



ORIGINAL ARTICLE

Effect of late preterm birth on longitudinal lung spirometry in school age children and adolescents

Sarah J Kotecha,¹ W John Watkins,² Shantini Paranjothy,² Frank D Dunstan,²
A John Henderson,³ Sailesh Kotecha¹

► Additional materials are published online only. To view these files please visit the journal online (<http://thorax.bmj.com/content/67/1.toc>).

¹Department of Child Health, School of Medicine, Cardiff University, Cardiff, UK

²Department of Primary Care and Public Health, School of Medicine, Cardiff University, Cardiff, UK

³School of Social and Community Medicine, University of Bristol, Bristol, UK

Correspondence to

Professor Sailesh Kotecha, Department of Child Health, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4XN, UK; kotechas@cardiff.ac.uk

Received 12 April 2011

Accepted 29 July 2011

Published Online First
27 September 2011

ABSTRACT

Background Rates of preterm birth have increased in most industrialised countries but data on later lung function of late preterm births are limited. A study was undertaken to compare lung function at 8–9 and 14–17 years in children born late preterm (33–34 and 35–36 weeks gestation) with children of similar age born at term (≥ 37 weeks gestation). Children born at 25–32 weeks gestation were also compared with children born at term.

Methods All births from the Avon Longitudinal Study of Parents and Children ($n=14\,049$) who had lung spirometry at 8–9 years of age ($n=6705$) and/or 14–17 years of age ($n=4508$) were divided into four gestation groups.

Results At 8–9 years of age, all spirometry measures were lower in the 33–34-week gestation group than in controls born at term but were similar to the spirometry decrements observed in the 25–32-week gestation group. The 35–36-week gestation group and term group had similar values. In the late preterm group, at 14–17 years of age forced expiratory volume in 1 s (FEV_1) and forced vital capacity (FVC) were not significantly different from the term group but FEV_1/FVC and forced expiratory flow at 25–75% FVC ($FEF_{25-75\%}$) remained significantly lower than term controls. Children requiring mechanical ventilation in infancy at 25–32 and 33–34 weeks gestation had in general lower airway function (FEV_1 and $FEF_{25-75\%}$) at both ages than those not ventilated in infancy.

Conclusions Children born at 33–34 weeks gestation have significantly lower lung function values at 8–9 years of age, similar to decrements observed in the 25–32-week group, although some improvements were noted by 14–17 years of age.

INTRODUCTION

In most industrialised countries the rate of preterm birth has risen markedly in recent years and late preterm births form a substantial proportion of these. In the USA the rate has increased from 9.5% in 1981 to 12.7% in 2005, with 20% of the preterm births occurring at 32–33 weeks gestation and 60–70% at 34–36 weeks gestation.¹ Therefore, even a small increased risk of later respiratory morbidity in infants born beyond 32 weeks gestation will contribute substantially to the number of individuals affected by respiratory disease associated with preterm birth. A workshop convened by the National Institute of Child Health and Human Development recognised the importance of infants

Key messages

What is the key question?

► We sought to compare lung function at 8–9 and 14–17 years in children born late preterm (33–34 and 35–36 weeks gestation) with children of similar age born at term (≥ 37 weeks gestation). We also compared children born at 25–32 weeks gestation with children born at term.

What is the bottom line?

► Children born at 33–34 weeks gestation have significantly lower lung function values at 8–9 years of age, similar to decrements observed in the 25–32-week group, although most of these differences were reduced by 14–17 years of age.

Why read on?

► The findings from this longitudinal population-based observational cohort study suggest that children born at 33–34 weeks gestation may be at risk of decreased lung function at 8–9 years of age. By 14–17 years of age there were improvements in FEV_1 .

born beyond 32 weeks gestation but identified the lack of outcome data for this population.² It is, however, increasingly recognised that late preterm infants have increased respiratory morbidity in the neonatal period, especially neonatal respiratory distress syndrome, transient tachypnoea of the newborn and pneumonia.³ Few studies have reported later respiratory outcomes in children born late preterm,^{4–5} with most studies on lung function outcomes following preterm birth focusing on extremely preterm infants of 32 weeks gestation or less, especially as they are at risk of developing chronic lung disease of prematurity (CLD).⁶

Prematurity is associated with delivery of an infant at early stages of lung development with the focus traditionally having been on those born at 32 weeks gestation or less.⁷ As these infants are born at the late canalicular and early sacular stages of lung development, they often require intervention especially to support their respiratory system. However, those born ‘late’ preterm (ie, 33–34 weeks gestation) are also at a vulnerable stage of lung development (the late sacular stage),⁸ and it is unknown if normal lung growth and

development continues normally in the extrauterine environment after preterm birth. Infants born at 35–36 weeks gestation are often termed 'near term' infants (eg, in the ECMO and Nitric Oxide clinical trials^{9–11}) and are often assumed to behave like term infants, but it is unclear if being born at an early stage of alveolar development also results in adverse longer term respiratory outcomes.

In this study we sought to compare lung function at 8–9 and 14–17 years in a population-based cohort of children born late preterm at 33–34 weeks gestation and near term at 35–36 weeks gestation with children of similar age born at term (≥ 37 weeks gestation). We also compared the results for these two gestation groups with children of similar age born extremely preterm at 25–32 weeks gestation. The longitudinal change in lung function between 8–9 and 14–17 years of age is also reported.

METHODS

Avon Longitudinal Study of Parents and Children (ALSPAC)

The study used data from the Avon Longitudinal Study of Parents and Children (ALSPAC) which has been described previously.¹² Additional detail is provided in the online supplement.

Gestation groups and perinatal data

Gestational age was based on maternal reporting and medical assessments.¹³ All children who had spirometry results at either 8–9 or 14–17 years of age were divided into four groups based on gestational age: 25–32 weeks, 33–34 weeks, 35–36 weeks and ≥ 37 –43 weeks (term control). Perinatal data on respiratory management were extracted from hospital records. Additional detail is provided in the online supplement.

Self-reported respiratory symptoms and tobacco smoke exposure

As part of ALSPAC studies, questionnaires at 7.6 years of age established if children 'ever had asthma diagnosed by a doctor together with symptoms and/or treatment in the previous 12 months'. Questionnaires at 13.8 years of age established if the children 'ever had asthma diagnosed by a doctor'. Regular smoking at age 14 was defined as the usual smoking of at least one cigarette per day or at least six cigarettes per week.

Lung spirometry

Spirometry was performed at 8–9 and 14–17 years of age as previously reported.¹⁴ The variables forced expiratory volume in 1 s (FEV₁), forced vital capacity (FVC) and forced expiratory flow at 25–75% of FVC (FEF_{25–75}) were recorded and transformed to SD scores adjusted for age, sex and height according to the method of Chinn *et al* in the ALSPAC cohort independently at 8–9 and 14–17 years of age.¹⁵ FEV₁/FVC and FEF_{25–75}/FVC were calculated as markers of airway obstruction and dysanapsis.¹⁶

Bronchial challenge

At 8–9 years of age children, unselected for asthma or wheezing, had a bronchial challenge test with methacholine using the method of Yan *et al*,¹⁷ as previously described.¹⁴ Additional detail is given in the online supplement.

Statistical analysis

Gestations of ≥ 44 weeks were considered implausible and thus excluded. The birth weight z-scores were calculated using the lms Growth program (Medical Research Council, UK) to take into account gestation, as the numbers in the earlier gestation groups were small and the birth weights of the ALSPAC cohort were

similar to published reference values (data not shown). Weight at the time of lung function testing was converted into a z-score within the ALSPAC cohort as the ALSPAC data showed a trend towards earlier growth (data not shown). Spirometry values are summarised using the mean with 95% CIs. For longitudinal changes in spirometry, only data from children who had spirometry at both time points were included in the analyses. To account for confounders the spirometry data, already adjusted for age, gender and height, were further adjusted using general linear modelling for smoking during pregnancy and social class status based on father's occupation, which are well-described confounders of lung function and preterm delivery. Missing values of each confounder were coded into a separate missing category. Adjusted scores were calculated and used in subsequent analyses. Longitudinal changes were assessed using paired t tests on the adjusted spirometry values. Differences between different gestational age groups and the term group and between ventilated and non-ventilated infants were tested using independent samples t tests. The results for ventilated and non-ventilated infants are given unadjusted for social class and smoking during pregnancy but are adjusted for age, gender and height. Bronchial hyper-responsiveness was compared between gestation groups by one-way analysis of variance. χ^2 Tests were used to compare asthma rates between gestation groups at 7.6 and 13.8 years of age. p values < 0.05 were considered statistically significant. All analyses were performed using SPSS V.16 (SPSS Inc).

RESULTS

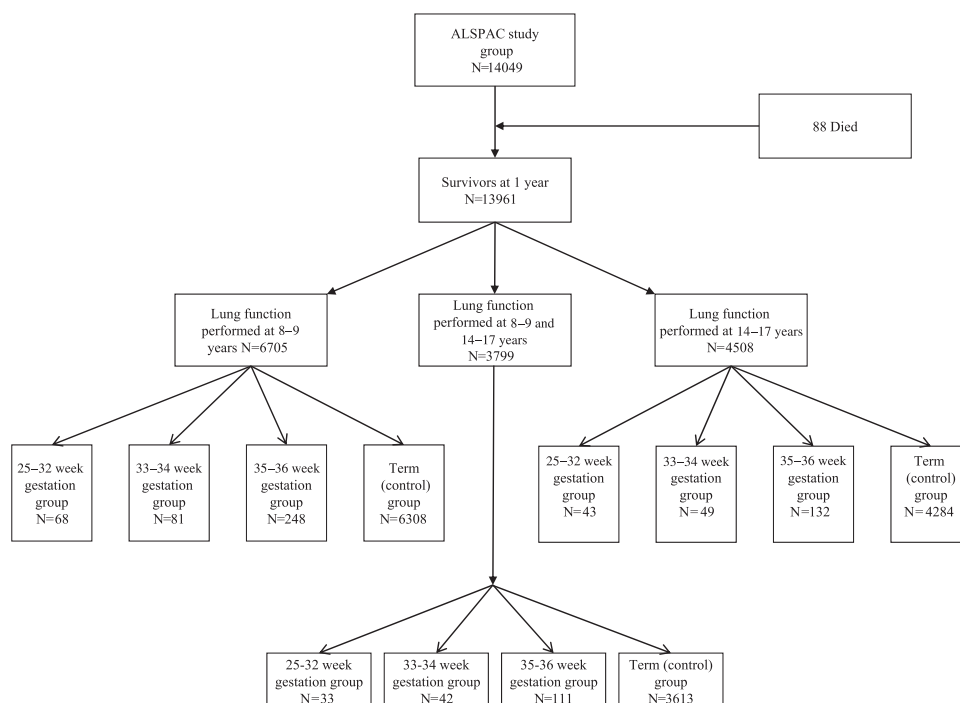
From 14 049 live births, spirometry data were available for 6705 (47.7%) children at 8–9 years of age and 4508 (32.1%) at 14–17 years of age, with measurements at both ages available for 3799 (27.0%) children (figure 1). Tables 1 and 2 show the characteristics of the children in each gestation group. A higher percentage of children in the 25–32-week gestation group were from lower social classes and had a higher prevalence of maternal smoking. A higher percentage of children who did not attend for spirometry were from a lower social class and had higher exposure to maternal smoking (table 3). 76% of the extremely preterm group and 17% of the late preterm group required mechanical ventilation but few ($\leq 1\%$) did so in the more mature groups (table 2).

Asthma in the previous year at 7.6 years of age was reported by more than 20% in the 25–32-week and 33–34-week gestation groups but by $< 15\%$ in the 35–36-week gestation and term groups, although these differences were not statistically significant. The rates of asthma ever at 13.8 years of age were similar (23–30%) in all groups. No differences were noted for smoking at 14 years of age between the gestational groups. Of the children who had any lung spirometry, 104 children reported regular smoking at 14 years of age (3 in the 35–36-week gestation group, 101 in the term group and none in those born at < 35 weeks gestation).

Association of preterm birth with lung function in later childhood

Spirometry data by gestation group at 8–9 and 14–17 years of age are shown in tables 4 and 5, respectively. At 8–9 years of age, all spirometry measures were significantly lower in the 33–34-week gestation group compared with the term group whereas values in the 35–36-week gestation group were similar to term infants. In the 25–32-week gestation group, all spirometry measures except FVC were significantly lower compared with the term group. The decrements of lung function were similar (table 4) between the 25–32-week and the 33–34-week

Figure 1 Study profile: number of children excluded from group analysed. ALSPAC, Avon Longitudinal Study of Parent and Children.



gestation groups. The approximate differences compared with term infants after adjustment for age, gender and height were -143 ml for FEV_1 , -98 ml for FVC and -253 ml/s for FEF_{25-75} in the 33–34-week gestation group and -121 ml for FEV_1 , -65 ml for FVC and -297 ml/s for FEF_{25-75} in the 25–32-week gestation group.

At 14–17 years of age, FEF_{25-75} , FEF_{25-75}/FVC and FEV_1/FVC remained significantly lower in the late preterm group than in the term group but FEV_1 and FVC were not significantly different. The approximate differences from those born at term after adjustment for age, gender and height were -20 ml for FEV_1 , 156 ml for FVC and -382 ml/s for FEF_{25-75} , in the 33–34-

Table 1 Characteristics of gestation groups

	25–32 weeks gestation	33–34 weeks gestation	35–36 weeks gestation	Term
Total liveborn	185	169	522	13148
Survivors at 1 year	160	165	519	13117
Had spirometry at 8–9 years	68 (43%)	81 (49%)	248 (48%)	6308 (48%)
Had spirometry at 14–17 years	43 (27%)	49 (30%)	132 (25%)	4284 (33%)
Had spirometry at 8–9 years and 14–17 years	33 (21%)	42 (25%)	111 (21%)	3613 (28%)
Total number in group who had any spirometry	78	88	269	6979
Male (%)	48 (62%)	54 (61%)	144 (54%)	3446 (49%)
Maternal smoking (%)	21/74 (28%)	20/84 (24%)	49/264 (19%)	1282/6826 (19%)
Manual workers (%)	27/54 (50%)	32/77 (42%)	108/255 (42%)	2439/6431 (38%)
Exclusively breast fed for at least 3 months (%)	6/75 (8%)	12/84 (14%)	25/266 (9%)	819/6834 (12%)
At spirometry at 8–9 years of age				
Mean (SD) age (years)	8.7 (0.4)	8.6 (0.3)	8.7 (0.3)	8.6 (0.3)
Mean (SD) height (cm)	130.9 (6.8)	130.9 (5.3)	132.8 (5.1)	132.4 (5.8)
Mean (SD) height (z-score)	-0.336 (1.225)	-0.252 (0.945)	0.036 (0.883)	0.011 (1.002)
Mean (SD) weight (kg)	28.1 (6.4)	28.3 (4.2)	30.4 (6.0)	30.3 (5.8)
Mean (SD) weight (z-score)	-0.439 (1.107)	-0.363 (0.738)	-0.046 (0.988)	-0.033 (0.990)
At spirometry at 14–17 years of age				
Mean (SD) age (years)	15.5 (0.2)	15.5 (0.4)	15.5 (0.3)	15.5 (0.3)
Mean (SD) height (cm)	166.7 (9.0)	168.2 (7.6)	170.0 (8.0)	169.2 (8.4)
Mean (SD) height (z-score)	-0.505 (1.237)	-0.202 (0.972)	0.028 (0.933)	0.007 (0.998)
Mean (SD) weight (kg)	56.3 (9.4)	60.6 (10.3)	60.6 (11.8)	61.6 (12.0)
Mean weight (z-score, SD)	-0.468 (0.898)	-0.090 (0.840)	-0.101 (1.006)	0.011 (1.001)
For children who had any spirometry				
Current asthma at 7.6 years of age based on reported doctor-diagnosed asthma ever with symptoms and/or treatment in previous 12 months (%)	11/52 (21%)	13/58 (22%)	22/182 (12%)	738/5221 (14%)
Current asthma at 13.8 years of age based on doctor-diagnosed asthma ever (%)	13/50 (26%)	17/57 (30%)	48/186 (26%)	1197/5227 (23%)

Some of the variables reported had missing data but the denominators are given.

Table 2 Neonatal characteristics of gestation groups

	25–32 weeks gestation	33–34 weeks gestation	35–36 weeks gestation	Term
Total liveborn	185	169	522	13148
Survivors at 1 year	160	165	519	13117
Total number in group who had any spirometry	78	88	269	6979
Mean (SD) birth weight of those who had any spirometry (g)	1437 (439.5)	2449 (416.8)	2634 (424.7)	3478 (476.6)
Mean (SD) birth weight of those who had any spirometry (z-score)	0.051 (1.297)	0.396 (1.267)	0.180 (1.066)	0.117 (0.976)
Mean (SD) gestation of those who had any spirometry (weeks)	29.7 (2.0)	33.6 (0.5)	35.7 (0.5)	39.7 (1.3)
Ventilated, n/N (%) [*]	59/78 (76)	15/88 (17)	2/269 (1)	12/6979 (0.2)
Mean (range) length of time of ventilation (h)	245 (1–1416)	77 (9–168)	15 (8–22)	107 (1–576)
Type of ventilation, n/N (%) [†] [*]				
CPAP	43/78 (55.1)	6/88 (6.8)	2/269 (0.7)	3/6979 (0.0)
Ventilated via ETT	41/78 (52.6)	7/88 (8.0)	0/269 (0)	11/6979 (0.2)
Other	21/78 (26.9)	6/88 (6.8)	1/269 (0.4)	5/6979 (0.1)
Surfactant treatment, n/N (%) [*]	35/78 (44.9)	6/88 (6.8)	0/269 (0)	1/6979 (0.01)
Received oxygen at any time, n/N (%) [*]	66/78 (85)	31/88 (35)	16/269 (6)	45/6979 (0.6)
Number (%) receiving oxygen treatment for: [*]				
<1 day	17/78 (21.8)	13/88 (14.8)	8/269 (3.0)	31/6986 (0.4)
2–4 days	11/78 (14.1)	8/88 (9.1)	7/269 (2.6)	9/6986 (0.1)
≥5 days	38/78 (48.7)	10/88 (11.4)	0/269 (0)	4/6986 (0.1)

Some of the variables reported had missing data.

^{*}It was assumed the answer was no for children who did not have a neonatal admissions questionnaire completed, hence percentages are calculated based on the total number in the group who had any spirometry, not the number who had neonatal admissions questionnaires completed.

[†]Some children may have received more than one mode of ventilation.

CPAP, continuous positive airway pressure; ETT, endotracheal tube.

week gestation group and –114 ml for FEV₁, –120 ml for FVC and –428 ml/s for FEF_{25–75} in the 25–32-week gestation group.

Changes in spirometry between 8–9 years and 14–17 years of age

Table 6 shows changes in spirometry for children with measurements at both time points. For the 33–34-week gestation group, significant increases were observed for FEV₁. In contrast, no significant changes were noted between the time periods for the 25–32-week, 35–36-week and term groups.

Effect of ventilation in the newborn period on later lung function

Table E1 (in the online supplement) shows spirometry for the extremely and late preterm children according to ventilation requirements in the newborn period. In general, children who

were ventilated in infancy had lower airway function (FEV₁ and FEF_{25–75}) at both ages compared with those not ventilated in infancy. More results are shown in the online supplement.

Bronchial responsiveness

No significant differences were found between the different gestation groups for the geometric mean provocative dose of methacholine causing a fall in FEV₁ of 20% or more (PD₂₀) in either responders or the whole study population (table 7).

DISCUSSION

This study shows that most lung function variables measured by spirometry at 8–9 years of age were significantly lower in infants born at 33–34 weeks gestation compared with term infants, with the effect sizes being very similar to those in the 25–32-week gestation group. By 14–17 years of age the differences in FEV₁ and FVC attenuated but differences with term controls persisted for FEF_{25–75%}, FEV₁/FVC and FEF_{25–75%}/FVC in the 33–34-week gestation group. In contrast, children born at 35–36 weeks had comparable lung function with children born at term at both follow-ups. While the numbers were small, the indications were that those in the 33–34-week gestation group who required mechanical ventilation fared worse than those who did not.

Late preterm infants of 33–34 weeks gestation are a growing population which is increasingly recognised to be at risk of morbidity in the neonatal period from respiratory diseases including respiratory distress syndrome,¹⁸ and in infancy from

Table 3 Characteristics of attenders and non-attenders

	Children who had any spirometry	Children who had no spirometry and were alive at 1 year
Number	7414	6547
Mean (SD) birth weight (g)	3412 (554)	3369 (565)
Mean (SD) gestation (weeks)	39.4 (1.9)	39.4 (1.9)
Male, N (%)	3692 (49.8)	3520 (53.8)
Maternal smoking, N (%)	1372/7248 (18.9)	2033/5896 (34.5)
Manual workers, N (%)	2606/6817 (38.2)	2407/4681 (51.4)
Exclusively breast fed for at least 3 months, N (%)	862/7259 (11.9)	470/5408 (8.7)

Table 4 Mean and SD of z-score of lung function measures adjusted for age, gender and height at age 8–9 years, with CIs for the difference of means (premature – term), unadjusted and adjusted for maternal smoking in pregnancy and social status

	25–32 weeks (n=65)		Term (n=6144)		Difference (95% CI) (unadjusted)	Difference (95% CI) (adjusted)	p Value for adjusted analysis
	Mean	SD	Mean	SD			
FEV ₁	–0.461	0.876	0.011	1.000	–0.472 (–0.716 to –0.228)	–0.487 (–0.791 to –0.184)	0.002
FVC	–0.204	0.950	0.005	1.001	–0.209 (–0.453 to 0.035)	–0.151 (–0.454 to 0.152)	0.329
FEF _{25–75}	–0.573	0.928	0.013	1.001	–0.586 (–0.829 to –0.342)	–0.627 (–0.930 to –0.325)	< 0.001
FEV ₁ /FVC	–0.362	1.183	0.009	0.997	–0.371 (–0.617 to –0.125)	–0.488 (–0.793 to –0.182)	0.002
FEF _{25–75} /FVC	–0.457	0.963	0.010	1.001	–0.467 (–0.713 to –0.221)	–0.512 (–0.818 to –0.207)	0.001
	33–34 weeks (n=79)		Term (n=6144)		Difference and 95% CI (unadjusted)	Difference and 95% CI (adjusted)	p Value for adjusted analysis
	Mean	SD	Mean	SD			
FEV ₁	–0.498	1.094	0.011	1.000	–0.509 (–0.733 to –0.285)	–0.485 (–0.724 to –0.246)	<0.001
FVC	–0.313	1.036	0.005	1.001	–0.318 (–0.540 to –0.096)	–0.352 (–0.589 to –0.115)	0.004
FEF _{25–75}	–0.425	1.095	0.013	1.001	–0.438 (–0.659 to –0.217)	–0.434 (–0.670 to –0.197)	< 0.001
FEV ₁ /FVC	–0.312	1.169	0.009	0.997	–0.320 (–0.545 to –0.096)	–0.239 (–0.477 to –0.001)	0.049
FEF _{25–75} /FVC	–0.271	1.059	0.010	1.001	–0.281 (–0.503 to –0.060)	–0.267 (–0.503 to –0.031)	0.027
	35–36 weeks (n=238)		Term (n=6144)		Difference and 95% CI (unadjusted)	Difference and 95% CI (adjusted)	p Value for adjusted analysis
	Mean	SD	Mean	SD			
FEV ₁	0.008	0.908	0.011	1.000	–0.003 (–0.133 to 0.127)	–0.012 (–0.147 to 0.123)	0.867
FVC	0.015	0.950	0.005	1.001	0.010 (–0.120 to 0.139)	0.009 (–0.125 to 0.143)	0.893
FEF _{25–75}	–0.010	0.879	0.013	1.001	–0.023 (–0.152 to 0.106)	–0.034 (–0.168 to 0.099)	0.616
FEV ₁ /FVC	0.004	0.906	0.009	0.997	–0.005 (–0.135 to 0.125)	–0.017 (–0.151 to 0.117)	0.805
FEF _{25–75} /FVC	–0.022	0.925	0.010	1.001	–0.033 (–0.162 to 0.097)	–0.037 (–0.171 to 0.096)	0.582

FEF_{25–75}, forced expiratory volume at 25–75% of FVC; FEV₁, forced expiratory volume in 1 s; FVC, forced vital capacity.

respiratory viruses such as respiratory syncytial virus.¹⁹ A large study of over 150 000 singleton births at 34–41 weeks gestation showed rates of severe respiratory disorders in neonates decreased with increasing gestational age (from 19.8% at 34 weeks to 0.28%

at 39–41 weeks).²⁰ Using the raised volume rapid thoracic compression technique, Friedrich *et al* reported decreased airway flows despite normal FVC at 4 and 16 months of age in healthy late preterm infants. However, it is unclear if decreased lung

Table 5 Mean and SD of z-scores of lung function measures adjusted for age, gender and height at age 14–17 years, with CIs for the difference of means (premature – term), unadjusted and adjusted for maternal smoking in pregnancy and social status

	25–32 weeks (n=42)		Term (n=4105)		Difference (95% CI) (unadjusted)	Difference (95% CI) (adjusted)	p Value for adjusted analysis
	Mean	SD	Mean	SD			
FEV ₁	–0.186	0.850	0.001	1.004	–0.186 (–0.494 to 0.121)	–0.145 (–0.534 to 0.244)	0.464
FVC	–0.151	1.034	–0.004	1.002	–0.148 (–0.452 to 0.156)	–0.041 (–0.427 to 0.346)	0.836
FEF _{25–75}	–0.438	0.899	0.011	1.001	–0.449 (–0.753 to –0.145)	–0.382 (–0.768 to 0.005)	0.053
FEV ₁ /FVC	–0.195	1.117	0.010	0.995	–0.205 (–0.517 to 0.106)	–0.150 (–0.535 to 0.236)	0.446
FEF _{25–75} /FVC	–0.260	1.006	0.013	1.000	–0.274 (–0.581 to 0.034)	–0.338 (–0.723 to 0.046)	0.084
	33–34 weeks (n=49)		Term (n=4105)		Difference and 95% CI (unadjusted)	Difference and 95% CI (adjusted)	p Value for adjusted analysis
	Mean	SD	Mean	SD			
FEV ₁	–0.022	1.135	0.001	1.004	–0.023 (–0.304 to 0.259)	–0.059 (–0.348 to 0.231)	0.691
FVC	0.169	0.905	–0.004	1.002	0.173 (–0.109 to 0.454)	0.127 (–0.161 to 0.415)	0.386
FEF _{25–75}	–0.279	1.143	0.011	1.001	–0.290 (–0.572 to –0.009)	–0.289 (–0.577 to –0.001)	0.049
FEV ₁ /FVC	–0.401	1.104	0.010	0.995	–0.411 (–0.693 to –0.130)	–0.379 (–0.666 to –0.092)	0.010
FEF _{25–75} /FVC	–0.415	0.984	0.013	1.000	–0.428 (–0.709 to –0.147)	–0.397 (–0.683 to –0.110)	0.007
	35–36 weeks (n=129)		Term (n=4105)		Difference and 95% CI (unadjusted)	Difference and 95% CI (adjusted)	p Value for adjusted analysis
	Mean	SD	Mean	SD			
FEV ₁	0.057	0.859	0.001	1.004	0.057 (–0.120 to 0.234)	0.094 (–0.089 to 0.276)	0.315
FVC	0.095	0.956	–0.004	1.002	0.099 (–0.077 to 0.274)	0.119 (–0.060 to 0.299)	0.191
FEF _{25–75}	–0.077	0.877	0.011	1.001	–0.088 (–0.263 to 0.087)	–0.056 (–0.235 to 0.124)	0.543
FEV ₁ /FVC	–0.079	1.047	0.010	0.995	–0.089 (–0.267 to 0.088)	–0.062 (–0.243 to 0.119)	0.502
FEF _{25–75} /FVC	–0.137	0.957	0.013	1.000	–0.151 (–0.326 to 0.024)	–0.133 (–0.311 to 0.045)	0.143

FEF_{25–75}, forced expiratory volume at 25–75% of FVC; FEV₁, forced expiratory volume in 1 s; FVC, forced vital capacity.

Table 6 Mean and SD of z-scores of lung function measures adjusted for age, gender and height at age 8–9 and 14–17 years, with CIs for the difference of means (14–17 – 8–9 years of age), unadjusted and adjusted for maternal smoking in pregnancy and social status

	25–32 weeks: age 8–9 (n=30)		25–32 weeks: age 14–17 (n=30)		Difference (95% CI) (unadjusted)	Difference (95% CI) (adjusted)	p Value for adjusted analysis
	Mean	SD	Mean	SD			
FEV ₁	–0.441	0.946	–0.318	0.932	0.123 (–0.258 to 0.504)	0.013 (–0.471 to 0.497)	0.958
FVC	–0.032	1.050	–0.052	0.918	–0.021 (–0.408 to 0.367)	–0.275 (–0.764 to 0.214)	0.271
FEF _{25–75}	–0.717	0.944	–0.572	0.909	0.145 (–0.202 to 0.491)	0.227 (–0.214 to 0.688)	0.313
FEV ₁ /FVC	–0.605	1.294	–0.438	1.132	0.167 (–0.211 to 0.546)	0.397 (–0.082 to 0.877)	0.104
FEF _{25–75} /FVC	–0.665	0.941	–0.509	0.949	0.156 (–0.184 to 0.496)	0.296 (–0.134 to 0.727)	0.177
	33–34 weeks: age 8–9 (n=42)		33–34 weeks: age 14–17 (n=42)		Difference (95% CI) (unadjusted)	Difference (95% CI) (adjusted)	p Value for adjusted analysis
	Mean	SD	Mean	SD			
FEV ₁	–0.483	1.039	–0.079	1.191	0.404 (0.082 to 0.726)	0.347 (0.015 to 0.680)	0.041
FVC	–0.090	0.938	0.130	0.960	0.220 (–0.108 to 0.547)	0.226 (–0.562 to 0.111)	0.188
FEF _{25–75}	–0.622	1.127	–0.336	1.173	0.287 (–0.006 to 0.580)	0.241 (–0.062 to 0.544)	0.119
FEV ₁ /FVC	–0.621	1.272	–0.432	1.156	0.190 (–0.130 to 0.510)	0.143 (–0.187 to 0.473)	0.395
FEF _{25–75} /FVC	–0.566	1.052	–0.445	1.017	0.121 (–0.166 to 0.408)	0.112 (–0.183 to 0.408)	0.456
	35–36 weeks: age 8–9 (n=107)		35–36 weeks: age 14–17 (n=107)		Difference (95% CI) (unadjusted)	Difference (95% CI) (adjusted)	p Value for adjusted analysis
	Mean	SD	Mean	SD			
FEV ₁	0.148	0.863	0.021	0.860	–0.127 (–0.332 to 0.079)	–0.130 (–0.343 to 0.084)	0.234
FVC	0.115	0.941	0.045	0.955	–0.070 (–0.275 to 0.135)	–0.097 (–0.308 to 0.115)	0.370
FEF _{25–75}	0.086	0.900	–0.076	0.888	–0.161 (–0.345 to 0.022)	–0.132 (–0.322 to 0.059)	0.175
FEV ₁ /FVC	0.006	0.911	–0.048	1.076	–0.054 (–0.258 to 0.150)	–0.007 (–0.218 to 0.205)	0.951
FEF _{25–75} /FVC	0.027	0.970	–0.104	0.992	–0.131 (–0.311 to 0.049)	–0.095 (–0.281 to 0.091)	0.318
	Term weeks: age 8–9 (n=3431)		Term weeks: age 14–17 (n=3431)		Difference (95% CI) (unadjusted)	Difference (95% CI) (adjusted)	p Value for adjusted analysis
	Mean	SD	Mean	SD			
FEV ₁	–0.004	1.006	0.010	1.002	0.014 (–0.021 to 0.051)	–0.013 (–0.060 to 0.034)	0.582
FVC	–0.019	1.003	0.005	1.001	0.025 (–0.012 to 0.061)	0.007 (–0.054 to 0.040)	0.774
FEF _{25–75}	0.014	0.997	0.010	1.001	–0.005 (–0.037 to 0.028)	–0.016 (–0.058 to 0.027)	0.468
FEV ₁ /FVC	0.016	1.003	0.003	0.982	–0.013 (0.049 to 0.023)	–0.006 (–0.053 to 0.041)	0.801
FEF _{25–75} /FVC	0.023	1.005	0.005	0.989	–0.017 (–0.049 to –0.014)	–0.011 (–0.052 to –0.031)	0.616

FEF_{25–75}, forced expiratory volume at 25–75% of FVC; FEV₁, forced expiratory volume in 1 s; FVC, forced vital capacity.

function persists into childhood and beyond. Our data show that the preterm group of 33–34 weeks gestation have lower spirometry measurements than those born at term at 8–9 years of age; however, improvement was observed for FEV₁ by 14–17 years of age but not for other measurements. The extremely preterm group had similar effect sizes but, in contrast, the 35–36-week gestation group had similar lung function to the term control group at both ages. Infants at 33–34 weeks gestation are born at the stage of sacral development⁸ and early delivery may alter their subsequent alveolar development, as has been suggested for extremely preterm infants.²¹ Interventions such as mechanical ventilation and oxygen therapy are well

known to promote pulmonary inflammation in infancy leading to abnormal growth and development.⁶ Although the numbers are small, especially as only 17% of the 33–34-week gestation group required mechanical ventilation in infancy, the data appear to support the notion that intervention is associated with worse spirometry at 8–9 years, which did not appear to improve to the same degree as for the non-ventilated group between the two ages (table E1 in online supplement). The interpretation is further complicated by the fact that mechanical ventilation was likely to be instituted in infants with underlying lung disease which may already identify them as candidates for future lung abnormalities. Other possible explanations for the poorer lung function in the

Table 7 Bronchial responsiveness at 8–9 years of age

	25–32 weeks	33–34 weeks	35–36 weeks	Term
Number who had lung function at 8–9 years	68	81	248	6308
Number (%) who were excluded or did not complete the challenge for the reasons listed below*	35 (51.5%)	34 (42.0%)	90 (36.3%)	2185 (34.6%)
Number (%) who underwent challenge	33 (48.5%)	47 (58.0%)	158 (63.7%)	4123 (65.4%)
Number (%) who responded to methacholine challenge	9 (27.3%)	9 (19.1%)	22 (13.9%)	646 (15.7%)
Geometric mean PD ₂₀ in responders (μmol)	1.20 (0.43, 3.35)	2.25 (1.08, 4.70)	2.27 (1.50, 3.42)	1.71 (1.57, 1.87)
Geometric mean PD ₂₀ for the whole population who underwent the challenge (μmol)	6.46 (4.15, 10.04)	8.86 (7.08, 11.09)	9.62 (8.65, 10.69)	8.96 (8.73, 9.19)

*Children did not start or complete the challenge for a number of reasons: the child refused whole/part of session, parent refused, parents did not attend, staff illness/difficulties, need doctor, child ill or taking medications, time restraint, equipment error/failure, staff error (general), child started coughing, child problems/difficulties with task, couldn't get reproducible FEV₁ <70% predicted.

33–34-week gestation group are an adverse prenatal environment associated with preterm delivery (eg, chorioamnionitis may affect later lung function)^{22–25} or dysregulated lung growth associated with premature cessation of late gestational development. The loss of late intrauterine lung maturation in the 33–34-week gestational group suggests that lung growth may be dysregulated after birth resulting in physiological differences compared with those completing term gestation.

We noted poorer lung function in childhood in the late preterm children which was similar to the extremely preterm group, but did not note significant differences in rates of asthma between the gestation groups. It would be interesting to investigate if the decreased FEV₁ can be reversed by bronchodilators as was recently shown for children who were born extremely preterm.²⁶ Furthermore, if there is plasticity of the airways, bronchodilators may be effective in this group of children, but whether they improve longer term airway function needs further investigation. Few studies have reported longitudinal lung function in preterm infants with most focusing on severe long-term respiratory consequences—that is, those who develop CLD. In a small number of children with CLD, often also called bronchopulmonary dysplasia, Filippone *et al* reported that airflow limitation tracked between 2 and 8.8 years of age did not show improvements between the two time points.²⁷ In contrast, Doyle and colleagues reported significant improvements in predicted values of lung function between 8 and 14 years of age in children with a birth weight <1501 g.²⁸ When we compared the preterm children who had spirometry at both time periods, our data support potential improvements in lung function in children born at 33–34 weeks gestation from 7–8 years of age to 14–17 years, especially for the larger airways. For the 25–32-week gestation group, relative improvements were noted in the z-scores for lung spirometry but these were not statistically significant. While tracking appears to be an attractive concept, our longitudinal data question the concept of tracking as lung function improved between 7–8 and 14–17 years of age. More recent imaging modalities such as the use of hyperpolarised helium may assist in investigating lung growth and development beyond infancy²⁹ to clarify if, indeed, continued growth and development occurs beyond infancy.

While we acknowledge that the data were not specifically collected to address the current hypothesis, the strength of this study is that it uses data from a large unselected population-based cohort with comprehensive longitudinal data with reasonable numbers of preterm-born children and a large number in the term control group. A further strength is that spirometry was measured using the same methods at the two time points and is an objective outcome. As with all cohort studies, a major weakness is the loss to follow-up and the resulting potential attrition bias. Fewer children who had mothers who smoked during pregnancy and a from lower social class attended for lung spirometry compared with the attenders. Since lower social class is associated with more preterm births, more maternal smoking and decreased attendance for spirometry, it is possible that a higher rate of attendance in these children would have increased the observed differences. In addition, ALSPAC did not measure the lung function of severely disabled children and, as these represent the severe end of morbidity associated with preterm delivery, the differences may have been attenuated by their exclusion. We included children from ethnic minorities in our analysis as there were very few of them (in the whole ALSPAC cohort only 5% of the children who reported their ethnic background reported being non-white), and the ALSPAC cohort is broadly representative of the population of the UK.¹²

In summary, our data show that, at 8–9 years of age, measures of airway function are reduced in children born at 33–34 weeks gestation compared with children born at term, and are of similar magnitude to those in the extremely preterm group who received a much higher proportion of respiratory intervention during the neonatal period. However, by 14–17 years of age, there was no significant difference in FEV₁ and FVC in the late preterm group compared with children born at term, although FEV₁/FVC and FEF_{25–75%} remained significantly lower than term controls. This raises the question about what occurs from a lung development point of view between 33–34 and 35–36 weeks of gestation. Now that there is a strong suspicion of respiratory physiological abnormalities in late preterm-born children, it will be interesting to study in more detail whether they are at greater risk of respiratory infections in infancy, have increased respiratory symptoms or indeed have limitation of exercise with or without exercise-induced bronchoconstriction.

Acknowledgements We are grateful to all the families who took part in this study, the midwives for their help in recruiting them and the whole ALSPAC team which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. We thank Dr David Herrick for his help in identifying relevant perinatal data.

Funding The Avon Longitudinal Study of Parents and Children (ALSPAC) receives core funding from the UK Medical Research Council, the Wellcome Trust and the University of Bristol. The lung function measures were supported by a grant from the UK Medical Research Council (G0401540). SJK was partially supported by Nutricia Research Foundation.

Competing interests None.

Ethics approval This study was conducted with the approval of the ALSPAC Law and Ethics Committee and the local research ethics committees.

Contributors All authors led by SK were involved in study design including hypothesis generation and execution. AJH is co-director of ALSPAC and was involved in collection of the data. FD and WJW were primarily responsible for statistical analyses. SJK and SK wrote the initial drafts of the paper which was contributed to by all authors. SK is the guarantor.

Provenance and peer review Not commissioned; externally peer reviewed.

Data Sharing Statement Data generated by the ALSPAC study are held as a national resource. Data are anonymised at entry and held separately from administrative data; therefore, identification of individual subjects' data is concealed from researchers. Access to data items is provided on application to the ALSPAC Executive Committee. The full ALSPAC collaboration policy, including details of data sharing arrangements, is published online at: <http://www.bris.ac.uk/alspac/sci-com/collab-policy/>.

REFERENCES

1. **Goldenberg RL**, Culhane JF, Iams JD, *et al*. Preterm birth 1. Epidemiology and causes of preterm birth. *Lancet* 2008;**371**:75–84.
2. **Raju TN**, Higgins RD, Stark AR, *et al*. Optimizing care and outcome for late-preterm (near-term) infants: a summary of the workshop sponsored by the National Institute of Child Health and Human Development. *Pediatrics* 2006;**118**:1207–14.
3. **The Consortium on Safe Labour**. Respiratory morbidity in late preterm births. *JAMA* 2010;**304**:419–25.
4. **Friedrich L**, Pitrez PM, Stein RT, *et al*. Growth rate of lung function in healthy preterm infants. *Am J Respir Crit Care Med* 2007;**176**:1269–73.
5. **Todisco T**, de Benedictis FM, Iannacci L, *et al*. Mild prematurity and respiratory functions. *Eur J Pediatr* 1993;**152**:55–8.
6. **Chakraborty M**, McGreal EP, Kotecha S. Acute lung injury in preterm newborn infants: mechanisms and management. *Paediatric Respir Rev* 2010;**11**:162–70.
7. **Colin AA**, McEvoy C, Castile RG. Respiratory morbidity and lung function in preterm infants of 32 to 36 weeks' gestational age. *Pediatrics* 2010;**126**:1115–28.
8. **Kotecha S**. Lung growth: implications for the newborn infants. *Arch Dis Child Fetal Neonatal Ed* 2000;**82**:F69–74.
9. **Bennett CC**, Johnson A, Field DJ, *et al*. UK collaborative randomised trial of neonatal extracorporeal membrane oxygenation: follow-up to age 4 years. *Lancet* 2001;**357**:1094–6.
10. **Cornfield DN**, Maynard RC, deRegnier RO, *et al*. Randomized, controlled trial of low-dose inhaled nitric oxide in the treatment of term and near-term infants with respiratory failure and pulmonary hypertension. *Pediatrics* 1999;**104**:1089–94.

11. **NINOS Study Group.** Inhaled nitric oxide in term and near-term infants: neuro-developmental follow-up of the Neonatal Inhaled Nitric Oxide study group (NINOS). *J Pediatr* 2000;**136**:611–17.
12. **Golding J,** Pembrey M, Jones R, *et al.* ALSPAC—the Avon Longitudinal Study of Parents and Children I. Study methodology. *Pediatr Perinat Epidemiol* 2001;**15**:74–87.
13. **Lawlor DA,** Cooper AR, Bain C, *et al.* Association of birth size and duration of breast feeding with cardiorespiratory fitness in childhood: findings from the Avon Longitudinal Study of Parents and Children (ALSPAC). *Eur J Epidemiol* 2008;**23**:411–22.
14. **Kotecha SJ,** Watkins WJ, Heron J, *et al.* Spirometric lung function in school-age children. Effect of intrauterine growth retardation and catch-up growth. *Am J Respir Crit Care Med* 2010;**181**:969–74.
15. **Chinn S,** Rona RJ. Height and age adjustment for cross sectional studies of lung function in children aged 6–11 years. *Thorax* 1992;**47**:707–14.
16. **Munkata M,** Ohe M, Homma Y, *et al.* Pulmonary dysanapsis, methacholine airway responsiveness and sensitization to airborne antigen. *Respirology* 1996;**2**:113–18.
17. **Yan K,** Salcome C, Woolcock AJ. Rapid method for measurement of bronchial responsiveness. *Thorax* 1983;**38**:760–5.
18. **Melamed N,** Klinger G, Tenenbaum-Gavish K, *et al.* Short-term neonatal outcome in low-risk, spontaneous, singleton, late preterm deliveries. *Obstet Gynecol* 2009;**114**:253–60.
19. **Boyce TG,** Mellen BG, Mitchel EF Jr, *et al.* Rates of hospitalization for respiratory syncytial virus infection among children in Medicaid. *J Pediatr* 2000;**137**:860–70.
20. **Gouyon JB,** Vintjeux A, Sagot P, *et al.* Neonatal outcome associated with singleton birth at 34–41 weeks of gestation. *Int J Epidemiol* 2010;**39**:769–76.
21. **Jobe AJ.** The new BPD: an arrest of lung development. *Pediatr Res* 1999;**46**:641–3.
22. **Maxwell NC,** Davies PL, Kotecha S. Antenatal infection and inflammation: what's new? *Curr Opin Infect Dis* 2006;**19**:253–8.
23. **Thomas W,** Speer CP. Chorioamnionitis: important risk factor or innocent bystander for neonatal outcome? *Neonatology* 2011;**99**:177–87.
24. **Watterberg KL,** Demers LM, Scott SM, *et al.* Chorioamnionitis and early lung inflammation in infants in whom bronchopulmonary dysplasia develops. *Pediatrics* 1996;**97**:210–15.
25. **Ogunyemi D,** Murillo M, Jackson U, *et al.* The relationship between placental histopathology findings and perinatal outcome in preterm infants. *J Matern Fetal Neonatal Med* 2003;**13**:102–9.
26. **Fawke J,** Lum S, Kirby J, *et al.* Lung function and respiratory symptoms at 11 years in children born extremely preterm. *Am J Respir Crit Care Med* 2010;**182**:237–45.
27. **Filippone M,** Sartor M, Zacchello F, *et al.* Flow limitation in infants with bronchopulmonary dysplasia and respiratory function at school age. *Lancet* 2003;**361**:753–4.
28. **Doyle LW,** Chavasse R, Ford GW, *et al.* Changes in lung function between 8 and 14 years in children with birth weight of less than 1,501g. *Pediatr Pulmonol* 1999;**27**:185–90.
29. **Sukstanskii AL,** Conradi MS, Yablonskiy DA. ³He lung morphometry technique: accuracy analysis and pulse sequence optimization. *J Magn Reson* 2010;**207**:234–41.

Journal club

Differentiation of malignant pleural mesothelioma from other pleural diseases

In this study, pleural effusion samples from 101 patients with suspected or newly diagnosed malignant pleural mesothelioma (MPM) or metastatic adenocarcinoma (ADCA) were collected between 1998 and 2010. Of the 101 patients, 65 were diagnosed with MPM, 25 with ADCA and 15 with benign pleural effusion. Cancer cells isolated from pleural effusion samples were subjected to genome-wide gene expression analysis, done with the help of microarrays and real-time PCR.

The study found 74 genes coding for markers that were overexpressed in MPM and 9 genes that were overexpressed in ADCA. The highest expression in MPM cells was the gene COL3A1, coding for type III collagen. Immunohistochemistry demonstrated 100% staining of MPM biopsy samples with antibodies specific for COL3A1 and no staining in ADCA.

Soluble markers such as CCL2 and galectin-3 were identified as useful markers for diagnosing MPM. CCL2 concentration was significantly higher in patients with MPM than in patients with ADCA or benign pleural effusion whereas galectin-3 was significantly lower in MPM compared with ADCA. CCL2 and COL3A1 can be used as positive markers, whereas galectin-3 can be used as a negative marker. Statistical analysis using receiver operating characteristic curve and sensitivity suggested that both these markers performed better than existing markers in differentiating MPM from other pleural diseases.

Although type III collagen, CCL2 and galectin-3 markers have been shown to be valuable diagnostic aids, a large multicentre study is required to validate these results and investigate other potential markers.

► **Gueugnon F,** Leclercq S, Blanquart C, *et al.* Identification of novel markers for the diagnosis of malignant pleural mesothelioma. *Am J Pathol* 2011;**178**:1033–42.

R Bhardwaj

Correspondence to Dr R Bhardwaj, Foundation Year 1, St James University Hospital, Leeds, UK; rahulbhardwaj@doctors.org.uk

Published Online First 17 May 2011

Thorax 2012;**67**:61. doi:10.1136/thoraxjnl-2011-200355

Appendix 1. Characteristics of 34 assessment tools in 48 studies

Tool	Item evaluation	Assessed knowledge, skills, and behaviour (Canadian Safety Competencies domains*)	Source	Study aim	Setting	Learners	Main findings
MEDICAL							
The Miller level "knows"							
A knowledge, skills, and attitudes questionnaire	28 items (5 MCQ, and 18 items of 5 point scale)	Medical error reporting, physician's attitude toward medical errors, tolerance of uncertainty, reaction to uncertainty, institutional safety culture (1, 6)	Madigosky WS, 2006 ²³	To evaluate patient safety and medical fallibility curriculum	Single US medical school	92 second-year medical students	Improvement in self-reported knowledge, and attitudes changed after training
			Hall LW, 2010 ²⁴	To evaluate patient safety training	Single US medical school	146 third-year medical students (81 study group, 65 control group)	Training increased student comfort in identifying the cause of an error.
Unnamed online test	26 items, MCQ	Patient safety, error prevention, system theory, United States health care system (1, 4)	Kerfoot BP, 2007 ²⁵	To evaluate a web-based training program	US web-based program	276 medical students (years 2 and 3) and 217 residents (PGY1 to 5)	Participants' test scores improved after completing the first module.
Unnamed questionnaire	14 items, MCQ	Error prevention, system theory (1, 4)	Kerfoot BP, 2007 ²⁶	To determine the current patient safety knowledge	7 US residencies and 2 US medical schools	268 second and third-year medical students and 372 medical residents (PGY 1 to 5)	Participants correctly answered a mean 58.4%. Patient safety knowledge levels varied significantly by participants' characteristics.
Unnamed questionnaire	MCQ	RCA analysis: defining violation errors, barriers to error, contributory factors, types of error (1)	Wallace LM, 2009 ²⁷	To evaluate a training program	UK national training program	374 health professionals immediately after the training, 350 participants 6 months after the training.	The RCA training enhanced knowledge of RCA.
The Miller level "knows how"							
Computer-based case management of USMLE Step 3	Free text entry: 25 cases and 11 cases	Patient management, errors in decision making (4)	Harik P, 2009 ²⁸	To examine potentially dangerous actions	US license examination	25,283 medical students	More than 20% of examinees ordered at least one dangerous action.
			Floreck LM, 2002 ²⁹	To examine potentially dangerous actions	US license examination	About 7,000 medical students	Examinees performed at least one non-indicated action in more than 40% of cases.

The Miller level "shows how"							
Unnamed checklist	10 items	Safe handover (4)	Klamen DL, 2009 ³⁰	To develop a training program	Single US medical school	69 second-year medical students	More than 80% of students could safely handover
Competencies assessed at multiple levels							
8 station OSCE	Checklist and written examinations (MCQ, short answer, scored based on original checklist)	Quality measurement, RCA, evidence-based medicine, insurance systems, team collaboration (only assess team competence), prescription errors, Nolan's model, negotiation (1, 2, 3, 4, 6)	Varkey P, 2008 ³¹	To develop an assessment tool	Single US quality improvement elective	9 medical residents (preventive medicine and endocrinology)	All faculty members agreed that the OSCE was realistic and capable of providing accurate assessments.
			Gupta P, 2009 ³²	To develop an assessment tool	Single US quality improvement elective	9 medical residents (preventive medicine and endocrinology)	Pilot study. All faculty members agreed that they were able to accurately assess competency in RCA using the tool.
6 station OSCE	Not mentioned	Interview: error case, health behaviour change case; simulations; video: teamwork, co-ordination of care; chart: adverse event case, deficient care case; written examination: behavioural skills, medication safety, system approach (1, 2, 3, 4, 6)	Singh R, 2009 ³³	To develop an assessment tool to evaluate a training program	US residency programs (patient safety curriculum)	59 medical residents (PGY1 to 3, 47 study group, 12 control group, family medicine)	Safety training increased systems thinking and inculcation of a culture of safety.
10 station OSCE	Checklist and written examinations (MCQ)	Accurate data gathering, accurate written communication, patient-centered communication, timely/effective patient care, team functioning/inter-professional communication, infection prevention (2, 3, 4, 6)	Wagner DP, 2009 ³⁴	To develop an assessment tool	64 US residency programs (patient safety curriculum)	235 medical residents (PGY1, multispecialty)	Significant performance deficits were noted. Trainee and administrator evaluation of the experience was positive.

ANESTHESIA

The Miller level "shows how"

The Anesthetists' Non-Technical Skills System (ANTS)	4 categories, 15 elements, 4 point scale	Non-technical skills: situation awareness, task management, team working, decision making (2, 3, 4)	Zausig YA, 2009 ³⁵	To compare non-technical skills and medical management skills after training	2 German university hospitals and 5 community hospitals	42 physicians (anesthesia)	A single session of training did not improve non-technical skills when compared with anesthesiologists who only received medical management training.
			Savoldelli GL, 2006 ³⁶	To evaluate the effect of debriefing	Single Canadian University	42 medical residents (PGY 1, 2, and 4, anesthesia)	Participants' non-technical skills did not improve in the control group, whereas the provision of oral feedback, either assisted or not assisted with videotape review, resulted in significant improvement ($P < 0.005$).
			Yee B, 2005 ³⁷	To evaluate a training program	Single Canadian University	20 medical residents (PGY 2 and 4, anesthesia)	A significant improvement in the non-technical skills of residents was demonstrated from their first to second session and from their first to third session (both $P < 0.005$).
			Welke TM, 2009 ³⁸	To evaluate a training program	Single Canadian University	30 medical residents (PGY1 and 2, anesthesia)	Posttest ($P = 0.009$) and retention ($P < 0.001$) performances of non-technical skills were significantly improved in the standardized multimedia instruction group compared with pretest.
			Müller MP, 2009 ³⁹	To evaluate a training program	5 ICUs at a single German university hospital	32 physicians (anesthesia)	ANTS scores as well as clinical performances were significantly better in the post-intervention scenario.
			Fletcher G, 2003 ⁴⁰	To develop an assessment tool	17 UK hospitals	50 physicians (anesthesia)	The results showed that the skills are observable and can be rated with acceptable levels of agreement and accuracy.
			Bruppacher HR, 2010 ⁴¹ (the "does" level)	To compare training programs	Single Canadian hospital	20 medical residents (PGY 4 or higher, anesthesia)	The simulation group scored significantly higher than the seminar group at both posttest and retention test.

Rating scale	3 categories, 5 point scale	Behaviour (non-technical skills), medical management, overall performance (2)	Weller JM, 2008 ⁴²	To develop an assessment tool	All anesthesia trainees in a region of New Zealand	20 medical residents (anesthesia)	Scores for behaviour were less dependent on the specific clinical context than the scores for overall performance and medical management.
Competencies assessed at multiple levels							
4 station: 1)Structured interview, 2)Portfolio review, 3)Presentation , 4)Simulation	Structured interview, portfolio, presentation, and simulation with checklist	Personal skills (communication, organization, working under pressure, situational awareness and decision making, team working), content of portfolio, global rating (2, 3, 4, 5)	Gale TC, 2010 ⁴³	To develop a selection centre for recruitment	Candidates attending interview for anesthesia training posts in the South West Peninsula Deanery in UK	224 candidates (anesthesia)	Candidates and assessors demonstrated strong approval of the selection centre with more than 70% of ratings 'good' or 'excellent'.
SURGERY The Miller level "shows how"							
The Non-technical Skills for Surgeons (NOTSS)	5 categories, 14 elements, 4 point scale	Non-technical skills: situation awareness, decision making, task management, leadership, communication and teamwork (2, 3, 4)	Yule S, 2006 ⁴⁴	To develop an assessment tool	UK	27 consultant surgeons	Five categories of non-technical skills were identified.
The Non-Technical Skills System (NOTECHS)	5 categories, 19 elements, 5 or 6 point scale	Non-technical skills: communication and interaction, situation awareness and vigilance, co-operation and team skills, leadership and managerial skills, decision making (revised NOTECH version) (2, 3, 4)	Black SA, 2010 ⁴⁵	To assess surgical competence in a simulated operating theatre	UK simulated operating theatre	30 physicians (surgery)	There was a significant difference in non-technical skill with ascending grade for both scenarios (P < 0.001). There was a highly significant correlation between technical and non-technical performance in both scenarios.
			Moorthy K, 2006 ⁴⁶	To evaluate a training program	UK simulated operating theatre	20 medical residents (surgery, 10 senior and 10 junior trainees)	While the senior trainees scored higher than the juniors for technical skills (P=0.001), there were no differences in human factors skills.

			Moorthy K, 2005 ⁴⁷	To develop a simulated environment	UK simulated operating theatre	26 surgical trainees (surgery, trainees were divided into 3 groups according to their level of experience)	While technical skills ratings discriminated between surgeons according to experience (P=0.002), there were no differences in terms of the checklist and teamwork skills (P=0.70).
			Undre S, 2007 ⁴⁸	To develop a training program	UK simulated operating theatre	20 teams (N=80, trainee surgeon, anesthetist, operating department practitioner, scrub nurse)	Leadership and decision making were scored lower than the other three non-technical skills (communication, team skills, and vigilance).
			Powers KA, 2008 ⁴⁹	To develop a simulation model	US simulated operating theatre	5 FLS-certified surgeon experts and 5 non-FLS certified novices	There was a significant difference between the two groups in their overall technical and non-technical skills.
The PAR matrix	18 skill-goals (the y-axis is divided into six levels and the x-axis into three columns: unsatisfactor y, competent, good)	Crew resource management skills, surgical technical skills (ergonomics) (2, 3, 5)	Papaspyros SC, 2009 ⁵⁰	To develop an assessment tool	Single UK hospital	14 medical residents (surgery)	The format of the PAR model is such that it allows trainer and trainee to objectively assess progress, identify deficiencies and strengths and formulate an appropriate plan of action.
Observational Teamwork Assessment for Surgery (OTAS)	5 elements, 7 point scale	Teamwork: shared monitoring, communication, co-operation, co-ordination, shared leadership (2)	Wetzel CM, 2010 ⁵¹	To investigate the effect of stress on performance	UK simulated operating theatre	30 physicians (surgery)	Coping significantly enhanced non-technical skills (beta, 0.302; 0.117-1.624, P= 0.03).

EMERGENCY/RESUSCITATION							
The Miller level "shows how"							
Stanford University for crisis management behaviours	11 items, 5 point scale	Orientation to case, inquiry/assertion, communication, feedback, leadership, group climate, anticipation/planning, workload distribution, vigilance and re-evaluation, overall leadership (1, 2, 3, 5)	Wallin CJ, 2007 ⁵²	To evaluate a training program	Swedish simulated emergency environment	15 medical students	Nine of 10 observed team skills improved significantly in response to practice, in parallel with a global rating of team skills.
A behaviour grading sheet	3 categories, 3 point scale (poor/fail, adequate, excellent)	Crisis management skills: teamwork, decision making, situation awareness (2, 4)	Knudson MM, 2008 ⁵³	To evaluate a training program	US simulated trauma emergency environment	18 medical residents (PGY2 to 3, surgery)	Simulation-based trained residents received higher overall scores and higher scores for crisis management skills compared with the didactic lecture-trained group, which was most evident in the scores received for the teamwork category (p=0.04).
Non-technical scorecard at simulated septic shock scenario	7 items, 5 point scale	Non-technical skills: anticipation and planning, communication, leadership, information transfer, task distribution, communication content, information use (2, 3, 4)	Williams JB, 2009 ⁵⁴	To assess reliability of real-time vs. videotaped evaluation	US simulated environment	26 medical residents (emergency and transitional)	Nland-Altman plot analysis of both conditions revealed substantial agreement between the real-time and videotaped review scores by reviewers.
A behaviourally anchored team skill rating scale	5 categories, 5 point scale	Assertiveness, decision making, leadership, communication, situation assessment (2, 4)	Wright MC, 2009 ⁵⁵	To evaluate the relationship between teamwork skills and clinical performance	US simulated environment	35 first-year medical students	There was moderate to high correlation between observer ratings of team skill and checklist-based measures of team performance.
Modified ANTS and Anti-Air Teamwork Observation Measure (ATOM)	7 categories	Modified ANTS: co-ordination, information exchange, use of authority, assessing capabilities, supporting behaviours; ATOM: shared mental model (2, 4)	Westli HK, 2010 ⁵⁶	To evaluate teamwork skills	Norwegian hospitals	27 trauma teams (139 clinicians, including surgeons, anesthesiologists, nurses, and radiographers, emergency)	Specific teamwork skills and behavioural markers were associated with indicators of good team performance.

The Ottawa Crisis Resource Management Global Rating Scale (Ottawa GRS) and The Ottawa Crew Resource Management checklist (Ottawa CRM checklist)	The Ottawa GRS: 5 categories, 7 point scale; the Ottawa CRM checklist: 5 categories, 12 items, 2 point scale ⁴²	Crew resource management skills: leadership, problem solving, situation awareness, resource utilization, communication (2, 3, 4)	Kim J, 2009 ⁵⁷	To compare assessment tools	Canadian university tertiary hospital	60 medical residents (PGY1 and 2, multispecialty)	Users indicated a strong preference for the Ottawa GRS given ease of scoring, presence of an overall score, and the potential for formative evaluation.
			Kim J, 2006 ⁵⁸	To develop an assessment tool	Canadian university tertiary hospital	60 first and third-year medical residents (multispecialty)	Construct validity was measured on the basis of content validity, response process, internal structure, and response to other variables.
Teamwork Behaviour Frequencies Form	Frequency (number of behaviours per minute)	Information sharing, inquiry, assertion, intentions shared and evaluation of plans, vigilance, workload (2, 3, 4, 5)	Thomas EJ, 2007 ⁵⁹	To evaluate a training program	US neonatal resuscitation program	51 medical residents (multispecialty)	The interns in the neonatal resuscitation program with team training exhibited more frequent team behaviours than interns in the control group.
Observational Skill-based Clinical Assessment tool for Resuscitation (OSCAR)	6 categories	Communication, co-operation, co-ordination, monitoring, leadership, decision making (2)	Walker S, 2010 ⁶⁰	To develop an assessment tool	UK	Multidisciplinary experts to assess content validity, 2 clinician observers used (resuscitation)	OSCAR is a psychometrically robust (reliable, content- and face- valid) tool for the assessment of team working skills in cardiac arrest events.
The Miller level "knows"							
Unnamed scale	36 items, 5 point scale	Team building, shared decision making in emergency situations (2)	Haller G, 2008 ⁶¹	To evaluate a training program	Swiss labour and delivery units of a university-affiliated hospital	239 multidisciplinary (midwives, nurses, physicians and technicians)	Except for seven items, the 36-item survey testing participants' learning demonstrated a significant change (P<0.05) towards better knowledge of teamwork and shared decision making after the training program.

NURSING							
The Miller level "knows"							
Quality Improvement Knowledge, Skills, and Attitudes (QulSKA)	73 items, (26 items: MCQ; 6 items: true-false; 31 items: 4 or 6 point scale)	QSEN competencies (quality improvement, safety, evidence-based practice, teamwork, patient-centered care, informatics) (1, 2, 3, 4)	Dycus P, 2009 ⁶²	To develop an assessment tool	US and Latin American nurses practicing at a hospital and affiliate partner sites	37 nurses (pediatric nursing)	The QulSKA inter-item correlation coefficient was 0.839 (P = .001). The mean knowledge score (based on 100) was 69.2 +/- 11.3.
The patient safety attitudes, skills, and knowledge scale (PS-ASK)	26 items, 5 point scale	Patient safety, patient safety practices relevant to nursing, evidence-based advanced practice nursing care (1, 4)	Schnall R, 2008 ⁶³	To develop an assessment tool	US nursing school	285 nursing students	The final instrument consists of 26 items and three separate scales: attitudes, skills, and knowledge.
The Miller level "knows how"							
Essay	N/A	Unit-based quality improvement, meaning of sentinel event, interdisciplinary team (1, 2, 6)	Murray ME, 2010 ⁶⁴	To describe patient safety education in nursing schools	US	-	The application assignment gives students an empowering experience of implementing a process change and measuring results.
Situation awareness check list	17 items	Situation awareness (4)	Cooper S, 2010 ⁶⁵	To examine the ability of nursing students in a simulated environment	Australian nursing school	112 nursing students	The mean situation awareness score across both scenarios was 59% (range 38-82%).
The Miller level "shows how"							
Clinical Simulation Evaluation Tool (CSET)	15 items, numeric scale	Safety, basic assessment, prioritization, problem-focused assessment, ensuring intervention, delegation, communication (3, 4)	Radhakrishnan K, 2006 ⁶⁶	To examine clinical performance	US simulation center	12 senior second degree BSN nursing students	Students who practiced with the simulation had significantly higher scores than the control group on patient identification and on assessing vital signs.
Knowledge, Skills, Attitudes criterion (Checklist) : QSEN	16 items, dichotomous scale	QSEN Competencies (quality improvement, safety, evidence-based practice, teamwork, patient-centered care, informatics) (1, 2, 3, 4)	Ironsides PM, 2009 ⁶⁷	To evaluate a training program	US simulation center	67 nursing students from 8 universities (acute care nursing)	The study showed that students' safety competencies improved significantly from the first to the second simulation.

Unnamed checklist	6 items	Patient-centered care of QSEN (1, 2, 3, 4, 6)	McKeon LM, 2009 ⁶⁸	To evaluate a training program	US university	53 nursing students (34 computer-based simulation, 31 traditional simulation)	Students in both intervention computer-based simulation and traditional simulation achieved similar competencies in patient-centered care.
Clinical Performance Evaluation Tool	6 categories, 3 choices (satisfactory, needs improvement, unsatisfactory)	Patient-centered care, teamwork and collaboration, evidence-based nursing, quality improvement, safety, informatics (2, 4)	Walsh T, 2010 ⁶⁹	To describe the development of an assessment tool	US core group of undergraduate faculty	Nurse	The tool more accurately appraised clinical capabilities by focusing on quality and safety in health care. Further research is needed for evaluation of reliability and validity.
The Miller level "does"							
Unnamed checklist	44 items, 6 point scale	Patient comfort and safety, time management, medication administration (4)	Fisher M, 2000 ⁷⁰	To develop an assessment tool	Australian tertiary and referral health service	30 nurses	The competency assessment tool has face validity and is internally consistent.

Abbreviations: MCQ=Multiple Choice Question; QSEN=Quality and Safety Education for Nurses

*Canadian Safety Competencies: 1) Contribute to a culture of patient safety, 2) Work in teams for patient safety, 3) Communicate effectively for patient safety, 4) Manage safety risks, 5) Optimize human and environmental factors, 6) Recognize, respond to and disclose adverse events

Effect of preterm birth on longitudinal lung spirometry in school age children and adolescents.

Sarah J Kotecha, W John Watkins, Shantini Paranjothy, Frank D Dunstan, A. John Henderson, Sailesh Kotecha.

ONLINE DATA SUPPLEMENT

Methods

ALSPAC

ALSPAC is a longitudinal birth cohort study of 14,541 women recruited during pregnancy, (with estimated dates of delivery between April 1991 and December 1992), and their 14,062 surviving offspring; data on the triplets and one set of quadruplet children were not made available, leaving 14049 live births.

Perinatal data

Perinatal data on respiratory management, including use of surfactant, mechanical ventilation or oxygen therapy, were extracted from the hospital records by trained midwives using a standard proforma for all deaths and discharges for all infants admitted to a neonatal unit or other hospital care during the first 4 weeks of life. We therefore assumed that any infant without a neonatal questionnaire was not admitted to hospital for medical care during the first four weeks of life.

Bronchial Challenge

The challenge ended when the maximum dose was reached (non-responders to the challenge), or when FEV₁ fell by 20% or more from the baseline reading (responders to the challenge).

Table E1 – Mean and SD of z-scores of lung function measures at 8-9 and 14-17 years of age for ventilated and not ventilated children adjusted for age, gender and height, unadjusted for maternal smoking in pregnancy and social status.

25-32 weeks 8-9 years						
	Ventilated n=50		Non-ventilated n=7		Difference and 95% CI (unadjusted)	p-value for unadjusted analysis
	Mean	SD	Mean	SD		
FEV₁	-0.583	0.872	-0.282	0.920	-0.302 (-1.011,0.408)	0.398
FVC	-0.234	0.971	-0.626	0.629	0.393 (-0.367,1.152)	0.305
FEF₂₅₋₇₅	-0.748	0.794	0.134	1.147	-0.882 (-1.561,-0.203)	0.012
25-32 weeks 14-17 years						
	Ventilated n=30		Non-ventilated n=7		Difference and 95% CI (unadjusted)	p-value for unadjusted analysis
	Mean	SD	Mean	SD		
FEV₁	-0.239	0.873	0.047	0.773	-0.286 (-1.018, 0.447)	0.434
FVC	-0.178	1.109	-0.267	0.878	0.090 (-0.825, 1.004)	0.843
FEF₂₅₋₇₅	-0.534	0.946	0.032	0.521	-0.565 (-1.322, 0.191)	0.138
33-34 weeks 8-9 years						
	Ventilated n=14		Non-ventilated n=46		Difference and 95% CI (unadjusted)	p-value for unadjusted analysis
	Mean	SD	Mean	SD		
FEV₁	-0.888	1.136	-0.400	1.144	-0.489 (-1.187, 0.209)	0.166
FVC	-0.472	1.065	-0.273	1.057	-0.200 (-0.847, 0.447)	0.539
FEF₂₅₋₇₅	-0.770	1.125	-0.360	1.030	-0.409 (-1.052, 0.234)	0.208
33-34 weeks 14-17 years						
	Ventilated n=9		Non-ventilated n=24		Difference and 95% CI (unadjusted)	p-value for unadjusted analysis
	Mean	SD	Mean	SD		
FEV₁	-0.645	1.147	0.171	1.089	-0.816 (-1.696, 0.064)	0.068
FVC	-0.079	0.922	0.270	0.950	-0.349 (-1.101, 0.403)	0.351
FEF₂₅₋₇₅	-1.024	1.079	-1.089	1.033	-0.935 (-1.768, -0.102)	0.029

There was missing data for some of the variables reported.

