Title: Lung function trajectories from pre-school age to adulthood and their associations with early life factors: a retrospective analysis of three population-based birth cohort studies.

Article Type: Article (Original Research)

Corresponding Author: Professor John Henderson,

Corresponding Author's Institution: University of Bristol

First Author: Danielle Belgrave

Order of Authors: Danielle Belgrave; Raquel Granell; Steve w Turner; John A Curtin; Iain E Buchan; Peter N LeSouëf; Angela Simpson; John Henderson; Adnan Custovic

Manuscript Region of Origin: UNITED KINGDOM

Abstract: Background: Maximal lung function in early adulthood is an important determinant of mortality and COPD. We investigated whether there are distinct trajectories of lung function during childhood, and whether these extend to adulthood and infancy.

Methods: To ascertain trajectories of forced expiratory volume in 1 second (FEV1), we studied two population-based birth cohorts with repeat spirometry from childhood into early adulthood (1046 and 1390 participants). To ascertain whether trajectories extend to early life, we used a third cohort with repeat lung function measures in infancy (V'maxFRC) and childhood (FEV1; n=196). We identified trajectories using latent profile modelling. We created an allele score to investigate genetic associations of trajectories, and constructed a multivariable model to identify their early-life predictors.

Results: We identified four childhood FEV1 trajectories: Persistently High; Normal; Below Average; and Persistently Low. Persistently Low trajectory (~5% of participants) was associated with persistent wheezing and asthma throughout follow-up. In genetic analysis, compared with the Normal trajectory, the pooled relative risk ratio (95%CI) per allele was: 0.96 (0.92, 1.01), P=0.10, for Persistently High; 1.01 (0.99, 1.03), P=0.49 for Below Average; and 1.05 (1.003, 1.09), P=0.036 for Persistently Low. The majority of children in low V'maxFRC trajectory in infancy (~75%) did not progress to low FEV1 trajectory in childhood. Early-life factors associated with Persistently Low trajectory included recurrent wheeze with severe exacerbations, early allergic sensitization, and tobacco smoke exposure.

Conclusion: Reducing childhood smoke exposure and minimising the risk of early-life sensitization and wheezing exacerbations may reduce the risk of diminished lung function in early adulthood.

Lung function trajectories from pre-school age to adulthood and their associations with early life factors: a retrospective analysis of three population-based birth cohort studies

Danielle C. M. Belgrave PhD[1], Raquel Granell PhD[2], Professor Steve W Turner MD[3], John A. Curtin PhD[4], Professor Iain E. Buchan PhD[5], Professor Peter N Le Souëf MD[6], Professor Angela Simpson MD[4§], Professor A. John Henderson MD[2§], Professor Adnan Custovic MD[1§]

[§]Equal contribution

[1] Section of Paediatrics, Department of Medicine, Imperial College London, UK

[2]Department of Population Health Sciences, Bristol Medical School, University of Bristol, UK

[3]Child Health, University of Aberdeen, Aberdeen

[4]Division of Infection, Immunity and Respiratory Medicine, School of Biological Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Science Centre, UK

[5]Health Informatics, Faculty of Biology, Medicine and Health, University of Manchester, UK

[6]School of Paediatrics and Child Health, University of Western Australia, Perth, Australia

Correspondence and requests for reprints:

Professor A. John Henderson MD, Department of Population Health Sciences, Bristol Medical School, University of Bristol, UK

Email: a.j.henderson@bristol.ac.uk

Abstract word count: 253

Word count: 3807

Contribution: AC, AJH and AS conceived the idea; DB, RG and IEB provided input on the methodology for analyzing the data; DB, RG and JAC carried out the analyses; AC, AJH, DB, SWT, PNLS and AS interpreted the data; all authors wrote the report

ABSTRACT

Background: Maximal lung function in early adulthood is an important determinant of mortality and COPD. We investigated whether there are distinct trajectories of lung function during childhood, and whether these extend to adulthood and infancy.

Methods: To ascertain trajectories of forced expiratory volume in 1 second (FEV_1), we studied two population-based birth cohorts with repeat spirometry from childhood into early adulthood (1046 and 1390 participants). To ascertain whether trajectories extend to early life, we used a third cohort with repeat lung function measures in infancy (V'_{maxFRC}) and childhood (FEV_1; n=196). We identified trajectories using latent profile modelling. We created an allele score to investigate genetic associations of trajectories, and constructed a multivariable model to identify their early-life predictors.

Results: We identified four childhood FEV_1 trajectories: Persistently High; Normal; Below Average; and Persistently Low. Persistently Low trajectory (~5% of participants) was associated with persistent wheezing and asthma throughout follow-up. In genetic analysis, compared with the Normal trajectory, the pooled relative risk ratio (95%CI) per allele was: 0.96 (0.92, 1.01), P=0.10, for Persistently High; 1.01 (0.99, 1.03), P=0.49 for Below Average; and 1.05 (1.003, 1.09), P=0.036 for Persistently Low. The majority of children in low V'_{maxFRC} trajectory in infancy (~75%) did not progress to low FEV_1 trajectory in childhood. Early-life factors associated with Persistently Low trajectory included recurrent wheeze with severe exacerbations, early allergic sensitization, and tobacco smoke exposure.

Conclusion: Reducing childhood smoke exposure and minimising the risk of early-life sensitization and wheezing exacerbations may reduce the risk of diminished lung function in early adulthood.

Research in context

Evidence before this study

Early life factors, such as weight at birth, are associated with low lung function and chronic obstructive pulmonary disease (COPD) in late adulthood. Lung function development reaches a plateau in early adult life, and low lung function in young adults is associated with early mortality from all causes. Failure to attain maximal lung function at its plateau is associated with COPD in later life, even when with physiological rate of decline of lung function is maintained. This strengthens the evidence that early life influences may be critical for normal lung function growth in childhood and COPD pathogenesis.

Added value of this study

Using data from population-based, birth cohort studies, we demonstrated four discrete trajectories of FEV_1 development from early childhood to young adulthood. Persistently low FEV_1 was associated with wheezing and asthma through childhood, tobacco smoke exposure and was predicted by severe recurrent wheezing and allergic sensitization by age 3 years. We were able to determine from a third independent cohort that the majority of infants with low infant lung function trajectory during the first year after birth transitioned to normal or above average FEV_1 trajectories

Implications of all the available evidence

A persistently low trajectory of FEV_1 development can be identified during childhood, which is associated with potentially modifiable influences in early childhood. This trajectory was replicated in a companion study that followed FEV_1 growth from childhood to late adulthood. Although perinatal factors are associated with low lung function during childhood and with later COPD, the majority of infants with low lung function trajectories during the first year appeared to recover to average or above average FEV_1 growth in later childhood. Interventions to maximise lung growth in early childhood may modify the risk of COPD in older age.

3

INTRODUCTION

Low forced expiratory volume in 1 second (FEV_1) in early adult age (at its physiological plateau) is as important in the genesis of chronic obstructive pulmonary disease (COPD) as a rapid decline of FEV_1 in later years.[1-3] Furthermore, low lung function in young adult life is associated with early mortality,[4] and low FEV_1 is a marker of premature death from all causes.[5] A recent study by Agusti *et al* has shown that low lung function in early adulthood (around the time of the physiological plateau) is associated with earlier onset of COPD, other chronic diseases, and death.[6] An accompanying Editorial highlighted the need to use longitudinal studies and modern statistical approaches to gain better understanding of the causal pathways, understand the critical windows for lung function development, and identify targets for early intervention.[4,5,7-11]

Hypothesis-driven analyses in unselected longitudinal cohorts and children with asthma have shown that lung function in adult life is in part determined by childhood events. The observation that FEV_1 was consistently lower amongst individuals with clinical phenotypes such as severe asthma,[8] and persistent wheezing[9] provided foundations for the concept that lung function tracks from school-age to adulthood.[9,10] This is supported by the data-driven analysis of spirometry records from childhood (age 11 years) to the fourth decade of life in Tucson cohort, which identified two distinct lung function trajectories (low and normal).[11] Persistently low trajectory comprised nearly 10% participants, and individuals in this class were predisposed to COPD later in life.

We hypothesized that a population comprises several distinct groups of individuals with similar lung function trajectories from pre-school age to adolescence, which extend to the physiological plateau in early adulthood, and that these trajectories are in part established at birth. However, factors which determine lung function growth may differ from those affecting the rate of decline after the plateau,[12] and inference from models that combine a limited number of measures in childhood (before peak lung

4

function is reached in early adulthood), with those in later adulthood (during the decline phase), may not adequately disaggregate the effects of risk factors which differentially impact growth and decline.

To address our hypotheses, and ascertain trajectories during the lung function growth and their risk factors, we combined data from three population-based birth cohorts in which lung function was assessed on multiple occasions from infancy, through early school age and adolescence, into early adulthood. We first set out to establish whether a method to categorise FEV_1 trajectories using latent profile analysis could be applied across different cohorts, and whether it yields consistent results across different populations. We proceeded to identify genetic and early-life factors associated with FEV_1 trajectories at the population level. Finally, we investigated whether FEV_1 trajectories track back to lung function patterns in infancy.

METHODS

Study design, setting and participants

We studied three population-based birth cohorts: Manchester Asthma and Allergy Study (MAAS)[13] and the Avon Longitudinal Study of Parents and Children (ALSPAC)[14] from the UK STELAR consortium[15], and the Australian Perth Infant Asthma Follow-up (PIAF) study.[16] All participants were recruited prenatally and followed prospectively. All studies were approved by research ethics committees. Informed consent was obtained from parents, and study subjects gave their consent when applicable.

Data sources/measurement

We investigated trajectories of FEV_1 from early school-age to adolescence in MAAS[13] and replicated these using data from ALSPAC from school age to early adultood[14]. We ascertained whether FEV_1 trajectories extended to lung function in early infancy using data from PIAF.[17]

We performed spirometry according to ATS/ERS criteria[18,19] at ages 5, 8, 11 and 16 years in MAAS; ages 8, 15 and 24 years in ALSPAC; and 6, 12 and 18 years in PIAF. We expressed FEV_1 as % predicted at each age.[20] In PIAF, we assessed infant lung function at ages 1, 6 and 12 months using the rapid thoracoabdominal compression test to determine the maximal flow at functional residual capacity (V'_{maxFRC}).[16,17] In MAAS we measured specific airways resistance (sR_{aw}) at age 3 years.[21] In MAAS and ALSPAC, we measured post-bronchodilator FEV_1 after administration of 400 μg of salbutamol (5, 11 and 16 years in MAAS, and 15 and 24 years in ALSPAC), and assessed assessed airway hyper-reactivity (AHR) using methacholine challenge.[22]

Details of clinical follow up, genotyping (MAAS and ALSPAC), and definitions of clinical outcomes (asthma,[23] wheeze phenotypes,[24] severe asthma exacerbations,[25] airway hyper-reactivity (AHR),[26] lower respiratory tract infections [LRTIs][27]), and environmental exposures are presented in the supplementary appendix.

Statistical Analysis

Longitudinal profile analysis to assess prototypical FEV_1 trajectories: We used a latent profile model to assign children to their most likely group based on the profile of FEV_1 over time. Latent profile modelling has been described extensively elsewhere.[28] Briefly, we assumed that each child belonged to one of a set of N latent profiles, the number or size of which were unknown *a priori*. We used two-level random intercept and random coefficients regression models to examine trajectory classes with no change over time, and under the assumption that the profile of FEV_1 changes linearly over time. The models were compared for goodness-of-fit using the Bayesian Information Criterion (BIC). For each child, the posterior probability of belonging to each of the classes was calculated, and children were assigned to the latent profile with the largest probability. Latent profile modelling was carried out using the gllamm (generalised linear latent and mixed models) package implemented in Stata (StataCorp, College Station, TX). Characteristics of trajectories were investigated using multinomial regression.

Genetic associations with FEV_1 trajectories: We constructed a weighted allele score from 77 SNPs found to have genome-wide evidence for association with FEV_1/FVC, or FEV_1 decline in adults[29] (Table S1, supplementary appendix). The dosage of the risk allele at each locus was multiplied by a SNP-specific weight (linear coefficient [beta] divided by the average of 77 linear coefficients), then averaged across SNPs. SNP-specific risk alleles and weights were based on the reported meta-analysis.[29] The units of the derived genetic score were per-beta per-SNP. Setting the 'Normal' trajectory as the baseline group, we used a multinomial logistic regression model to investigate the risk of the membership of other trajectories per 1-unit increase of the genetic score. The independent associations of the resulting genetic risk score with FEV_1 trajectories in MAAS and ALSPAC were meta-analysed with fixed and random effect models using the method of Mantel and Haenszel (*metan* command in Stata). Using multinomial logistic regression models, we also calculated the relative risk ratios for 'Persistently Low'

7

(vs. 'Normal' trajectory) in extreme quantiles of the allelic risk score in two cohorts independently (deciles in ALSPAC; eight quantiles in MAAS).

Identification of the early-life predictors of FEV$_1$ trajectories: We constructed a multivariable regression model to identify early-life predictors of subsequent latent profiles of FEV$_1$. We carried out these analyses in MAAS, in which we had richer data collected in early life. We used receiver operating characteristic (ROC) curves to evaluate the predictive ability of these predictors.

Role of the funding source

The study sponsors had no role in study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the paper for publication.

RESULTS

We included 1046 participants from MAAS, 1390 from ALSPAC, and 196 from PIAF. Participant flow is presented in the Supplementary appendix.

FEV$_1$ trajectories from early school-age to young adulthood (MAAS and ALSPAC)

We selected 4-class model in MAAS and ALSPAC (please see Supplementary appendix and Table S2). The posterior probability of class membership was high in both cohorts (Table S3). The proportion of participants and the mean FEV$_1$ over time in each of the classes were similar across the cohorts (Table S4). We labelled the classes (FEV$_1$ trajectories) as: (1) Persistently High; (2) Normal; (3) Below Average; and (4) Persistently Low (Figure 1). Children within these four trajectories had stable lung function that tracked from age 5 to age 16 years (MAAS), and from ages 8 to 24 years (ALSPAC), with no overlap in FEV$_1$ at any time. The solutions remained stable and results unchanged when we included only children who had spirometry on at least 3 follow-ups (Tables S5-6).

Characteristics of children in each of the FEV$_1$ trajectories are shown in Tables S7 and S8. Persistently Low trajectory was consistently associated with wheeze and asthma throughout the follow-up in both cohorts; in MAAS, almost half the children in this trajectory had asthma, and in ALSPAC this was approximately one third, significantly higher than in the rest of the population. The proportion of children with LRTIs in the first 3 years (data from medical records, available only in MAAS[27]) was significantly higher in the Persistently Low trajectory (32.3%) compared to the Normal trajectory (12.1%). Post-bronchodilator FEV$_1$ and FEV$_1$/FVC followed similar patterns as the baseline FEV$_1$ in both cohorts. In MAAS, sR$_{aw}$ at age 3 years was significantly higher in Persistently Low compared to other trajectories, indicating diminished lung function (P=0.001). Persistently Low trajectory was significantly associated with AHR in both cohorts. There was no association between trajectory membership and

9

allergic sensitisation in adolescence in MAAS, or at age 7 years in ALSPAC. Sensitisation in early childhood in MAAS was significantly more common in Persistently Low trajectory.

There was no difference in gender between trajectories. Birth weight was significantly higher in the Persistently High trajectory in ALSPAC, but not in MAAS. In MAAS, children in the Persistently High trajectory were less likely to be exposed to environmental tobacco smoke (ETS) compared to all others (Figure 2a). Children continuously exposed to ETS from birth to age 16 years were at increased risk of belonging to the Below Average and Persistently Low trajectories. We observed similar trends in ALSPAC, with Persistently High trajectory having the lowest proportion of ETS-exposed participants (Figure 2b).

Genetic associations

Genotyping data was available in 927 participants in ALSPAC, and 902 in MAAS. The weighted genetic score based on 77 SNPs was inversely associated with the order of FEV_1 trajectories (Persistently High to Persistently Low) in both cohorts (Table S9). The meta-analysis is shown in Figure 3: compared with the Normal trajectory, the pooled relative risk ratio (95%CI) per allele was: 0.97 (0.92, 1.01), P=0.13, for Persistently High; 1.01 (1.00, 1.03), P=0.49 for Below Average; and 1.05 (1.00, 1.09), P=0.036 for Persistently Low. Results using the inverse variance method were very similar. Risk of Persistently Low FEV_1 (vs. Normal) in 10th decile of risk score (vs. 1st decile) was 2.03 (0.62, 6.67), p=0.24 in ALSPAC. In MAAS, the risk of Persistently Low FEV_1 in the highest compared with the lowest of eight quantiles was 4.19 (0.83-21.05), p=0.08.

Do FEV_1 trajectories extend back to lung function in early life (PIAF)?

In PIAF, a 3-class model provided the optimal solution for FEV_1 (Above Average, Normal and Below Average, Figure 4a), and a 2-class model for V'_{maxFRC} (Above Average and Below Average, Figure 4b). Among children in the Below Average FEV_1 trajectory, V'_{maxFRC} was significantly lower compared to other

10

trajectories at age 1 month (P=0.05) and 6 months (P=0.018), but not at age 1 year (Table S10). Although children in the Below Average V'_{maxFRC} trajectory were significantly more likely to remain in the Below Average FEV_1 trajectory (OR=3.63, 95%CI 1.21–10.90, P=0.022), the majority of these children (~75%) transitioned to the Above Average (12.2%) or Normal (63.4%) FEV_1 trajectories in later childhood (Figure 4, panel c).

Early-life predictors of FEV_1 trajectories (MAAS)

Children with recurrent severe wheeze exacerbations by age 3 years, with allergic sensitization at age 3 years, and who have been exposed to ETS by age 3 years, were more likely to belong to the Persistently Low trajectory (Table S10). This set of variables provided a good predictive ability in MAAS for discriminating between subjects in Persistently Low from Persistently High trajectory (Figure 5, Area Under the ROC Curve-AUROC=90.7%, sensitivity=81.8%; specificity=80.95%), and a fair predictive ability for discriminating between subjects in Persistently Low from all others (Figure S1, AUROC=77.2%, sensitivity=81.8%; specificity=67.8%).

We found an association between low birth weight and low childhood FEV_1 trajectories. In MAAS, which included no preterm infants, low birth weight (LBW) at term (<2500g) was associated with Peristently Low FEV_1 (Odds ratio 3.39 [95% CI 0.98-11.77], p=0.055). In ALSPAC, LBW adjusted for preterm (<37 weeks) gestation at birth was associated with Persistently Low FEV_1 (OR 1.98 [0.65-5.99], p=0.226, and Below Average FEV_1 (OR 1.84 [0.91-3.73], p=0.09).

DISCUSSION

In two population-based birth cohorts, using data-driven analyses we identified four distinct lung function trajectories extending from early school age, through mid-school age, into adolescence and early adulthood. Subjects assigned to the Persistently Low trajectory, representing approximately 1 in 20 participants, had low lung function in the third decade of life, around the time of the physiological plateau. Individuals in this trajectory had FEV_1 below 80% predicted from early childhood to the point beyond which lung function starts to decline with aging.[3,20] In the Persistently Low trajectory, both baseline and post-bronchodilator FEV_1 was up to 25% lower, and FEV_1/FVC up to 10% lower than the next lowest category. This tracking of lung function from early school-age is consistent with previous reports from population-based studies[9,10], and provides confirmatory empirical evidence of this phenomenon based on a hypothesis-free approach. In addition to using data-driven analyses, another novel feature of the present study was the extension of lung function trajectories to early childhood, where spirometry is not practicable, and the observation that "low" lung function recovered in the majority of cases. At group level, children in the Low FEV_1 trajectory in PIAF had significantly diminished lung function in infancy (measured by V'_{maxFRC}), which is very similar to the finding in MAAS that, as a group, those in Persistently Low FEV_1 trajectory had significantly diminished lung function at age 3 years (measured by sR_{aw}). However, comparison of lung function trajectories in infancy and in school-age in PIAF has shown that although infants in the Below Average V'~~maxFRC~~ trajectory were significantly more likely than children in the higher V'_{maxFRC} categories to be classified in the Below Average FEV_1 trajectory in childhood, the majority of these children (~75%) transitioned to the Above Average (12.2%) or Normal (63.4%) FEV_1 trajectories. These findings are consistent with previous results from PIAF obtained using standard hypothesis-testing analyses,[17,30] and support the concept of postnatal "recovery" of low neonatal lung function. The application of the same data-driven approach to longitudinal analysis as used in the MAAS, ALSPAC and Tasmanian Longitudinal Health Study (TAHS),[31] allows some inference to

the general population on the patterns of lung function from one month of age to the sixth decade. Disaggregation of this population into discrete trajectories suggests that the opportunity for lung function recovery extends beyond the neonatal period.

Our data revealed no evidence of groups of children with declining FEV_1. A report from TAHS published in this issue of the journal, which used a similar statistical approach to discovery of lung function trajectories and spanned childhood (7-18 years) and later adulthood (45-55 years), described six trajectories, four of which were strikingly similar to ours.[31] One further trajectory with below average FEV_1 by age 18, and accelerated decline in later adulthood after reaching the plateau, could not be captured within our data set. This trajectory was characterized by both childhood factors and personal smoking in adulthood, which were independent risk factors. Together, our studies and TAHS elucidate the developmental patterns of lung function through the life-course, confirm their link with COPD, and provide evidence that different factors affect lung function growth from those affecting lung function decline. We have shown that Persistently Low trajectory is determined by both genetic factors and early-life events. We showed relatively weak evidence that SNPs discovered through GWAS of adult lung function[29] were associated with a low childhood trajectory of FEV_1. This suggests an influence of at least some of the included variants on lung function in early life but, as they were derived from studies of adults, it is likely that they consisted of a combination of factors influencing lung function growth and decline. A recent study of adults with repeat lung function measures showed that genetic variants associated with cross-sectional lung function in adults were not associated with longitudinal decline, suggesting that some of the observed association may be explained by developmental influences[32]. We did not have sufficient statistical power in this sample to investigate the individual contributions of risk alleles, nor to construct a *de novo* GWAS.

Individuals with Persistently Low FEV_1 in childhood have low maximally attained lung function in early adulthood, and are consequently at higher risk of COPD[31], with or without additional adverse exposures

13

in adulthood. Personal smoking acts as an independent risk factor which contributes to accelerated lung function decline in middle adulthood, predominantly among individuals with below average lung function.[31] In this group, in addition to early-life risk factors contributing to the diminished lung function at physiological plateau, adverse exposures in adult life lead to COPD development. Childhood asthma and early-life LRTIs are associated with low lung function trajectories and higher COPD risk in both scenarios. We found an association between asthma reported at all available ages and persistently Low FEV_1 trajectory. This is consistent with previously reported associations between persistent asthma in childhood and subsequent lung function,[8, 9] but we acknowledge the potential for misclassification of symptoms associated with pre-existing airway obstruction and asthma diagnosis, particularly when relying on self-report from questionnaires for the latter. This limits interpretation of the causal pathway between asthma in childhood and low lung function trajectories in later life.

Although children with severe wheeze and early sensitization had lower FEV_1 throughout childhood, their FEV_1 remained relatively stable. In contrast, a recent study of FEV_1 trajectories between ages 7 and 26 years among patients with childhood asthma described two trajectories characterized by decline (Normal growth and early decline, and Reduced growth with early decline).[33] However, direct comparison with our findings is not possible, as our cohorts are population-based, and we used data-driven approach rather than predetermined categorization to identify trajectories. In MAAS, similar proportions of children (approximately one third) in Persistently Low trajectory had persistent wheeze and transient early wheeze, whist <10% had late-onset wheezing. This is in contrast with the Australian high-risk cohort, which suggested that only persistent wheeze is associated with reduced growth in FEV_1 over adolescence, but that transient wheeze is a benign condition with no sequelae.[34] In MAAS, an important predictor of subsequent Persistently Low FEV_1 trajectory was severe exacerbation of wheezing in early life, rather than only wheezing persistence. Low birth weight, which is associated with tobacco some exposure during pregnancy and early onset wheezing in infancy was associated with low

childhood FEV$_1$ trajectories in our study independently of preterm delivery. This may be explained by restricted intrauterine growth influencing airway development and subsequent lung function[35] with an associated tendency to wheezing in the context of LRTIs.

Persistently Low lung function trajectory was also identified in the Tucson cohort (comprising nearly 10% participants),[11] and is remarkably similar to our Persistently Low trajectory, which comprised a sizeable proportion (3.4%-6.7%) of our study populations. Assuming that 13.3% of the population develop COPD (http://www.hse.gov.uk/Statistics/causdis/copd/copd.pdf), and that among these, subjects in the Persistently Low trajectory would likely be diagnosed with COPD,[34] the calculated population attributable risk is 22.9%-46.7%. This also represents the estimated reduction in COPD which would occur if early interventions improve lung function for this group. To inform potential strategies for improving lung function in early adulthood, we have identified early-life risk factors associated with Persistently Low trajectory, which included early-life ETS exposure, recurrent severe exacerbations of wheezing, and early allergic sensitisation. A recent study suggested that adult smoking, but not parental smoking history, can negatively influence adult lung function trajectory and modify the adverse impact of early-life exposures.[35] Our and TAHS data demonstrate that both are important. We identified an association between the absence of early-life tobacco smoke exposure and Persistently High trajectory, which In TAHS had markedly lower risk of COPD at age 55 years.[31] Thus, while tobacco smoke exposure will not necessarily lead to a disease, the absence of this exposure is associated with exceptionally good lung function and subsequent health. The risk of being in one of the diminished lung function profiles was decreased by 76% among children who were not exposed to parental smoking, suggesting that ETS exposure in early childhood may be causally related to COPD development, and that this may be mediated via the effect of early-life exposure on subsequent trajectories of lung function into adulthood.

Strengths and limitations

One of the limitations of our study is that it was not possible to have identical measures of lung function through the life-course in our cohorts. We did not measure spirometry before age five years, since forced expiratory manoeuvres are difficult to carry out in young children. We used different cohorts and age-appropriate measures of pulmonary function to make inference about the life-course (rapid thoraco-abdominal compression in infancy[16], plethysmography in pre-school age[36], spirometry from school age onwards), which may represent the best available approximation of what might be happening from birth to adulthood. Another limitation is that the questions we used contained small differences in wording.[37] ALSPAC included pre-term babies, which may explain slightly higher proportion of children in the Persistently Low trajectory in ALSPAC (6.7%, compared to 3.4% in MAAS). One of the strengths of this study is the ability to disambiguate the temporal structure of profiles during childhood, allowing us to identify patterns during the lung function growth phase. It is striking that although we modelled data collected at different time points, and separately for each cohort, the results indicated almost identical latent profiles. In PIAF, we did not observe Persistently Low FEV_1 trajectory. This is not surprising, given a relatively small proportion of children in this trajectory (~5%), and a considerably smaller size of PIAF compared to MAAS and ALSPAC.

We wish to emphasise that our current study is exploratory in nature. Therefore, the results should be interpreted with caution, and need to be validated in a larger population. We are also cognizant of the potential effect of the withdrawal/missing data rates. We carried out our analyses under the assumption that data was missing at random. To test this assumption, we carried out analyses using both children with at least three data points available, and those with at least one time point. Both analyses gave consistent optimal goodness-of-fit using the BIC, and the child class assignments were stable across the two analyses. This suggests that the missing at random assumption was plausible,

given that if children with missing data points were not missing at random, there would have been a higher mismatch between classes.

Conclusion

Using data from population-based birth cohort studies, we demonstrated four discrete trajectories of FEV_1 development from early childhood to young adulthood. Interventions to reduce the risk of Persistently Low FEV_1 trajectory should start in early childhood, and should include aggressive strategies to reduce tobacco smoke exposure, and novel strategies to prevent early-life sensitization and exacerbations of early-childhood wheezing.

REFERENCES

1.   Lange P, Celli B, Agusti A, et al. Lung-Function Trajectories Leading to Chronic Obstructive Pulmonary Disease. *N Engl J Med* 2015; **373**(2): 111-22.

2.   Martinez FD. Early-Life Origins of Chronic Obstructive Pulmonary Disease. *N Engl J Med* 2016; **375**(9): 871-8.

3.   Postma DS, Bush A, van den Berge M. Risk factors and early origins of chronic obstructive pulmonary disease. *Lancet* 2015; **385**(9971): 899-909.

4.   Vasquez MM, Zhou M, Hu C, Martinez FD, Guerra S. Low Lung Function in Young Adult Life Is Associated with Early Mortality. *American journal of respiratory and critical care medicine* 2017; **195**(10): 1399-401.

5.   Young RP, Hopkins R, Eaton TE. Forced expiratory volume in one second: not just a lung function test but a marker of premature death from all causes. *European Respiratory Journal* 2007; **30**(4): 616-22.

6.   Agusti A, Noell G, Brugada J, Faner R. Lung function in early adulthood and health in later life: a transgenerational cohort analysis. *The Lancet Respiratory medicine* 2017; **5**(12): 935-45.

7.   Taylor-Robinson DC, Stanojevic S. Low lung function in early adulthood: morbidity and death. *The Lancet Respiratory medicine* 2017; **5**(12): 910-1.

8.   Tai A, Tran H, Roberts M, et al. Outcomes of childhood asthma to the age of 50 years. *J Allergy Clin Immunol* 2014; **133**(6): 1572-8 e3.

9.   Sears MR, Greene JM, Willan AR, et al. A longitudinal, population-based, cohort study of childhood asthma followed to adulthood. *New England Journal of Medicine* 2003; **349**(15): 1414-22.

10.   Phelan PD, Robertson CF, Olinsky A. The Melbourne Asthma Study: 1964-1999. *J Allergy Clin Immunol* 2002; **109**(2): 189-94.

11.   Berry CE, Billheimer D, Jenkins IC, et al. A Distinct Low Lung Function Trajectory from Childhood to the Fourth Decade of Life. *American journal of respiratory and critical care medicine* 2016; **194**(5): 607-12.

12.   Grol MH, Gerritsen J, Vonk JM, et al. Risk factors for growth and decline of lung function in asthmatic individuals up to age 42 years. A 30-year follow-up study. *American journal of respiratory and critical care medicine* 1999; **160**(6): 1830-7.

13.   Custovic A, Simpson BM, Murray CS, Lowe L, Woodcock A. The National Asthma Campaign Manchester Asthma and Allergy Study. *Pediatr Allergy Immunol* 2002; **13 Suppl 15**: 32-7.

14.   Boyd A, Golding J, Macleod J, et al. Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol* 2013; **42**(1): 111-27.

15.   Custovic A, Ainsworth J, Arshad H, et al. The Study Team for Early Life Asthma Research (STELAR) consortium 'Asthma e-lab': team science bringing data, methods and investigators together. *Thorax* 2015; **70**(8): 799-801.

16.   Young S, Le Souef PN, Geelhoed GC, Stick SM, Turner KJ, Landau LI. The influence of a family history of asthma and parental smoking on airway responsiveness in early infancy. *N Engl J Med* 1991; **324**(17): 1168-73.

17.   Turner S, Fielding S, Mullane D, et al. A longitudinal study of lung function from 1 month to 18 years of age. *Thorax* 2014; **69**(11): 1015-20.

18.   Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. *European respiratory journal* 2005; **26**(2): 319-38.

19.   Beydon N, Davis SD, Lombardi E, et al. An official American Thoracic Society/European Respiratory Society statement: pulmonary function testing in preschool children. *American journal of respiratory and critical care medicine* 2007; **175**(12): 1304-45.

20.   Stanojevic S, Wade A, Cole TJ, et al. Spirometry centile charts for young Caucasian children: the Asthma UK Collaborative Initiative. *American journal of respiratory and critical care medicine* 2009; **180**(6): 547-52.

21.   Lowe L, Murray CS, Custovic A, et al. Specific airway resistance in 3-year-old children: a prospective cohort study. *Lancet* 2002; **359**(9321): 1904-8.

22.   Crapo RO, Casaburi R, Coates AL, et al. Guidelines for methacholine and exercise challenge testing-1999. This official statement of the American Thoracic Society was adopted by the ATS Board of Directors, July 1999. *American journal of respiratory and critical care medicine* 2000; **161**(1): 309-29.

23.   Lodrup Carlsen KC, Roll S, Carlsen KH, et al. Does pet ownership in infancy lead to asthma or allergy at school age? Pooled analysis of individual participant data from 11 European birth cohorts. *PLoS One* 2012; **7**(8): e43214.

24.   Lowe LA, Simpson A, Woodcock A, Morris J, Murray CS, Custovic A. Wheeze phenotypes and lung function in preschool children. *American journal of respiratory and critical care medicine* 2005; **171**(3): 231-7.

25.   Reddel HK, Taylor DR, Bateman ED, et al. An official American Thoracic Society/European Respiratory Society statement: asthma control and exacerbations: standardizing endpoints for clinical asthma trials and clinical practice. *American journal of respiratory and critical care medicine* 2009; **180**(1): 59-99.

26.   Nicolaou NC, Lowe LA, Murray CS, Woodcock A, Simpson A, Custovic A. Exhaled breath condensate pH and childhood asthma: unselected birth cohort study. *American journal of respiratory and critical care medicine* 2006; **174**(3): 254-9.

27.   Semic-Jusufagic A, Belgrave D, Pickles A, et al. Assessing the association of early life antibiotic prescription with asthma exacerbations, impaired antiviral immunity, and genetic variants in 17q21: a population-based birth cohort study. *The Lancet Respiratory medicine* 2014; **2**(8): 621-30.

28.   Oberski D. Mixture Models: Latent Profile and Latent Class Analysis. . In: Robertson J KM, ed. Modern Statistical Methods for HCI Human–Computer Interaction Series Springer, Cham; 2016.

29.   Wain LV, Shrine N, Artigas MS, et al. Genome-wide association analyses for lung function and chronic obstructive pulmonary disease identify new loci and potential druggable targets. *Nat Genet* 2017; **49**(3): 416-25.

30.   Turner SW, Palmer LJ, Rye PJ, et al. Infants with flow limitation at 4 weeks: outcome at 6 and 11 years. *American journal of respiratory and critical care medicine* 2002; **165**(9): 1294-8.

31.   Bui DS BJ, Lodge C L, Lowe A, Perret J, Bowatte G, Thompson BR, Hamilton GS, Frith PA, James AL, Johns DP, Thomas PS, Jarvis D, Svanes C, Morrison S, Feather I, Allen KJ, Wood-Baker R, Bui MQ, Russell M, Abramson MJ, Walters EH, Matheson MC, Dharmage SC. Lung function trajectories from childhood to the six decade, their associated childhood factors and consequences: a population based longitudinal cohort. *Submitted* 2017.

32.   John C, Soler Artigas M, Hui J, Nielsen SF, Rafaels N, Paré PD, et al. Genetic variants affecting cross-sectional lung function in adults show little or no effect on longitudinal lung function decline. *Thorax* 2017;**72**(5):400-408.

33.   McGeachie MJ, Yates KP, Zhou X, et al. Patterns of growth and decline in lung function in persistent childhood asthma. *The New England Journal of Medicine* 2016; (19): 1842.

34.   Lodge CJ, Lowe AJ, Allen KJ, et al. Childhood wheeze phenotypes show less than expected growth in FEV1 across adolescence. *American journal of respiratory and critical care medicine* 2014; **189**(11): 1351-8.

35.   den Dekker HT, Jaddoe VWV, Reiss IK, de Jongste JC, Duijts L. Fetal and Infant Growth Patterns and Risk of Lower Lung Function and Asthma. The Generation R Study. *Am J Respir Crit Care Med* 2018;**197**(2):183-192.

34.	Vogelmeier CF, Criner GJ, Martinez FJ, et al. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease 2017 Report. GOLD Executive Summary. *American journal of respiratory and critical care medicine* 2017; **195**(5): 557-82.

35.	Allinson JP, Hardy R, Donaldson GC, Shaheen SO, Kuh D, Wedzicha JA. Combined Impact of Smoking and Early Life Exposures on Adult Lung Function Trajectories. *American journal of respiratory and critical care medicine* 2017.

36.	Belgrave DC, Buchan I, Bishop C, Lowe L, Simpson A, Custovic A. Trajectories of lung function during childhood. *American journal of respiratory and critical care medicine* 2014; **189**(9): 1101-9.

37.	Belgrave DC, Granell R, Simpson A, et al. Developmental profiles of eczema, wheeze, and rhinitis: two population-based birth cohort studies. *PLoS medicine* 2014; **11**(10): e1001748.

LEGENDs FOR FIGURES

Figure 1: Mean % Predicted FEV_1 over time by FEV_1 trajectory:

   a) MAAS

   b) ALSPAC

Figure 2: % Children exposed to Environmental Tobacco Smoke over time by FEV_1 trajectory:

   a) MAAS

   b) ALSPAC

Figure 3: Meta-analysis (random effects model) of weighted genetic score based on 77 SNPs with FEV_1

trajectories

Figure 4: Lung function trajectories in PIAF:

   a) Mean % Predicted FEV_1 over time by FEV_1 trajectory

   b) Mean % Predicted V'_{maxFRC} over time by V'_{maxFRC} trajectory

   c) Membership of FEV_1 trajectory in relation to the V'_{maxFRC} trajectories

Figure 5:   Receiver Operating Characteristic Curve showing predictive ability of a logistic regression

analysis using the Persistently Low *vs*. Persistently High trajectory as the outcome (or predicted)

measure with the following covariates: (1) recurrent wheeze with severe exacerbations by age 3 years,

(2) at least one positive skin test by age 3 years, and (3) Environmental Tobacco Smoke exposure by age

3 years

Lung function trajectories from pre-school age to adulthood and their associations with early life factors: a retrospective analysis of three population-based birth cohort studies

Danielle C. M. Belgrave PhD[1], Raquel Granell PhD[2], Professor Steve W Turner MD[3], John A. Curtin PhD[4], Professor Iain E. Buchan PhD[5], Professor Peter N Le Souëf MD[6], Professor Angela Simpson MD[4§], Professor A. John Henderson MD[2§], Professor Adnan Custovic MD[1§]

[§]Equal contribution

[1] Section of Paediatrics, Department of Medicine, Imperial College London, UK

[2]Department of Population Health Sciences, Bristol Medical School, University of Bristol, UK

[3]Child Health, University of Aberdeen, Aberdeen

[4]Division of Infection, Immunity and Respiratory Medicine, School of Biological Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Science Centre, UK

[5]Health Informatics, Faculty of Biology, Medicine and Health, University of Manchester, UK

[6]School of Paediatrics and Child Health, University of Western Australia, Perth, Australia

Correspondence and requests for reprints:

Professor A. John Henderson MD, Department of Population Health Sciences, Bristol Medical School, University of Bristol, UK

~~Tel: +44 117 331 0071,~~ Email: a.j.henderson@bristol.ac.uk

Abstract word count: 253

Word count: 3807

Contribution: AC, AJH and AS conceived the idea; DB, RG and IEB provided input on the methodology for analyzing the data; DB, RG and JAC carried out the analyses; AC, AJH, DB, SWT, PNLS and AS interpreted the data; all authors wrote the report

ABSTRACT

Background: Maximal lung function in early adulthood is an important determinant of mortality and COPD. We investigated whether there are distinct trajectories of lung function during childhood, and whether these extend to adulthood and infancy.

Methods: To ascertain trajectories of forced expiratory volume in 1 second (FEV_1), we studied two population-based birth cohorts with repeat spirometry from childhood into early adulthood (1046 and 1390 participants). To ascertain whether trajectories extend to early life, we used a third cohort with repeat lung function measures in infancy (V'_{maxFRC}) and childhood (FEV_1; n=196). We identified trajectories using latent profile modelling. We created an allele score to investigate genetic associations of trajectories, and constructed a multivariable model to identify their early-life predictors.

Results: We identified four childhood FEV_1 trajectories: Persistently High; Normal; Below Average; and Persistently Low. Persistently Low trajectory (~5% of participants) was associated with persistent wheezing and asthma throughout follow-up. In genetic analysis, compared with the Normal trajectory, the pooled relative risk ratio (95%CI) per allele was: 0.96 (0.92, 1.01), P=0.10, for Persistently High; 1.01 (0.99, 1.03), P=0.49 for Below Average; and 1.05 (1.003, 1.09), P=0.036 for Persistently Low. The majority of children in low V'_{maxFRC} trajectory in infancy (~75%) did not progress to low FEV_1 trajectory in childhood. Early-life factors associated with Persistently Low trajectory included recurrent wheeze with severe exacerbations, early allergic sensitization, and tobacco smoke exposure.

Conclusion: Reducing childhood smoke exposure and minimising the risk of early-life sensitization and wheezing exacerbations may reduce the risk of diminished lung function in early adulthood.

Research in context

Evidence before this study

Early life factors, such as weight at birth, are associated with low lung function and chronic obstructive pulmonary disease (COPD) in late adulthood. Lung function development reaches a plateau in early adult life, and low lung function in young adults is associated with early mortality from all causes. Failure to attain maximal lung function at its plateau is associated with COPD in later life, even when with physiological rate of decline of lung function is maintained. This strengthens the evidence that early life influences may be critical for normal lung function growth in childhood and COPD pathogenesis.

Added value of this study

Using data from population-based, birth cohort studies, we demonstrated four discrete trajectories of FEV_1 development from early childhood to young adulthood. Persistently low FEV_1 was associated with wheezing and asthma through childhood, tobacco smoke exposure and was predicted by severe recurrent wheezing and allergic sensitization by age 3 years. We were able to determine from a third independent cohort that the majority of infants with low infant lung function trajectory during the first year after birth transitioned to normal or above average FEV_1 trajectories

Implications of all the available evidence

A persistently low trajectory of FEV_1 development can be identified during childhood, which is associated with potentially modifiable influences in early childhood. This trajectory was replicated in a companion study that followed FEV_1 growth from childhood to late adulthood. Although perinatal factors are associated with low lung function during childhood and with later COPD, the majority of infants with low lung function trajectories during the first year appeared to recover to average or above average FEV_1 growth in later childhood. Interventions to maximise lung growth in early childhood may modify the risk of COPD in older age.

INTRODUCTION

Low forced expiratory volume in 1 second (FEV_1) in early adult age (at its physiological plateau) is as important in the genesis of chronic obstructive pulmonary disease (COPD) as a rapid decline of FEV_1 in later years.[1-3] Furthermore, low lung function in young adult life is associated with early mortality,[4] and low FEV_1 is a marker of premature death from all causes.[5] A recent study by Agusti *et al* has shown that low lung function in early adulthood (around the time of the physiological plateau) is associated with earlier onset of COPD, other chronic diseases, and death.[6] An accompanying Editorial highlighted the need to use longitudinal studies and modern statistical approaches to gain better understanding of the causal pathways, understand the critical windows for lung function development, and identify targets for early intervention.[4,5,7-11]

Hypothesis-driven analyses in unselected longitudinal cohorts and children with asthma have shown that lung function in adult life is in part determined by childhood events. The observation that FEV_1 was consistently lower amongst individuals with ~~certain~~ clinical phenotypes (such as severe asthma,[8] and persistent wheezing[9]) provided foundations for the concept that lung function tracks from school-age to adulthood.[9,10] This is supported by the data-driven analysis of spirometry records from childhood (age 11 years) to the fourth decade of life in Tucson cohort, which identified two distinct lung function trajectories (low and normal).[11] Persistently low trajectory comprised nearly 10% participants, and individuals in this class were predisposed to COPD later in life.

We hypothesized that a population comprises several distinct groups of individuals with similar lung function trajectories from pre-school age to adolescence, which extend to the physiological plateau in early adulthood, and that these trajectories are in part established at birth. However, factors which determine lung function growth may differ from those affecting the rate of decline after the plateau,[12] and inference from models that combine a limited number of measures in childhood (before peak lung

4

function is reached in early adulthood), with those in later adulthood (during the decline phase), may not adequately disaggregate the effects of risk factors which differentially impact growth and decline.

To address our hypotheses, and ascertain trajectories during the lung function growth and their risk factors, we combined data from three population-based birth cohorts in which lung function was assessed on multiple occasions from infancy, through early school age and adolescence, into early adulthood. We first set out to establish whether a method to categorise FEV_1 trajectories using latent profile analysis could be applied across different cohorts, and whether it yields consistent results across different populations. We proceeded to identify genetic and early-life factors associated with FEV_1 trajectories at the population level. Finally, we investigated whether FEV_1 trajectories track back to lung function patterns in infancy.

METHODS

Study design, setting and participants

We studied three population-based birth cohorts: Manchester Asthma and Allergy Study (MAAS)[13] and the Avon Longitudinal Study of Parents and Children (ALSPAC)[14] from the UK STELAR consortium[15], and the Australian Perth Infant Asthma Follow-up (PIAF) study.[16] All participants were recruited prenatally and followed prospectively. All studies were approved by research ethics committees. Informed consent was obtained from parents, and study subjects gave their consent when applicable.

Data sources/measurement

We investigated trajectories of FEV_1 from early school-age to adolescence in MAAS[13] and replicated these using data from ALSPAC from school age to early adultood[14]. We ascertained whether FEV_1 trajectories extended to lung function in early infancy using data from PIAF.[17]

We performed spirometry according to ATS/ERS criteria[18,19] at ages 5, 8, 11 and 16 years in MAAS; ages 8, 15 and 24 years in ALSPAC; and 6, 12 and 18 years in PIAF. We expressed FEV_1 as % predicted at each age.[20] In PIAF, we assessed infant lung function at ages 1, 6 and 12 months using the rapid thoracoabdominal compression test to determine the maximal flow at functional residual capacity (V'_{maxFRC}).[16,17] In MAAS we measured specific airways resistance (sR_{aw}) at age 3 years.[21] In MAAS and ALSPAC, we measured post-bronchodilator FEV_1 after administration of 400 μg of salbutamol (5, 11 and 16 years in MAAS, and 15 and 24 years in ALSPAC), and assessed assessed airway hyper-reactivity (AHR) using methacholine challenge.[22]

Details of clinical follow up, genotyping (MAAS and ALSPAC), and definitions of clinical outcomes (asthma,[23] wheeze phenotypes,[24] severe asthma exacerbations,[25] airway hyper-reactivity (AHR),[26] lower respiratory tract infections [LRTIs][27]), and environmental exposures are presented in the supplementary appendix.

Statistical Analysis

Longitudinal profile analysis to assess prototypical FEV_1 trajectories: We used a latent profile model to assign children to their most likely group based on the profile of FEV_1 over time. Latent profile modelling has been described extensively elsewhere.[28] Briefly, we assumed that each child belonged to one of a set of N latent profiles, the number or size of which were unknown *a priori*. We used two-level random intercept and random coefficients regression models to examine trajectory classes with no change over time, and under the assumption that the profile of FEV_1 changes linearly over time. The models were compared for goodness-of-fit using the Bayesian Information Criterion (BIC). For each child, the posterior probability of belonging to each of the classes was calculated, and children were assigned to the latent profile with the largest probability. Latent profile modelling was carried out using the gllamm (generalised linear latent and mixed models) package implemented in Stata (StataCorp, College Station, TX). Characteristics of trajectories were investigated using multinomial regression.

Genetic associations with FEV_1 trajectories: We constructed a weighted allele score from 77 SNPs found to have genome-wide evidence for association with FEV_1/FVC, or FEV_1 decline in adults[29] (Table S1, supplementary appendix). The dosage of the risk allele at each locus was multiplied by a SNP-specific weight (linear coefficient [beta] divided by the average of 77 linear coefficients), then averaged across SNPs. SNP-specific risk alleles and weights were based on the reported meta-analysis.[29] The units of the derived genetic score were per-beta per-SNP. Setting the 'Normal' trajectory as the baseline group, we used a multinomial logistic regression model to investigate the risk of the membership of other trajectories per 1-unit increase of the genetic score. The independent associations of the resulting genetic risk score with FEV_1 trajectories in MAAS and ALSPAC were meta-analysed with fixed and random effect models using the method of Mantel and Haenszel (*metan* command in Stata). Using multinomial logistic regression models, we also calculated the relative risk ratios for 'Persistently Low'

(vs. 'Normal' trajectory) in extreme quantiles of the allelic risk score in two cohorts independently

(deciles in ALSPAC; eight quantiles in MAAS).

Identification of the early-life predictors of FEV$_1$ trajectories: We constructed a multivariable regression

model to identify early-life predictors of subsequent latent profiles of FEV$_1$. We carried out these

analyses in MAAS, in which we had richer data collected in early life. We used receiver operating

characteristic (ROC) curves to evaluate the predictive ability of these predictors.

Role of the funding source

The study sponsors had no role in study design; in the collection, analysis, and interpretation of data; in

the writing of the report; or in the decision to submit the paper for publication.

RESULTS

We included 1046 participants from MAAS, 1390 from ALSPAC, and 196 from PIAF. Participant flow is presented in the Supplementary appendix.

FEV$_1$ trajectories from early school-age to young adulthood (MAAS and ALSPAC)

We selected 4-class model in MAAS and ALSPAC (please see Supplementary appendix and Table S2). The posterior probability of class membership was high in both cohorts (Table S3). The proportion of participants and the mean FEV$_1$ over time in each of the classes were similar across the cohorts (Table S4). We labelled the classes (FEV$_1$ trajectories) as: (1) Persistently High; (2) Normal; (3) Below Average; and (4) Persistently Low (Figure 1). Children within these four trajectories had stable lung function that tracked from age 5 to age 16 years (MAAS), and from ages 8 to 24 years (ALSPAC), with no overlap in FEV$_1$ at any time. The solutions remained stable and results unchanged when we included only children who had spirometry on at least 3 follow-ups (Tables S5-6).

Characteristics of children in each of the FEV$_1$ trajectories are shown in Tables S7 and S8. Persistently Low trajectory was consistently associated with wheeze and asthma throughout the follow-up in both cohorts; in MAAS, almost half the children in this trajectory had asthma, and in ALSPAC this was approximately one third, significantly higher than in the rest of the population. The proportion of children with LRTIs in the first 3 years (data from medical records, available only in MAAS[27]) was significantly higher in the Persistently Low trajectory (32.3%) compared to the Normal trajectory (12.1%). Post-bronchodilator FEV$_1$ and FEV$_1$/FVC followed similar patterns as the baseline FEV$_1$ in both cohorts. In MAAS, sR$_{aw}$ at age 3 years was significantly higher in Persistently Low compared to other trajectories, indicating diminished lung function (P=0.001). Persistently Low trajectory was significantly associated with AHR in both cohorts. There was no association between trajectory membership and

9

allergic sensitisation in adolescence in MAAS, or at age 7 years in ALSPAC. Sensitisation in early childhood in MAAS was significantly more common in Persistently Low trajectory.

There was no difference in gender between trajectories. Birth weight was significantly higher in the Persistently High trajectory in ALSPAC, but not in MAAS. In MAAS, children in the Persistently High trajectory were less likely to be exposed to environmental tobacco smoke (ETS) compared to all others (Figure 2a). Children continuously exposed to ETS from birth to age 16 years were at increased risk of belonging to the Below Average and Persistently Low trajectories. We observed similar trends in ALSPAC, with Persistently High trajectory having the lowest proportion of ETS-exposed participants (Figure 2b).

Genetic associations

Genotyping data was available in 927 participants in ALSPAC, and 902 in MAAS. The weighted genetic score based on 77 SNPs was inversely associated with the order of FEV_1 trajectories (Persistently High to Persistently Low) in both cohorts (Table S9). The meta-analysis is shown in Figure 3: compared with the Normal trajectory, the pooled relative risk ratio (95%CI) per allele was: 0.97 (0.92, 1.01), P=0.13, for Persistently High; 1.01 (1.00, 1.03), P=0.49 for Below Average; and 1.05 (1.00, 1.09), P=0.036 for Persistently Low. Results using the inverse variance method were very similar. Risk of Persistently Low FEV_1 (vs. Normal) in 10th decile of risk score (vs. 1st decile) was 2.03 (0.62, 6.67), p=0.24 in ALSPAC. In MAAS, the risk of Persistently Low FEV_1 in the highest compared with the lowest of eight quantiles was 4.19 (0.83-21.05), p=0.08.

Do FEV_1 trajectories extend back to lung function in early life (PIAF)?

In PIAF, a 3-class model provided the optimal solution for FEV_1 (Above Average, Normal and Below Average, Figure 4a), and a 2-class model for V'_{maxFRC} (Above Average and Below Average, Figure 4b). Among children in the Below Average FEV_1 trajectory, V'_{maxFRC} was significantly lower compared to other

trajectories at age 1 month (P=0.05) and 6 months (P=0.018), but not at age 1 year (Table S10). Although children in the Below Average V'_{maxFRC} trajectory were significantly more likely to remain in the Below Average FEV_1 trajectory (OR=3.63, 95%CI 1.21–10.90, P=0.022), the majority of these children (~75%) transitioned to the Above Average (12.2%) or Normal (63.4%) FEV_1 trajectories in later childhood (Figure 4, panel c).

Early-life predictors of FEV_1 trajectories (MAAS)

Children with recurrent severe wheeze exacerbations by age 3 years, with allergic sensitization at age 3 years, and who have been exposed to ETS by age 3 years, were more likely to belong to the Persistently Low trajectory (Table S10). This set of variables provided a good predictive ability in MAAS for discriminating between subjects in Persistently Low from Persistently High trajectory (Figure 5, Area Under the ROC Curve-AUROC=90.7%, sensitivity=81.8%; specificity=80.95%), and a fair predictive ability for discriminating between subjects in Persistently Low from all others (Figure S1, AUROC=77.2%, sensitivity=81.8%; specificity=67.8%).

We found an association between low birth weight and low childhood FEV_1 trajectories. In MAAS, which included no preterm infants, low birth weight (LBW) at term (<2500g) was associated with Peristently Low FEV_1 (Odds ratio 3.39 [95% CI 0.98-11.77], p=0.055). In ALSPAC, LBW adjusted for preterm (<37 weeks) gestation at birth was associated with Persistently Low FEV_1 (OR 1.98 [0.65-5.99], p=0.226, and Below Average FEV_1 (OR 1.84 [0.91-3.73], p=0.09).

DISCUSSION

In two population-based birth cohorts, using data-driven analyses we identified four distinct lung function trajectories extending from early school age, through mid-school age, into adolescence and early adulthood. Subjects assigned to the Persistently Low trajectory, representing approximately 1 in 20 participants, had low lung function in the third decade of life, around the time of the physiological plateau. Individuals in this trajectory had FEV_1 below 80% predicted from early childhood to the point beyond which lung function starts to decline with aging.[3,20] In the Persistently Low trajectory, both baseline and post-bronchodilator FEV_1 was up to 25% lower, and FEV_1/FVC up to 10% lower than the next lowest category. This tracking of lung function from early school-age is consistent with previous reports from population-based studies[9,10], and provides confirmatory empirical evidence of this phenomenon based on a hypothesis-free approach. In addition to using data-driven analyses, another novel feature of the present study was the extension of lung function trajectories to early childhood, where spirometry is not practicable, and the observation that "low" lung function recovered in the majority of cases. At group level, children in the Low FEV_1 trajectory in PIAF had significantly diminished lung function in infancy (measured by V'_{maxFRC}), which is very similar to the finding in MAAS that, as a group, those in Persistently Low FEV_1 trajectory had significantly diminished lung function at age 3 years (measured by sR_{aw}). However, comparison of lung function trajectories in infancy and in school-age in PIAF has shown that although infants in the Below Average V'_{maxFRC} trajectory were significantly more likely than children in the higher V'_{maxFRC} categories to be classified in the Below Average FEV_1 trajectory in childhood, the majority of these children (~75%) transitioned to the Above Average (12.2%) or Normal (63.4%) FEV_1 trajectories. These findings are consistent with previous results from PIAF obtained using standard hypothesis-testing analyses,[17,30] and support the concept of postnatal "recovery" of low neonatal lung function. The application of the same data-driven approach to longitudinal analysis as used in the MAAS, ALSPAC and Tasmanian Longitudinal Health Study (TAHS),[31] allows some inference to

the general population on the patterns of lung function from one month of age to the sixth decade. Disaggregation of this population into discrete trajectories suggests that the opportunity for lung function recovery extends beyond the neonatal period.

Our data revealed no evidence of groups of children with declining FEV_1. A report from TAHS published in this issue of the journal, which used a similar statistical approach to discovery of lung function trajectories and spanned childhood (7-18 years) and later adulthood (45-55 years), described six trajectories, four of which were strikingly similar to ours.[31] One further trajectory with below average FEV_1 by age 18, and accelerated decline in later adulthood after reaching the plateau, could not be captured within our data set. This trajectory was characterized by both childhood factors and personal smoking in adulthood, which were independent risk factors. Together, our studies and TAHS elucidate the developmental patterns of lung function through the life-course, confirm their link with COPD, and provide evidence that different factors affect lung function growth from those affecting lung function decline. We have shown that Persistently Low trajectory is determined by both genetic factors and early-life events. We showed relatively weak evidence that SNPs discovered through GWAS of adult lung function[29] were associated with a low childhood trajectory of FEV_1. This suggests an influence of at least some of the included variants on lung function in early life but, as they were derived from studies of adults, it is likely that they consisted of a combination of factors influencing lung function growth and decline. A recent study of adults with repeat lung function measures showed that genetic variants associated with cross-sectional lung function in adults were not associated with longitudinal decline, suggesting that some of the observed association may be explained by developmental influences[32]. We did not have sufficient statistical power in this sample to investigate the individual contributions of risk alleles, nor to construct a *de novo* GWAS.

Individuals with Persistently Low FEV_1 in childhood have low maximally attained lung function in early adulthood, and are consequently at higher risk of COPD[31], with or without additional adverse exposures

in adulthood. Personal smoking acts as an independent risk factor which contributes to accelerated lung function decline in middle adulthood, predominantly among individuals with below average lung function.[31] In this group, in addition to early-life risk factors contributing to the diminished lung function at physiological plateau, adverse exposures in adult life lead to COPD development. Childhood asthma and early-life LRTIs are associated with low lung function trajectories and higher COPD risk in both scenarios. We found an association between asthma reported at all available ages and persistently Low FEV_1 trajectory. This is consistent with previously reported associations between persistent asthma in childhood and subsequent lung function,[8, 9] but we acknowledge the potential for misclassification of symptoms associated with pre-existing airway obstruction and asthma diagnosis, particularly when relying on self-report from questionnaires for the latter. This limits interpretation of the causal pathway between asthma in childhood and low lung function trajectories in later life.

Although children with severe wheeze and early sensitization had lower FEV_1 throughout childhood, their FEV_1 remained relatively stable. In contrast, a recent study of FEV_1 trajectories between ages 7 and 26 years among patients with childhood asthma described two trajectories characterized by decline (Normal growth and early decline, and Reduced growth with early decline).[33] However, direct comparison with our findings is not possible, as our cohorts are population-based, and we used data-driven approach rather than predetermined categorization to identify trajectories. In MAAS, similar proportions of children (approximately one third) in Persistently Low trajectory had persistent wheeze and transient early wheeze, whist <10% had late-onset wheezing. This is in contrast with the Australian high-risk cohort, which suggested that only persistent wheeze is associated with reduced growth in FEV_1 over adolescence, but that transient wheeze is a benign condition with no sequelae.[34] In MAAS, an important predictor of subsequent Persistently Low FEV_1 trajectory was severe exacerbation of wheezing in early life, rather than only wheezing persistence. Low birth weight, which is associated with tobacco some exposure during pregnancy and early onset wheezing in infancy was associated with low

childhood FEV_1 trajectories in our study independently of preterm delivery. This may be explained by restricted intrauterine growth influencing airway development and subsequent lung function[35] with an associated tendency to wheezing in the context of LRTIs.

Persistently Low lung function trajectory was also identified in the Tucson cohort (comprising nearly 10% participants),[11] and is remarkably similar to our Persistently Low trajectory, which comprised a sizeable proportion (3.4%-6.7%) of our study populations. Assuming that 13.3% of the population develop COPD (http://www.hse.gov.uk/Statistics/causdis/copd/copd.pdf), and that among these, subjects in the Persistently Low trajectory would likely be diagnosed with COPD,[34] the calculated population attributable risk is 22.9%-46.7%. This also represents the estimated reduction in COPD which would occur if early interventions improve lung function for this group. To inform potential strategies for improving lung function in early adulthood, we have identified early-life risk factors associated with Persistently Low trajectory, which ~~in addition to genetic factors~~ included early-life ETS exposure, recurrent severe exacerbations of wheezing, and early allergic sensitisation. A recent study suggested that adult smoking, but not parental smoking history, can negatively influence adult lung function trajectory and modify the adverse impact of early-life exposures.[35] Our and TAHS data demonstrate that both are important. We identified an association between the absence of early-life tobacco smoke exposure and Persistently High trajectory, which In TAHS had markedly lower risk of COPD at age 55 years.[31] Thus, while tobacco smoke exposure will not necessarily lead to a disease, the absence of this exposure is associated with exceptionally good lung function and subsequent health. The risk of being in one of the diminished lung function profiles was decreased by 76% among children who were not exposed to parental smoking, suggesting that ETS exposure in early childhood may be causally related to COPD development, and that this may be mediated via the effect of early-life exposure on subsequent trajectories of lung function into adulthood.

Strengths and limitations

One of the limitations of our study is that it was not possible to have identical measures of lung function through the life-course in our cohorts. We did not measure spirometry before age five years, since forced expiratory manoeuvres are difficult to carry out in young children. We used different cohorts and age-appropriate measures of pulmonary function to make inference about the life-course (rapid thoraco-abdominal compression in infancy[16], plethysmography in pre-school age[36], spirometry from school age onwards), which may represent the best available approximation of what might be happening from birth to adulthood. Another limitation is that the questions we used contained small differences in wording.[37] ALSPAC included pre-term babies, which may explain slightly higher proportion of children in the Persistently Low trajectory in ALSPAC (6.7%, compared to 3.4% in MAAS). One of the strengths of this study is the ability to disambiguate the temporal structure of profiles during childhood, allowing us to identify patterns during the lung function growth phase. It is striking that although we modelled data collected at different time points, and separately for each cohort, the results indicated almost identical latent profiles. In PIAF, we did not observe Persistently Low FEV_1 trajectory. This is not surprising, given a relatively small proportion of children in this trajectory (~5%), and a considerably smaller size of PIAF compared to MAAS and ALSPAC.

We wish to emphasise that our current study is exploratory in nature. Therefore, the results should be interpreted with caution, and need to be validated in a larger population. We are also cognizant of the potential effect of the withdrawal/missing data rates. We carried out our analyses under the assumption that data was missing at random. To test this assumption, we carried out analyses using both children with at least three data points available, and those with at least one time point. Both analyses gave consistent optimal goodness-of-fit using the BIC, and the child class assignments were stable across the two analyses. This suggests that the missing at random assumption was plausible,

given that if children with missing data points were not missing at random, there would have been a higher mismatch between classes.

Conclusion

Using data from population-based birth cohort studies, we demonstrated four discrete trajectories of FEV_1 development from early childhood to young adulthood. Interventions to reduce the risk of Persistently Low FEV_1 trajectory should start in early childhood, and should include aggressive strategies to reduce tobacco smoke exposure, and novel strategies to prevent early-life sensitization and exacerbations of early-childhood wheezing.

Authors' contributions: AC, AJH and AS conceived the idea; DB, RG and IEB provided input on the methodology for analyzing the data; DB, RG and JAC carried out the analyses; AC, AJH, DB, SWT, PNLS and AS interpreted the data; all authors wrote the report. The corresponding author (AJH) had access to all the data in the study and takes final responsibility for the decision to submit for publication.

Declaration of interests: None of the authors declare any relevant conflicts in relation to the content of this paper.

REFERENCES

1.   Lange P, Celli B, Agusti A, et al. Lung-Function Trajectories Leading to Chronic Obstructive Pulmonary Disease. *N Engl J Med* 2015; **373**(2): 111-22.

2.   Martinez FD. Early-Life Origins of Chronic Obstructive Pulmonary Disease. *N Engl J Med* 2016; **375**(9): 871-8.

3.   Postma DS, Bush A, van den Berge M. Risk factors and early origins of chronic obstructive pulmonary disease. *Lancet* 2015; **385**(9971): 899-909.

4.   Vasquez MM, Zhou M, Hu C, Martinez FD, Guerra S. Low Lung Function in Young Adult Life Is Associated with Early Mortality. *American journal of respiratory and critical care medicine* 2017; **195**(10): 1399-401.

5.   Young RP, Hopkins R, Eaton TE. Forced expiratory volume in one second: not just a lung function test but a marker of premature death from all causes. *European Respiratory Journal* 2007; **30**(4): 616-22.

6.   Agusti A, Noell G, Brugada J, Faner R. Lung function in early adulthood and health in later life: a transgenerational cohort analysis. *The Lancet Respiratory medicine* 2017; **5**(12): 935-45.

7.   Taylor-Robinson DC, Stanojevic S. Low lung function in early adulthood: morbidity and death. *The Lancet Respiratory medicine* 2017; **5**(12): 910-1.

8.   Tai A, Tran H, Roberts M, et al. Outcomes of childhood asthma to the age of 50 years. *J Allergy Clin Immunol* 2014; **133**(6): 1572-8 e3.

9.   Sears MR, Greene JM, Willan AR, et al. A longitudinal, population-based, cohort study of childhood asthma followed to adulthood. *New England Journal of Medicine* 2003; **349**(15): 1414-22.

10.   Phelan PD, Robertson CF, Olinsky A. The Melbourne Asthma Study: 1964-1999. *J Allergy Clin Immunol* 2002; **109**(2): 189-94.

11.   Berry CE, Billheimer D, Jenkins IC, et al. A Distinct Low Lung Function Trajectory from Childhood to the Fourth Decade of Life. *American journal of respiratory and critical care medicine* 2016; **194**(5): 607-12.

12.   Grol MH, Gerritsen J, Vonk JM, et al. Risk factors for growth and decline of lung function in asthmatic individuals up to age 42 years. A 30-year follow-up study. *American journal of respiratory and critical care medicine* 1999; **160**(6): 1830-7.

13.   Custovic A, Simpson BM, Murray CS, Lowe L, Woodcock A. The National Asthma Campaign Manchester Asthma and Allergy Study. *Pediatr Allergy Immunol* 2002; **13 Suppl 15**: 32-7.

14.   Boyd A, Golding J, Macleod J, et al. Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol* 2013; **42**(1): 111-27.

15.   Custovic A, Ainsworth J, Arshad H, et al. The Study Team for Early Life Asthma Research (STELAR) consortium 'Asthma e-lab': team science bringing data, methods and investigators together. *Thorax* 2015; **70**(8): 799-801.

16.   Young S, Le Souef PN, Geelhoed GC, Stick SM, Turner KJ, Landau LI. The influence of a family history of asthma and parental smoking on airway responsiveness in early infancy. *N Engl J Med* 1991; **324**(17): 1168-73.

17.   Turner S, Fielding S, Mullane D, et al. A longitudinal study of lung function from 1 month to 18 years of age. *Thorax* 2014; **69**(11): 1015-20.

18.   Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. *European respiratory journal* 2005; **26**(2): 319-38.

19.   Beydon N, Davis SD, Lombardi E, et al. An official American Thoracic Society/European Respiratory Society statement: pulmonary function testing in preschool children. *American journal of respiratory and critical care medicine* 2007; **175**(12): 1304-45.

20.   Stanojevic S, Wade A, Cole TJ, et al. Spirometry centile charts for young Caucasian children: the Asthma UK Collaborative Initiative. *American journal of respiratory and critical care medicine* 2009; **180**(6): 547-52.

21.   Lowe L, Murray CS, Custovic A, et al. Specific airway resistance in 3-year-old children: a prospective cohort study. *Lancet* 2002; **359**(9321): 1904-8.

22.   Crapo RO, Casaburi R, Coates AL, et al. Guidelines for methacholine and exercise challenge testing-1999. This official statement of the American Thoracic Society was adopted by the ATS Board of Directors, July 1999. *American journal of respiratory and critical care medicine* 2000; **161**(1): 309-29.

23.   Lodrup Carlsen KC, Roll S, Carlsen KH, et al. Does pet ownership in infancy lead to asthma or allergy at school age? Pooled analysis of individual participant data from 11 European birth cohorts. *PLoS One* 2012; **7**(8): e43214.

24.   Lowe LA, Simpson A, Woodcock A, Morris J, Murray CS, Custovic A. Wheeze phenotypes and lung function in preschool children. *American journal of respiratory and critical care medicine* 2005; **171**(3): 231-7.

25.   Reddel HK, Taylor DR, Bateman ED, et al. An official American Thoracic Society/European Respiratory Society statement: asthma control and exacerbations: standardizing endpoints for clinical asthma trials and clinical practice. *American journal of respiratory and critical care medicine* 2009; **180**(1): 59-99.

26.   Nicolaou NC, Lowe LA, Murray CS, Woodcock A, Simpson A, Custovic A. Exhaled breath condensate pH and childhood asthma: unselected birth cohort study. *American journal of respiratory and critical care medicine* 2006; **174**(3): 254-9.

27.   Semic-Jusufagic A, Belgrave D, Pickles A, et al. Assessing the association of early life antibiotic prescription with asthma exacerbations, impaired antiviral immunity, and genetic variants in 17q21: a population-based birth cohort study. *The Lancet Respiratory medicine* 2014; **2**(8): 621-30.

28.   Oberski D. Mixture Models: Latent Profile and Latent Class Analysis. . In: Robertson J KM, ed. Modern Statistical Methods for HCI Human–Computer Interaction Series Springer, Cham; 2016.

29.   Wain LV, Shrine N, Artigas MS, et al. Genome-wide association analyses for lung function and chronic obstructive pulmonary disease identify new loci and potential druggable targets. *Nat Genet* 2017; **49**(3): 416-25.

30.   Turner SW, Palmer LJ, Rye PJ, et al. Infants with flow limitation at 4 weeks: outcome at 6 and 11 years. *American journal of respiratory and critical care medicine* 2002; **165**(9): 1294-8.

31.   Bui DS BJ, Lodge C L, Lowe A, Perret J, Bowatte G, Thompson BR, Hamilton GS, Frith PA, James AL, Johns DP, Thomas PS, Jarvis D, Svanes C, Morrison S, Feather I, Allen KJ, Wood-Baker R, Bui MQ, Russell M, Abramson MJ, Walters EH, Matheson MC, Dharmage SC. Lung function trajectories from childhood to the six decade, their associated childhood factors and consequences: a population based longitudinal cohort. *Submitted* 2017.

32.   John C, Soler Artigas M, Hui J, Nielsen SF, Rafaels N, Paré PD, et al. Genetic variants affecting cross-sectional lung function in adults show little or no effect on longitudinal lung function decline. *Thorax* 2017;**72**(5):400-408.

33.   McGeachie MJ, Yates KP, Zhou X, et al. Patterns of growth and decline in lung function in persistent childhood asthma. *The New England Journal of Medicine* 2016; (19): 1842.

34.   Lodge CJ, Lowe AJ, Allen KJ, et al. Childhood wheeze phenotypes show less than expected growth in FEV1 across adolescence. *American journal of respiratory and critical care medicine* 2014; **189**(11): 1351-8.

35.   den Dekker HT, Jaddoe VWV, Reiss IK, de Jongste JC, Duijts L. Fetal and Infant Growth Patterns and Risk of Lower Lung Function and Asthma. The Generation R Study. *Am J Respir Crit Care Med* 2018;**197**(2):183-192.

34.   Vogelmeier CF, Criner GJ, Martinez FJ, et al. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease 2017 Report. GOLD Executive Summary. *American journal of respiratory and critical care medicine* 2017; **195**(5): 557-82.

35.   Allinson JP, Hardy R, Donaldson GC, Shaheen SO, Kuh D, Wedzicha JA. Combined Impact of Smoking and Early Life Exposures on Adult Lung Function Trajectories. *American journal of respiratory and critical care medicine* 2017.

36.   Belgrave DC, Buchan I, Bishop C, Lowe L, Simpson A, Custovic A. Trajectories of lung function during childhood. *American journal of respiratory and critical care medicine* 2014; **189**(9): 1101-9.

37.   Belgrave DC, Granell R, Simpson A, et al. Developmental profiles of eczema, wheeze, and rhinitis: two population-based birth cohort studies. *PLoS medicine* 2014; **11**(10): e1001748.

LEGENDs FOR FIGURES

Figure 1: Mean % Predicted FEV_1 over time by FEV_1 trajectory:

a) MAAS

b) ALSPAC

Figure 2: % Children exposed to Environmental Tobacco Smoke over time by FEV_1 trajectory:

a) MAAS

b) ALSPAC

Figure 3: Meta-analysis (random effects model) of weighted genetic score based on 77 SNPs with FEV_1 trajectories

Figure 4: Lung function trajectories in PIAF:

a) Mean % Predicted FEV_1 over time by FEV_1 trajectory

b) Mean % Predicted V'_{maxFRC} over time by V'_{maxFRC} trajectory

c) Membership of FEV_1 trajectory in relation to the V'_{maxFRC} trajectories

Figure 5: Receiver Operating Characteristic Curve showing predictive ability of a logistic regression analysis using the Persistently Low *vs*. Persistently High trajectory as the outcome (or predicted) measure with the following covariates: (1) recurrent wheeze with severe exacerbations by age 3 years, (2) at least one positive skin test by age 3 years, and (3) Environmental Tobacco Smoke exposure by age 3 years

We thank the Editors and Reviewers for their comments on our manuscript. We have addressed the remaining specific comments below in a point by point response.

Editorial comments:

1. Please remove corr author's phone numbers
Response 1. *Completed*

2. The figures provided still aren't editable I'm afraid.
Response 2. *The figures have been uploaded as .EPS files with the revised manuscript. Would the editors like us to submit these in a different format?*

Reviewers' comments:

Reviewer #1: Authors have addressed all my comments satisfactorily. I do not have any further one other than congratulate them on a fine piece of work.

Response: *We thank the reviewer for these comments; no specific response required*

Reviewer #2: The authors have satisfactorily addressed the comprehensive issues raised by the reviewers.

Response: *We thank the reviewer for these comments; no specific response required*

Reviewer #3

Comment 1: There is missing data, some of the analysis was based on small sample size, the authors have performed multiple significance tests and there are strong assumptions made by the latent models. The authors mention missing data in their limitations. I think they should say a bit more about the exploratory nature of the analyses and hence any results obtained should be interpreted cautiously and perhaps needs to be validated in a larger study.

Response 1: *As per reviewer's suggestion, we have now Included the following statement into the discussion: "We wish to emphasise that our current study is exploratory in nature. Therefore, the results should be interpreted with caution, and need to be validated in a larger population." (please see page 16, revised manuscript).*

We would like to note that our manuscript has been submitted to the journal for a possible joint publication with a complementary manuscript ("Childhood predictors and adult COPD risk of lung function trajectories from the first to the sixth decade") from the Tasmanian Longitudinal Health Study (TAHS). TAHS modelled trajectories of FEV_1 measured at 7, 13, 18, 45, 50 and 53 years among 2438 subjects. The results in TAHS indicated almost identical latent lung function profiles. Together, our studies elucidate with remarkable consistency the developmental patterns of lung function through the life-course. Thus, we would argue that our findings have been validated in a larger study. We discuss this in detail on page 13.

As a further validation to reassure the reviewer about the consistency of our findings across different populations, we have applied the same latent profile model we used to investigate trajectories of FEV_1 in MAAS and ALSPAC to recently collected lung function data from 827 participants in the third STELAR cohort (SEATON from Aberdeen). This data was not available at the time of the submission of our manuscript. Although we modelled data collected at different time points (5, 10 and 15 years),

the results in SEATON indicated almost identical latent profiles as those in MAAS and ALSPAC (please see Figure R1).
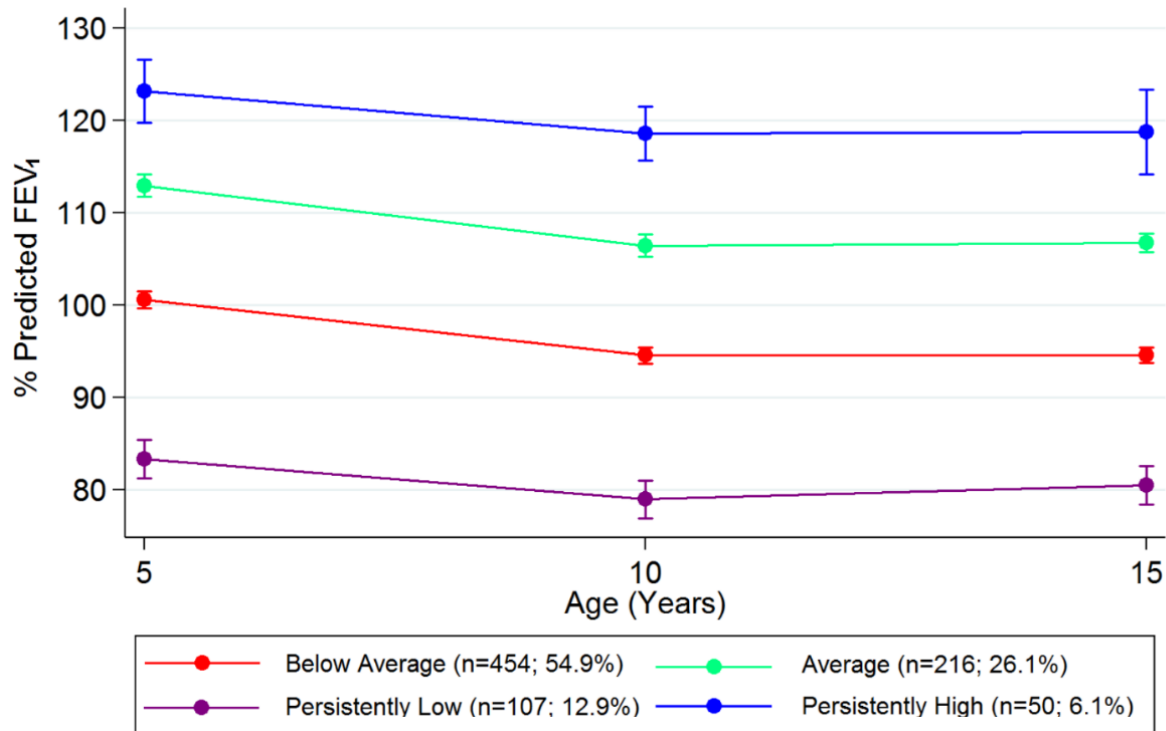

Figure R1. Mean % Predicted FEV$_1$ over time by FEV$_1$ trajectory in SEATON cohort

Comment 2: They have not dealt with my comments regarding model assumption and missing data satisfactorily. There are more assumptions that latent class random effects models make. How were they investigated, besides looking at model fit? I don't think they have investigated predictors of missingness. There is over reliance on the BIC statistic.

Response 2: We carried out our analyses under the assumption that data was missing at random. To test this assumption, we carried out analyses using both children with at least three data points available, and those with at least one time point. Both analyses gave consistent optimal goodness-of-fit using the BIC, and the child class.

We agree that we relied on BIC, but this is due to an artefact of the package used in Stata which is based on log likelihood. We have tried to compensate for this by testing numbers of classes on subsets of this data, which is a form of cross-validation.

As per our response to the original C1 by the reviewer, the first set of assumptions we made were that FEV$_1$ profiles remained constant over time; we used two-level random intercept model to test this assumption. We then assumed that FEV$_1$ trajectory changes during childhood. This was investigated using a random coefficients regression under the assumption that FEV$_1$ profile changes linearly over time. We have extended the methods to include this description.

We have investigated predictors of missingness, but have not included this data in the manuscript, as the results are already very busy, and these results may distract from the main message. Missing FEV$_1$ data was not associated with gender, allergy, or parental history of asthma or atopy. However, children with missing data in both cohorts were more likely to be exposed to environmental tobacco smoke, and come from families with lower socio-economic status. These results are shown in the Table R1 below, and give us confidence to assume that study participants with missing lung function data will not have higher or lower FEV$_1$ values than those with observed data.
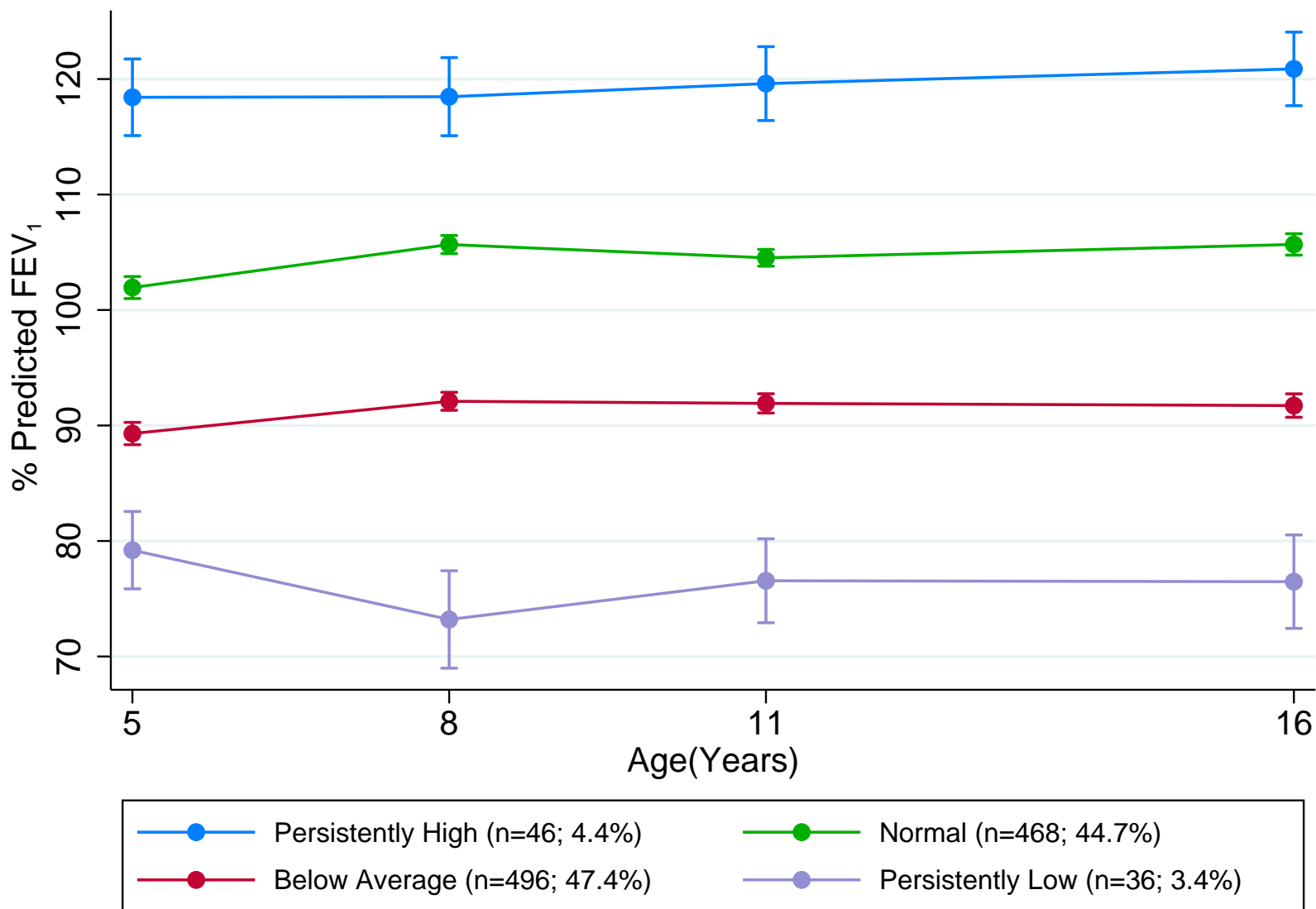
Table R1: Summary of investigation of factors associated with probability of missing FEV_1

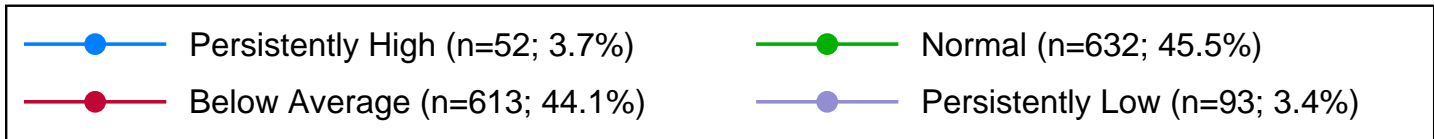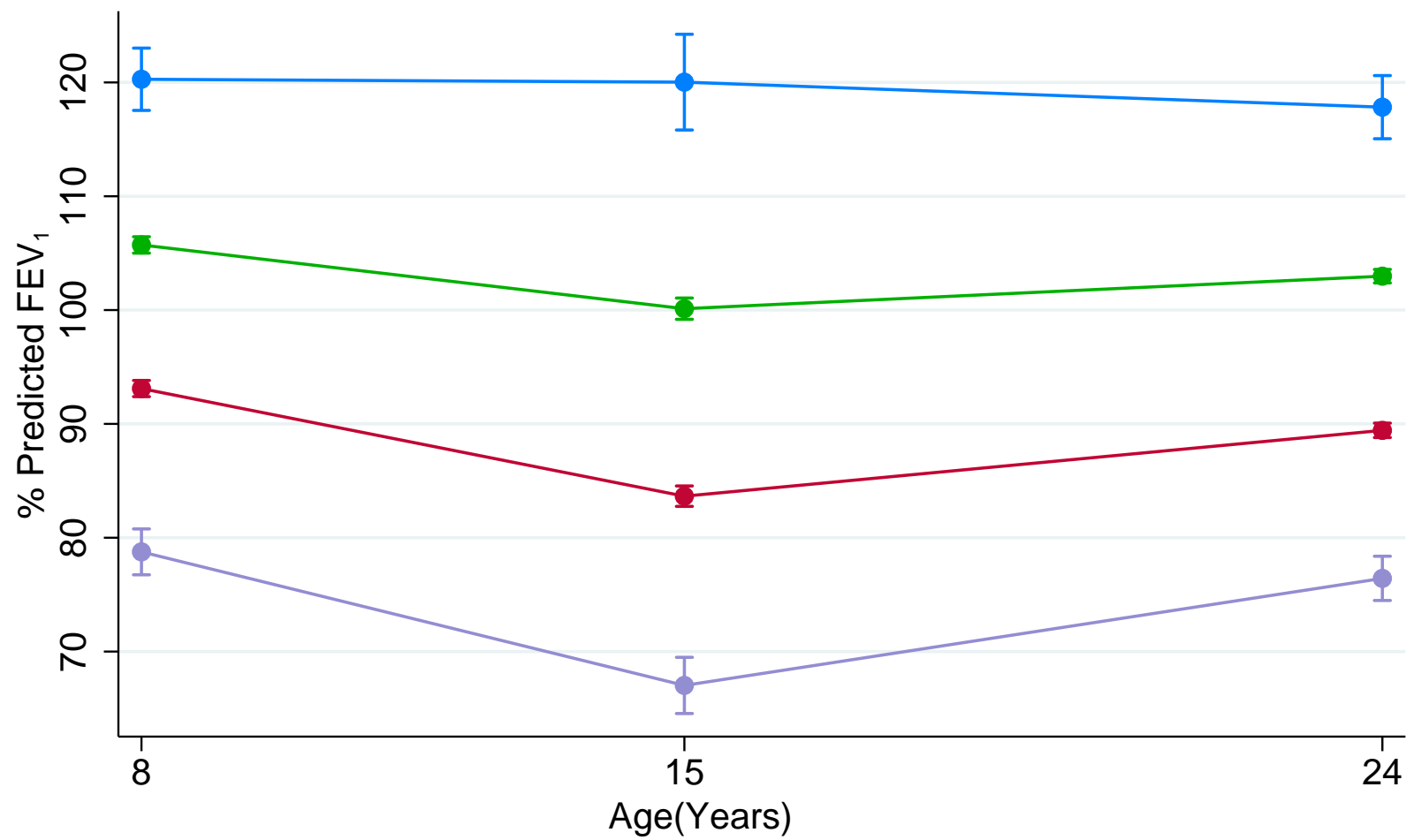| Predictor	Main effect OR (SE)	P value
MAAS		
Sex (Male)	1.200 (0.098)	0.10
BMI	0.877 (0.008)	<0.001
Asthma	1.035 (0.135)	0.80
Use of Asthma medication	1.000 (0.132)	1.00
Current Wheeze	0.906 (0.116)	0.44
Sensitization (SPT)	0.930 (0.132)	0.61
Any smoking exposure	1.305 (0.151)	0.022
Maternal Asthma	0.966 (0.125)	0.79
Paternal Asthma	1.094 (0.161)	0.54
Maternal atopy (SPT)	0.960 (0.104)	0.71
Paternal atopy (SPT)	1.048 (0.118)	0.68
Socio-economic Status (Baseline= Managerial Group)		
Intermediate	1.101 (0.203)	0.601
Routine	2.332 (0.519)	<0.001
Not Working	3.819 (1.491)	0.001
ALSPAC		
Sex (Male)	1.210 (0.019)	p<0.001
BMI at 7 years	1.007 (0.005)	0.14
Asthma (5 to 15 years)	1.116 (0.038)	0.004
Sensitization (SPT) at 7 years	0.989 (0.025)	0.66
Any smoking exposure (birth to 18 years)	1.947 (0.032)	p<0.001
Maternal asthma	1.004 (0.031)	0.90
Paternal asthma	0.969 (0.032)	0.33
Social class (ref= non-manual)	1.415 (0.020)	p<0.001

Necessary Additional Data

[Click here to download Necessary Additional Data: Supplementary Appendix unmarked.docx](#)
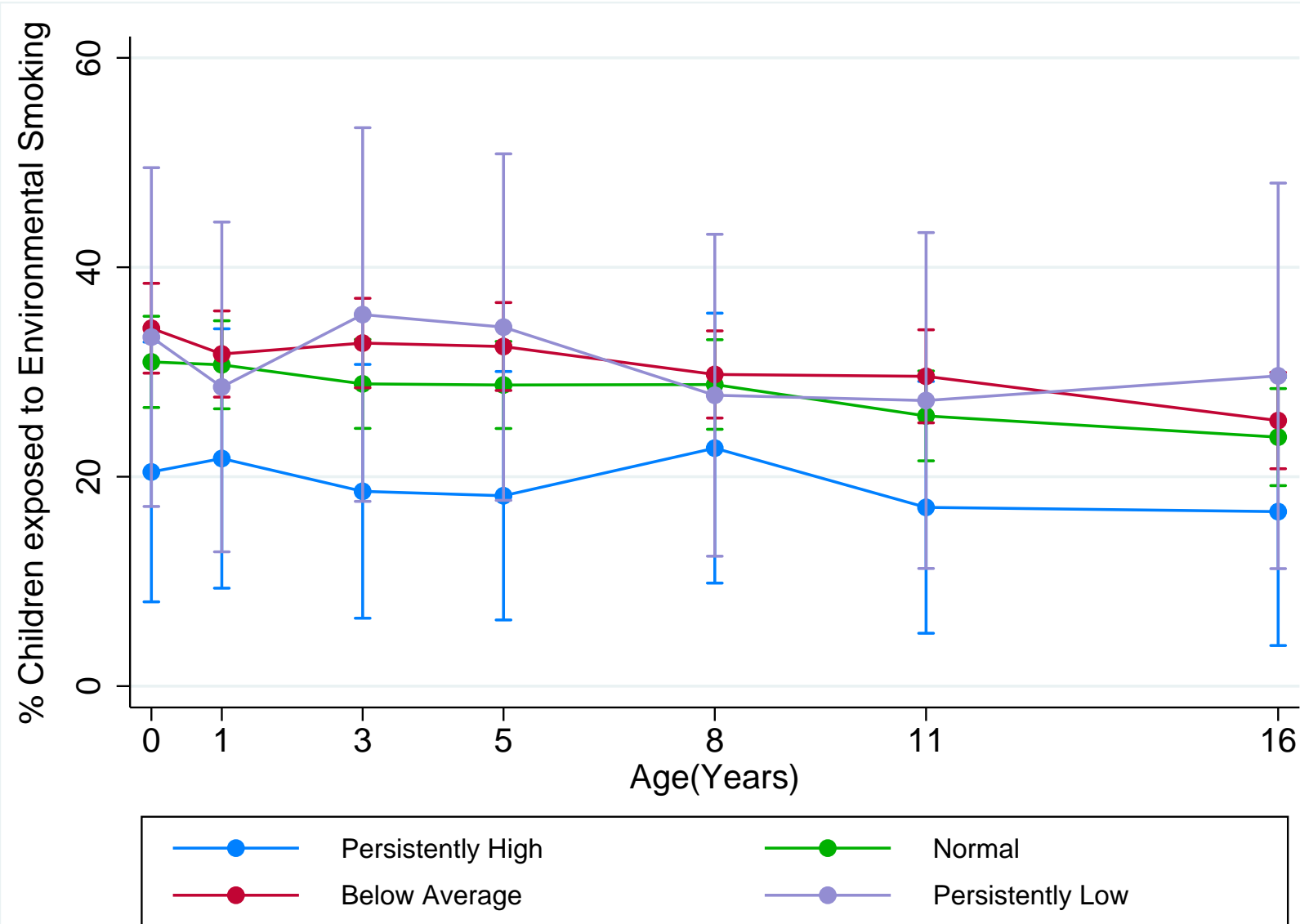
Figure

Figure

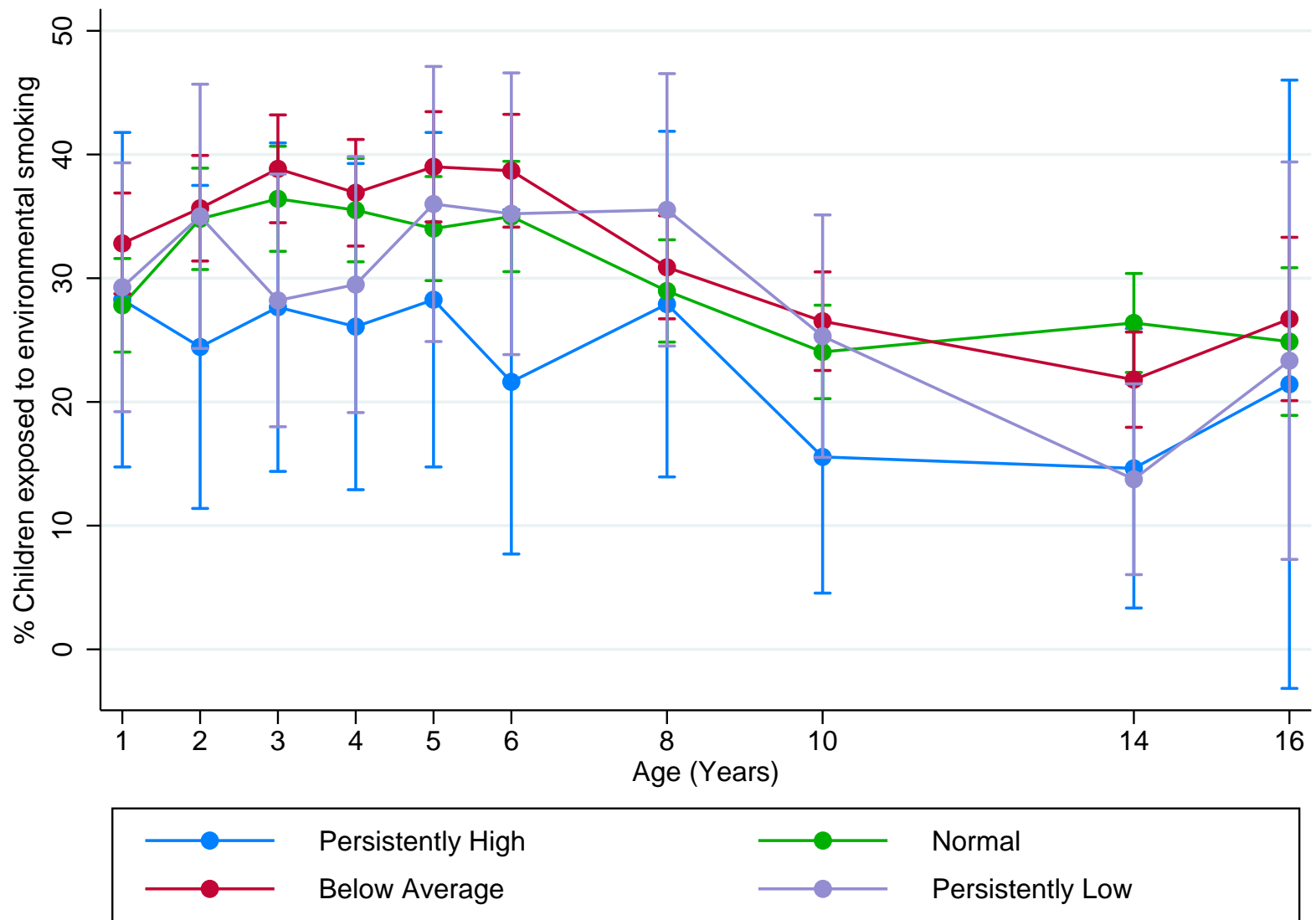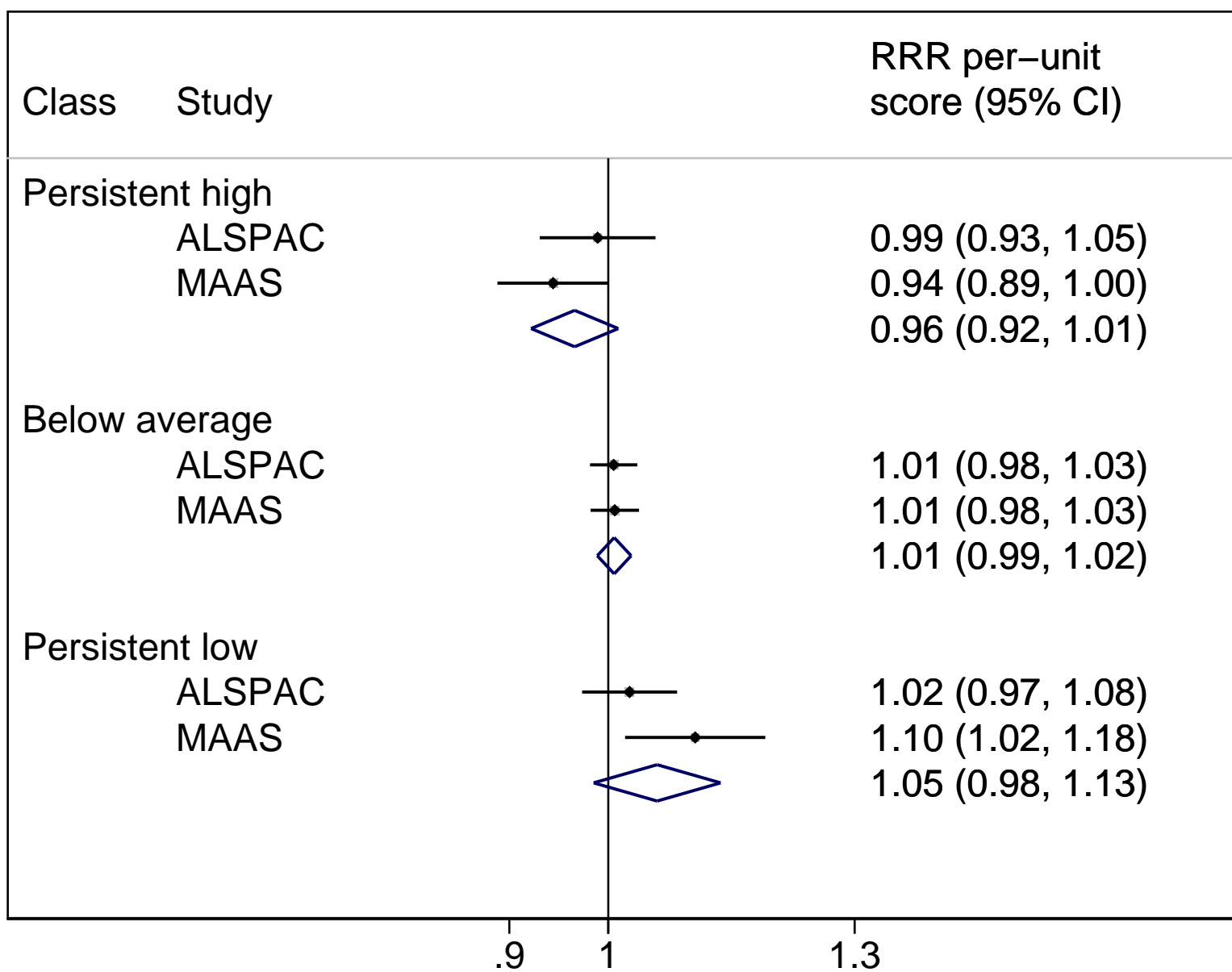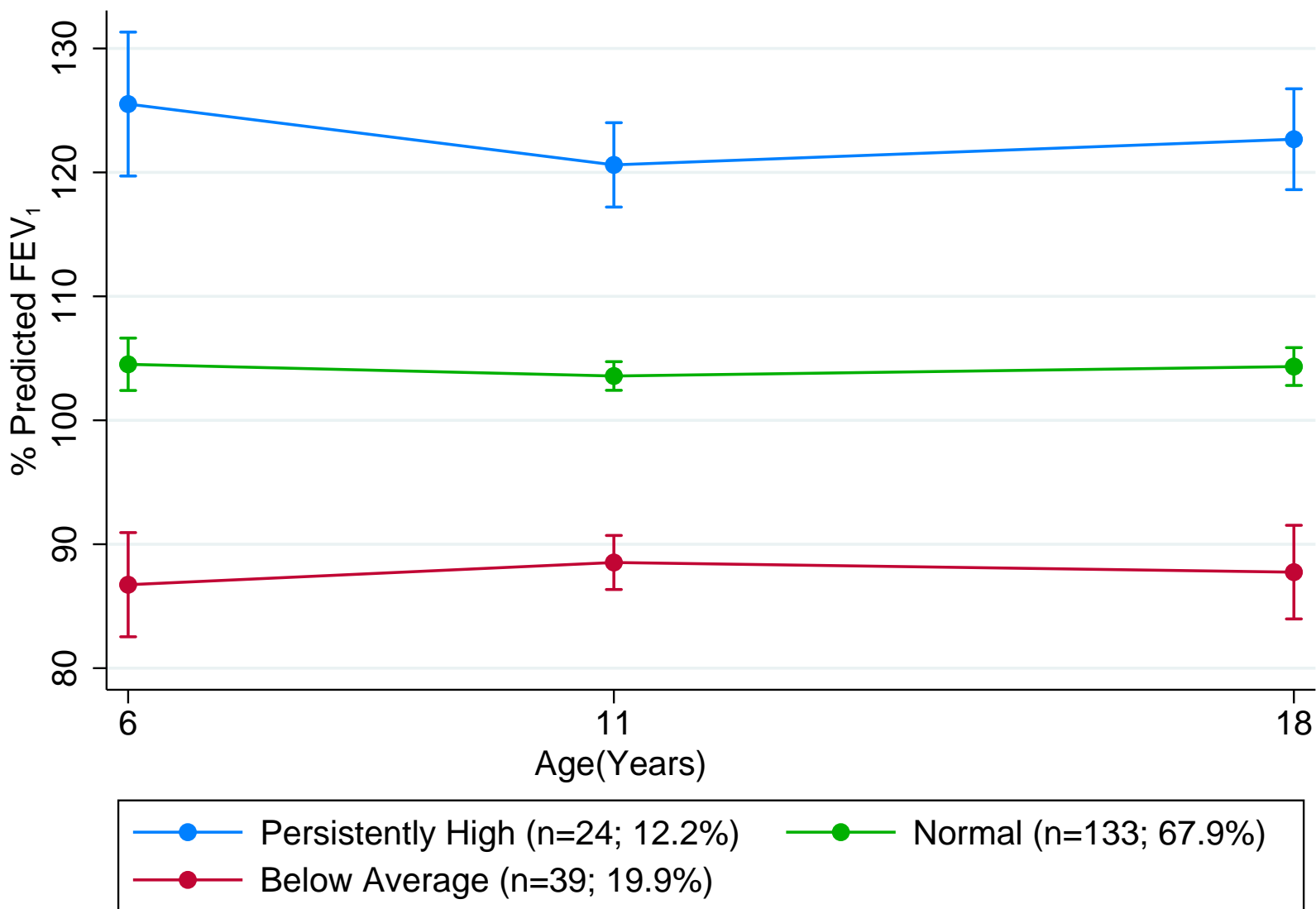**Figure**

Figure

Figure

Class	Study		RRR per-unit score (95% CI)
Persistent high			
	ALSPAC		0.99 (0.93, 1.05)
	MAAS		0.94 (0.89, 1.00)
			0.96 (0.92, 1.01)
Below average			
	ALSPAC		1.01 (0.98, 1.03)
	MAAS		1.01 (0.98, 1.03)
			1.01 (0.99, 1.02)
Persistent low			
	ALSPAC		1.02 (0.97, 1.08)
	MAAS		1.10 (1.02, 1.18)
			1.05 (0.98, 1.13)

.9   1   1.3

Figure

Figure

Figure

Figure

Area under ROC curve = 0.9069