



Does DNA methylation mediate the association of age at puberty with forced vital capacity or forced expiratory volume in 1 s?

Liang Li¹, Hongmei Zhang ¹, John W. Holloway ², Susan Ewart³, Caroline L. Relton⁴, S. Hasan Arshad^{5,6} and Wilfried Karmaus ¹

¹Division of Epidemiology, Biostatistics, and Environmental Health, School of Public Health, University of Memphis, Memphis, TN, USA. ²Human Development and Health, Faculty of Medicine, University of Southampton, Southampton, UK. ³College of Veterinary Medicine, Michigan State University, East Lansing, MI, USA. ⁴MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK. ⁵Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK. ⁶The David Hide Asthma and Allergy Research Centre, St Mary's Hospital, Newport, Isle of Wight, UK.

Corresponding author: Hongmei Zhang (h Zhang@memphis.edu)



Shareable abstract (@ERSpublications)

The association between age at pubertal onset and lung function parameters FVC or FEV₁ in young adulthood is not likely to be mediated by DNA methylation in peripheral blood

<https://bit.ly/31G8hDi>

Cite this article as: Li L, Zhang H, Holloway JW, *et al.* Does DNA methylation mediate the association of age at puberty with forced vital capacity or forced expiratory volume in 1 s? *ERJ Open Res* 2022; 8: 00476-2021 [DOI: 10.1183/23120541.00476-2021].

Copyright ©The authors 2022

This version is distributed under the terms of the Creative Commons Attribution Non-Commercial Licence 4.0. For commercial reproduction rights and permissions contact permissions@ersnet.org

Received: 25 July 2021
Accepted: 30 Dec 2021

Abstract

Background Age of pubertal onset is associated with lung function in adulthood. However, the underlying role of epigenetics as a mediator of this association remains unknown.

Methods DNA methylation (DNAm) in peripheral blood was measured at age 18 years in the Isle of Wight birth cohort (IOWBC) along with data on age of pubertal events, forced vital capacity (FVC) and forced expiratory volume in 1 s (FEV₁) at 26 years. Structural equation models were applied to examine mediation effects of DNAm on the association of age at pubertal events with FVC and FEV₁. Findings were further tested in the Avon Longitudinal Study of Parents and Children (ALSPAC) cohort.

Results In the IOWBC, for females, 21 cytosine-phosphate-guanine sites (CpGs) were shown to mediate the association of age at puberty with FVC or FEV₁ at 26 years ($p < 0.05$). In males, DNAm at 20 CpGs was found to mediate the association of age at puberty with FVC ($p < 0.05$). At almost all these CpGs, indirect effects (effects of age at pubertal events on FVC or FEV₁ via DNAm) contributed a smaller portion to the total effects compared to direct effects (*e.g.* at cg08680129, ~22% of the estimated total effect of age at menarche on FVC at age 26 was contributed by an indirect effect). Among the IOWBC-discovered CpGs available in ALSPAC, none of them was replicated in ALSPAC ($p > 0.05$).

Conclusions Our findings suggest that post-adolescence DNAm in peripheral blood is likely not to mediate the association of age at pubertal onset with young adulthood FVC or FEV₁.

Introduction

Lung function parameters, including forced vital capacity (FVC) and forced expiratory volume in 1 s (FEV₁), are shown to be associated with the age of menarche [1–4]. Subjects with early menarche have lower FVC and FEV₁ in adulthood [1, 2, 4, 5]. A recent Mendelian randomisation study suggests that pubertal timing in general rather than the time of a single pubertal event is associated with lung function [2], implying the necessity to examine the contribution of other pubertal events onset to lung function development in both males and females. Our recent study has shown that age at puberty was associated with FVC in young adulthood in both sexes [6].

A covalent addition of a methyl group to the DNA at a cytosine residue followed by a guanine (CpG site), *i.e.* DNA methylation (DNAm) at CpG sites, is an important epigenetic mechanism that regulates gene



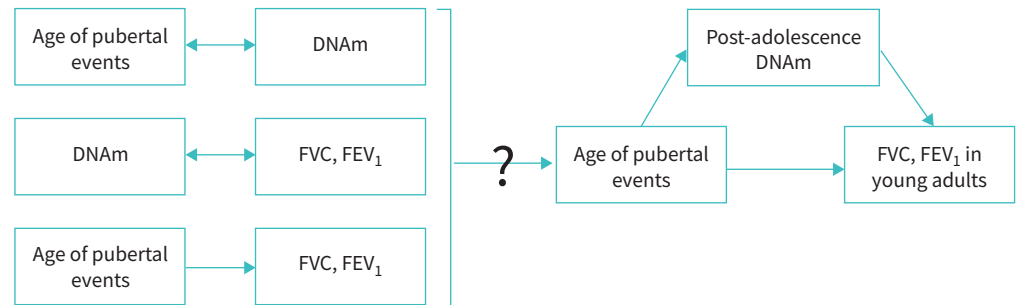


FIGURE 1 Rationale of assessing potential mediation effects of post-adolescence DNA methylation (DNAm) on the association of age of pubertal events with lung function parameters. FVC: forced vital capacity; FEV₁: forced expiratory volume in 1 s.

transcription. Multiple studies have demonstrated that DNAm, as an epigenetic mechanism, links ageing, environmental exposure and lifestyle factors to disease and health status [7–10]. Recent studies examined such impact of DNAm *via* mediation analyses, which offers an opportunity to better understand the role that DNA methylation plays between exposures and health outcomes [11–14]. Studies have shown that DNAm is associated with various traits including age of pubertal onset [15–19]. For females, one study suggested that DNA methylation level might be modulated by age at menarche with later puberty associated with lower DNAm [18]. An association between DNA methylation patterns and lung function measures has also been demonstrated [20, 21].

Given the association between age at pubertal events and lung function, and the association of DNAm with both age of pubertal events and with lung function, this study aimed to ascertain whether DNAm at post-adolescence or close to post-adolescence mediates the association between age of pubertal events and young adulthood lung function parameters FVC and FEV₁ (figure 1). To test this, we applied mediation analyses using structural equation modelling [22], using data in the Isle of Wight birth cohort (IOWBC; discovery cohort) [23] and the Avon Longitudinal Study of Parents and Children (ALSPAC) birth cohort (replication cohort) [24–26]. Mediation analysis is a statistical model with a minimum of three variables, in which there is one hypothesised mediator variable (M) transmitting the effect of independent (X) to the dependent variable (Y) [22]. In our study, mediation analysis allows evaluation of indirect and direct effects of an exposure (*e.g.* age at puberty) on an outcome of interest (*e.g.* a lung function parameter) (figure 1), where the expected value of the mediator (*e.g.* DNAm) and the outcome can be expressed as a function of the exposure *via* a generalised linear model [27].

Methods

Study population in IOWBC

A population-based birth cohort study was initiated in 1989 on the Isle of Wight (IOWBC), UK, to prospectively study the natural history of allergic diseases and associated risk factors [23]. After exclusion of adoptions, perinatal deaths and refusal to follow-up, 1456 newborns were enrolled in the study. In total, 1261 children were followed up at 10, 18 or 26 years (supplementary figure S1). The study was approved by the Isle of Wight Local Research Ethics Committee (06/Q1701/34). The internal review board of the University of Memphis approved the project (FWA00006815). At each follow-up, demographic information was collected, allergic status was assessed, and height, weight, and FVC and FEV₁ were measured.

Age at onset of pubertal events

The age at onset of pubertal events was obtained through questionnaires when subjects were interviewed at age 18 years. Questions relating to pubertal events came from the National Institute of Child and Human Development (NICHD) questionnaire from the Study of Early Child Care and Youth Development based on the Pubertal Development Scale (PDS) [28]. For females, pubertal events included body hair growth, breast growth, menarche, skin changes and growth spurt, and for males, pubertal events including body hair growth, facial hair growth, voice deepening, skin changes and growth spurt were the focus. The detailed measurements of pubertal events are described elsewhere [29].

Lung function assessment and potential confounders

Assessment of lung function at 26 years of age was conducted using the KoKo spirometry software package on portable desktop device (PDS instrumentation, Louisville, KY, USA) according to the guidelines of the American Thoracic Society and the European Respiratory Society [30]. The lung function measurements included in this study were FVC and FEV₁.

Low birthweight (<2500 g), maternal smoking during pregnancy, asthma status at age 10 years, body mass index (BMI) at age 10 years, socioeconomic status, atopy at age 10 years and personal smoking at age 18 years were included in the analyses as potential confounders (supplementary material section S2).

DNAm and cell type composition

In the IOWBC, DNA was extracted from whole blood using a standard procedure, and DNAm was measured using Infinium HumanMethylation450 or MethylationEPIC BeadChips. Details of DNAm assessment and preprocessing are described in supplementary material section S2. CpGs on autosome chromosomes and M values, calculated as base 2 logit transformed DNAm β values at each CpG, were used in the analyses. In this study, DNAm at age 18 years was included in the analyses.

Heterogeneity in white blood cell composition may confound DNA methylation measurements [31]. Proportions of six cell types, CD4⁺ T-cells, natural killer cells, neutrophils, B-cells, monocytes and eosinophils, were inferred using the R package *minfi* [32, 33] and included in the analyses as confounders.

Statistical analyses

Given the differences in age of pubertal onset and in lung function growth between males and females [4, 6], all analyses were stratified by sex. To assess whether the subsample in the study was representative of the IOWBC, one sample t-tests were applied to the key variables, including ages of pubertal events onset, and lung function measures FVC and FEV₁ at age 26 years.

The study aimed to examine whether DNAm at age 18 years mediates the association between age of puberty and age 26-year FVC or FEV₁ in young adults (figure 1). A set of analyses was carried out to achieve this goal (supplementary figure S2). Firstly, the association of FVC and FEV₁ with age of pubertal event onset was assessed using linear regressions. Only pubertal events that passed this screening were included in subsequent analyses. Next, we screened CpG sites from two directions in an effort to include as many CpGs as possible in the assessment of their mediation effects while avoiding significant power loss should we analyse at the genome-scale. One direction is to identify candidate CpGs potentially associated with age at onset of a pubertal event (forward screening), and the other direction is to screen

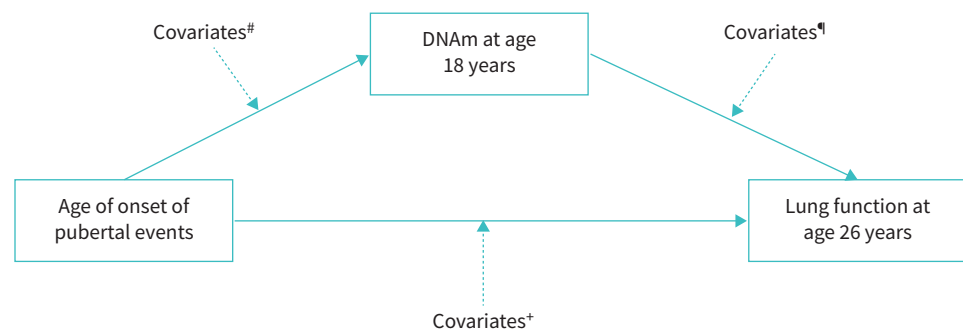


FIGURE 2 Diagram of DNA methylation (DNAm) pathway at cytosine-phosphate-guanine sites (CpGs) at age 18 years. The principle of mediation analysis is to analyse that an initial variable X may influence an outcome variable Y through a mediating variable M or multiple mediators. In our study, X is age of onset of puberty; Y is lung function parameters at age 26 years; and M is DNAm at age 18 years. The direct effect of age of puberty on lung function at age 26 years is assessed by regressing lung function parameters at age 26 years on age at puberty, and the regression coefficient represents the direct effect. Indirect effect is assessed by age at puberty indirectly affecting lung function parameters at age 26 years through DNAm at age 18 years. A total effect is the sum of direct and indirect effects. #: asthma status, height, body mass index (BMI), socioeconomic status (SES) and atopy status at age 10 years, low birthweight status, maternal smoking during pregnancy, and cell types; ¶: smoking at age 18 years; +: asthma status, height, BMI, SES and atopy status at age 10 years, low birthweight status, and maternal smoking during pregnancy.

CpGs associated with the two lung function parameters (backward screening). Screening was carried out using the R package *ttScreening* [34], and CpGs showing association with age at puberty or lung function at age 26 were included in subsequent path analyses. To assess the potential of mediation effects of DNAm, PROC CALIS (SAS 9.4) was applied to evaluate direct and indirect effects (*via* DNAm) of ages at pubertal events on FVC or FEV₁ (figure 2); statistical significance level was set at 0.05.

Replication study in ALSPAC

Pubertal events showing statistically significant indirect effects through DNAm on FVC or FEV₁ were further tested in the ALSPAC birth cohort [24–26] (supplementary material section S2). Complete data on demographics, ages of puberty onset (menarche and body hair growth for females, and body hair growth and voice deepening for males), DNAm at ages 15 or 17 years, along with FVC and FEV₁ at age 24 years were analysed (n=353 females and 335 males) [35]. The same path analysis modelling with comparable covariates was applied. Statistical significance was set at p-value <0.05.

Results

Population characteristics

The IOWBC subsamples reasonably represented the complete cohort, indicated by statistically insignificant differences between the subsamples and the complete cohort on the variables of interest (all p-values ≥0.05, table 1).

Initial screening for pubertal events

Based on linear regressions, ages at menarche and age at body hair growth in females were associated with both FVC and FEV₁, and for males, age of body hair growth, growth spurt, voice deepening and facial hair growth were shown to be associated with FVC only (supplementary table S1). Age of onset for these pubertal events was included in the subsequent mediation analyses.

Mediation analysis through DNAm at CpGs that passed forward screening

In total, 321 unique CpGs for females and 489 unique CpGs for males survived forward screening and were included in path analyses (supplementary material section S3 and figure 3). Results from path analyses in IOWBC indicated that, for females, at 6 and 7 CpGs, DNAm mediated the association of age at menarche with FVC and FEV₁ at age 26 years, respectively (p<0.05; supplementary table S2A). For FEV₁ at age 26 years, 4 CpGs showed a mediation effect on the association of age at body hair growth

TABLE 1 Comparison of the analytical subsample (n=434) with the whole Isle of Wight birth cohort (IOWBC) (n=1261) on ages of pubertal events and lung function

Variables	Cohort samples	Subsamples	p-value
Subjects n	1261	434	
Age of pubertal events years			
Female			
Breast growth	12.43±1.58	12.29±1.60	0.20
Body hair growth	12.25±1.43	12.21±1.42	0.66
Growth spurt	12.52±1.70	12.36±1.80	0.21
Skin changes	13.11±1.50	13.08±1.59	0.79
Menarche	12.72±1.42	12.60±1.37	0.18
Male			
Body hair growth	13.41±1.37	13.26±1.40	0.14
Growth spurt	13.72±1.67	13.54±1.63	0.14
Voice deepening	14.24±1.24	14.07±1.19	0.05
Facial hair growth	15.38±1.16	15.32±1.09	0.45
Skin changes	13.99±1.38	13.94±1.29	0.61
Lung function L (age 26 years)			
Female			
FVC	4.24±0.54	4.29±0.55	0.24
FEV ₁	3.42±0.43	3.45±0.42	0.45
Male			
FVC	5.85±0.82	5.88±0.82	0.89
FEV ₁	4.61±0.72	4.66±0.72	0.50
Data presented as mean±SD. FVC: forced vital capacity; FEV ₁ : forced expiratory volume in 1 s.			

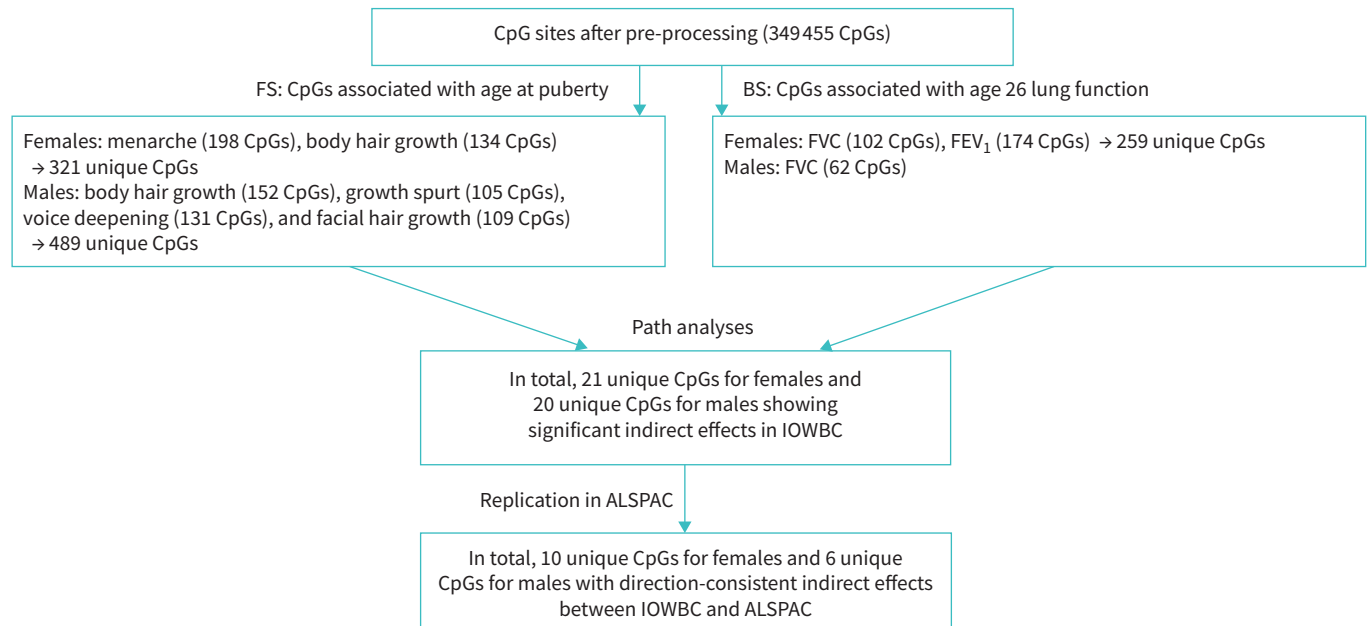


FIGURE 3 Flow chart of identified cytosine-phosphate-guanine (CpG) sites from screening CpGs through mediation analysis. FVC: forced vital capacity; FEV₁: forced expiratory volume in 1 s; FS: forward screening; BS: backward screening; IOWBC: Isle of Wight birth cohort; ALSPAC: Avon Longitudinal Study of Parents and Children.

with FEV₁ ($p < 0.05$; supplementary table S2A), but we did not identify any mediating CpGs for FVC at age 26 years. In total, 14 unique CpGs were identified. At these CpGs, of the total effects (direct + indirect effects), direct effects accounted for a large fraction of the total effect, *e.g.* at cg08680129, ~78% of the estimated total effect of age at menarche on FVC at age 26 was contributed by direct effect, and an indirect effect through DNAm only contributed a small portion (supplementary table S2A and figure S3). The small contributions from indirect effects were also reflected by the small ratios of indirect effect over direct effect, *e.g.* at cg08680128, the ratio is $0.04/0.14 = 2/7$. It is worth noting that using ratios of indirect effect over direct effect is especially informative when the two effects were in opposite directions (*e.g.* a negative direct effect but positive indirect effect), in which case the ratio was calculated using absolute values of estimated coefficients.

In males, through path analyses, we found that DNAm at 7, 3, 2 and 4 candidate CpGs mediated the association of ages of body hair growth, growth spurt, voice deepening and facial hair growth, respectively, with FVC at age 26 years (p -values < 0.05) (supplementary table S2A). In total, 16 unique CpGs demonstrated statistically significant mediation effects. At 10 of the 16 identified CpGs, ratios of indirect effect over direct effect were < 0.5 , indicating an overall small contribution of DNAm as a mediator on the association of age of pubertal events with FVC.

Mediation analysis through DNAm at CpGs that passed backward screening

Path analyses were applied to 321 unique CpGs that passed backward screening (259 in females; supplementary material section S3 and figure 3). In IOWBC, for females, DNAm at three of the 259 candidate CpGs was shown to mediate the association between age at menarche and FEV₁ at age 26 years ($p < 0.05$) (supplementary table S2B). Mediation effects were not identified at any of the candidate CpGs for FVC at age 26 years. For age at body hair growth, DNAm at three and four CpGs mediated its association with FVC and FEV₁, respectively ($p < 0.05$) (supplementary table S2B). Altogether, seven unique CpGs showed statistically significant mediation effects, but in most mediation analyses, magnitudes of indirect effects were lower than a half of that for direct effect (in eight of the 10 mediation analyses; supplementary table S2B).

In males, for ages of body hair growth, growth spurt, voice deepening and facial hair growth, mediation effects of DNAm on the association with FVC at age 26 years were found at one CpG site of the 62 candidate CpGs for each of the four pubertal events (p -values < 0.05) (supplementary table S2B). At all

these four CpGs, contributions of indirect effects to total effects were small, ranging from 25% to 38% (supplementary table S2B) and the ratios of indirect over direct effects were all <0.6 .

Comparison between forward screening and backward screening

In females, no common CpGs were observed between the 321 CpGs that survived forward screening and the 259 CpGs that survived backward screening. Similarly, for males, no common CpGs were identified between 489 CpGs that passed forward screening and the 62 CpGs that passed backward screening. When comparing common genes to which these CpGs were mapped between forward screening and backward screening, seven and one common genes were observed in females and males, respectively.

Further assessment of the IOWBC-discovered CpGs in ALSPAC

In females, of the 21 discovered CpGs, indirect effects at 10 CpGs (tables 2 and 3) were consistent with those observed in IOWBC, but none of them were statistically significant. In males, ages of body hair growth and voice deepening were available in ALSPAC. Of the 11 IOWBC-discovered CpGs for these two pubertal events, consistent mediation effects were observed at six CpGs but were statistically insignificant (tables 2 and 3). Overall, indirect effects of these two pubertal events were negligible in ALSPAC, consistent with the findings in the IOWBC. Detailed results are given in the supplementary material section S4.

Discussion

Combining findings in both the discovery and replication cohorts, post-adolescence or close-to-post-adolescence DNA methylation showed negligible mediation effects on the association of age at pubertal events with young adulthood FVC or FEV₁. Specifically, indirect effects only represented a small proportion of total effects of age at pubertal events on the two lung function measurements in young adults. Overall, the findings in both cohorts did not support a role for DNAm in mediating the association of pubertal events with FVC or FEV₁.

Our findings were drawn based on a rigorous study design with the potential to maximise statistical power. We used a comprehensive screening process to screen pubertal events and CpG sites to maximise statistical power. The two ways of screening for CpG sites, age at puberty \rightarrow DNAm (forward), DNAm \rightarrow lung function (backward), allowed us to include as many potentially informative CpG sites (with respect to potential mediation effects) as possible. These carefully designed screening procedures provided us with an opportunity to identify mediation CpGs, if any, with high power. Even with this carefully designed screening process, our data in both cohorts did not identify CpGs showing strong potential as mediators. In addition, the findings between the discovery and replication cohort were consistent and not affected by cell type compositions. Cell type could introduce multicollinearity in path analyses [36], which could lead to an apparent reversal of the direction of indirect effects and changes in the size of indirect effects. In IOWBC, the direction of associations with or without cell type compositions was intact for the indirect effects. For the indirect effects estimated in the ALSPAC cohort, although the direction of associations changed at two CpGs when removing the cell types in the models, since the size of indirect effects was very trivial regardless of the involvement of cell types, we expect the change in direction of association was due to random error.

Although our data did not support the role of DNAm as a mediator between age at puberty and young adult lung function parameters, we cannot rule out the existence of CpGs (but we have not discovered) acting as mediators. One explanation of the lack of association might be that current methods for arrays (Illumina 450K and EPIC arrays) cover only 2–4% of CpG sites across the whole genome [37, 38]. In addition, we used DNAm at CpGs from peripheral blood. DNAm shows strong tissue-specific effects and DNAm in another tissue (*e.g.* lung) might mediate the association of pubertal event timing with lung function [39]. Our mediation analyses assumed a linear relationship between exposure and outcome variables. However, biological relationships can be complex and nonlinear with unknown forms. More advanced analytical approaches in this area are certainly needed.

Our study has multiple strengths. The first is the utilisation of discovery and replication cohorts, and the results from these two independent cohorts suggest that our identified insignificant mediation effects of DNAm were potentially conclusive. Second, to avoid reverse causation, the mediation analyses were built upon a longitudinal setting, offering an insight to future investigations on possible “causal paths” and helping open the “black box” from age of pubertal events onset to lung function later in life [40]. Third, the carefully designed screening process in the detection of candidate pubertal events and candidate CpGs offered a strong potential to improve statistical power and at the same time reduced the risk of excluding potentially important epigenetic mediators. *Via* currently available epigenetic assessment platforms of Illumina, results from our carefully designed study support the absence of epigenetic mediation on the

TABLE 2 Associations of age of puberty with FVC and FEV₁ at age 26 years through DNA methylation at CpG sites that survived forward screening at age 18 years in the Isle of Wight birth cohort (IOWBC), further tested in the Avon Longitudinal Study of Parents and Children (ALSPAC) cohort (only direction-consistent indirect effects included)

Pubertal events (sex)	Lung function at age 26 years	CpGs	Total effect [#] (IOWBC)		Direct effect [#] (IOWBC)		Indirect effect [#] (IOWBC)		Total effect (ALSPAC)		Direct effect (ALSPAC)		Indirect effect (ALSPAC)	
			Est [¶]	p-value [¶]	Est	p-value	Est	p-value	Est	p-value	Est	p-value	Est	p-value
Menarche (female)	FVC	cg15711028	0.19	0.01	0.15	0.04	0.04	0.04	0.21	8.00×10 ⁻⁴	0.21	9.33×10 ⁻⁴	2.68×10 ⁻³	0.51
		cg08680129	0.18	0.01	0.14	0.06	0.04	0.03	0.22	5.62×10 ⁻⁴	0.20	1.28×10 ⁻³	1.41×10 ⁻²	0.15
		cg01315063	0.19	0.01	0.13	0.07	0.06	0.02	0.21	7.98×10 ⁻⁴	0.21	8.82×10 ⁻⁴	1.74×10 ⁻³	0.59
	FEV ₁	cg17394530	0.23	9.22×10 ⁻⁴	0.28	5.54×10 ⁻⁵	-0.05	0.02	0.23	3.61×10 ⁻⁴	0.23	3.44×10 ⁻⁴	-4.66×10 ⁻⁴	0.90
		cg08680129	0.24	5.92×10 ⁻⁴	0.20	0.01	0.05	0.02	0.23	3.01×10 ⁻⁴	0.22	5.81×10 ⁻⁴	1.02×10 ⁻²	0.25
		cg01315063	0.24	5.85×10 ⁻⁴	0.20	0.01	0.04	0.05	0.23	3.81×10 ⁻⁴	0.23	4.40×10 ⁻⁴	2.58×10 ⁻³	0.54
Body hair growth (female)	FEV ₁	cg16332936	0.14	0.11	0.09	0.30	0.05	0.04	0.18	3.10×10 ⁻³	0.18	4.29×10 ⁻³	7.04×10 ⁻³	0.34
		cg04827020	0.17	0.04	0.12	0.16	0.05	0.03	0.18	3.30×10 ⁻³	0.17	5.91×10 ⁻³	1.13×10 ⁻²	0.17
Body hair ⁺ growth (male)	FVC	cg13636404	0.14	0.10	0.08	0.37	0.06	0.02	1.0×10 ⁻³	0.99	-4.1×10 ⁻⁴	1.00	1.40×10 ⁻³	0.71
		cg04130455	0.15	0.08	0.23	0.01	-0.08	0.01	3.9×10 ⁻³	0.95	0.01	0.93	-1.78×10 ⁻³	0.83
		cg21276535	0.16	0.05	0.11	0.20	0.06	0.02	4.8×10 ⁻³	0.94	-2.3×10 ⁻⁴	1.00	5.04×10 ⁻³	0.36
		cg27626110	0.11	0.21	0.17	0.05	-0.06	0.04	-7.3×10 ⁻⁴	0.99	-5.6×10 ⁻⁴	0.99	-1.70×10 ⁻⁴	0.96
Voice deepening (male)	FVC	cg01482375	0.17	0.04	0.12	0.16	0.05	0.04	-0.02	0.77	-0.03	0.68	7.88×10 ⁻³	0.28

The complete results are in supplementary table S2A. p-values <0.05 are in bold, representing statistically significant indirect effects in IOWBC. FVC: forced vital capacity; FEV₁: forced expiratory volume in 1 s; CpG: cytosine-phosphate-guanine. [#]: total effect is the sum of direct and indirect effects of age of pubertal events on lung function at age 26 years; indirect effect is the effect of age of pubertal events on lung function *via* DNA methylation at CpGs; all the effects are represented by regression coefficients. [¶]: Est: regression coefficient estimates with unit in litres (L); it represents expected lung function change for 1-year increase in pubertal age. ⁺: age at body hair growth is identified by age at attainment of Tanner stage >2 in ALSPAC cohort.

TABLE 3 Associations of age of puberty with FVC and FEV₁ at age 26 years through DNA methylation at CpG sites that survived backward screening at age 18 years in the Isle of Wight birth cohort (IOWBC), further tested in the Avon Longitudinal Study of Parents and Children (ALSPAC) cohort (only direction-consistent indirect effects included)

Pubertal events (sex)	Lung function at age 26 years	CpGs	Total effect [#] (IOWBC)		Direct effect [#] (IOWBC)		Indirect effect [#] (IOWBC)		Total effect (ALSPAC)		Direct effect (ALSPAC)		Indirect effect (ALSPAC)	
			Est [¶]	p-value [¶]	Est	p-value	Est	p-value	Est	p-value	Est	p-value	Est	p-value
Body hair growth (female)	FVC	cg04988476	0.14	0.09	0.1	0.22	0.04	0.05	0.14	0.02	0.14	0.02	1.55×10 ⁻³	0.64
Body hair growth (female)	FEV ₁	cg25601319	0.12	0.15	0.08	0.36	0.04	0.03	0.19	2.52×10 ⁻³	0.19	2.50×10 ⁻³	5.74×10 ⁻⁶	1.00
		cg04988476	0.13	0.11	0.09	0.28	0.04	0.04	0.19	2.23×10 ⁻³	0.19	2.52×10 ⁻³	2.39×10 ⁻³	0.54
		cg07143296	0.13	0.12	0.1	0.25	0.04	0.05	0.18	2.97×10 ⁻³	0.18	3.52×10 ⁻³	3.86×10 ⁻³	0.52
Body hair ⁺ growth (male)	FVC	cg06998210	0.13	0.13	0.09	0.32	0.05	0.05	0.005	0.94	4.31×10 ⁻³	0.95	1.03×10 ⁻³	0.79

The complete results are in supplementary table S2B. p-values <0.05 are in bold, representing statistically significant indirect effects in IOWBC. FVC: forced vital capacity; FEV₁: forced expiratory volume in 1 s; CpG: cytosine-phosphate-guanine. [#]: total effect is the sum of direct and indirect effects of age of pubertal events on lung function at age 26 years; indirect effect is the effect of age of pubertal events on lung function *via* DNA methylation at CpGs; all the effects are represented by regression coefficients. [¶]: Est: regression coefficient estimates with unit in litres (L); it represents expected lung function change for 1-year increase in pubertal age. ⁺: age at body hair growth is identified by age at attainment of Tanner stage >2 in ALSPAC cohort.

association of age at puberty with lung function measurements FVC and FEV₁, based on DNAm measured in whole blood. Further assessment is certainly deserved, accompanied by a much denser assessment of DNAm such as pyrosequencing, and by novel analytical approaches such as analysing a set of CpGs in one region as a unified entity. However, genome-scale pyrosequencing with all CpG sites covered is challenging. Novel analytical tools are greatly needed and will allow an in-depth assessment of the role of DNAm in the association of age at puberty with lung function measurements. Furthermore, instead of DNAm in peripheral blood, using DNAm measured in lung tissues or tissues in the respiratory system will enable us to assess whether our insignificant findings were due to tissue-specificity and in the meantime to better reveal underlying biological mechanisms from pubertal events to lung function in young adulthood. Also, in the screening process based on data in the IOWBC, some pubertal events were not associated with FEV₁ or FVC. The lack of association between age at those pubertal events and lung function parameters was likely due to the variations caused by recall bias. In IOWBC, ages of onset for each pubertal event were obtained by retrospective assessment at age 18 through questionnaire. The timing order of different pubertal ages in IOWBC had been evaluated in our earlier study, and we showed that the timing order was consistent with other studies [29]. However, compared to pubertal events such as menarche that have less recall bias on age of onset [41], ages at other pubertal events, *e.g.* skin changes, were likely to have more recall bias, potentially leading to insignificant associations with FEV₁ or FVC.

Conclusion

Our study suggests that DNA methylation in peripheral blood in post- or late-adolescence might not play a role in mediating the association between age of pubertal events and lung function in young adults.

Provenance: Submitted article, peer reviewed.

Acknowledgements: We are extremely grateful to all participants and their families who are followed up in the Isle of Wight 1989 birth cohort and the Avon Longitudinal Study of Parents and Children cohort, which include interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, midwives and nurses. We also appreciate the hard efforts of these two research teams in maintaining the cohorts and collecting data. Our special thanks also go to the High Performance Computing facility provided by the University of Memphis.

Author contributions: The original concept and design were initiated by Hongmei Zhang and Wilfried Karmaus. Data analyses were performed by Liang Li and Hongmei Zhang. Data interpretation and manuscript writing was led by Liang Li and Hongmei Zhang. Critical comments on the manuscript and final approval of the manuscript were given by all the authors: Liang Li, Hongmei Zhang, John W. Holloway, Susan Ewart, Caroline L. Relton, S. Hasan Arshad and Wilfried Karmaus.

Conflict of interest: None declared.

Support statement: This work was supported by the National Institutes of Health research funds R01AI121226 (H. Zhang and J.W. Holloway) and R01 AI091905 (W. Karmaus). The 10-year follow-up of IOWBC was funded by National Asthma Campaign, UK (Grant No. 364) and the 18-year follow-up by R01HL082925 (S.H. Arshad). The UK Medical Research Council and Wellcome (grant ref. 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. A comprehensive list of grants funding is available on the ALSPAC website (<http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf>). The sponsor had no role in study design, data collection and analysis, or the preparation of the manuscript. This publication is the work of the authors and H. Zhang will serve as guarantor for the contents of this paper. Funding information for this article has been deposited with the Crossref Funder Registry.

References

- 1 Macsali F, Real FG, Plana E, *et al.* Early age at menarche, lung function, and adult asthma. *Am J Respir Crit Care Med* 2011; 183: 8–14.
- 2 Gill D, Sheehan NA, Wielscher M, *et al.* Age at menarche and lung function: a Mendelian randomization study. *Eur J Epidemiol* 2017; 32: 701–710.
- 3 Murri V, Antoniazzi F, Piazza M, *et al.* Lung function in women with idiopathic central precocious puberty: a pilot study. *Horm Res Paediatr* 2017; 87: 95–102.
- 4 Mahmoud O, Granell R, Tilling K, *et al.* Association of height growth in puberty with lung function: a longitudinal study. *Am J Respir Crit Care Med* 2018; 198: 1539–1548.
- 5 Campbell B, Simpson JA, Bui DS, *et al.* Early menarche is associated with lower adult lung function: a longitudinal cohort study from the first to sixth decade of life. *Respirology* 2020; 25: 289–297.

- 6 Li L, Zhang H, Holloway JW, *et al.* Pubertal onset with adulthood lung function mediated by height growth in adolescence. *ERJ Open Res* 2020; 6: 00535-02020.
- 7 Li S, Hursting SD, Davis BJ, *et al.* Environmental exposure, DNA methylation, and gene regulation: lessons from diethylstilbesterol-induced cancers. *Ann N Y Acad Sci* 2003; 983: 161–169.
- 8 Zhao W, Ammous F, Ratliff S, *et al.* Education and lifestyle factors are associated with DNA methylation clocks in older African Americans. *Int J Environ Res Public Health* 2019; 16: 3141.
- 9 Needham BL, Smith JA, Zhao W, *et al.* Life course socioeconomic status and DNA methylation in genes related to stress reactivity and inflammation: The Multi-Ethnic Study of Atherosclerosis. *Epigenetics* 2015; 10: 958–969.
- 10 Smith JA, Zhao W, Wang X, *et al.* Neighborhood characteristics influence DNA methylation of genes involved in stress response and inflammation: The Multi-Ethnic Study of Atherosclerosis. *Epigenetics* 2017; 12: 662–673.
- 11 Tobi EW, Slieker RC, Luijk R, *et al.* DNA methylation as a mediator of the association between prenatal adversity and risk factors for metabolic disease in adulthood. *Sci Adv* 2018; 4: eaa04364.
- 12 Fujii R, Sato S, Tsuboi Y, *et al.* DNA methylation as a mediator of associations between the environment and chronic diseases: a scoping review on application of mediation analysis. *Epigenetics* 2021; in press [<https://doi.org/10.1080/15592294.2021.1959736>].
- 13 McDade TW, Ryan C, Jones MJ, *et al.* Social and physical environments early in development predict DNA methylation of inflammatory genes in young adulthood. *Proc Natl Acad Sci USA* 2017; 114: 7611–7616.
- 14 Barker ED, Walton E, Cecil CAM. Annual Research Review: DNA methylation as a mediator in the association between risk exposure and child and adolescent psychopathology. *J Child Psychol Psychiatry* 2018; 59: 303–322.
- 15 Lomniczi A, Loche A, Castellano JM, *et al.* Epigenetic control of female puberty. *Nat Neurosci* 2013; 16: 281–289.
- 16 Luo L, Yao Z, Ye J, *et al.* Identification of differential genomic DNA methylation in the hypothalamus of pubertal rat using reduced representation bisulfite sequencing. *Reprod Biol Endocrinol* 2017; 15: 81.
- 17 Almstrup K, Lindhardt Johansen M, Busch AS, *et al.* Pubertal development in healthy children is mirrored by DNA methylation patterns in peripheral blood. *Sci Rep* 2016; 6: 28657.
- 18 Demetriou CA, Chen J, Polidoro S, *et al.* Methylome analysis and epigenetic changes associated with menarcheal age. *PLoS One* 2013; 8: e79391.
- 19 Chen S, Refaey H, Mukherjee N, *et al.* Age at onset of different pubertal signs in boys and girls and differential DNA methylation at age 10 and 18 years: an epigenome-wide follow-up study. *Hum Reprod Open* 2020; 2020: hoaa006.
- 20 Machin M, Amaral AF, Wielscher M, *et al.* Systematic review of lung function and COPD with peripheral blood DNA methylation in population based studies. *BMC Pulm Med* 2017; 17: 54.
- 21 Imboden M, Wielscher M, Rezwan FI, *et al.* Epigenome-wide association study of lung function level and its change. *Eur Respir J* 2019; 54: 1900457.
- 22 MacKinnon DP, Fairchild AJ, Fritz MS. Mediation analysis. *Annu Rev Psychol* 2007; 58: 593–614.
- 23 Arshad SH, Holloway JW, Karmaus W, *et al.* Cohort Profile: The Isle Of Wight Whole Population Birth Cohort (IOWBC). *Int J Epidemiol* 2018; 47: 1043–1044i.
- 24 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: 111–127.
- 25 Fraser A, Macdonald-Wallis C, Tilling K, *et al.* Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol* 2013; 42: 97–110.
- 26 Northstone K, Lewcock M, Groom A, *et al.* The Avon Longitudinal Study of Parents and Children (ALSPAC): an update on the enrolled sample of index children in 2019. *Wellcome Open Res* 2019; 4: 51.
- 27 Albert JM, Nelson S. Generalized causal mediation analysis. *Biometrics* 2011; 67: 1028–1038.
- 28 Petersen AC, Crockett L, Richards M, *et al.* A self-report measure of pubertal status: reliability, validity, and initial norms. *J Youth Adolesc* 1988; 17: 117–133.
- 29 Yousefi M, Karmaus W, Zhang H, *et al.* Relationships between age of puberty onset and height at age 18 years in girls and boys. *World J Pediatr* 2013; 9: 230–238.
- 30 American Thoracic Society. Standardization of Spirometry, 1994 Update. *Am J Respir Crit Care Med* 1995; 152: 1107–1136.
- 31 Adalsteinsson BT, Gudnason H, Aspelund T, *et al.* Heterogeneity in white blood cells has potential to confound DNA methylation measurements. *PLoS One* 2012; 7: e46705.
- 32 Aryee MJ, Jaffe AE, Corrada-Bravo H, *et al.* Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* 2014; 30: 1363–1369.
- 33 Houseman EA, Accomando WP, Koestler DC, *et al.* DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* 2012; 13: 86.
- 34 Ray MA, Tong X, Lockett GA, *et al.* An efficient approach to screening epigenome-wide data. *Biomed Res Int* 2016; 2016: 2615348.

- 35 Harris PA, Taylor R, Thielke R, *et al.* Research electronic data capture (REDCap): a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009; 42: 377–381.
- 36 Barton SJ, Melton PE, Titcombe P, *et al.* In epigenomic studies, including cell-type adjustments in regression models can introduce multicollinearity, resulting in apparent reversal of direction of association. *Front Genet* 2019; 10: 816.
- 37 Fan S, Li C, Ai R, *et al.* Computationally expanding Infinium HumanMethylation450 BeadChip array data to reveal distinct DNA methylation patterns of rheumatoid arthritis. *Bioinformatics* 2016; 32: 1773–1778.
- 38 Pidsley R, Zotenko E, Peters TJ, *et al.* Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol* 2016; 17: 208.
- 39 de Vries M, van der Plaats DA, Nedeljkovic I, *et al.* From blood to lung tissue: effect of cigarette smoke on DNA methylation and lung function. *Respir Res* 2018; 19: 212.
- 40 Hafeman DM, Schwartz S. Opening the Black Box: a motivation for the assessment of mediation. *Int J Epidemiol* 2009; 38: 838–845.
- 41 Coleman L, Coleman J. The measurement of puberty: a review. *J Adolesc* 2002; 25: 535–550.