

ASSOCIATION OF PREMATURITY, LUNG DISEASE AND BODY SIZE WITH LUNG VOLUME AND VENTILATION INHOMOGENEITY IN UNSEDATED NEONATES: A MULTICENTRE STUDY

^{1,2}Georg Hülskamp, ¹Sooky Lum, ¹Janet Stocks, ³Angie Wade; ¹Ah-fong Hoo, ⁴Kate Costeloe, ⁵Jane Hawdon, ⁶Kandadai Deeptha and ⁶J. Jane Pillow

¹Portex Anaesthesia, Intensive Therapy and Respiratory Medicine Unit, UCL, Institute of Child Health and Great Ormond Street Hospital for Children NHS Trust, London, UK;

²Department of Paediatrics, Westfälische Wilhelms-Universität, Münster, Germany;

³Paediatric Epidemiology and Biostatistics Unit, UCL, Institute of Child Health, London, UK;

⁴Academic Unit of Paediatrics, Barts and the London School of Medicine and Dentistry, Neonatal Unit, Homerton University Hospital, London, UK;

⁵Neonatal Unit, University College London Hospital, London, UK;

⁶School of Women's and Infants' Health, University of Western Australia, Perth, Australia;

Corresponding author:

Georg Hülskamp MD

Department of Paediatrics, Universitätsklinikum, Westfälische Wilhelms-Universität and Department of Paediatrics, Clemenshospital Münster, Duesbergweg 124, D-48153 Münster, Germany

Tel: +49 (0) 251 9762601, Fax: +49 (0) 251 9762612,

Email: g.huelskamp@clemenshospital.de

Running Head: Newborn Lung Volume and Ventilation Inhomogeneity

Word Count: 3107

ABSTRACT:

Background: Previous studies have suggested that preterm birth with or without subsequent chronic lung disease is associated with reduced functional residual capacity (FRC) and increased ventilation inhomogeneity in the neonatal period. We aimed to establish whether such findings are associated with the degree of prematurity, neonatal respiratory illness, and disproportionate somatic growth.

Methods: Multiple breath washout measurements using an ultrasonic flowmeter were obtained from 219 infants on 306 test occasions, during the first few months of life, at three neonatal units in the United Kingdom and Australia. Tests were performed during unsedated sleep in clinically stable infants (assigned to four exclusive diagnostic categories: term controls, preterm controls, respiratory distress syndrome and chronic lung disease). The determinants of neonatal lung function were assessed using multivariable, multilevel modelling.

Results: After adjustment for age and body proportions, the factors gestation, intrauterine growth restriction and days of supplemental oxygen were all significantly associated with a reduced FRC. In contrast, increased ventilation inhomogeneity (elevated Lung Clearance Index) was only significantly associated with duration of supplemental oxygen. After adjusting for continuous variables, diagnostic category made no further contribution to the models. Despite using identical techniques, unexpected inter-centre differences occurred, associated with the equipment used; these did not alter the negative association of preterm delivery and disease severity with lung function outcomes.

Conclusion: Reduction in FRC is independently associated with prematurity, intrauterine growth restriction and severity of neonatal lung disease. Determinants of lung function shortly after birth are highly complex in different disease groups. (247 words)

Keywords: functional residual capacity; infant; Lung Clearance Index; bronchopulmonary dysplasia; chronic lung disease

INTRODUCTION

Abnormal lung development in preterm infants due to intrauterine and early postnatal factors[1] is characterised by impaired alveolarisation and dysmorphic vasculogenesis.[2] These features are the pathological hallmarks of chronic lung disease of infancy (CLD – frequently referred to as Bronchopulmonary Dysplasia). Despite increasing survival of extremely preterm infants, the incidence of CLD remains high.[3] While infants are generally categorised as having CLD or not based on the need for mechanical ventilation and/or supplemental oxygen (O₂) at 36 weeks (w) postmenstrual age (PMA), the reality is that a continuum of disease severity is observed.[4-7] Accurate and non-invasive bedside techniques that evaluate the functional consequences of these structural changes are urgently required.[3,8]

Measurements of lung volume are relevant for assessing lung growth and development and for interpreting volume-dependent lung function parameters. Functional Residual Capacity (FRC), the resting lung volume at end expiration, is the only ‘static’ lung volume that can be readily assessed in non-cooperative infants and very young children. As summarised recently,[9] FRC is initially low in CLD but subsequently normalises or becomes elevated in later infancy, secondary to airway obstruction. The association between these findings and the degree of prematurity, the disproportionate somatic growth pattern that may accompany severe illness or preterm delivery and the severity of neonatal lung disease has not been investigated thoroughly. Furthermore, the lack of commercially available equipment for assessing lung function in small, unselected infants has limited widespread use of these techniques.

Infant whole body plethysmography[10] is unsuitable for assessing FRC at the bedside or in unstable or ventilated infants. An alternative method, the multiple-breath inert gas washout (MBW) technique, has recently gained popularity due to the potential for simultaneous evaluation of FRC and ventilation inhomogeneity.[8,11,12] A prototype, commercially available ultrasonic flowmeter based on that described by Buess et al.[13] has been used to measure FRC and indices of ventilation inhomogeneity in experimental animal models[14] and in non-ventilated[15] and ventilated[14] infants and children. Accuracy of FRC measured using the ultrasonic flowmeter (FRC_{US}) has been confirmed *in vitro*[15,16] and *in vivo*.[17] The use of this device to evaluate the impact of antecedent factors on lung development in preterm infants has not been reported.

The primary aim of this study was to assess the association of gestation, intrauterine growth and respiratory illness with lung volume and ventilation inhomogeneity during the neonatal period. The hypothesis to be tested was that: 1) preterm delivery (*per se*); 2) the degree of prematurity and intrauterine growth pattern; and 3) the severity of any neonatal respiratory disease are independently associated with lower FRC and reduced efficiency of ventilation during the first months of life after resolution of any acute respiratory illness.

METHODS

Study design

This study was designed as an evaluation of lung function in unsedated preterm and term infants during the first months of life. Studies were approved by the local research ethics committees and written, informed parental consent was obtained prior to lung function testing.

Study Subjects

Infants were recruited from the neonatal units of Homerton University Hospital (HUH) and University College London Hospital (UCLH), in London, United Kingdom, and King Edward Memorial Hospital (KEMH), Perth, Australia, using established recruitment protocols.[10,18] Infants were classified according to gestation and clinical status into four exclusive diagnostic categories:

Fullterm Controls (FTC): ≥ 37 weeks (w) gestational age (GA) with no signs or history of respiratory illness prior to testing.

Preterm Controls (PTC): < 37 w GA with minimal requirements for ventilatory support (i.e. no invasive ventilation, maximum 48 hours (h) support with nasal continuous positive airway pressure (CPAP), no supplemental inspiratory Oxygen (O_2) after 48h postnatal age and no signs or history of respiratory illness prior to testing).

Respiratory Distress Syndrome (RDS) < 37 w GA with a history of resolved respiratory distress syndrome and no requirement for supplemental FiO_2 at 36w postmenstrual age (PMA).

Chronic Lung Disease (CLD): requirement for supplemental FiO_2 or assisted ventilation at 36w PMA.[19,20]

Clinical and demographic data were recorded from hospital notes and parental questionnaires. Weight and length were assessed using standardised methods;[21] results were converted into Z-scores based on postnatal age adjusted for prematurity.[22] Body mass index (BMI: $\text{weight} \cdot \text{length}^{-2}$) was calculated as an index of disproportionate growth.

Lung Function Tests

Measurements were performed on clinically stable, unsedated infants in supine position during quiet sleep at least 30 minutes post-feed. Preterm infants were studied at least once prior to initial hospital discharge, generally between 34-38w PMA. When feasible, these infants returned for follow-up studies at approximately 44w PMA in order to match test age more closely with term controls who were only studied once, usually within 4 w to 8 w of birth (i.e. 44-48w PMA).

Tidal volume (V_T), Respiratory rate (RR), FRC and indices of ventilation inhomogeneity (Lung Clearance Index (LCI) and first and second Moment Ratios (M_1/M_0 , M_2/M_0), which are different methods of expressing ventilation efficiency)[12] were measured using the MBW technique with 4% SF_6 as a tracer gas using a commercially available prototype ultrasonic flow-meter (Ecomedics AG, Duernten, Switzerland).[15,17] Attempts were made to collect three acceptable MBW manoeuvres in each infant within each test occasion (See OLS).

Statistical Models

Standard software packages were used for data inspection, distribution and descriptive statistics (SPSS for Windows, v15.0, SPSS Inc.). In infants in whom more than one test had been performed, presentation of results in tabular form was limited to data from the last test occasion (i.e one result

per subject). All technically valid data from all test occasions were used in multilevel modelling, to take account of the correlation between repeat measurements from the same child. Univariable regression analysis established the association between each outcome measure and likely explanatory variables. Where potentially significant associations ($p < 0.05$) existed, these were further explored using multivariable, multilevel linear regression modelling (MLwiN, version 2.12; Institute of Education, UK), using a two-level model (centre; subject) and stepwise approach to multivariable model development that aimed to understand the mechanistic basis of differences observed (see OLS).

RESULTS

Population Characteristics

Technically acceptable data were obtained from 219 infants on 306 test occasions. Forty-eight infants were tested on 2 occasions, 15 on 3 occasions and 3 infants on 4 occasions. Overall failure rate was 10% (see OLS). As expected, there were significant differences in background characteristics and ventilatory support according to diagnostic category (Table 1, and Figure OLS 1).

Test Occasion Details

PMA at time of testing ranged from 39 to 47w among the fullterm infants, which overlapped with that for all those born preterm (Figure OLS 1D). Despite this overlap, the preterm groups (controls, CLD and RDS) were less mature, shorter and lighter at time of test when compared with term controls (Table 2, Figure OLS 2 and OLS 3). There was evidence of growth restriction among all the preterm groups, in whom length and weight Z-scores were significantly lower than in term infants. This discrepancy was particularly marked in those with CLD who were very short for their weight (Figure OLS 3). Reflecting this growth pattern, CLD infants also had a higher BMI at time of test compared with PTC or RDS groups.

Table 1: Population characteristics according to diagnostic category

	FTC	PTC	RDS	CLD
Total n (% male)	64 (44)	59 (42)	54 (57)	42 (57)
-HUH: n (% male)	25 (40)	22 (50)	7 (42)	17 (47)
-UCLH: n (% male)	4 (25)	10 (50)	5 (60)	7 (71)
-KEMH: n (% male)	35 (48)	27 (33)	42 (60)	18 (61)
Gestation (w)	39.7 (1.2)	33.5 (1.9)	29.6 (2.3) ^a	26.6 (2.1) ^{b,c}
Birthweight (kg)	3.36 (0.41)	1.85 (0.50)	1.32 (0.45) ^a	0.89 (0.29) ^{b,c}
Birthweight Z-score[22]	- 0.12 (0.90)	- 0.78 (1.19)	- 0.34 (1.18)	- 0.58 (0.29)
Ethnicity (% Caucasian)	73 %	70 %	85 %	64 % ^c
In utero smoke exposure (%)	13 %	22 %	21 %	12 %
Antenatal steroids (%)	-	62 %	87 % ^a	86 % ^b
PROM (%)	-	19 %	15 %	26 %
Surfactant	-	-	83 %	95 %
Total respiratory support (d) ^{*,d}	-	0.0 (0.0; 0.0)	4.8 (1.3; 21) ^a	42 (16; 70) ^{b,c}
Total invasive ventilation (d) [*]	-	-	0.6 (0.2; 1.0)	4.8 (0.6; 39) ^c
O ₂ requirement (d) [*]	-	0.0 (0.0; 0.0)	1.5 (0.0; 16)	88 (75; 156) ^{b,c}

FTC, full term controls; PTC, preterm controls; RDS, Respiratory Distress Syndrome; CLD, Chronic Lung Disease; HUH, Homerton University Hospital; UCLH, University College London Hospital; KEMH, King Edward Memorial Hospital; PROM, prolonged rupture of membranes; d, days. Invasive ventilation indicates tracheal tube in place.

Results expressed mean (SD) if not stated otherwise

*Median (Interquartile Range)

^asignificant difference between PTC and RDS,

^bsignificant difference between PTC and CLD,

^csignificant difference between RDS and CLD (see OLS for details)

^dtotal respiratory support (d): total number of days the child received either invasive (ie intubated) or non-invasive (CPAP, continuous positive airway pressure) ventilatory support.

Table 2: Test occasion details and main results

	FTC	PTC	RDS	CLD
Number of infants	64	59	54	42
Number of tests ^a	64	94	86	62
1 test occasion: n	64	34	30	25
2 test occasions: n	-	17	17	14
3 test occasions: n	-	6	6	3
4 test occasions: n	-	2	1	-
Details from last test occasion in each infant				
PMA (w)	44.2 (1.4)	40.2 (4.1)	39.7 (4.3)	41.3 (3.8)
Weight (kg)	4.37 (0.54)	3.10 (1.00)	2.91 (0.90)	3.16 (0.91)
Weight Z-score[22]	0.09 (0.78)	-0.92 (1.08)	-1.11 (1.06)	-1.26 (1.31)
Length (cm)	54.6 (2.8)	48.7 (4.3)	47.8 (4.6)	47.5 (3.9)
Length Z-score[22]	0.29 (1.12)	-1.06 (1.32)	-1.38 (1.39)	-2.2 (1.35)
BMI	14.6 (1.3)	12.7 (2.3)	12.4 (1.8)	13.7 (2.0)
BMI Z-score	-0.55 (1.10)	-0.54 (0.98)	-0.34 (1.18)	0.16 (1.27)
FRC (mL)	79.6 (14.5)	64.3 (15.6)	59.2 (16.6)	53.5 (15.3)
FRC (mL·kg ⁻¹)	18.4 (3.6)	22.1 (6.0)	21.2 (5.4)	17.5 (4.8)
LCI	7.17 (0.54)	7.14 (0.88)	6.89 (0.74)	7.56 (1.12)
M ₁ /M ₀	2.12 (0.21)	2.08 (0.26)	1.99 (0.20)	2.22 (0.32)
M ₂ /M ₀	8.14 (2.47)	7.70 (2.18)	6.97 (1.48)	8.96 (2.82)
V _T (mL)	29.7 (5.9)	21.8 (7.0)	19.1 (7.0)	18.7 (6.1)
V _T (mL·kg ⁻¹)	6.8 (1.3)	7.2 (1.2)	6.6 (1.5)	6.0 (1.0)
RR (min ⁻¹)	47 (12)	53 (13)	53 (14)	59 (16)
MV (mL·kg ⁻¹)	311 (57)	362 (87)	336 (74)	345 (83)

PMA, postmenstrual age; V_T, tidal volume; BMI, body mass index; RR, Respiratory rate; MV, minute volume; FRC, functional residual capacity; LCI, lung clearance index; M₁/M₀ and M₂/M₀, 1st and 2nd moment ratios. Results are expressed as mean (SD). ^a total number of tests used for multivariable and multilevel analysis

Lung Function Outcomes

Within-test variability, as measured by average SD for differences between replicate measurements on the same test occasion was 0.56 (LCI), 0.15 (M_1/M_0) and 1.26 (M_2/M_0) respectively. No significant differences in variability between fullterm and preterm infants were observed for indices of ventilation inhomogeneity. By contrast, the average SD for FRC in preterm infants (1.88 mL/kg) was significantly higher (p 0.004) than in fullterm infants (1.37 mL/kg).

Summary of univariable analysis

As expected, on univariable analysis body size and age at test were significantly associated with FRC, V_T and RR, as were a wide range of potential determinants, including gestational age (GA), birthweight and birthweight Z-score, severity of neonatal respiratory disease (both duration of O_2 requirements and total ventilation days (d)) and maternal smoking during pregnancy. By contrast, indices of ventilation inhomogeneity (LCI, M_1/M_0 and M_2/M_0) were not associated with either current age or body size or with any other variable except disease severity, for which there was a very small but significant contribution. Sex and ethnic group had no effect on any outcome variable except RR and MV, which was lower amongst Caucasian infants (see table OLS 1).

Functional Residual Capacity

Figure 1 shows both absolute FRC and FRC/kg body weight according to diagnostic category. (Scatter plots of FRC in relation to age, weight, length and BMI are shown in Figure OLS 4). Multivariable analysis demonstrated that, after adjustment for centre, length and BMI, PMA and weight at test no longer had any effect on FRC. After correcting for body size, being preterm resulted in a significant decrement of 0.5 mL/w GA while intrauterine growth restriction was associated with an average reduction of 1.9 mL per unit reduction in birth weight Z-score. FRC was further reduced by 0.04 mL/d of supplemental O_2 (Table OLS 2). Having these factors accounted for, diagnostic category had no additional effect. However, there were additional equipment differences. When compared with FRC data from HUH collected using Equipment 2, those from UCLH and KEMH (Equipment 1) were significantly lower ($p < 0.0001$) by an average of 8.6 mL. Despite these equipment variations, the strength of associations between FRC and length, BMI, GA, birth weight Z-score and days of supplemental O_2 remained unchanged (see OLS).

Ventilation Inhomogeneity

Figure 2 shows LCI according to disease category, with similar plots for moment ratios shown in Figure OLS 6. As expected, indices of ventilation inhomogeneity were not associated with age or body size at test, derived values already being 'internally corrected' by adjusting for FRC. On multivariable analysis, the only factor significantly associated with LCI was disease severity, being 0.003 higher/day supplemental O_2 (Table OLS 3, Figure OLS 5). The association with O_2 remained significant after adjusting for FRC. LCI data from KEMH and UCLH (Equipment 1) was 0.34 lower than that from HUH ($p < 0.0001$).

Tidal Breathing Indices

Respiratory rate fell with increasing age (Table OLS 6). After adjusting for BMI and age at test, ethnic group no longer had any significant influence on RR. Slightly lower rates were found in boys, whereas higher rates were associated with intrauterine growth restriction and disease severity but not with prematurity. No equipment differences were observed for RR.

V_T : After adjusting for current age and body size, V_T was significantly and positively associated with gestational age and intra-uterine growth (Table OLS 7), but not with either disease severity or

diagnostic category. Equipment effects were observed, with V_T from HUH being on average 2.6 mL higher ($p < 0.0001$) than that from KEMH and UCLH.

Minute Volume: After adjusting for current weight, Equipment set was the only factor significantly associated with minute volume on multivariable analysis, this being on average 96 mL higher for Equipment set 2 (HUH) (Table OLS 8).

DISCUSSION

This multicentre investigation represents the largest of its kind and the first to use sophisticated statistical modelling to explore the mechanistic basis for changes in infant lung function due to immaturity and/or lung disease after accounting for known determinants such as age and body size. By enrolling a large cohort of newborns, including term and preterm controls as well as those recovering from respiratory disease, we were able to adopt a more clinically plausible approach using continuous measures of immaturity and disease severity, rather than simply categorising infants according to somewhat arbitrary definitions. After adjusting for equipment, age and body size at test, reductions in gestation and birth weight Z-score were both associated with a lower FRC, together with a small but significant contribution of disease severity (supplemental O_2 days). These findings support contemporary concepts of disturbed lung growth after preterm birth, and suggest that a significant proportion of the abnormal lung function associated with chronic lung disease can be accounted for by the degree of prematurity and intrauterine growth restriction. In contrast, pre and postnatal growth patterns had little impact on LCI, which was only affected by duration of O_2

Strengths and Limitations

Important strengths of this study include the large study group, multicentre recruitment and inclusion of prospective preterm and fullterm controls. Extensive quality control in accordance with international standards was applied to data collection and analysis, with sophisticated and robust statistical methodology being used to explore the complex interactions within the data. Although we summarised our data according to diagnostic categories to facilitate comparison with previous literature, we addressed the potential problems arising from such categorisation by using continuous variables of disease severity and immaturity during multivariable analysis.

The inter-centre differences observed in this study were unexpected, since the same investigators were involved at all three sites, and we had used the same methods and type of equipment. Further investigations using multilevel modelling indicated that these centre differences were primarily associated with which set of equipment was used. Given that FRC was ~three times V_T (Table 2), there appears to be internal consistency, in that an over-estimation of 2.6 mL in V_T would multiply to one of around 8 mL for FRC. We can only surmise the presence of subtle differences in hardware or software of this prototype USFM device that, despite extensive investigation, we were unable to detect (see OLS for details). Fortunately, sample size was sufficiently large to allow us to model for these differences, which did not affect the overall conclusions of the study. However, such differences could have confounded a smaller study or one relying on normative data collected elsewhere; this highlights the need to recruit prospective controls in such studies.

Determinants of neonatal lung function:

The rapid somatic and lung growth that occur in the first months of life must be properly accounted for if the impact of either prematurity or lung disease on lung function is to be determined reliably. Height (or length) is known to be the major determinant of FRC,[23] as confirmed in this study. Lung function is commonly corrected for differences in body size during the neonatal period simply by dividing by body weight (see Figure 1B). In the presence of low weight Z-scores, this practice will over-estimate reported values. By contrast, in the presence of high 'weight-for-length' or age,

values will be under-estimated. The combination of both situations, as seen among the CLD infants (Table 2) can totally confound interpretation of results. In this study, adjustment for age and body size was achieved by repeating tests in the preterm infants to overlap with the age at which fullterm infants were studied and by using appropriate statistical models. Inappropriate adjustment for growth could be responsible for much of the conflicting evidence regarding neonatal lung function.[9,24]

Although birth weight is a known determinant of lung function throughout life,[25-29] this is a complex factor which encompasses both maturation and adequacy of intrauterine growth. In our study, the combination of gestation and birth weight Z-score explained more variability than birth weight alone, and provided some insight into mechanisms underlying the variability of neonatal lung function. Our findings show that maturity at birth and intrauterine growth are independently predictive of tidal volume and FRC soon after birth, and that intrauterine growth also influences respiratory rate.

Comparison with published literature:

The multivariable statistical approach used in this study differs substantially from previous reports that have relied on t-tests or ANOVA.[8,30,31] In addition, in contrast to previous publications, by entering each category into the models as separate factors, the patterns of association could emerge without prior assumptions being made as to the nature of the differences expected. In this study, V_T , RR and FRC were primarily influenced by markers of maturity, intrauterine growth, disease severity and postnatal anthropometry. FRC is frequently used to monitor lung disease during infancy and early childhood. Reports of a lower FRC at term/near term for infants with CLD when compared with healthy controls in the pre-surfactant era[32] have been reconfirmed by several studies in the post-surfactant era.[8,30,33] More recent studies have shown that preterm birth *per se* and lung disease requiring treatment may have independent effects on lung volume,[31,34] although restricted sample sizes have often limited power to fully elucidate factors influencing neonatal lung function. Like Hjalmarsen et al,[31] we observed lower FRC in preterm than term controls at a comparative postmenstrual age, although differences are not as obvious when FRC is plotted against length (Figure OLS 4). A more recent study presented findings on infants with CLD and RDS, as well as controls[8] but did not report length or any differences in the proportionality of growth between groups with which to compare our data. Whereas we acknowledge the ideal of matching for postmenstrual age, this was impractical. We thus prioritised modelling for other highly variable factors likely to have a greater effect on lung function than PMA *per se*.

The recent interest in assessing ventilation inhomogeneity in preterm infants[8,11,30,31,35] is not surprising, given that neonatal respiratory disease is dominated by abnormalities of the lung periphery and interrupted acinar development.[4,36] Increases in either LCI or moment ratios signify reduced ventilation efficiency. Our observation that duration of O_2 is positively associated with increased ventilation inhomogeneity is consistent with previous studies, [8] although the effect size was small. Importantly, we did not find evidence to support a preliminary report that lung clearance index decreased with advancing postnatal age during the neonatal period.[37] LCI is corrected for FRC, hence the lack of an association between LCI and body size is not surprising. The indices we studied are also influenced by breathing patterns,[38] particularly dead space/tidal volume and V_T /FRC ratios. In contrast to its proven utility when detecting early lung disease in sedated infants with CF[39], indices of ventilation inhomogeneity may be less useful when assessing the effect of prematurity or lung disease in small unsedated neonates.

Conclusions

Sophisticated analysis of this large study has highlighted the complexity of factors contributing to lung function outcomes after preterm birth. Lung volume and tidal breathing parameters are

influenced by intrauterine growth, maturity at birth, anthropometry and disease severity, whereas disease severity alone has a small but significant influence on ventilation inhomogeneity. These findings support contemporary concepts of disturbed lung growth after preterm birth, and suggest that a significant proportion of the abnormal lung function associated with chronic lung disease can be accounted for by the degree of prematurity and intrauterine growth. Disproportionate growth patterns following preterm delivery raise concerns about potential misinterpretation of studies using simplistic normalisation of variables such as FRC/kg to compare disparate subject groups. The contribution of equipment differences to most of the outcome measures further highlights the importance of recruiting a prospective control group for such investigations. Only by using multivariable analysis were we able to account for the multiple complex and interacting determinants of lung function and distinguish the effects of prematurity and lung disease from those of growth and development.

ACKNOWLEDGEMENTS

The authors thank the babies and their parents who participated in this study, and the staff of the neonatal units at University College London Hospital, Homerton University Hospital and King Edward Memorial Hospital. We also thank Ms Anne Cantarella, Ms Kerry Williams, Dr David Baldwin and Dr Sanjay Sinhal, who provided valuable assistance in subject recruitment and obtaining the measurements.

COMPETING INTERESTS

The authors (JS, SL, JP) received some research discount on equipment supplied by EcoMedics in association with undertaking studies contributing to the development of the EcoMedics infant multiple breath washout equipment and software.

FUNDING

Supported by: the Innovative Medizinische Forschung, University of Münster and the Gesellschaft für Pädiatrische Pneumologie, Germany (GH), a Neil Hamilton Fairley NHMRC Postdoctoral Fellowship and NHMRC Career Development Award (JP) and Portex Ltd. (JS). Research at the Institute of Child Health and Great Ormond Street Hospital for Children NHS Trust benefits from R&D funding received from the NHS Executive. Research at King Edward Memorial Hospital was supported by grants from the Women & Infants Research Foundation and the Raine Medical Research Foundation.

DECLARATION

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence (or non exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd and its Licensees to permit this article (if accepted) to be published in [THORAX] editions and any other BMJPG Ltd products to exploit all subsidiary rights, as set out in our licence (<http://thorax.bmj.com/ifora/licence.pdf>).

FIGURE LEGENDS

Figure 1: Functional Residual Capacity according to diagnostic category

o: term control (FTC; n = 64); **Δ**: preterm control (PTC; n = 59); **□**: preterm respiratory distress syndrome (RDS; n = 54); **x**: preterm chronic lung disease (CLD; n = 42). In infants in whom more than one test had been performed, presentation of results in the figure is limited to data from the last test occasion (i.e one result per subject). While there were significant differences between fullterm and preterm groups due to differences in body size at time of test (Fig 1A, Table 2), simply expressing results as FRC/kg as shown in Figure 1B, would have led to misleading conclusions due to the disproportionate growth patterns among the preterm groups (Table 2) (see text for details).

Figure 2: Lung clearance index (LCI) according to diagnostic category

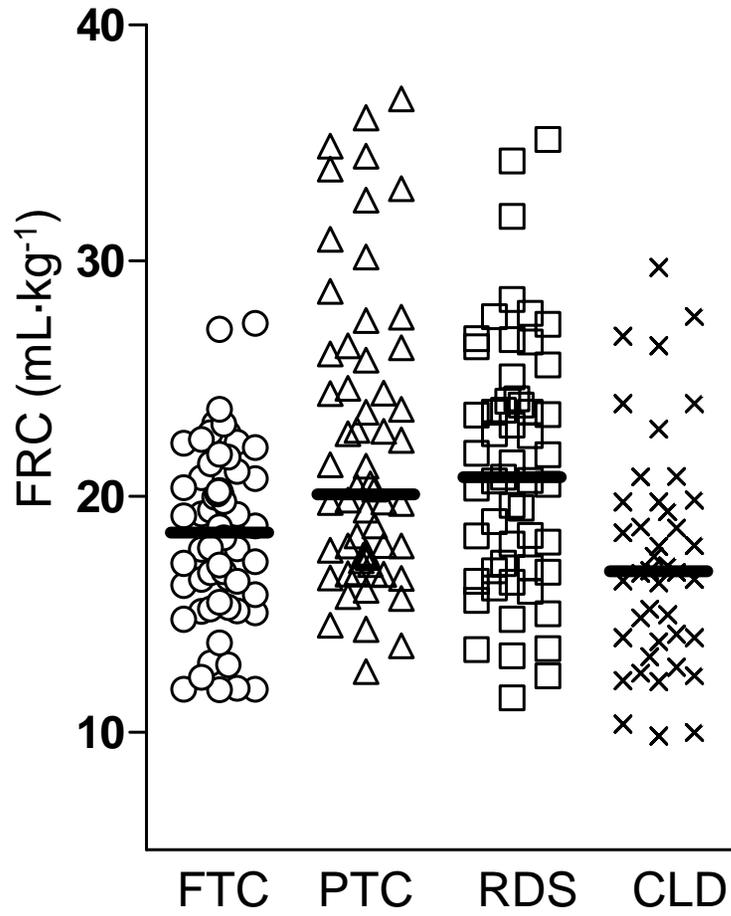
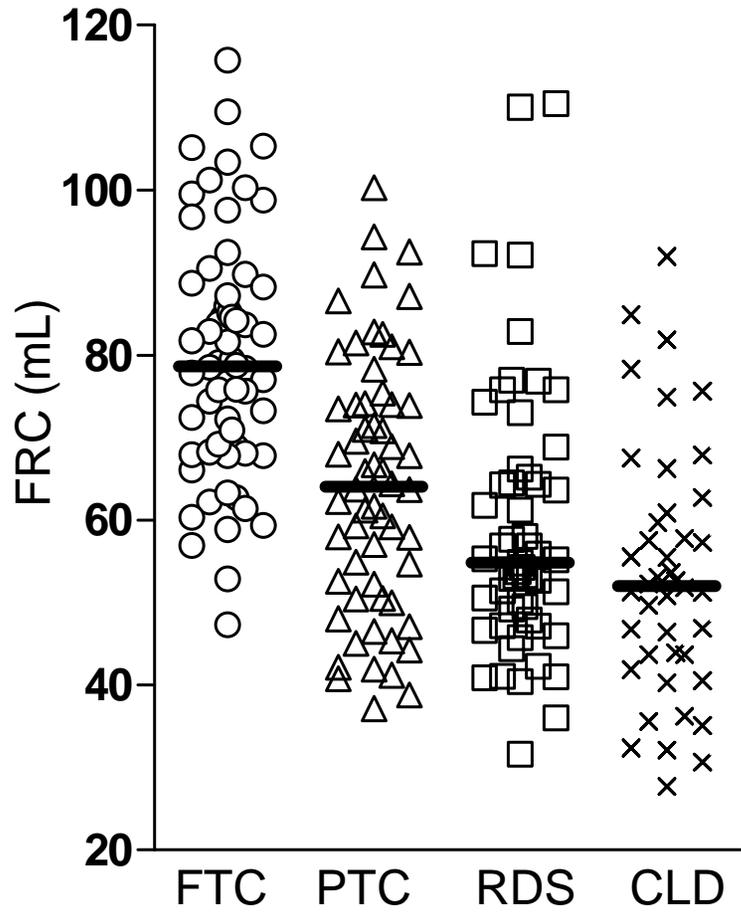
o: term control (FTC; n = 64); **Δ**: preterm control (PTC; n = 59); **□**: preterm respiratory distress syndrome (RDS; n = 54); **x**: preterm chronic lung disease (CLD; n = 42). In infants in whom more than one test had been performed, presentation of results in the figure is limited to data from the last test occasion (i.e one result per subject).

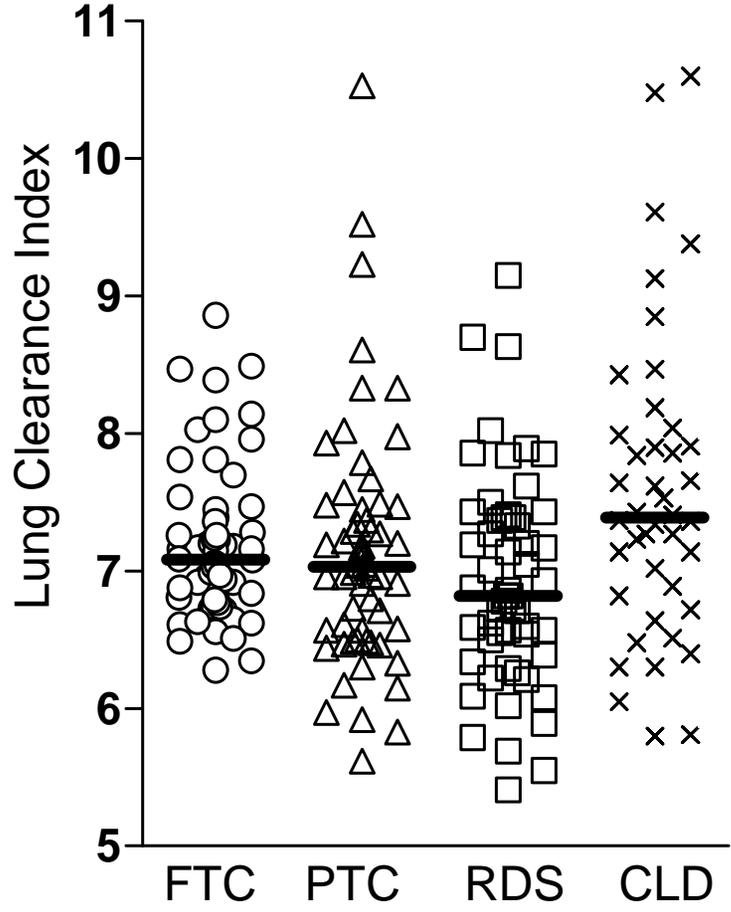
REFERENCES

1. Jobe AH, Ikegami M. Prevention of bronchopulmonary dysplasia. *Curr Opin Pediatr* 2001; **13**: 124-129.
2. Coalson JJ. Pathology of new bronchopulmonary dysplasia. *Semin Neonatol* 2003; **8**: 73-81.
3. Allen J, Zwerdling R, Ehrenkranz R *et al*. Statement on the care of the child with chronic lung disease of infancy and childhood. *Am J Respir Crit Care Med* 2003; **168**: 356-396.
4. Jobe AH, Bancalari E. Bronchopulmonary dysplasia. *Am J Respir Crit Care Med* 2001; **163**: 1723-1729.
5. Ehrenkranz RA, Walsh MC, Vohr BR *et al*. Validation of the National Institutes of Health consensus definition of bronchopulmonary dysplasia. *Pediatrics* 2005; **116**: 1353-1360.
6. Walsh MC, Yao Q, Gettner P *et al*. Impact of a physiologic definition on bronchopulmonary dysplasia rates. *Pediatrics* 2004; **114**: 1305-1311.
7. Walsh MC, Szeffler S, Davis J *et al*. Summary proceedings from the bronchopulmonary dysplasia group. *Pediatrics* 2006; **117**: S52-S56.
8. Hjalmarson O, Sandberg KL. Lung function at term reflects severity of bronchopulmonary dysplasia. *J Pediatr* 2005; **146**: 86-90.
9. Hülkamp G, Pillow JJ, Dinger J *et al*. Lung function tests in neonates and infants with chronic lung disease of infancy: functional residual capacity. *Pediatr Pulmonol* 2006; **41**: 1-22.
10. Hülkamp G, Hoo AF, Ljungberg H *et al*. Progressive decline in plethysmographic lung volumes in infants: physiology or technology? *Am J Respir Crit Care Med* 2003; **168**: 1003-1009.
11. Pillow JJ, Frerichs I, Stocks J. Lung function tests in neonates and infants with chronic lung disease: global and regional ventilation inhomogeneity. *Pediatr Pulmonol* 2006; **41**: 105-121.
12. Latzin P, Thamrin C, Kraemer R. Ventilation inhomogeneities assessed by the multibreath washout (MBW) technique. *Thorax* 2008; **63**: 98-99.
13. Buess C, Pietsch P, Guggenbuhl W *et al*. A pulsed diagonal-beam ultrasonic airflow meter. *J Appl Physiol* 1986; **61**: 1195-1199.
14. Schibler A, Henning R. Measurement of functional residual capacity in rabbits and children using an ultrasonic flow meter. *Pediatr Res* 2001; **49**: 581-588.
15. Schibler A, Hall GL, Businger F *et al*. Measurement of lung volume and ventilation distribution with an ultrasonic flow meter in healthy infants. *Eur Respir J* 2002; **20**: 912-918.
16. Wauer J, Leier TU, Henschen M *et al*. In vitro validation of an ultrasonic flowmeter in order to measure the functional residual capacity in newborns. *Physiol Meas* 2003; **24**: 355-365.

17. Pillow JJ, Ljungberg H, Hülskamp G *et al.* Functional residual capacity measurements in healthy infants: ultrasonic flow meter versus a mass spectrometer. *Eur Respir J* 2004; **23**: 768.
18. Lum S, Hoo AF, Dezateux C *et al.* The association between birthweight, sex, and airway function in infants of nonsmoking mothers. *Am J Respir Crit Care Med* 2001; **164**: 2078-2084.
19. Shennan AT, Dunn MS, Ohlsson A *et al.* Abnormal pulmonary outcomes in premature infants: prediction from oxygen requirement in the neonatal period. *Pediatrics* 1988; **82**: 527-532.
20. Charafeddine L, D'Angio CT, Phelps DL. Atypical chronic lung disease patterns in neonates. *Pediatrics* 1999; **103**: 759-765.
21. Gaultier C, Fletcher M, Beardsmore C *et al.* Respiratory function measurements in infants: Measurement conditions. *Am J Respir Crit Care Med* 1995; **151**: 2058-2064.
22. Freeman JV, Cole TJ, Chinn S *et al.* Cross sectional stature and weight reference curves for the UK, 1990. *Arch Dis Child* 1995; **73**: 17-24.
23. Stocks J, Quanjer PH. Reference values for residual volume, functional residual capacity and total lung capacity ATS Workshop on Lung Volume Measurements. Official statement of the European Respiratory Society. *Eur Respir J* 1995; 492-506.
24. Hülskamp G, Pillow JJ, Stocks J. Lung Function Testing in Acute Neonatal Respiratory Disorders and Chronic Lung Disease of Infancy: A Review Series. *Pediatr Pulmonol* 2005; **40**: 467-470.
25. Sharma P, McKay K, Rosenkrantz TS *et al.* Comparisons of mortality and pre-discharge respiratory outcomes in small-for-gestational-age and appropriate-for-gestational-age premature infants. *BMC Pediatr* 2004; **4**: 9.
26. Torrance HL, Voorbij HA, Wijnberger LD, van Bel F, Visser G. Lung maturation in small for gestational age fetuses from pregnancies complicated by placental insufficiency or maternal hypertension. *Early Hum Dev* 2008; **84**: 465-469.
27. Hoo AF, Stocks J, Lum S *et al.* Development of lung function in early life: influence of birth weight in infants of nonsmokers. *Am J Respir Crit Care Med* 2004; **170**: 527-533.
28. Chan KN, Noble-Jamieson CM, Elliman A *et al.* Lung function in children of low birth weight. *Arch Dis Child* 1989; **64**: 1284-1293.
29. Canoy D, Pekkanen J, Elliott P *et al.* Early growth and adult respiratory function in men and women followed from the fetal period to adulthood. *Thorax* 2007; **62**: 396-402.
30. Shao H, Sandberg K, Hjalmarson O. Impaired gas mixing and low lung volume in preterm infants with mild chronic lung disease. *Pediatr Res* 1998; **43**: 536-541.
31. Hjalmarson O, Sandberg K. Abnormal lung function in healthy preterm infants. *Am J Respir Crit Care Med* 2002; **165**: 83-87.

32. Gerhardt T, Hehre D, Feller R *et al*. Serial determination of pulmonary function in infants with chronic lung disease. *J Pediatr* 1987; **110**: 448-456.
33. de Winter JP, Merth IT, Brand R *et al*. Functional residual capacity and static compliance during the first year in preterm infants treated with surfactant. *Am J Perinatol* 2000; **17**: 377-384.
34. Kavvadia V, Greenough A, Itakura Y *et al*. Neonatal lung function in very immature infants with and without RDS. *J Perinat Med* 1999; **27**: 382-387.
35. Shao H, Sandberg K, Sjoqvist BA *et al*. Moment analysis of multibreath nitrogen washout in healthy preterm infants. *Pediatr Pulmonol* 1998; **25**: 52-58.
36. Coalson JJ, Winter V, DeLemos RA. Decreased alveolarization in baboon survivors with bronchopulmonary dysplasia. *Am J Respir Crit Care Med* 1995; **152**: 640-646.
37. Pillow JJ, Hülskamp G, Stocks J. Lung clearance index changes with maturation in preterm neonates. *Eur Respir J* 26[Suppl 49], 6s. 2005.
38. Habib RH, Lutchen KR. Moment analysis of a multibreath nitrogen washout based on an alveolar gas dilution number. *Am Rev Respir Dis* 1991; **144**: 513-519.
39. Lum S, Gustafsson P, Ljungberg H *et al*. Early detection of cystic fibrosis lung disease: multiple-breath washout versus raised volume tests. *Thorax* 2007; **62**: 341-347.





ASSOCIATION OF PREMATURITY, LUNG DISEASE AND BODY SIZE WITH LUNG VOLUME AND VENTILATION INHOMOGENEITY IN UNSEDATED NEONATES: A MULTICENTRE STUDY

ONLINE SUPPLEMENT

Georg Hülskamp, Sooky Lum, Janet Stocks, Angie Wade, Ah-fong Hoo, Kate Costeloe, Jane Hawdon, Kandadai Deeptha and J. Jane Pillow

Study Subjects

Data were collected in the two London units (Homerton University Hospital, HUH and University College London Hospital, UCLH) between December 2001 and August 2003, and between April 2005 to August 2006 at King Edward Memorial Hospital, KEMH in Perth, Australia. Information was stored on a protected relational database adhering to standard and institutional data protection requirements. According to power calculations after allowing for some drop out of individuals prior to follow-up studies, we aimed to recruit up to 60 (minimum 40) subjects into each diagnostic category. Results of 18 measurements in 14 healthy infants have been reported previously as part of a methodological study.[1]

Diagnostic Categories

Infants were classified according to gestation and clinical status as follows:

1. **Fullterm Controls (FTC):** ≥ 37 weeks (w) gestational age (GA) with no signs or history of respiratory illness prior to testing.
2. **Preterm Controls (PTC):** < 37 w GA with minimal requirements for ventilatory support (i.e. no invasive ventilation, maximum 48 hours (h) support with nasal continuous positive airway

pressure (CPAP), no supplemental inspiratory Oxygen (O₂) after 48 h postnatal age and no signs or history of respiratory illness prior to testing).

3. **Respiratory Distress Syndrome (RDS):** <37 w GA with a history of resolved respiratory distress syndrome and no requirement for supplemental FiO₂ at 36 w postmenstrual age (PMA).
4. **Chronic Lung Disease (CLD):** requirement for supplemental FiO₂ or assisted ventilation at 36 w PMA.[2,3]

Exclusion criteria:

- Congenital malformations of the thorax;
- Major cardiac malformation (isolated small atrial or ventricular septal defects and haemodynamically insignificant patent ductus arteriosus could be included);
- Congenital neuromuscular disorders;
- Meconium aspiration syndrome;
- Isolated persistent pulmonary hypertension of the newborn (PPHN);
- Recruitment contra-indicated on psycho-social grounds or
- Lack of written informed parental consent.

Definitions used for diagnostic categories:

A major challenge when trying to determine the impact of neonatal lung disease or its associated treatment lies in the difficulty of defining an appropriate control group. Term healthy controls are easily identified, but do not control for the effect of preterm birth on lung development. Defining a gestation-matched “healthy” preterm control group is virtually impossible, since many infants at the lower extremes of gestation are routinely ventilated and given exogenous surfactant at birth even in the

absence of overt lung disease. Ventilation is known to be a risk factor for lung disease[4] and even brief periods of large tidal volume ventilation at resuscitation[5] or gentle ventilation lasting several hours[6] may be associated with lung injury. For this reason, infants were only classified as ‘preterm controls’ if they did not receive invasive ventilation and received supplemental O₂ for less than the first 48 h of life. The boundary between PTC and RDS groups or RDS and CLD groups are similarly indistinct and highlight the difficulties associated with the use of categorical definitions of disease severity. We defined CLD as a mandatory requirement for increased O₂ or mechanical ventilation at 36 w PMA irrespective of the type of CLD.[3] RDS was defined as those needing support in excess of that permitted for inclusion in the preterm control category, but not requiring supplemental O₂ or mechanical ventilation at 36 w PMA. This approach distinguishes the infant who is weaned from oxygen at 35 w and 6 days (d) PMA as less severe than one weaned 1 d later, when there is little clinical difference between the two. Although a semi-quantitative categorisation of CLD according to O₂ requirements at 36w PMA has been suggested,[7] for the purposes of this study, we addressed the potential problems arising from inappropriate categorisation by using continuous variables of disease severity and immaturity during multivariable analysis. Data were also summarised according to the classical diagnostic categories to facilitate comparison with previous literature.

Lung Function Tests

Test Conditions and Measurements

Once clinically stable, preterm infants were studied on at least one, and up to four, occasions in hospital prior to discharge. When geographically and practically feasible, these infants returned for follow-up studies at approximately 44 w PMA (as part of an ongoing longitudinal study), which allowed measurements to be obtained at a similar PMA to those born full term. Due to the rapid discharge from hospital following full term deliveries and practical constraints for new parents, recruitment of term

infants did not occur until after discharge home. These term infants were studied on one occasion only within 4 w to 8 w of birth.

Infants were tested when clinically stable and free from any infections for at least 3 weeks. All measurements were performed in supine, unsedated infants at least 30 minutes following a feed and during quiet sleep. Nasal cannulae for oxygen supply were removed (O₂ being supplemented through the equipment where necessary), whereas nasogastric feeding tubes were left *in situ* and closed. Potential leak around feeding tubes was sealed with therapeutic putty. Ideally, nasogastric tubes should be removed prior to lung function tests, since their presence may alter the child's nasal and hence total airway resistance.[8] However, their effect on FRC and ventilation inhomogeneity is likely to be far less. For the purposes of this study, it was felt that any benefits from removing would be outweighed by the disturbance caused to the infant, and potential loss to recruitment if parents were concerned about having to re-pass a recently changed tube. Measurements in all but 21 infants (42 tests) were obtained without a nasogastric tube *in situ*.

Tidal breathing, FRC and indices of ventilation inhomogeneity were measured by the Multiple Breath Washout technique (MBW) with 4 % SF₆ as a tracer gas using a commercially available prototype ultrasonic flow-meter (Spiroson®, Ecomedics AG, Duernten, Switzerland) with acquisition and analysis software (Wbreath v3.16.11.0, Ndd Medizintechnik AG, Zurich, Switzerland), described previously.[1,9] A deadspace reducer was used, reducing the internal dead space of the flowmeter to 1.2 mL. A disposable bacterial filter (Spirette™, EcoMedics AG, Duernten, Switzerland) surrounded the dead space reducer. The absence of leak from the sensor was confirmed up to 2 kPa during equipment calibration.[10] A size 0/0 (5 mL effective dead space) or size 0/1 (7.5 mL effective deadspace) Laerdal face mask (Laerdal Medical AS, Stavanger, Norway) was used as the interface between the sensor and the patient airway. A bias flow of 200 mL·sec⁻¹ was used.

During data collection, flow-volume loops were observed to ensure there was no leak. Data were recorded at ambient temperature, pressure and saturation with flow and volume being converted to body temperature and pressure, saturated conditions during data analysis. Tidal breathing data (tidal volume, V_T and respiratory rate, RR) were computed from recordings obtained during quiet tidal breathing, at least 30 s after initial mask placement and prior to wash-in with 4 % SF₆.

Washout (with air) was commenced after equilibrium was achieved. The washout was continued until the tracer gas had been eliminated from the lungs. The process was repeated until 3 acceptable MBW washouts had been recorded, or the infant awoke. Quality control criteria for technical acceptability were established prospectively. FRC and indices of ventilation including lung clearance index (LCI) and the first and second to zeroeth moment ratios (M_1/M_0 and M_2/M_0) were computed.

Data Acquisition and Analysis

Calculation of Tidal Breathing, Lung Volume and Ventilation Inhomogeneity Indices:

Data were collected and immediately stored in “raw” format, and modified as outlined below using algorithms within the WBreath software: The Spiroson® ultrasonic flow meter allows assessment of molar mass in the range of 20-45 g·mol⁻¹ with a precision of 0.01 g·mol⁻¹. Molar mass (MM in g·mol⁻¹) of the gas in the main stream of the flowmeter can be computed from the transit time and its value is directly proportional to the density of the medium. The density of the medium, however, is dependent on the precise temperature along the sound transmission path. Due to the complexity of the determination of the temperature along the sound transmission path and because minor changes in temperature are present during tidal breathing, the changes of temperature along the sound path need to be simulated using a temperature model. The temperature model uses the measured flow, instrumental

dead space and the preset temperature (at body temperature and ambient pressure, and saturated with water vapor (BTPS) correction parameters) of the inspiratory and expiratory gas to calculate the temperature changes along the sound transmission path. In the current study, the effects of within-cycle changes in temperature on molecular mass were accounted for[11] using an updated temperature model.[12] This model assumes a lower dead space for the direct “flow path” between the airway opening and the midsensor point, compared to the larger “effective” deadspace of the mask (50% of mask volume) which is subtracted from the initial FRC at the midsensor point ($FRC_{\text{midsensor}}$) to obtain the FRC at the airway opening (patient mouth/nose).

The analysis of MBW data also included BTPS correction of flow and correction for a side chamber reservoir effect. $FRC_{\text{midsensor}}$ was computed from the cumulative volume of expired SF_6 divided by the difference between end-tidal gas concentration at the start and completion of the washout. The volume of exhaled tracer gas was automatically corrected for re-inspired gas between the bias flow and the sensor mid-point and for the dead space between the sensor mid-point and the mask. The “effective” dead space[13] of the mask used was subtracted from the midsensor value to obtain the functional residual capacity at the airway opening (FRC_{ao}).

Lung Clearance Index (LCI) was computed as the number of volume turnovers (cumulative expired volume divided by $FRC_{\text{midsensor}}$) required to reduce the end-tidal tracer gas concentration to 1/40th of the concentration at the start of washout.[14] The first and second moment ratios (M_1/M_0 and M_2/M_0) were computed as the ratio of the first to zeroeth and second to zeroeth moments respectively.[15] The zeroeth moment was computed from the area under the washout curve as a function of the lung volume turnover number. The first and second moments were calculated from the area under the zeroeth and first moment curves (respectively) plotted against the lung turnover number.

To assess any potential inter-observer differences, blinded cross analysis of 31 sets of data from 11 infants was undertaken by the three principal investigators (JP, GH, SL). One of the investigators (JP) was involved with data collection and analysis from all three sites.

Quality Control During Data Collection and Analysis

Data were accepted for analysis if obtained during regular breathing, with a visually stable end-expiratory level[16] for at least five breaths prior to commencement of washout and if there was no evidence of leak (no sudden change in inspired molar mass or flow). Traces were not accepted for analysis if a sigh was present in the six breaths immediately prior to, or the 10 breaths immediately following commencement of the washout. Technically acceptable washouts had at least 10 breaths after tracer gas concentration had fallen below 2.5 % of the starting concentration. A minimum of 2 technically acceptable washout traces were required for final acceptance of data into the database.

Statistical Models

From previous experience, we estimated that 80 % of the observed between-subject variability for any specific parameter of lung function could be accounted for by factors such as sex, body size and age at test.[17,18] While precise estimates of variability will vary according to disease state, age, measurement condition (sedated or not) and so forth, this provides some guidance when attempting to predict required sample size. Using this information 160 subjects (subdivided into 4 diagnostic subgroups) would provide 90 % power at the 5 % significance level to detect differences due to diagnostic group that would account for at least 2 % of the total variability (i.e. 10 % of the remaining variability after controlling for known confounders).[19]

Standard software packages were used for data inspection, distribution and descriptive statistics (SPSS for Windows, v15.0, SPSS Inc.). All technically valid data from all test occasions were used in

multilevel modelling, to take account of the correlation between repeat measurements from the same child. Prior to modelling, all potential univariable associations were plotted and inspected to assess linearity, so that data could be transformed if necessary. Univariable regression analysis was used to establish the association between each outcome measure (FRC, LCI, moment ratios and tidal breathing variables) and likely explanatory variables (sex, ethnicity, antenatal smoke exposure, age and body proportions at time of test; maturity at birth and intrauterine growth (gestation, birth weight and birth weight Z-score); and measures of disease severity expressed either as continuous variables (days with respiratory support, O₂ or invasive ventilation) or categorically (according to diagnostic group: term control, preterm control, RDS or CLD) and equipment. Where potentially significant associations ($p < 0.05$) existed, these were further explored using multivariable, multilevel linear regression modelling (MLwiN, version 2.12; Institute of Education, UK). These highly flexible models adjust for the correlated nature of repeated measurements in individuals within each centre and allow inclusion of variable numbers of measurements per child to provide the most precise characterization of changes over time.[20-23]

To tease out the mechanistic basis for differences in lung function outcomes, a stepwise approach was used in developing the multivariable models: each dependent variable (FRC, LCI etc.) was adjusted for potential determinants such as PMA, body size (body weight, length and BMI). The extent to which prematurity (as reflected by GA), intra-uterine growth restriction (birth weight Z-score) or disease severity (d of supplemental O₂ and/or total d of mechanical ventilation) were associated with the remaining inter-subject variability was ascertained, before adding diagnostic category (term control, preterm control, RDS and CLD) to establish whether using such categories had any additional effects. Given the bimodal nature of the distribution, PMA was further evaluated: examination of the residuals did not show any pattern with PMA, indicating that it was adequately modelled with a single linear

term. In all the models, GA and a measure of disease severity (O₂ days or total mechanical ventilation days) were included, irrespective of whether these factors added significantly to the model in order to show their association with the respective dependent variable. The CLD group was used as reference for diagnostic category in order to see the added impact of severe lung disease compared to the RDS, preterm control and term control groups. Potential centre and equipment effects were also examined in each of the models.

RESULTS

Population characteristics

Measurements were attempted in 232 individuals on 340 test occasions. Technically acceptable data were obtained from 219 infants on 306 test occasions. The 10 % failure rate was primarily due to failure to sleep or insufficient/poor quality data. 48 infants (17 PTC, 17 RDS, 14 CLD) were tested on 2 occasions, 15 infants (6 PTC, 6 RDS, 3 CLD) on 3 occasions and 3 infants (2 PTC, 1 RDS) on 4 occasions. The median (range) interval between repeat tests was 7 (1 - 85) d. There were no significant differences in background characteristics between preterm infants who were studied on several occasions as compared with those in whom only a single test was obtained (data not shown).

Details at birth are presented in Figure OLS 1 and summarised in Table 1, main manuscript. As expected, there were marked differences in birth weight and GA between groups, those with CLD predominantly having a birth weight <1.5 kg and a GA <30 w. Within the preterm groups, GA and birth weight were significantly higher in the PTC than in RDS (95 % CI of difference 3.0 w; 4.9 w, and 0.31 kg; 0.73 kg respectively), and in the RDS group when compared with those with CLD (95 % CI: 1.8 w; 3.8 w GA, 0.20 kg; 0.65 kg birth weight). Preterm groups also had lower birth weight Z-scores (all preterm infants versus FTC: -0.46 [-0.74; -0.17]). While not significantly different between diagnostic groups, the lowest birth weight Z-score was observed in PTC, despite this group having a higher gestation. This observation is not surprising, as intrauterine stress may precipitate preterm delivery and reduce the need for respiratory support at birth[24] but cause structural alterations to the developing lung.[25] A slightly higher proportion (34 %) of PTC were small-for-gestational age (as defined as birth weight below the 10th centile for GA or -1.3 Z scores) compared to those with RDS (24%) or CLD (26%) (Figure OLS 1B).

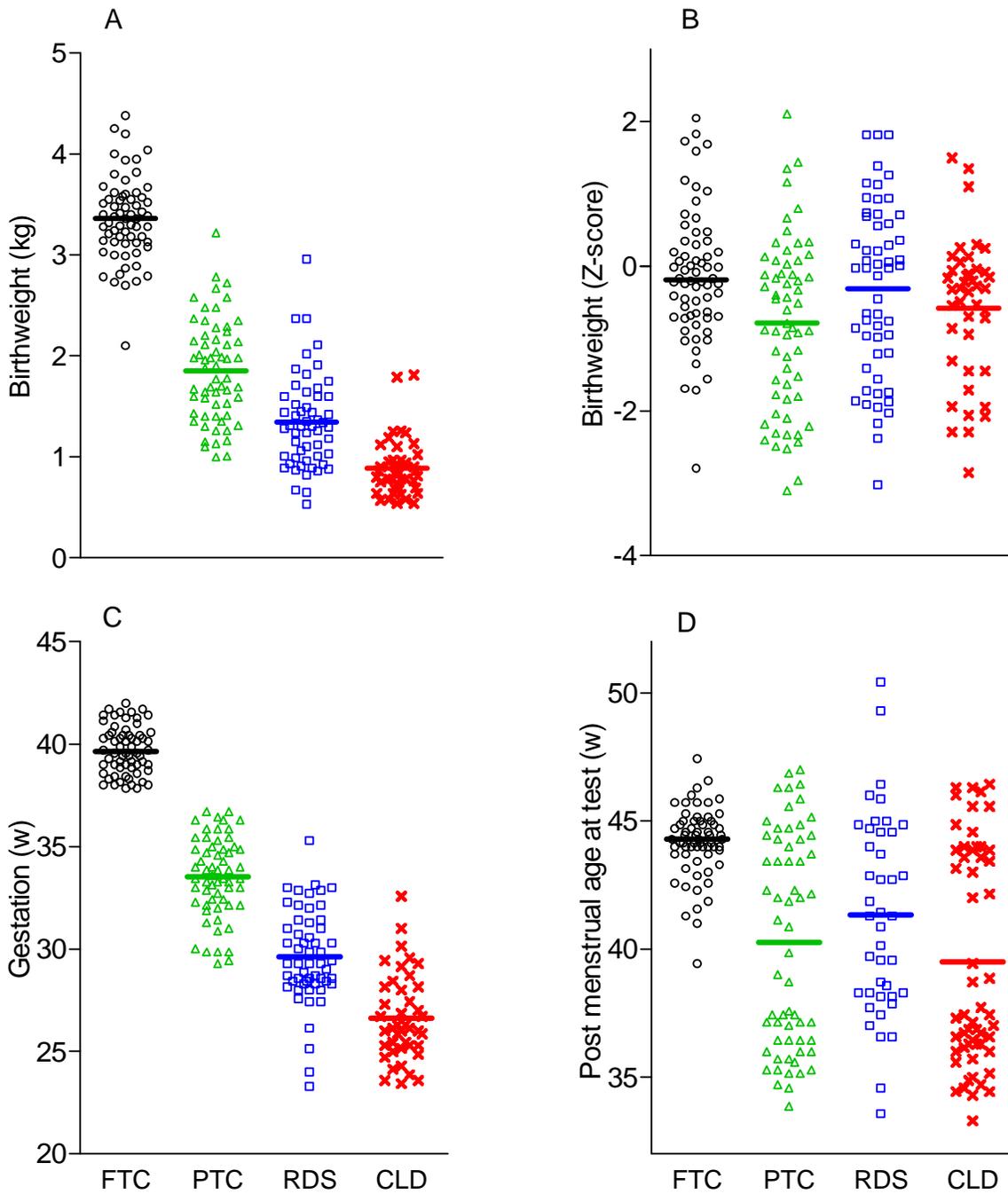


Figure OLS 1: Population characteristics at birth and postmenstrual age at time of test according to diagnostic category. **Legend:** ○ – FTC, fullterm controls; △ – PTC, preterm controls; □ – RDS, respiratory distress syndrome; × – CLD, chronic lung disease. PMA at time of test is only displayed for the last test occasion in those with repeated measurements. Horizontal bars represent median values.

Non-Caucasian ethnic background (primarily Afro-Caribbean) was more common in infants recruited from HUH (61 %) than UCLH (18 %) [95 % CI difference: 21 %; 61 %] or KEMH (9 %) [95 % CI: 32 %; 56 %]. When compared with those with CLD, a significantly higher proportion of infants with RDS were born to Caucasian mothers (19 [2; 35] %). This largely reflected that 41/53 (77 %) of infants with RDS were recruited from the predominantly Caucasian population at KEMH rather than any ethnic susceptibility to CLD. Infants with RDS were much more difficult to study at the two specialised neonatal units in London since they were frequently discharged back to their local district general hospitals before measurements could be undertaken. As expected, there were significant differences in total d of ventilatory support and supplemental O₂ between groups, the mean [95 % difference] between those with CLD and RDS being 32 [24; 40] d and 122 [95;149] d respectively.

Test Occasion Details

PMA at time of testing ranged from 39 w to 47 w among the fullterm infants, which overlapped with that for the preterm infants (Figure OLS 1D). Although all infants were studied before 50 w postmenstrual age (10 w corrected postnatal age) with considerable overlap between groups, the logistics of when infants were discharged home or back to their local hospital, together with a wide range of pre- and postnatal growth patterns, made it impossible to match age and body size at time of test between groups. Distribution of test age for preterm controls and RDS centred on 36 w postmenstrual age pre-discharge and around 43 w postmenstrual age for follow up studies, whereas that for the CLD infants was more scattered, reflecting the age at which they became clinically stable enough to study (Figures OLS 2). Despite this overlap, the preterm groups were less mature, shorter, and lighter at time of test when compared with term controls (Table 2 main text, Figures OLS 2 and OLS 3).

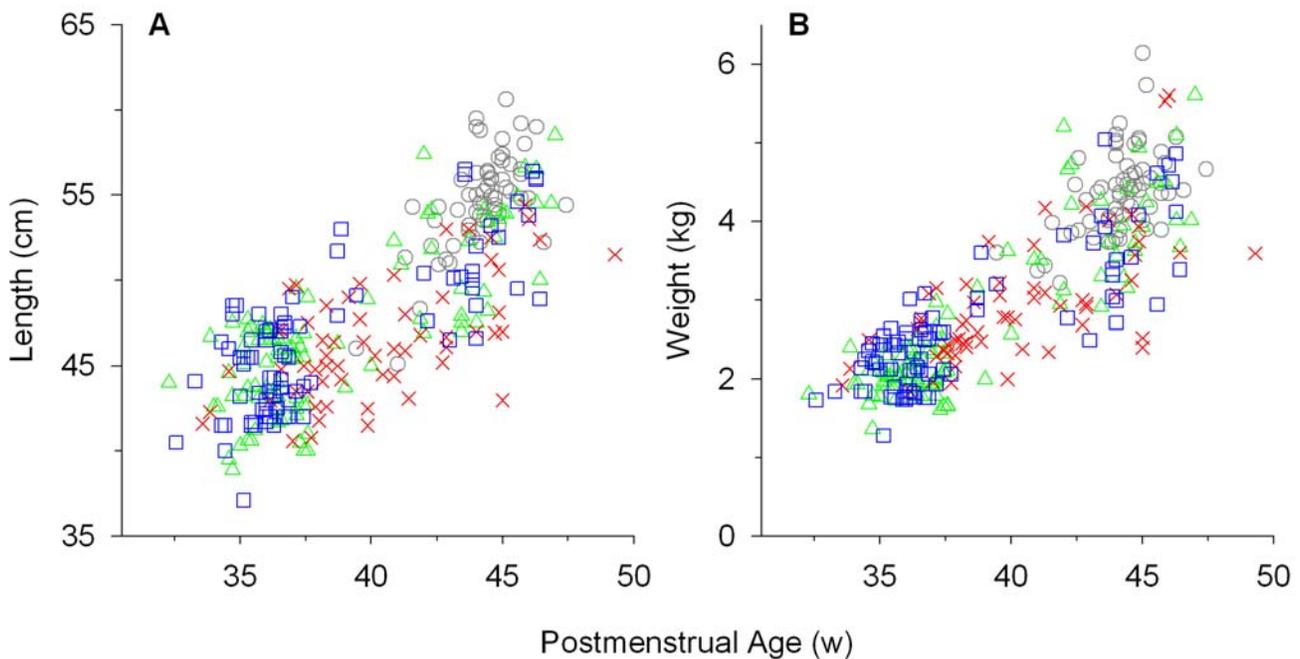


Figure OLS 2: Body size relative to postmenstrual age on all test occasions: A) Length and B) Weight.

Legend: ○ – FTC, fullterm controls; ▲ – PTC, preterm controls; □ – RDS, respiratory distress syndrome; × – CLD, chronic lung disease. Distribution of test age for preterm controls and RDS centred on 36 w pre-discharge and around 43 w for follow up studies, whereas that for the CLD infants was more scattered, reflecting the age at which they became clinically stable enough to study.

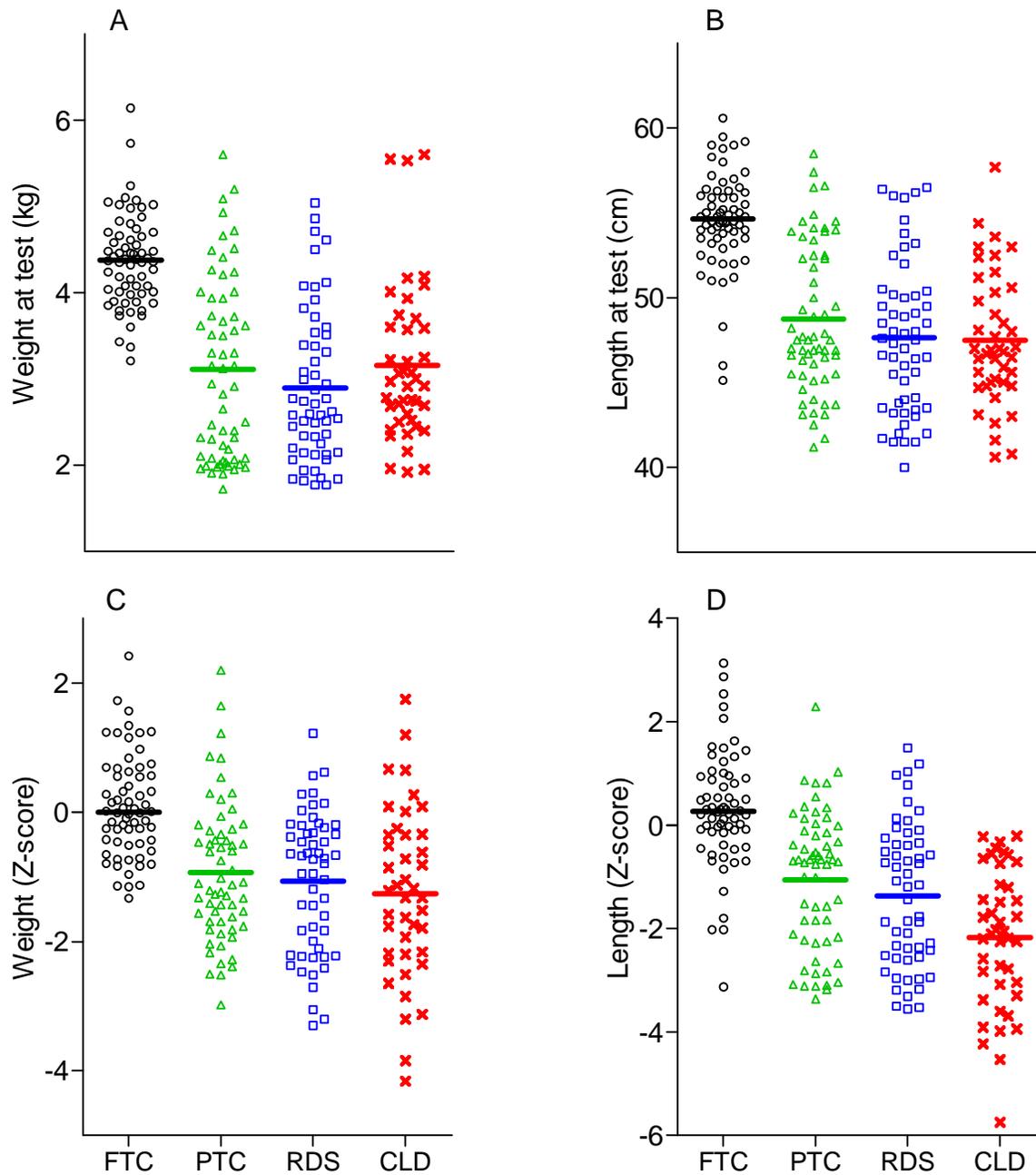


Figure OLS 3: Details at time of test according to diagnostic category

Legend: ○ – FTC, fullterm controls; △ – PTC, preterm controls; □ – RDS, respiratory distress syndrome; × – CLD, chronic lung disease.

There was evidence of postnatal growth restriction among all the preterm groups, in whom weight and length Z-scores were significantly lower than in term infants (Figures OLS 3C and OLS 3D). This discrepancy was particularly marked in those with CLD in whom length Z-score was significantly lower than in either the PTC [95% CI: -1.8, -0.4] or RDS [-1.5; -0.1] groups (Figure OLS 3D), indicating disproportionate growth in the CLD group who were very short for their weight. Consequently, BMI was higher in CLD infants at time of test compared with PTC [95% CI: 0.1, 2.0 kg·m⁻²] and RDS groups [95% CI: 0.3; 2.3 kg·m⁻²].

Lung Function Results:

Summary of univariable analysis

On univariable regression analysis, known determinants such as current body size (weight, length and BMI), age at test (PMA) and potential determinants such as the degree of prematurity (GA) and intrauterine growth (birth weight, birth weight Z-score); severity of neonatal respiratory disease (duration of supplemental O₂, ventilatory support, administration of surfactant) and maternal smoking during pregnancy were all significantly associated with FRC, V_T and RR. By contrast, indices of ventilation inhomogeneity (LCI, M₁/M₀ and M₂/M₀) were not associated with either current age or body size, or with any other variables except disease severity (duration of oxygen requirements and total days of ventilation) for which there was a very small but significant contribution. Sex and ethnic group had no effect on any of the outcome variables, with the exception of respiratory rate, which was lower amongst Caucasian infants (see table OLS 1 for details).

Table OLS 1: Univariable analysis of potential determinants on lung function indices

	FRC	LCI	M ₁ /M ₀	M ₂ /M ₀	RR	V _T	MV
GA (w)	↑	-	-	-	↓	↑	↑
Birthweight (kg)	↑	-	-	-	↓	↑	↑
Birthweight Z-score	↑	-	-	-	↓	↑	↑
Boy	-	-	-	-	-	-	-
Duration of O ₂ requirement (d)	↓	↑	↑	↑	↑	↓	↓
Total ventilation (d)	↓	↑	↑	↑	↑	↓	↓
Maternal smoking during pregnancy	↓	-	-	-	↑	↓	↓
PMA (w)	↑	-	-	-	↓	↑	↑
Length (cm)	↑	-	-	-	↓	↑	↑
Weight (kg)	↑	-	-	-	↓	↑	↑
BMI (kg·m ⁻²)	↑	-	-	-	↓	↑	↑
Equipment 2 (used at HUH)	↑	↑	↑	↑	-	↑	↑
Ethnic group (White)	-	-	-	-	↓	-	↓

↑ denotes significant positive association per unit increase (significance level set at $p < 0.05$)

↓ denotes significant negative association per unit increase

- denotes non-significant association

HUH : Homerton University Hospital, London

Functional Residual Capacity

The higher absolute values of FRC in term controls than in any of the preterm groups seen in Figure 1 (main paper) are at least partially due to differences in body size at test (Table OLS 2, Figure OLS 3).

Correction for such differences simply by dividing FRC by weight (Figure 1B main text) is common

practice, but is inappropriate, due to the growth restriction among all the preterm groups, compounded by disproportional growth in those recovering from CLD, who were very short for their weight.

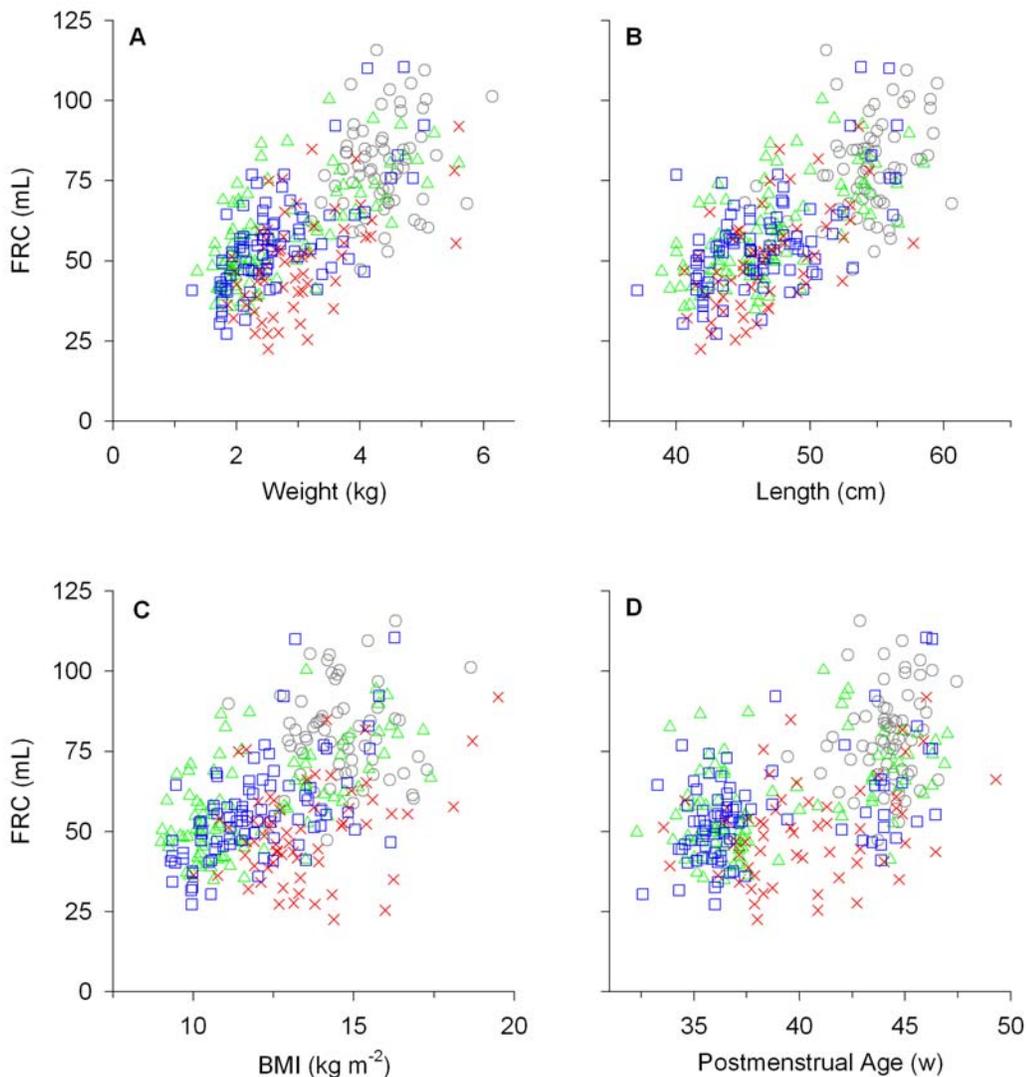


Figure OLS 4. Functional Residual Capacity according to diagnostic groups:

Data shown in relation to A) weight B) length, C) BMI, and D) postmenstrual age at time of test

Legend: ○ – FTC, fullterm controls; △ – PTC, preterm controls; □ – RDS, respiratory distress syndrome; × – CLD, chronic lung disease. Tests in preterm controls and RDS infants were clustered around 37 or at 44 w postmenstrual age (the latter being equivalent to 4 w corrected postnatal age, similar to term controls), whilst the postmenstrual age of CLD infants at test was more scattered. Clustering is less apparent when FRC is plotted against length.

Similarly, Figure OLS 4 D highlights the potential to misinterpret findings when comparisons are based on corrected age alone.

Multivariable, multilevel models

A high degree of collinearity between potential confounding factors influencing the outcome variables was expected. For example, postnatal age, weight and length at time of test are each predictive of lung function, but are also highly correlated with each other. Generally, once one of these variables was included, addition of the collinear factor did not account for additional variability in the data.

The relationship between FRC and both body size and age at test is shown in Figure OLS 4. Once adjusted for length and BMI, the factors PMA and weight at test no longer had an effect on FRC. After correcting for body size, the effect of being preterm resulted in a mean decrement of 0.67 mL (95 % CI: 0.24 mL, 1.10 mL; $p = 0.002$) per w GA, while intrauterine growth restriction was associated with a reduction of 2.1 mL (0.62 mL, 3.67 mL; $p = 0.006$) per reduction in BW Z-score. Additionally FRC was reduced by 0.03 mL (-0.06 mL, -0.01 mL; $p = 0.02$) per day of supplemental oxygen.

Although days of supplemental O₂ or ventilatory support made a similar contribution to the variability of FRC when examined individually, once supplementary O₂ had been added, there was no further contribution from days of ventilatory support. Days on supplemental O₂ was therefore retained in all models to reflect degree of disease severity. Once these factors had been taken into account, diagnostic category (FTC, PTC, RDS, CLD) had no additional effect.

The following worked examples show predicted FRC for two infants with different neonatal respiratory history.

Example 1: FRC for a term control infant

with test length of 54.6 cm; BMI 14.6; birth weight z-score -0.12 and GA 39.7 w was predicted to be:

$$\text{FRC} = -55.9 + [1.53 \times \text{length}] + [1.82 \times \text{BMI}] + [2.14 \times \text{birth weight Z-score}] + [0.67 \times \text{GA}]$$

$$\text{i.e FRC} = -55.9 + 83.5 (\text{length}) + 26.6 (\text{BMI}) - 0.26 (\text{birth weight Z-score}) + 26.6 (\text{GA}) =$$

80.5 mL

Example 2: FRC for a preterm infant

with neonatal lung disease with an identical length and BMI at test, but with a birth weight z-score: -0.59; GA: 26.8 w; and 135 d supplemental O₂, was predicted to be:

$$\text{FRC} = -55.9 + [1.53 \times \text{length}] + [1.82 \times \text{BMI}] + [2.14 \times \text{birth weight Z-score}] + [0.67 \times \text{GA}] - [0.03 \times \text{O}_2 \text{ d}]$$

$$\text{i.e. FRC} = -55.9 + 83.5 (\text{length}) + 26.6 (\text{BMI}) - 1.26 (\text{birth weight Z-score}) + 18.0 (\text{GA})$$

$$- 4.1 (\text{O}_2) = \mathbf{66.8 \text{ mL}}$$

After adjusting for these factors, centre differences were observed such that, when compared with data from HUH, measured values of FRC were significantly lower at UCLH and KEMH ($p < 0.0001$), by an average of 11.4 mL (95 % CI: -16.2 mL, -6.6 mL) and 7.8 mL (-11.3 mL, -4.3 mL) respectively.

Adding centre as a level in the model (random effect) gave similar relationships between FRC and the other factors and the between-centre variability was not significant. However, since centre was confounded with equipment used (UCLH and KEMH both using Equipment 1 while HUH used Equipment 2), a model that incorporated the three centres as factors was investigated. This approach was not significantly better than a model with only a binary variable representing the difference in equipment, suggesting that the centre differences was most likely due to equipment variation. Thus FRC data collected using equipment 2 (i.e. at HUH) were, on average, 8.6 mL higher compared to data

collected at UCLH and KEMH with Equipment 1 (Table OLS 1). There was no significant interaction between centre or equipment and any of the explanatory variables (length, BMI, Birthweight Z-score, gestational age and oxygen requirement). Neither was there significant interaction between equipment type and diagnostic group after accounting for the other variables, and the strength of the associations of each of the determinants on FRC (Table OLS 2) remained unchanged. The same approach (i.e. with centre entered as a level in the model (random effect)) was undertaken for subsequent analysis of the remaining dependent variables

Table OLS 2: FRC adjusted coefficients (with centre included as a level in the model)

	Coefficient	95 % CI	p-value
Length (cm)	1.63	1.20, 2.06	<0.0001
BMI (kg· m ⁻²)	1.57	0.72, 2.42	0.0003
Birth weight Z-score	1.94	0.49, 3.39	0.009
GA (w)	0.47	0.05, 0.89	0.029
Duration of O ₂ (d)	-0.04	-0.07, -0.02	0.001
Equipment 2	8.64	5.28, 11.99	<0.0001

Coefficients adjusted for repeat measurements (see text for details);

Equipment 1 used at UCLH and KEMH; Equipment 2: used at HUH.

Ventilation Inhomogeneity*Lung Clearance Index (LCI):*

There was a significant association between LCI and days of supplemental O₂ (Table OLS 3). Once supplementary O₂ had been taken into account, addition of diagnostic category showed no additional effect on LCI. According to analysis strategy, GA was retained in the model but was not found to be associated with LCI. As for FRC, there were equipment differences, with LCI data from KEMH and UCLH (equipment 1) showing, on average, a reduction of 0.34 compared with data collected at HUH (equipment 2).

Table OLS 3: LCI adjusted coefficients

	Coefficient	95 % CI	p-value
Duration O ₂ (d)	0.003	0.001, 0.005	0.003
GA (w)	0.007	-0.017, 0.031	0.660
Equipment 2	0.34	0.112, 0.560	0.002

Coefficients adjusted for repeat measurements (see text for details);

Equipment 1: used at UCLH and KEMH; Equipment 2: used at HUH.

The relationship between LCI and days of supplemental O₂ is shown in Figure OLS 5. Similar patterns were observed for M₁/ M₀ and M₂/ M₀ (data not shown).

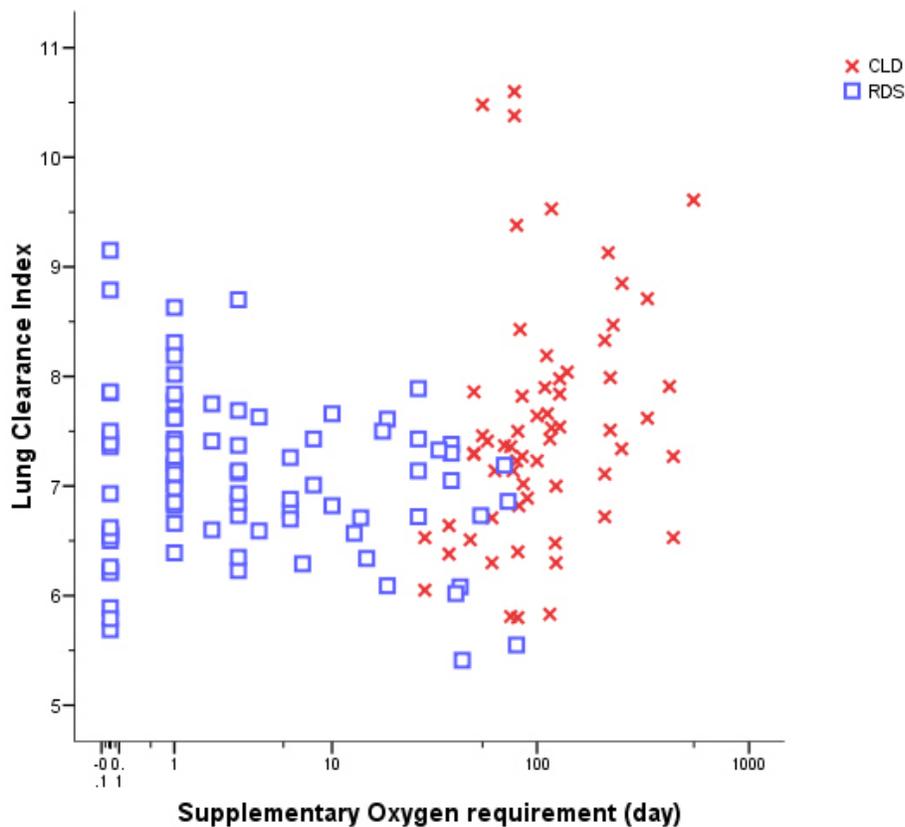


Figure OLS 5 Lung clearance index plotted against supplementary oxygen requirement in infants with RDS and CLD.

Note log scale on the x-axis.

Moment Ratios:

As for LCI, disease severity (O₂ days) was significantly associated with M₁/M₀ (Table OLS 4).

However this difference was very small and unlikely to be of clinical significance. Although moment ratios appeared higher among those with CLD (Figure OLS 6), diagnostic category did not contribute significantly to the model after accounting for GA and days of supplemental oxygen.

Table OLS 4: 1st Moment Ratio (M_1/M_0) adjusted coefficients

	Coefficient	95 % CI	p-value
Duration O ₂ (d)	0.001	0.001, 0.001	<0.0001
GA (w)	0.005	-0.003, 0.013	0.211
Equipment 2	0.122	0.055, 0.189	0.0003

Coefficients adjusted for repeat measurements (see text for details);

Equipment 1 used at UCLH and KEMH; Equipment 2: used at HUH.

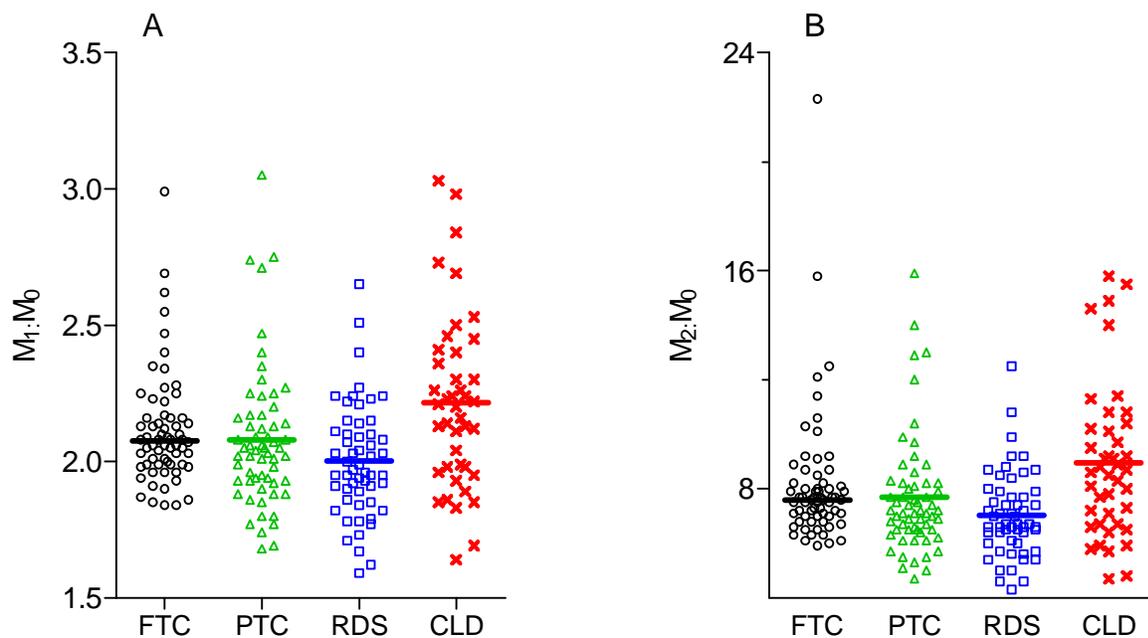


Figure OLS 6: Moment Ratios as a measure of Ventilation Inhomogeneity according to diagnostic groups.

Legend: ○ – FTC, fullterm controls; △ – PTC, preterm controls; □ – RDS, respiratory distress syndrome; × – CLD, chronic lung disease.

Similar equipment differences were seen as for LCI, with HUH data (equipment 2) being significantly higher than those from UCLH and KEMH (Table OLS 3). This did not however, influence the relative contributions of supplemental oxygen and GA.

Results for multivariable and multilevel modelling for M_2/M_0 were similar to those for M_1/M_0 , as shown in Table OLS 5 and Figure OLS 6.

Table OLS 5: 2nd Moment Ratio (M_2/M_0) adjusted coefficients

	coefficient	95 % CI	p-value
Duration O ₂ (d)	0.009	0.005, 0.013	<0.0001
GA (w)	0.057	0.033, -0.008	0.08
Boy	-0.594	-1.133, -0.055	0.03
Equipment 2	0.963	0.357, 1.569	0.002

Coefficients adjusted for repeat measurements (see text for details).

Equipment 1 used at UCLH and KEMH; Equipment 2: used at HUH.

Respiratory Rate

As expected, RR decreased with increasing age (by approximately 2 breaths per w advancing age) and was also slightly lower in boys and in those with better intra-uterine growth, i.e. higher birth weight Z-score (Table OLS 6). Higher BMI, as was seen in infants with CLD and disproportionately short length for weight, was associated with an increase in RR. Duration of supplemental O₂ was also associated with significantly increased rates, although the impact was relatively small; administration of 100 d supplemental O₂ only being associated with an increase of 3.2 breaths/min after accounting for other factors. Neither GA nor diagnostic category contributed to the model, and there were no equipment effects.

Table OLS 6: RR adjusted coefficients

	Coefficient	95 % CI	p-value
PMA (w)	-1.875	-2.34, -1.37	<0.0001
BMI (kg· m ⁻²)	1.928	0.95, 2.91	0.0001
Birth weight Z-score	-2.712	-4.27, -1.15	0.0007
Duration O ₂ (d)	0.032	0.005, 0.059	0.022
GA (w)	-0.287	-0.689, 0.115	0.162
Boy	-3.6	-6.9, -0.3	0.03

Coefficients adjusted for repeat measurements (see text for details)

Tidal Volume

The major determinants of V_T were age and body size at test, with gestational age and intra-uterine growth also showing a positive association (Table OLS 7). Higher BMI was associated with a significant reduction in V_T , which again could reflect disproportionate growth patterns in some of these children. Supplementary oxygen had a weak and non-significant association with V_T , but was kept in the model according to analysis strategy.

Table OLS 7: V_T adjusted coefficients

	Coefficient	95 % CI	p-value
PMA (w)	0.68	0.45, 0.91	<0.0001
Weight (kg)	4.75	3.19, 6.32	<0.0001
BMI (kg/m ²)	-0.85	-1.30, -0.39	0.0003
GA (w)	0.20	0.06, 0.33	0.004
BW Z-score	0.76	0.20, 1.33	0.008
Duration O ₂ (d)	-0.01	-0.018, -0.002	0.012
Equipment 2	2.61	1.57, 3.66	<0.0001

Coefficients adjusted for repeat measurements (see text for details);

Equipment 1 used at UCLH and KEMH; Equipment 2: used at HUH.

Equipment effects were seen with V_T data from HUH using equipment 2 being significantly higher by an average of 2.6 mL (95 % CI: 1.6 mL, 3.7 mL; $p < 0.0001$) compared to KEMH and UCLH data.

Minute Volume

Once weight and gestation were added to the model, none of the other variables, which were significantly associated with minute volume during univariable analysis, contributed to the model (Table OLS 8). Similarly, there was only a weak association between supplementary oxygen and minute volume.

Table OLS 8: Minute Volume (mL) adjusted coefficients

	coefficient	95 % CI	p-value
Weight (kg)	204	173, 236	<0.0001
GA (w)	5.8	-2.8, 14.5	0.19
Duration O ₂ (d)	0.16	-0.37, 0.68	0.56
Equipment 2	95.8	25.4, 166.1	0.008

Coefficients adjusted for repeat measurements (see text for details);

Equipment 1 used at UCLH and KEMH; Equipment 2: used at HUH.

The inter-centre differences observed in this study were unexpected, since the same protocol, and supposedly the same investigators at all three sites had used the same equipment and software. These differences persisted after accounting for any differences in population (ethnic group, diagnostic category) or characteristics at time of test. An identical protocol for calibration and data collection was used in all three centres and we discounted inter-observer variability as an explanation for any discrepancies after examining and cross analysing a large subset of results from each centre. The equipment used at UCLH was subsequently transferred to Australia where it was used for data collection at KEMH, which would explain similarities between these two centres. When data were

analysed using a two-level approach and clustered on centre level first and then on the individual level, with equipment set introduced as a binary variable, we found significant differences in outcome variables depending on which device had been used. We can only assume there must have been some subtle differences in either hardware or software between the two devices that were undetectable until this very large data set was collated and subjected to multivariable analysis. Fortunately, each diagnostic group was studied in all three centres and sample size was sufficiently large to allow us to model these inter-centre differences, which did not affect the overall conclusions of the study. However, such differences could have confounded a smaller study or one relying on normative data collected elsewhere, and highlights the need to recruit a control group in similar studies.

REFERENCES

1. Pillow JJ, Ljungberg H, Hulskamp G *et al.* Functional residual capacity measurements in healthy infants: ultrasonic flow meter versus a mass spectrometer. *Eur Respir J* 2004; **23**: 768.
2. Shennan AT, Dunn MS, Ohlsson A *et al.* Abnormal pulmonary outcomes in premature infants: prediction from oxygen requirement in the neonatal period. *Pediatrics* 1988; **82**: 527-532.
3. Charafeddine L, D'Angio CT, Phelps DL. Atypical chronic lung disease patterns in neonates. *Pediatrics* 1999; **103**: 759-765.
4. Van Marter LJ, Allred EN, Pagano M *et al.* Do clinical markers of barotrauma and oxygen toxicity explain interhospital variation in rates of chronic lung disease? The Neonatology Committee for the Developmental Network. *Pediatrics* 2000; **105**: 1194-1201.
5. Hillman NH, Moss TJ, Kallapur SG *et al.* Brief, large tidal volume ventilation initiates lung injury and a systemic response in fetal sheep. *Am J Respir Crit Care Med* 2007; **176**: 575-581.
6. Jobe AH, Kramer BW, Moss TJ *et al.* Decreased indicators of lung injury with continuous positive expiratory pressure in preterm lambs. *Pediatr Res* 2002; **52**: 387-392.
7. Jobe AH, Bancalari E. Bronchopulmonary dysplasia. *Am J Respir Crit Care Med* 2001; **163**: 1723-1729.
8. Stocks J. Effect of nasogastric tubes on nasal resistance during infancy. *Arch Dis Child* 1980; **55**: 17-21.

9. Schibler A, Hall GL, Businger F *et al.* Measurement of lung volume and ventilation distribution with an ultrasonic flow meter in healthy infants. *Eur Respir J* 2002; **20**: 912-918.
10. Frey U, Stocks J, Sly P *et al.* Specifications for signal processing and data handling used for infant pulmonary function testing. *Eur Respir J* 2000; **16**: 1016-1022.
11. Schibler A, Henning R. Measurement of functional residual capacity in rabbits and children using an ultrasonic flow meter. *Pediatr Res* 2001; **49**: 581-588.
12. Latzin P, Sauteur L, Thamrin C *et al.* Optimized temperature and deadspace correction improve analysis of multiple breath washout measurements by ultrasonic flowmeter in infants. *Pediatr Pulmonol* 2007; **42**: 888-897.
13. Frey U, Stocks J, Coates A *et al.* Standards for infant respiratory function testing: Specifications for equipment used for infant pulmonary function testing. *Eur Respir J* 2000; **16**: 731-740.
14. Larsson A, Jonmarker C, Werner O. Ventilation inhomogeneity during controlled ventilation. Which index should be used? *J Appl Physiol* 1988; **65**: 2030-2039.
15. Saidel GM, Saniie J, Chester EH. Lung washout during spontaneous breathing: parameter estimation with a time-varying model. *Comput Biomed Res* 1980; **13**: 446-457.
16. Hulskamp G, Hoo AF, Ljungberg H *et al.* Progressive decline in plethysmographic lung volumes in infants: physiology or technology? *Am J Respir Crit Care Med* 2003; **168**: 1003-1009.
17. Ranganathan S, Dezateux CA, Bush A *et al.* Airway function in infants newly diagnosed with cystic fibrosis. *Lancet* 2001; **358**: 1964-1965.

18. Lum S, Hoo AF, Dezateux C *et al.* The association between birthweight, sex, and airway function in infants of nonsmoking mothers. *Am J Respir Crit Care Med* 2001; **164**: 2078-2084.
19. Cohen J. *Multiple Regression and Correlation Analysis. Statistical Power Analysis for the Behavioral Sciences.* 2nd ed. Lawrence Erlbaum Associates, 1998: 407.
20. Edwards LJ. Modern statistical techniques for the analysis of longitudinal data in biomedical research. *Pediatr Pulmonol* 2000; **30**: 330-344.
21. Schluchter MD, Konstan MW, Drumm ML *et al.* Classifying severity of cystic fibrosis lung disease using longitudinal pulmonary function data. *Am J Respir Crit Care Med* 2006; **174**: 780-786.
22. Alexander FE, Anderson TJ, Brown HK *et al.* 14 years of follow-up from the Edinburgh randomised trial of breast-cancer screening. *Lancet* 1999; **353**: 1903-1908.
23. Brown HK, Prescott RJ. *Applied Mixed Models in Medicine.* John Wiley and sons, 1999:
24. Canoy D, Pekkanen J, Elliott P *et al.* Early growth and adult respiratory function in men and women followed from the fetal period to adulthood. *Thorax* 2007; **62**: 396-402.
25. Lipsett J, Tamblyn M, Madigan K *et al.* Restricted fetal growth and lung development: a morphometric analysis of pulmonary structure. *Pediatr Pulmonol* 2006; **41**: 1138-1145.