

## ARTICLE OPEN

## Home environment: respiratory and allergic phenotypes from birth to age six in the PELAGIE cohort

Katharina Apel<sup>1,2,3</sup>, Nathalie Costet<sup>1</sup>, Anthony Chapron<sup>2,3</sup>, Sylvaine Cordier<sup>1</sup>, Christine Monfort<sup>1</sup>, Cécile Chevrier<sup>1</sup> and Fabienne Pelé<sup>1,2,3</sup>

Childhood asthma and allergies are particularly prevalent diseases. Our objective is to identify respiratory and allergic phenotypes from birth to 6 years of age, and to explore their environmental determinants, especially those related to the home environment. Data on respiratory and allergic health outcomes and domestic environmental exposure were collected for 935 mother–infant pairs from a longitudinal mother–child cohort based on mothers, included before 19 weeks of gestation in Brittany between 2002 and 2006. Information was obtained by self-administered questionnaires completed by parents at inclusion, delivery, and when the child was 2 and 6 years old. Kml3D clustering was used to describe profiles of children who shared similar trajectories of symptoms as phenotypes. Association with environmental determinants was estimated by polytomous logistic regression. Five phenotypes were identified: a reference group characterized by low symptom levels (31.1%), a transient cough phenotype (36.5%), an eczema/cough phenotype (12.3%), a wheeze/cough phenotype (11.8%), and finally a mixed phenotype (8.0%). The wheeze/cough profile was associated with postnatal exposure to glues used in renovation activities (aOR 2.3 [1.2–4.7]), and the mixed phenotype with postnatal exposure to paint (aOR 2.1 [1–4.5]). The phenotypes observed showed some consistencies with those seen in previous studies. Some exposures associated with respiratory/allergic phenotypes observed in this study are avoidable. If confirmed by further research including interventional trials, home-based environmental counseling could be a possible prevention target for primary care professionals.

*npj Primary Care Respiratory Medicine* (2019)29:29; <https://doi.org/10.1038/s41533-019-0141-y>

## INTRODUCTION

The prevalence of childhood asthma and allergies varies greatly worldwide:<sup>1</sup> at age 6–7, 2.8–37.6% of children have had asthma symptoms in the past year and 2–22.3% eczema symptoms. Asthma is the most frequent chronic disease in childhood and represents a major burden for patients, their families, and society.<sup>2</sup>

Asthma and allergic disorders commonly appear in infancy, but their pathogenesis and interplay remain incompletely understood.<sup>3</sup> To date, only three studies have collected longitudinal data for multiple 12-month profiles of respiratory and allergic symptoms.<sup>4–6</sup> Rancière et al.<sup>4</sup> simultaneously observed profiles of wheezing, dry night cough, eczema, and rhinitis symptoms from ages 0 to 4 in 2522 children from the PARIS (Pollution and Asthma Risk: an Infant Study) birth cohort. Two transient (rhinitis or wheeze) and two persistent (cough/rhinitis or dermatitis) phenotypes were identified, with a low symptoms reference group. Panico et al.<sup>6</sup> analyzed wheezing and proxies for atopy (eczema and/or hay fever ever) in 11,632 children from the Millennium Cohort Study at ages 3, 5, and 7 and obtained a 4-trajectory clustering: a group with low symptom levels, another with low wheeze and high atopic symptoms, a third with high levels of both wheeze and atopic symptoms, and finally a group with high levels of wheeze only. Belgrave et al.<sup>5</sup> studied wheeze, eczema, and rhinitis at ages 1, 3, 5, 7, and 11 in 9801 children from the ALSPAC (The Avon Longitudinal Study of Parents and Children) and MAAS (Manchester Asthma and Allergy Study) birth cohorts and identified eight developmental profiles: no disease, atopic march, persistent eczema and wheeze, persistent eczema with

later-onset rhinitis, persistent wheeze with later-onset rhinitis, transient wheeze, eczema only, and rhinitis only.

The development and expression of asthma/allergic disorders depend on gene–environment interactions.<sup>7</sup> Worldwide variations in prevalence rates of asthma and allergies call attention to the potential role of the environment, as genetic differences alone cannot explain this heterogeneity.<sup>1</sup> The literature has identified numerous avoidable determinants for respiratory and allergic disorders in childhood, in particular, the indoor environment,<sup>8–10</sup> because young children usually spend most of their time at home.<sup>9,11,12</sup> The objective of this study was to describe the joint trajectory of asthma, eczema, and rhinitis symptoms at three time points, at ages 1, 2, and 6 years, in a population-based cohort conducted in Brittany (France) from 2002 to 2006. We explored the associations of these phenotypes with some avoidable indoor environmental factors in early life.

## RESULTS

## Participants and descriptive data

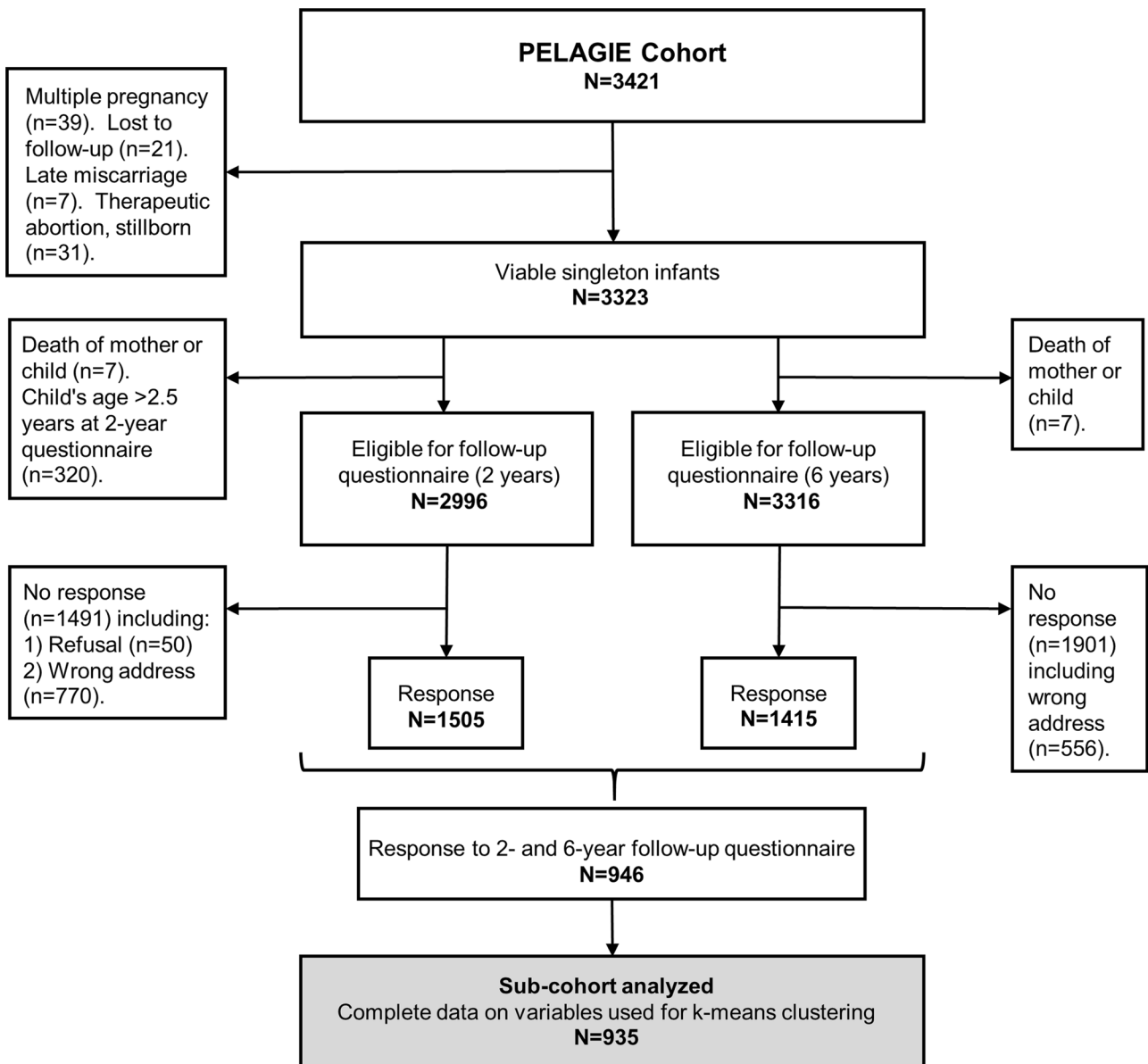
The cluster analysis, included 935 of the 3323 eligible mother–child couples from the PELAGIE (Perturbateurs Endocriens: Étude Longitudinale sur les Anomalies de la Grossesse, l'Infertilité et l'Enfance) mother–child cohort (Fig. 1).

Compared with the nonrespondents, respondents were older ( $p < 0.001$ ), had higher education level ( $p < 0.001$ ), and were less often nulliparous ( $p = 0.008$ ) or a smoker at inclusion ( $p < 0.001$ ) (Table 1).

<sup>1</sup>Univ Rennes, Inserm, EHESP, Irset (Institut de recherche en santé, environnement et travail) - UMR\_S 1085, F-35000 Rennes, France; <sup>2</sup>Univ Rennes, Department of General Practice, F-35000, France and <sup>3</sup>Univ Rennes, CHU Rennes, Inserm, CIC 1414 [(Centre d'Investigation Clinique de Rennes)], F-35000 Rennes, France  
Correspondence: Katharina Apel ([katharina.apel@univ-rennes1.fr](mailto:katharina.apel@univ-rennes1.fr))

Received: 18 October 2018 Accepted: 2 July 2019

Published online: 25 July 2019



**Fig. 1** Flow chart of the PELAGIE mother-child cohort. The absence to follow-up at different time points figures on the left (Inclusion, 2-year follow-up) and the right (6-year follow-up) side of the flow chart. Complete data for this analysis were available for 935 mother-child couples

The participating mothers had a mean age of 30.8 years (SD 4.1), and 69.8% had at least some post-secondary education. At inclusion, 22.0% smoked, and 14.7% drank some alcohol. Most of the children were full-term (96.8%) and born by vaginal delivery (83.2%); 51.3% were boys (Table 1). There was a family history of allergy in 249 (26.3%) mother-child pairs. Fewer than half of the children had been exclusively breastfed (41.1%) during the first 3 months. At age 2, 34.8% had been exposed to environmental tobacco smoke (ETS), 44.4% to bleach, 45.0% to glue, and 48.5% to alkyl/acrylic paints used for renovation activities in the dwellings. Half of the children had been exposed to pets at age 2, 40.5% slept on a mattress bought at least 3 years ago, and 41.2% in a bedroom aired less than once a day. Overall, 55.0% of the dwellings used individual heating devices fueled by gas/fuel oil/wood or chimneys (Table 2).

#### Outcome data

The best partition for the cluster analysis was a five-cluster solution (Fig. 2). The prevalence of respiratory/allergic symptoms

(wheeze, rash, rhinitis, and cough) at age 1, 2, and 6 years of the child (except cough, recorded only at age 2 and 6 years) were used to characterize and label each cluster obtained (supplementary Table 1). At time point 1, 4% of the subjects were assigned to a cluster we labeled “reference” phenotype. The prevalence of respiratory/allergic symptoms for this phenotype was quite low and ranged from 0.0 to 13.7%, except for cough, which was prevalent in 24.4% of the children at 6 years of age only. The “transient cough” phenotype (36.4%) had cough in the first 2 years (100%) but had recovered by age 6 to reach 25.2%, almost identical to the “reference” at that age. In the “eczema/cough” phenotype (12.3%), 94.8% of the children experienced itchy rash in early life, which decreased by age 6 to 33.0%. Cough was more frequent than in the reference group during the first 2 years and was still reported in 34.2% of these children at age 6. The “wheeze/cough” phenotype showed a very substantial decrease in the prevalence of wheezing, from 77.3% at age 2 to 22.7% at age 6. Cough also decreased in this phenotype, from 95.5% at age 2 to 29.1% at age 6. The itchy rash prevalence in this phenotype was

**Table 1.** Respondents and nonrespondents to the two-year follow-up questionnaire from the PELAGIE mother–child cohort

Variable	Analyzed subcohort <i>n</i> = 935 <i>N</i> (%)	Nonrespondents <i>n</i> = 2370 <i>N</i> (%)	<i>p</i> -Value <sub>1</sub>
<i>Maternal BMI at inclusion (kg/m<sup>2</sup>)<sub>2</sub></i>			
<18.5	62 (6.6)	181 (7.7)	0.4
18.5 ≤ BMI <25	720 (76.3)	1741 (74.2)	
25 ≤ BMI <30	122 (12.9)	302 (12.9)	
≥30	40 (4.2)	123 (5.2)	
<i>Maternal age at inclusion (years)</i>			
<30	423 (44.7)	1269 (53.5)	<0.0001
30 ≤ age < 35	371 (39.2)	790 (33.3)	
≥35	152 (16.1)	311 (13.1)	
<i>Nulliparous<sub>3</sub></i>			
Yes	386 (40.9)	1083 (45.9)	0.008
No	559 (59.2)	1276 (54.1)	
<i>Maternal level of education<sub>4</sub></i>			
Primary or secondary education	124 (13.1)	500 (21.2)	<0.0001
Baccalaureate	161 (17.0)	462 (19.6)	
Higher education	660 (69.8)	1401 (59.1)	
<i>Maternal smoking at inclusion<sub>5</sub></i>			
Non smoker	732 (78.0)	1623 (69.3)	<0.0001
<10 cig/day	142 (15.1)	469 (20.0)	
≥10 cig/day	64 (6.8)	249 (10.6)	
<i>Maternal alcohol consumption at inclusion<sub>6</sub></i>			
Yes	138 (14.7)	345 (14.8)	1.0
No	798 (85.2)	1193 (85.2)	
<i>Gender</i>			
Male	485 (51.3)	1186 (50.0)	0.5
Female	461 (48.7)	1184 (50.0)	
<i>Cesarean delivery (vs. vaginal)<sub>7</sub></i>			
Yes	156 (16.8)	411 (17.8)	0.5
<i>Preterm birth (&lt;37 weeks)<sub>8</sub></i>			
Yes	30 (28.6)	93 (4.0)	0.3
No	912 (96.1)	2262 (96.0)	
<i>Low birth weight (&lt;10% for gestational age)<sub>9</sub></i>			
Yes	59 (6.3)	156 (6.6)	0.7
No	885 (93.8)	2212 (93.4)	

BMI body mass index (kg/m<sup>2</sup>)

1—chi-square test; Missing data: 2—*n* = 25; 3—*n* = 12; 4—*n* = 8; 5—*n* = 37; 6—*n* = 42; 7—*n* = 73; 8—*n* = 19; 9—*n* = 4

**Table 2.** Characteristics of the analyzed subcohort (*n* = 935)

	Number (%) or mean (SD)
<i>Maternal characteristics</i>	
<i>Maternal BMI at inclusion</i>	
<18.5	62 (6.6)
18.5–25	720 (76.3)
25–30	122 (12.9)
≥30	40 (4.2)
Maternal age at inclusion (years)	30.8 (4.1)
Nulliparous (vs. multiparous)	386 (40.9)
<i>Maternal level of education</i>	
Primary or secondary education	124 (13.1)
Baccalaureate	161 (17.0)
Higher education	660 (69.8)
Maternal smoking at inclusion (vs. no)	206 (22.0)
Maternal alcohol consumption at inclusion (vs. no)	138 (14.7)
Family history of allergy (yes) <sub>1</sub>	249 (26.3)
<i>Child's characteristics</i>	
Male gender	485 (51.3)
Cesarean delivery (yes)	156 (16.8)
Exclusive breastfeeding during the first 3 months (yes)	389 (41.1)
Preterm birth (<37 weeks)	30 (3.2)
Low birth weight (<10% for gestational age)	59 (6.3)
Height (cm)	49.8 (2.2)
Weight (g)	3419 (484)
<i>Environmental characteristics</i>	
ETS at home since birth (yes)	329 (34.8)
Time spent in rooms cleaned with bleach (yes)	420 (44.4)
Use of glues for renovation activities in the first 2 years of life (vs. no)	326 (45.0)
Use of alkyd/acrylic paints for renovation activities in the first 2 years of life (vs. no)	459 (48.5)
Pets at home since birth	472 (50.0)
Child's mattress ≥ 3 years	383 (40.5)
Ventilation of the child's bedroom less than on time/day (vs. no)	390 (41.2)
Individual heating using gas/fuel oil/wood/chimneys (vs. no)	520 (55.0)

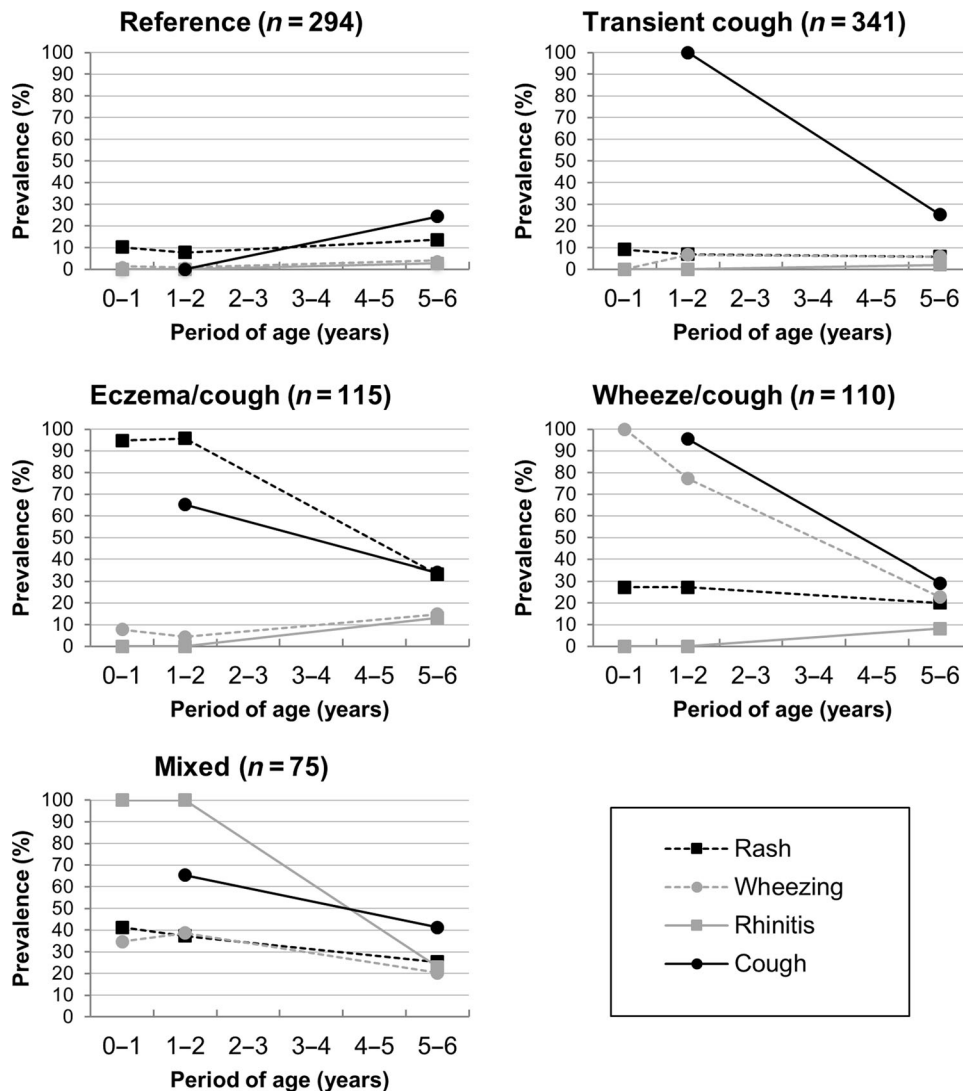
SD standard deviation, BMI body mass index, ETS environmental tobacco smoke

Missing values: 1—*n* = 137, *n* < 40 for other variables

27.3 at age 1 and 2 and 20.0% at age 6. In the “mixed” phenotype, all respiratory and allergic symptoms exceeded prevalence rates from the reference group. Eczema symptoms were more frequent than in any but the eczema/cough phenotype, and wheezing more frequent over time than in any but the wheezing/cough phenotype. Cough at age 6 was more prevalent (42.0%) than in either the wheeze/cough or transient cough phenotype.

The description of other respiratory and allergic health outcomes is available in Supplementary Table 2. At age 2, 64.9 and 72.9% children in the eczema/cough and wheeze/cough phenotypes, respectively, had had otitis media. Most of the children who had ever been awakened by breathlessness belonged to the wheeze/cough phenotype (15.5%), and most of those who had

been awakened by respiratory discomfort belonged to the mixed phenotype (41.9%). The mixed phenotype also contained the most children with hay fever (27.0%) and other allergic rhinitis (24.3%). Table 3 summarizes the associations between physician-diagnosed diseases and phenotypes. At the age 6 of the child, 107 (11.4%), 212 (22.6%), and 57 (6.1%) parents declared that a doctor had already diagnosed their child with asthma, eczema, or food allergy, respectively. The highest proportion of the 107 children with physician-diagnosed asthma belonged to the wheeze/cough phenotype (29%). Comparably 29.3% of the 212 children diagnosed with eczema were allocated to the eczema/cough phenotype, but 23.1% were defined to the reference and the transient cough phenotypes. At time point 2, 8 (1%) of children diagnosed with food allergy had been attributed to the eczema/



**Fig. 2** Description of the five-cluster solution obtained in the cluster analysis. Each graph corresponds to one cluster. Asthma and allergy symptoms (rash, wheezing, cough, and rhinitis symptoms) figure as prevalence at each time point (1, 2, and 6 years) and serve to characterize each cluster. The “reference group” had overall low asthma/allergy symptoms

cough phenotype. Logistic regression models showed that the wheeze/cough phenotype was associated only with physician-diagnosed asthma. Both the eczema/cough and mixed phenotypes were associated with all three physician-diagnosed conditions of interest (asthma, eczema, and food allergies). The transient cough phenotype, however, did not differ significantly from the reference group for any of these conditions. Other respiratory and/or allergic symptoms reported by parents are shown in Supplementary Table 2.

#### Main results

Factors associated with phenotypes are shown in Table 4. The risk of transient cough and eczema/cough phenotypes were lower in children born to mothers aged 35 years and older and in those with exclusive breastfeeding for 3 months. Family related factors associated with the wheeze/cough phenotype were multiparity, maternal smoking at inclusion, and family history of allergy; the latter was also associated with the mixed phenotype. Children born to mothers aged at least 35 years had a reduced risk of belonging to the wheeze/cough phenotype. An environmental determinant associated with this phenotype was exposure to glues for indoor renovation activities in early life. Children with the

wheeze/cough phenotype were at reduced risk at age 2 of sleeping on a mattress bought at least 3 years earlier. The only environmental determinant associated with the mixed phenotype was renovation-related use of alkyd/acrylic paints in early life. The sensitivity analysis (for preterm birth >37 weeks of gestation and birth weight < 10th percentile) did not modify the adjusted model.

#### DISCUSSION

This study identified five phenotypes of respiratory and allergic symptoms over the first 6 years of life in a population-based mother–child cohort. Phenotypes differed in their physician-diagnosed asthma/allergic diseases and parent-reported comorbidities. We observed an association between some avoidable domestic environmental determinants with respiratory/allergic phenotypes from birth to age 6, specifically, early postnatal exposure to glues and paints.

One of this study’s strengths is its use of an unsupervised statistical method to describe the coevolution of a set of respiratory and allergic symptoms. Unsupervised clustering techniques are particularly appropriate to describe patterns of evolution in a set of variables measured repeatedly like in birth

**Table 3.** Physician-diagnosed respiratory and allergic diseases in phenotypes

	Asthma <sub>1</sub>		Eczema <sub>2</sub>		Food allergy <sub>3</sub>	
	N = 107 (11.4%)		N = 212 (22.6%)		N = 57 (6.1%)	
	OR (CI 95%)	%	OR (CI 95%)	%	OR (CI 95%)	%
<i>Phenotypes</i>						
Reference	ref.	12.2	ref.	23.1	ref.	17.5
Transient cough	1.6 (0.8–3.1)	21.5	0.8 (0.5–1.3)	23.1	0.9 (0.4–2.3)	19.3
Eczema/cough	<b>4.0 (1.9–8.4)</b>	16.8	<b>6.0 (3.7–9.6)</b>	29.3	<b>4.5 (2.0–10.3)</b>	28.1
Wheeze/cough	<b>8.4 (4.2–16.9)</b>	29	1.4 (0.8–2.5)	11.8	1.6 (0.6–4.6)	10.5
Mixed	<b>8.0 (3.7–16.9)</b>	18.7	<b>2.2 (1.2–4.0)</b>	10.9	<b>5.4 (2.3–13.2)</b>	21.1

CI confidence interval, OR odds ratio, ref reference

The ORs are calculated in relation to the reference phenotype. Bold aORs differ significantly from the reference ( $p < 0.05$ ). Missing values: 1— $n = 32$ ;  $n = 83$ :  $n = 5$

**Table 4.** Factors associated with respiratory/allergic phenotypes from birth to age 6 in the PELAGIE cohort

	Phenotypes					p-Value
	Reference	Transient cough	Eczema/cough	Wheeze/cough	Mixed	
	N = 229	N = 238	N = 87	N = 82	N = 58	
	aOR (95% CI)	aOR (95% CI)	aOR (95% CI)	aOR (95% CI)	aOR (95% CI)	
<i>Maternal and family characteristics</i>						
Maternal BMI before pregnancy $\geq 30$ (vs. $<30$ )	ref.	1.8 (0.6–5.0)	2.9 (0.9–9.7)	0.8 (0.2–2.7)	2.4 (0.6–9.0)	0.33
Maternal age at inclusion $\geq 35$ years (vs. $<35$ years)	ref.	<b>0.5 (0.3–0.8)</b>	0.5 (0.3–1.1)	<b>0.4 (0.2–0.9)</b>	0.4 (0.2–1.0)	0.03
Parous (vs. nulliparous) at inclusion	ref.	1.3 (0.9–2.1)	1.0 (0.5–1.8)	<b>3.1 (1.6–5.7)</b>	1.0 (0.5–1.9)	0.01
Maternal smoking at inclusion <sub>1</sub>	ref.	0.9 (0.5–1.6)	0.8 (0.4–1.7)	<b>2.4 (1.2–4.9)</b>	1.1 (0.5–2.5)	0.06
Family history of allergy <sub>1</sub>	ref.	1.4 (0.9–2.3)	1.5 (0.8–2.7)	<b>2.2 (1.2–4.0)</b>	<b>2.5 (1.3–4.8)</b>	0.06
<i>Child's characteristics</i>						
Male gender	ref.	1.2 (0.8–1.8)	1.3 (0.8–2.1)	1.5 (0.9–2.6)	1.2 (0.7–2.1)	0.6
Cesarean delivery (vs. vaginal)	ref.	1.1 (0.7–1.9)	0.6 (0.3–1.4)	1.4 (0.7–2.9)	1.0 (0.4–2.3)	0.6
Exclusive breastfeeding $\geq 3$ months <sub>1</sub>	ref.	1.0 (0.7–1.4)	<b>0.5 (0.3–0.9)</b>	1.1 (0.7–2.0)	1.0 (0.6–1.4)	0.1
<i>Environmental characteristics</i>						
ETS at home since birth (yes)	ref.	0.9 (0.6–1.4)	0.9 (0.5–1.6)	1.2 (0.7–2.3)	1.3 (0.7–2.6)	0.7
Time spent in rooms cleaned with bleach <sub>1</sub>	ref.	1.1 (0.7–1.6)	1.5 (0.9–2.6)	1.4 (0.8–2.5)	1.5 (0.8–2.7)	0.3
Use of glues for renovation activities ( $<age 2$ ) <sub>1</sub>	ref.	1.3 (0.8–2.2)	1.0 (0.5–2.0)	<b>2.3 (1.2–4.7)</b>	0.8 (0.4–1.7)	0.1
Use of glycerophthalic/acrylic paints for renovation activities in the first two years of life <sub>1</sub>	ref.	1.1 (0.7–1.7)	1.6 (0.8–3.0)	0.8 (0.3–1.5)	<b>2.1 (1.0–4.4)</b>	0.1
Pets at home since birth	ref.	0.7 (0.5–1.1)	1.0 (0.6–1.7)	0.8 (0.4–1.3)	0.9 (0.5–1.6)	0.5
Child's mattress $\geq 3$ years old	ref.	0.8 (0.5–1.3)	0.8 (0.5–1.5)	<b>0.5 (0.3–0.9)</b>	1.1 (0.5–2.0)	0.2
Child's bedroom aerated less than once a day <sub>1</sub>	ref.	1.0 (0.7–1.4)	0.9 (0.5–1.5)	0.9 (0.5–1.7)	0.9 (0.5–1.6)	1.0
Individual heating using gas/fuel oil/wood/chimneys <sub>1</sub>	ref.	0.8 (0.6–1.2)	1.1 (0.7–1.8)	0.9 (0.6–1.6)	0.8 (0.5–1.5)	1.0

CI confidence interval, aOR adjusted odds ratio

Multinomial logistic regression model [the dependent variable is the five-cluster solution variable obtained in the cluster analysis, (ref = Reference phenotype)]. All the independent variables used in the final model are shown in this table. R-square of the full model was 0.14. Bold aORs differ significantly from the reference ( $p < 0.05$ ), 1: vs. no

cohorts. Kml3D allows clustering individuals on several trajectory variables jointly and takes into account their possible coevolution while liberating the observer from clinical a priori. Meanwhile, some subjective choices had to be made, basically the number of clusters tested, which was of necessity due to our sample size. We were able to explore symptoms in very early and late preschoolers over a relatively long period and past age 5, the threshold when wheezing starts to be an indicator of asthma. The use of binary data in the  $k$ -means algorithm might have led to clustering errors in the construction of phenotypes, but these errors are likely to be nonsystematic.

The most important limitation involves the measurement of symptoms on which phenotypes were built. Wheezing was

assessed with items derived from ISSAC (The International Study of Asthma and Allergies in Childhood) questionnaire,<sup>3</sup> validated on 7 year old children. It is uncertain what kinds of respiratory noises parents qualify as wheezing.<sup>13</sup> The questionnaire is nonetheless widely used in epidemiologic studies in younger children<sup>14</sup> and was used exclusively in two<sup>4,6</sup> of the three<sup>4–6</sup> articles to which we compare this study. Eczema symptoms might be overestimated. Because we did not assess typical localizations of itchy rash, and did not dispose of any eczema-medication data, noneczema symptoms such as xeroderma might have been assessed and could explain a certain degree of misclustering of children with eczema allocated to the reference group and the transient cough phenotype. Different cough measures were used



for the follow up of the 2 and 6 year olds. Since we did not dispose of any specific item on cough in the 2 first years of life, we supposed that parents who declared having had a child with bronchiolitis/bronchitis, had had a child with cough. This could be source of misclustering. Data on asthma/allergic symptoms were not collected between ages 2 and 6, so that we lack data about relevant time periods in symptom evolution, in particular, age 3.<sup>15,16</sup> Correct asthma medication might lead to less symptoms what for our symptom trajectories might be imprecise. But our phenotypes were consistent with most of the physician-diagnosed conditions and other respiratory and allergic symptoms, which reinforces their likelihood. Asthma medication was also consistent with phenotypes since it was more frequent in the wheeze/cough and mixed phenotypes. Misclustering due to under-declared or correctly treated symptoms are unlikely to be differential and would, as the case may be, let the association observed tend towards zero. We, therefore, do not believe that it affected associations observed.

Another limitation is attrition. The questionnaires were self-administered, sent by the postal service, and lengthy (>1 h for completion) due to the broad research scope covered by the cohort. Each of these factors may have discouraged participation and induced bias in association estimates. Because the outcomes measured in this paper do not concern behavioral variables, we believe, in accordance with Greene et al.<sup>17</sup> that follow-up bias is likely to be small.

Three previous studies have used unsupervised techniques to study multiple longitudinal respiratory- and allergy-related symptoms and to identify phenotypes.<sup>4–6</sup> Comparison appears thorny; not all studies used only the ISAAC questionnaire,<sup>5</sup> they considered different variables for phenotype identification, and health outcomes were not measured at the same time points. As Pinart et al.<sup>18</sup> recently underlined, efforts to standardize the reporting of allergic disease phenotypes in cohorts are needed. However, similarities are apparent, especially for phenotypes. Our wheeze/cough phenotype was exclusively associated with physician-diagnosed asthma and similar to the transient wheeze phenotype identified by Rancière et al.<sup>4</sup> Likewise, two of the studies<sup>4,5</sup> identified a group with high levels of eczema symptoms and low levels of other disorders, strongly associated with physician-diagnosed eczema.<sup>4</sup> Similarities also exist between the persistent sensitization-associated cough/rhinitis phenotype in Rancière et al.<sup>4</sup> and our mixed phenotype: cough and rhinitis were predominant, moderate wheeze and eczema symptoms remained, and it was associated with all the physician-diagnosed conditions tested (asthma, eczema, and food allergy).

The prenatal and early postnatal period is an important window of vulnerability for developing immune and respiratory systems, which are sensitive to environmental chemicals.<sup>19,20</sup> Our wheeze/cough phenotype was associated with early-life exposure to glue, and Rancière et al.<sup>4</sup> observed associations between transient wheezing and postnatal exposure to new particle board furniture. Our mixed phenotype showed associations with early exposure to alkyd/acrylic paints used for interior home redecoration. Glues, particle board furniture, and paints can emit chemical compounds that, inhaled or by direct skin contact, can induce adverse health effects. Early exposure to chemicals used for indoor renovation activity is suspected of harmful effects, including the development of asthma<sup>11,21</sup> and eczema.<sup>22</sup> We did not, however, measure these compounds, and their effects are not clearly understood.

The causal relation between postnatal exposure to ETS and respiratory outcomes in children is well established.<sup>9,23</sup> It was studied in two of the papers cited above<sup>4,6</sup> and observed by Rancière et al.<sup>6</sup> in two of their phenotypes, rhinitis and dermatitis. A lack of statistical power due to our smaller sample might explain the absence of association in our study.

Exposure to bleach was not associated with any phenotype. Studies of household cleaning products and their effect on

asthma/allergy in children are scarce and conflicting. Maternal exposure to a set of domestic chemicals including bleach has been associated with persistent wheezing in offspring at age 3.<sup>24</sup> More recent focus on the specific use of bleach during pregnancy has shown no association with wheezing at age 1,<sup>25</sup> while a cross-sectional study found 10–13-year-old children living in homes cleaned with bleach at least weekly less likely to have asthma, eczema, or sensitization to aeroallergens.<sup>26</sup> No phenotypes were associated with exposure to pets at home, whose respiratory/allergic health effects on children are currently the subject of debate.<sup>9,27</sup> Surprisingly, children from our wheeze/cough phenotype were less likely to sleep on a mattress supposed to be laden with house dust mite (HDM) than the reference group. Rancière et al.<sup>4</sup> observed a positive association between their dermatitis phenotype and the presence of a used mattress (older than 3 years at birth). Reverse causation due to the purchase of a new mattress in symptomatic children explaining our surprising association is improbable because it concerned only 19 children. HDM are thought to be implicated in the development of allergic asthma in children at high risk,<sup>28</sup> with a nonlinear relation: both very high and very low exposure to HDM are probably protective against asthma and HDM sensitization at age 5.<sup>29</sup> We treated the age of the mattress as a surrogate for HDM exposure in early life, but we cannot quantify the actual exposure, nor compare our sample to a population at high risk for asthma/allergy.

We do not believe that phenotypes identified represent clinical entities that clinicians should relate to. The development and expression of asthma/allergic disorders depend on gene-environment interactions. From the point of view of the primary care giver, if one focuses on potential prevention targets, to date, genetic predisposition to respiratory/allergic diseases is inaccessible to intervention. This paper aimed to describing respiratory/allergic symptoms in the first 6 years of life. It was quite correctly in line with the currently admitted concept of allergic march. We generated hypotheses on the potentially harmful effect on respiratory and allergic health in children aged 0–6 of some chemical exposures that are directly related to parental behavior (renovation and redecoration) and thus avoidable. If confirmed in observational studies and interventional home-based trials, environmental counseling by primary care givers, because of their close relationship to families and potential knowledge of domestic behavior, could be particularly useful in prevention strategies for respiratory and allergic diseases in childhood.

## METHODS

### Study design, participants, and settings

PELAGIE, described previously,<sup>30,31</sup> is a population-based mother–child cohort that enrolled 3421 women in pregnancy (before 19 weeks' gestation) in Brittany between 2002 and 2006. Data were collected at birth and when the child was 2 and 6 years old. We studied a subgroup of 935 mother–child pairs for whom respiratory and allergic data were available at both follow ups (Fig. 1). Participants provided written informed consent for data collection, and a national INSERM (Institut National de la Santé Et de la Recherche Médicale) ethics committees approved the study procedures.

### Methods of data collection

Data were collected by a self-administered questionnaire filled out mostly by the mother at inclusion and when the child was aged 2 and 6 years, as well as medical reports at delivery (obstetric data collected by midwives and physical examination of the newborn by a pediatrician).

### Variables

The ISAAC questionnaire<sup>32</sup> assesses four respiratory/allergic symptoms: wheezing, eczema, rhinitis, and cough. Wheezing was measured at ages 1, 2 (2 years questionnaire), and 6 (6 years questionnaire). At age 1 and 2, wheezing was defined as at least one episode of whistling in the chest or

an asthma attack. Children not reported to wheeze but with physician-diagnosed asthma were also considered wheezers. Wheezing or whistling in the child's chest and/or a wheezy sound there during or after exercise during the last 12 months defined wheezing at age 6. Eczema was defined at ages 1, 2, and 6 by a positive answer to the question: "Has your child ever had an itchy rash which was coming and going in the last twelve months?" Rhinitis symptoms were assessed for the first 2 and the sixth years of life and defined by positive answers to ISAAC questions: "Has your child ever had a problem with sneezing, or a runny, or a blocked nose when he/she DID NOT have a cold or the flu?" and "In the past 12 months, has this nose problem been accompanied by itchy-watery eyes?" The PELAGIE questionnaire did not include any specific item to assess child's cough at the 2-year follow-up. Cough during the first 2 years of life was defined by a positive answer to the question: "Since birth, has your child ever had bronchitis or bronchiolitis?" At age 6, cough was the occurrence of dry night cough in the past 12 months.

To confirm phenotypes, we assessed parent-reported comorbidities such as otitis media at age 2 or ever awakening due to breathlessness or respiratory discomfort and hay fever or other allergic rhinitis at age 6. They were also described against physician-diagnosed asthma, eczema, and food allergies as at the 6-year follow-up we also asked the parents whether a physician had ever diagnosed those disorders.

We assessed indoor environmental factors known or suspected to be involved with respiratory and allergic symptoms.

At age 2, we assessed: ETS, defined as maternal and/or paternal consumption of at least one cigarette per day at home since birth (yes/no); whether child spends time in rooms cleaned with bleach (yes/no); renovation activities in the dwelling since birth that used glues or either alkyd or acrylic paints (yes/no); usual home heating, defined by individual heating by at least one of the following: gas/fuel oil/wood or chimneys (yes/no).

At age 2, we assessed: pets at home, that is, at least one of the following: cat/dog/rodent/bird/farm animal (yes/no) since birth; daily airing of the child's bedroom (yes/no); and HDM exposure (yes/no), which was evaluated indirectly by the age of the child's mattress considered old if purchased at least 3 years earlier.

## Statistical methods

Phenotype identification relied on the joint temporal evolution of wheezing, cough, eczema, and rhinitis symptoms. We used the Kml3D unsupervised clustering method<sup>33</sup> designed to study the joint longitudinal evolution of several variables over time. This method applies the *k*-means algorithm on longitudinal data (it divides the original sample in clusters of individuals with similar symptom trajectories). Moreover, several outcomes/symptom trajectories are considered simultaneously to compute interindividual distances. Each individual is represented by a matrix with his/her levels of different symptoms (lines) at different time points (columns). The multidimensional Euclidean metrics was used to calculate interindividual distances (Euclidean norm of the distance matrices). As the *k*-means algorithm requires predefining a number of clusters, we ran the *k*-means algorithm 20 times to obtain, respectively 3, 4, and 5 clusters. These numbers of clusters, based on an a priori choice were reasonable regarding our sample size and expected interpretable clinical phenotypes.<sup>4–6</sup> The choice of the optimal clustering was based on the Davies Bouldin cluster validity index.<sup>34</sup> For each clustering (different numbers of clusters), a proximity indicator reflecting the average within-cluster distance/between-cluster distance is computed. The optimal clustering minimizes this proximity indicator. The resulting clusters were then interpreted as clinical phenotypes based on the mean symptom trajectories observed in each cluster and our medical expertise. The complete cluster analysis was performed with the R software package Kml3D.<sup>33</sup>

Associations between phenotypes and covariates were estimated with univariate polytomous logistic regression models. We used single random imputation by the mean or median when missing values for covariates did not exceed 5.0%. The final regression model included the lifestyle factors associated with phenotypes in the univariate analysis and all the environmental factors, regardless of the strength of their association with phenotypes. Multivariate regression models were adjusted for maternal body mass index (BMI  $\geq 30$  kg/m<sup>2</sup>: yes/no) and age at inclusion, parity, maternal smoking habits at inclusion (at least one cigarette per day: yes/no), family history of allergy (mothers' history of allergy and/or maternal grandparent's history of asthma: yes/no), child's gender, vaginal delivery (yes/no), and exclusive breastfeeding for at least 3 months (yes/no). A sensitivity analysis, included preterm children and those with a birth weight <10th centile expected for gestational age. The prevalence of these

two characteristics was low and thus not appropriate for use in the adjusted regression model. Logistic regression analyses were performed with SAS 9.4, and results were expressed as odds ratios and their 95% confidence intervals (95% CI). The level of statistical significance was 0.05. Our data were analyzed in 2017.

## Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## DATA AVAILABILITY

The data from the PELAGIE cohort are not freely available as they are the property of our research institute. Access might be asked by contacting the corresponding author.

## CODE AVAILABILITY

Codes used for the cluster analysis can be accessed by contacting the corresponding author.

## ACKNOWLEDGEMENTS

We are grateful to the families that participated in the study. We also thank Jo Ann Cahn for her careful revision of the paper. The financial support from the National Institute for Public Health Surveillance (InVS), the Ministry of Labor, the French Agency for Food, Environmental and Occupational Health and Safety (ANSES), and the French National Research Agency (ANR) is acknowledged.

## AUTHOR CONTRIBUTIONS

K.A. drafted the paper with the input from F.P.; the study was planned by S.C., C.C., and C.M. K.A., F.P., C.C., N.C., and A.C. participated in editing the text and interpretation of the data. K.A., F.P., N.C., and C.M. take responsibility for the integrity of the data and K.A., F.P., and N.C. for the accuracy of the data analysis. All authors had full access to all the data in the study and edited the final text. All authors read and approved the final version of the paper and agreed to submit for publication.

## ADDITIONAL INFORMATION

**Supplementary Information** accompanies the paper on the *npj Primary Care Respiratory Medicine* website (<https://doi.org/10.1038/s41533-019-0141-y>).

**Competing interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## REFERENCES

- Asher, M. I. et al. Worldwide time trends in the prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and eczema in childhood: ISAAC Phases One and Three repeat multicountry cross-sectional surveys. *Lancet* **368**, 733–743 (2006).
- Sennhauser, F. H., Braun-Fahrlander, C. & Wildhaber, J. H. The burden of asthma in children: a European perspective. *Paediatr. Respir. Rev.* **6**, 2–7 (2005).
- Bousquet, J. et al. Birth cohorts in asthma and allergic diseases: report of a NIAID, NHLBI, MeDALL joint workshop. *J. Allergy Clin. Immunol.* **133**, 1535–1546 (2014).
- Rancière, F., Nikasinovic, L., Bousquet, J. & Momas, I. Onset and persistence of respiratory/allergic symptoms in preschoolers: new insights from the PARIS birth cohort. *Allergy* **68**, 1158–1167 (2013).
- Belgrave, D. C. M. et al. Developmental profiles of eczema, wheeze, and rhinitis: two population-based birth cohort studies. *PLoS Med.* **11**, 1–14 (2014).
- Panico, L., Stuart, B., Bartley, M. & Kelly, Y. Asthma trajectories in early childhood: identifying modifiable factors. *PLoS ONE* **9**, 1–11 (2014).
- Nieto, A. et al. Allergy and asthma prevention 2014. *Pediatr. Allergy Immunol.* **25**, 516–533 (2014).
- Deng, Q., Lu, C., Ou, C., Chen, L. & Yuan, H. Preconceptional, prenatal and postnatal exposure to outdoor and indoor environmental factors on allergic diseases/symptoms in preschool children. *Chemosphere* **152**, 459–467 (2016).
- Heinrich, J. Influence of indoor factors in dwellings on the development of childhood asthma. *Int. J. Hyg. Environ. Health* **214**, 1–25 (2011).

10. Chin, J.-Y. et al. Levels and sources of volatile organic compounds in homes of children with asthma. *Indoor Air* **24**, 403–415 (2014).
11. Mendell, M. J. Indoor residential chemical emissions as risk factors for respiratory and allergic effects in children: a review. *Indoor Air* **17**, 259–277 (2007).
12. A. Farrow, H. T. Time Spent in the Home by Different Family Members. *Environ. Technol.* **18**, 605–613 (1997).
13. Mellis, C. Respiratory noises: how useful are they clinically? *Pediatr. Clin. North Am.* **56**, 1–17 (2009).
14. Garcia-Aymerich, J. et al. Phenotyping asthma, rhinitis and eczema in MeDALL population-based birth cohorts: an allergic comorbidity cluster. *Allergy* **70**, 973–984 (2015).
15. Martinez, F. D. et al. Asthma and wheezing in the first six years of life. *N. Engl. J. Med.* **332**, 133–138 (1995).
16. Chen, Q. et al. Using latent class growth analysis to identify childhood wheeze phenotypes in an urban birth cohort. *Ann. Allergy Asthma Immunol.* **108**, 311–315.e1 (2012).
17. Greene, N., Greenland, S., Olsen, J. & Nohr, E. A. Estimating bias from loss to follow-up in the Danish National Birth Cohort. *Epidemiology* **22**, 815–822 (2011).
18. Pinart, M. et al. Systematic review on the definition of allergic diseases in children: the MeDALL study. *Int. Arch. Allergy Immunol.* **168**, 110–121 (2015).
19. Dietert, R. R. et al. Workshop to identify critical windows of exposure for children's health: immune and respiratory systems work group summary. *Environ. Health Perspect.* **108**, 483 (2000).
20. Kajekar, R. Environmental factors and developmental outcomes in the lung. *Pharmacol. Ther.* **114**, 129–145 (2007).
21. Diez, U. et al. Redecoration of apartments promotes obstructive bronchitis in atopy risk infants—results of the LARS study. *Int. J. Hyg. Environ. Health* **206**, 173–179 (2003).
22. Herbarth, O. et al. Association between indoor renovation activities and eczema in early childhood. *Int. J. Hyg. Environ. Health* **209**, 241–247 (2006).
23. Burke, H. et al. Prenatal and passive smoke exposure and incidence of asthma and wheeze: systematic review and meta-analysis. *Pediatrics* **129**, 735–744 (2012).
24. Henderson, J. et al. Associations of wheezing phenotypes in the first 6 years of life with atopy, lung function and airway responsiveness in mid-childhood. *Thorax* **63**, 974–980 (2008).
25. Casas, L. et al. The use of household cleaning products during pregnancy and lower respiratory tract infections and wheezing during early life. *Int. J. Public Health* **58**, 757–764 (2013).
26. Nickmilder, M., Carbonnelle, S. & Bernard, A. House cleaning with chlorine bleach and the risks of allergic and respiratory diseases in children. *Pediatr. Allergy Immunol.* **18**, 27–35 (2007).
27. Apfelbacher, C., Frew, E., Xiang, A., Apfel, A. & Smith, H. Assessment of pet exposure by self report in epidemiological studies of allergy and asthma: a systematic review. *J. Asthma* **0**, 1–34 (2015).
28. Gaffin, J. M., Kanchnongkittiphon, W. & Phipatanakul, W. Perinatal and early childhood environmental factors influencing allergic asthma immunopathogenesis. *Int. Immunopharmacol.* **22**, 21–30 (2014).
29. Tovey, E. R., Almqvist, C., Li, Q., Crisafulli, D. & Marks, G. B. Nonlinear relationship of mite allergen exposure to mite sensitization and asthma in a birth cohort. *J. Allergy Clin. Immunol.* **122**, 114–118.e5 (2008).
30. Petit, C. et al. Association of environmental insecticide exposure and fetal growth with a bayesian model including multiple exposure sources: the PELAGIE mother-child cohort. *Am. J. Epidemiol.* **175**, 1182–1190 (2012).
31. Chevrier, C. et al. Urinary biomarkers of prenatal atrazine exposure and adverse birth outcomes in the PELAGIE birth cohort. *Environ. Health Perspect.* **119**, 1034–1041 (2011).
32. Asher, M. I. et al. International study of asthma and allergies in childhood (ISAAC): rationale and methods. *Eur. Respir. J.* **8**, 483–491 (1995).
33. Genolini, C. et al. Kml3D: a non-parametric algorithm for clustering joint trajectories. *Comput. Methods Prog. Biomed.* **109**, 104–111 (2013).
34. Davies, D. L. & Bouldin, D. W. A cluster separation measure. *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-1**, 224–227 (1979).



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019