D, Alfirevic Z, et al. Effects of gestational age at birth on health outcomes at 3 and 5 years of age: population based cohort study. *BMJ*. 2012; 344:e896. [PubMed: 22381676]
10. Tate AR, Dezateux C, Cole TJ, Davidson L. Factors affecting a mother's recall of her baby's birth weight. *Int J Epidemiol*. 2005; 34(3):688–95. Epub 2005/03/02. [PubMed: 15737964]
11. Adegboye ARA, Heitmann BL. Accuracy and correlates of maternal recall of birthweight and gestational age. *BJOG*. 2008; 115(7):886–93. [PubMed: 18485168]
12. Catov JM, Newman AB, Kelsey SF, Roberts JM, Sutton-Tyrrell KC, Garcia M, et al. Accuracy and reliability of maternal recall of infant birth weight among older women. *Ann Epidemiol*. 2006; 16(6):429–31. Epub 2005/11/11. [PubMed: 16280248]
13. Lule SA, Webb EL, Ndibazza J, Nampijja M, Muhangi L, Akello F, et al. Maternal recall of birthweight and birth size in Entebbe, Uganda. *Trop Med Int Health*. 2012; 17(12):1465–9. Epub 2012/09/22. [PubMed: 22994260]
14. Jaspers M, de Meer G, Verhulst FC, Ormel J, Reijneveld SA. Limited validity of parental recall on pregnancy, birth, and early childhood at child age 10 years. *J Clin Epidemiol*. 2010; 63(2):185–91. Epub 2009/08/18. [PubMed: 19682856]
15. Mbuagbaw L, Gofin R. Can recall of birth size be used as a measure of birthweight in Cameroon? *Paediatr Perinat Epidemiol*. 2010; 24(4):383–9. Epub 2010/07/14. [PubMed: 20618728]
16. Lum S, Sonnappa S, Wade A, Harding S, Wells J, Treleavan P, et al. Exploring ethnic differences in lung function: the Size and Lung function In Children (SLIC) study protocol and feasibility. 2014 Available from: <http://discovery.ucl.ac.uk/1417500/>.
17. Government CaL. , editor. English Indices of Deprivation 2010. UK: 2011.
18. Taylor-Robinson DC, Smyth R, Diggle PJ, Whitehead M. A longitudinal study of the impact of social deprivation and disease severity on employment status in the UK cystic fibrosis population. *PLoS One*. 2013; 8(8):e73322. [PubMed: 24009747]
19. Currie CE, Elton RA, Todd J, Platt S. Indicators of socioeconomic status for adolescents: the WHO Health Behaviour in School-aged Children Survey. *Health Educ Res*. 1997; 12(3):385–97. [PubMed: 10174221]
20. Andersen A, Krolner R, Currie C, Dallago L, Due P, Richter M, et al. High agreement on family affluence between children's and parents' reports: international study of 11-year-old children. *J Epidemiol Community Health*. 2008; 62(12):1092–4. [PubMed: 18413436]

21. DfCaL G. English indices of deprivation 2010. 2011
22. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986; 1(8476):307–10. [PubMed: 2868172]
23. IBM Corp. IBM SPSS Statistics for Windows, Version 20.0. IBM Corp.; Armonk, NY: 2011.
24. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: 2013.
25. Poulsen G, Kurinczuk JJ, Wolke D, Boyle EM, Field D, Alfirevic Z, et al. Accurate reporting of expected delivery date by mothers 9 months after birth. *J Clin Epidemiol*. 2011; 64(12):1444–50. [PubMed: 21684117]
26. Harding S, Enayat ZE, Cassidy A, Read UM, Ferrell C, Molaodi OR. The effect of flexible settings on participation rates, data quality, and costs in a longitudinal study of ethnically diverse young adults: the London Determinants of young Adult Social wellbeing and Health (DASH) study. *The Lancet*. 2013; 382:S41.

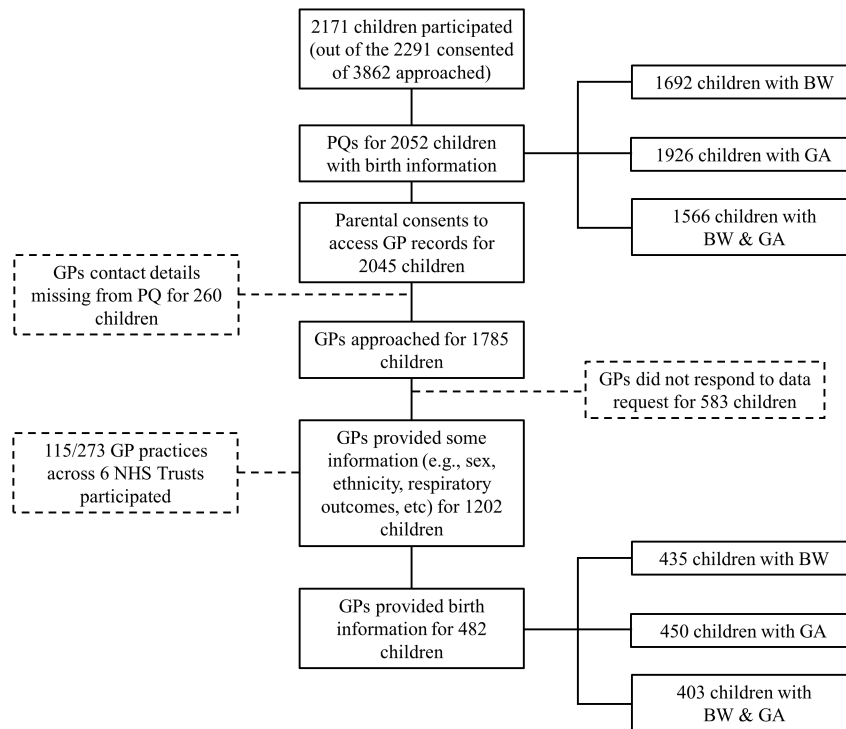


Figure 1. Study participation and birth data retrieval from parental recall and General Practitioner
Abbreviations: BW: Birthweight; GA: Gestational age; GP: General Practitioner; NHS: National Health Service; PQ: parental questionnaire.
376 children had paired information (i.e. PQ & GP) for BW; 407 children had paired information for GA; 322 children had paired information for BW & GA.

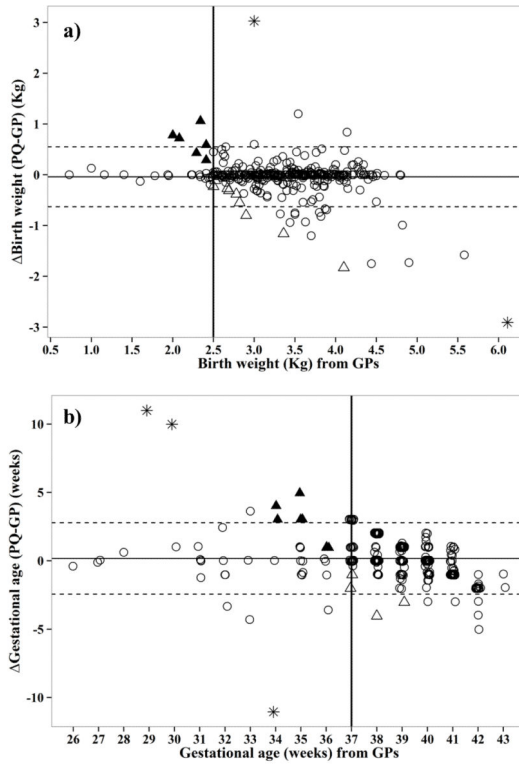


Figure 2.

Difference in a) birthweight and b) gestational age between PQ and GP data vs. GP data.

For clarity, GP data were used as baseline and are thus plotted on the x-axis, rather than the mean of GP and PQ data. Solid horizontal line represents the bias (i.e. mean difference) of the two methods while dotted lines represent the 95% limits of agreement (LoA) between the two methods. Bold solid vertical lines indicate critical cut-offs of <2.5 kg and <37 weeks which were used to categorise children born low birthweight or preterm respectively according to GPs.

▲ symbols indicate children who would have been ‘misclassified’ as having normal birthweight (n=6) or fullterm (n=8) if based on PQs rather than GP records.

△ symbols indicate children who would be potentially ‘misclassified’ to low birth weight (n=11) or preterm (n=4) if based on PQs rather than GP data. The outliers indicated by * which obviously indicate misreporting by either PQ or GPs have been excluded from the analyses.

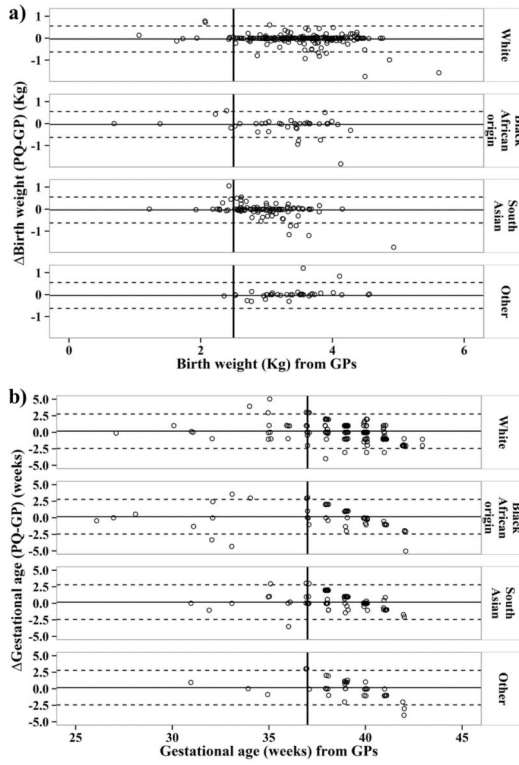


Figure 3.

Differences between PQ and GP data with respect to a) birth weight and b) gestational age according to ethnicity.

Solid horizontal lines represent the bias (i.e. mean difference) between the parental and GP data. Dashed lines represent the 95% limits of agreement between the two methods for the overall population. Bold solid vertical line indicate critical cut-offs of <2.5 kg for BW and <37 weeks for GA which were used to categorise children who were born low birth weight or preterm according to GPs. Points indicating extreme misclassification were excluded from this plot and analyses.

Table 1

Factors associated with GPs non-response upon request for information.

		Number of GPs contacted [⊥]	% non-response	Univariable models	
				OR (95% CI)	p [†]
Sex					
	<i>Girls</i>	954	34%	Baseline	
	<i>Boys</i>	831	32%	0.92 (0.75; 1.12)	0.40
Ethnicity					
	<i>White</i>	637	20%	Baseline	
	<i>Black-African origin</i>	458	38%	2.49 (1.89; 3.26)	<0.0001
	<i>South Asian</i>	474	43%	3.12 (2.39; 4.07)	<0.0001
	<i>Other</i>	216	36%	2.25 (1.60; 3.16)	<0.0001
Born in UK					
	<i>Yes</i>	1514	32%	Baseline	
	<i>No</i>	243	41%	1.52 (1.15; 2.01)	0.003
Dominant language in family					
	<i>English</i>	828	28%	Baseline	
	<i>Other</i>	360	37%	1.56 (1.20; 2.03)	0.001
Family's IMD domain [‡]					
	<i>1st quintile (least deprived)</i>	90	12%	Baseline	
	<i>2nd quintile</i>	258	32%	3.35 (1.69; 6.63)	<0.0001
	<i>3rd quintile</i>	253	34%	3.70 (1.87; 7.32)	<0.0001
	<i>4th quintile</i>	548	38%	4.33 (2.25; 8.32)	<0.0001
	<i>5th quintile (most deprived)</i>	632	31%	3.23 (1.68; 6.20)	<0.0001
FAS ^{‡, †}					
	<i>High (5-6)</i>	393	27%	Baseline	
	<i>Moderate (2-4)</i>	1135	33%	1.34 (1.04; 1.72)	0.03
	<i>Low (0-1)</i>	137	39%	1.74 (1.16; 2.62)	0.008

Abbreviations: GP=General Practitioner; IMD=Index of multiple deprivation; FAS=Family affluent score, OR (95% CI)=Odds ratio (95% Confidence interval).

[⊥]No feedback was received from GPs for 583/1785 (33%) children for whom there was parental consent to access records (see Figure 1).

[†]P-values derived through univariable logistic regression models to evaluate the factors related with the likelihood of non-response.

[‡]Detailed information regarding the IMD distribution of income and GPs domain and the individual components for FAS is presented in Supplementary Table S1.

[†]FAS was grouped in three categories due to the small sample size in the lower scores.

Table 2

Factors associated with missing birth data from GP records received.

	Number of GPs who responded [⊥]	% missing birth data	Univariable models		Multivariable model	
			OR (95% CI)	p [‡]	OR (95% CI)	p [‡]
Child's age (per year)			1.20 (1.12; 1.30)	<0.0001	1.14 (1.04; 1.26)	0.007
Sex						
<i>Girls</i>	634	61%	baseline		baseline	
<i>Boys</i>	568	58%	0.89 (0.71; 1.12)	0.33	1.12 (0.82; 1.53)	0.49
Ethnicity						
<i>White</i>	511	53%	baseline		baseline	
<i>Black-African origin</i>	284	75%	2.61 (1.90; 3.59)	<0.0001	1.55 (0.95; 2.53)	0.08
<i>South Asian</i>	268	55%	1.08 (0.80; 1.45)	0.63	0.50 (0.33; 0.77)	0.002
<i>Other</i>	139	65%	1.63 (1.10; 2.40)	0.01	1.60 (0.97; 2.64)	0.07
Born in UK						
<i>Yes</i>	1037	56%	baseline		baseline	
<i>No</i>	143	86%	4.88 (3.00; 7.96)	<0.0001	5.00 (2.59; 9.65)	<0.0001
Dominant language in family						
<i>English</i>	600	50%	baseline		baseline	
<i>Other</i>	226	72%	2.58 (1.86; 3.59)	<0.0001	1.71 (1.12; 2.62)	0.01
Family's IMD domain [‡]						
<i>1st quintile (least deprived)</i>	79	39%	baseline		baseline	
<i>2nd quintile</i>	176	39%	0.98 (0.57; 1.68)	0.93	1.10 (0.57; 2.10)	0.78
<i>3rd quintile</i>	167	41%	1.09 (0.63; 1.88)	0.76	0.86 (0.44; 1.67)	0.65
<i>4th quintile</i>	342	68%	3.31 (2.00; 5.49)	<0.0001	2.51 (1.31; 4.80)	0.006
<i>5th quintile (most deprived)</i>	436	73%	4.17 (2.54; 6.87)	<0.0001	2.17 (1.15; 4.12)	0.02
FAS ^{‡, †}						
<i>High (5-6)</i>	286	42%	baseline		baseline	
<i>Moderate (2-4)</i>	757	64%	2.43 (1.84; 3.20)	<0.0001	1.59 (1.10; 2.30)	0.01
<i>Low (0-1)</i>	83	78%	5.00 (2.82; 8.86)	<0.0001	2.86 (1.27; 6.43)	0.01

Abbreviations: GP=General Practitioner; IMD=Index of multiple deprivation; FAS=Family affluent score, OR (95% CI): Odds ratio (95% Confidence interval).

[⊥]GPs who responded could not provide any data on BW or GA for 720/1202 (60%) children (see Figure 1).

[‡]P-values derived through univariable or multivariable logistic regression models to evaluate the factors related with the likelihood of missing birth data. Multivariable model was adjusted for age, sex, ethnicity, country of birth, language, family's IMD domain and FAS.

[‡]Detailed information regarding the IMD distribution of income and GPs domain and the individual components for FAS is presented in Supplementary Table S1.

[†]FAS was grouped in three categories due to the small sample size in the lower scores.

Table 3

Factors associated with the likelihood of parental 'misclassification' of child's birth weight.

Univariable multinomial regression models [†]	Underestimation (n=70) [‡]		Overestimation (n=45) [‡]	
	OR (95% CI)	P	OR (95% CI)	P
Child's Age (per year)	1.14 (0.97; 1.34)	0.13	1.17 (0.96; 1.42)	0.13
Sex (baseline: girls)				
Boys	1.23 (0.73; 2.10)	0.44	1.09 (0.58; 2.05)	0.80
Ethnicity (baseline: white)				
Black-African origin	2.48 (1.12; 5.51)	0.03	1.55 (0.53; 4.51)	0.42
South Asian	1.57 (0.84; 2.92)	0.16	1.57 (0.75; 3.26)	0.23
Other	0.65 (0.21; 1.98)	0.45	0.97 (0.31; 3.04)	0.96
Born in UK (baseline: Yes)				
No	0.82 (0.17; 3.87)	0.80	1.29 (0.27; 6.19)	0.75
Dominant language in family (baseline: English)				
Other	1.42 (0.64; 3.13)	0.38	0.82 (0.27; 2.51)	0.73
Family's IMD domain (baseline: 1st & 2nd quintile (least deprived)) [§]				
3rd quintile	1.76 (0.80; 3.88)	0.16	0.94 (0.38; 2.33)	0.90
4th quintile	3.24 (1.51; 6.92)	0.002	1.76 (0.76; 4.08)	0.19
5th quintile (most deprived)	4.22 (2.00; 8.91)	<0.0001	1.80 (0.75; 4.28)	0.19
FAS (baseline: High 5-6)				
Moderate (2-4)	1.66 (0.92; 3.00)	0.10	1.17 (0.60; 2.3)	0.65
Low (0-1)	11 (1.9; 65)	0.007	13 (2.2; 77)	0.004

Abbreviations: OR: Odds ratio, GP: General Practitioner; IMD = Index of multiple deprivation; FAS = Family affluent score.

[§]Parental 'misclassification' was defined as a difference in child's BW of more than 0.10 kg compared with the GP records. For the purpose of this analysis, it was presumed that GP data would be the more accurate, but as mentioned in the discussion this assumption was not necessarily always correct

[†]Modelling was based on 376 cases which paired data were available.

[‡]The middle category (i.e. those neither under nor over estimated by PQ) was used as the baseline against which the other two were compared.

[§]The 1st and 2nd quintile of IMD were grouped together due to the small sample size in the 1st quintile.

^{||}FAS was grouped in three categories due to the small sample size in the lower scores.