Received: 18 February 2022 | Revised: 27 July 2022 | Accepted: 13 August 2022

DOI: 10.1111/all.15511

Allergy EUROPEAN JOURNAL OF ALLERGY AND CLINICAL IMMUNOLOGY EAACI WILEY

ORIGINAL ARTICLE

Allergic disease trajectories up to adolescence: Characteristics, early-life, and genetic determinants

Anna Kilanowski[1,2,3] | Elisabeth Thiering[1,3] | Gang Wang[4,5] | Ashish Kumar[5] | Sara Kress[6,7] | Claudia Flexeder[1,8,9] | Carl-Peter Bauer[10] | Dietrich Berdel[11] | Andrea von Berg[11] | Anna Bergström[12,13] | Monika Gappa[14] | Joachim Heinrich[9,15] | Gunda Herberth[16] | Sibylle Koletzko[17,18] | Inger Kull[5,19,20] | Erik Melén[5] | Tamara Schikowski[6] | Annette Peters[1,21] | Marie Standl[1,8]

[1]Institute of Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany

[2]Institute for Medical Information Processing, Biometry, and Epidemiology, Pettenkofer School of Public Health, LMU Munich, Munich, Germany

[3]Division of Metabolic and Nutritional Medicine, Dr. von Hauner Children's Hospital, University of Munich Medical Center, Munich, Germany

[4]Department of Integrated Traditional Chinese and Western Medicine, West China Hospital, Sichuan University, Chengdu, Sichuan, China

[5]Department of Clinical Science and Education, Södersjukhuset, Karolinska Institutet, Stockholm, Sweden

[6]Leibniz Research Institute for Environmental Medicine, Düsseldorf, Germany

[7]Medical Research School Düsseldorf, Heinrich Heine University, Düsseldorf, Germany

[8]German Center for Lung Research (DZL), Munich, Germany

[9]Institute and Clinic for Occupational, Social and Environmental Medicine, University Hospital, LMU Munich, Munich, Germany

[10]Department of Pediatrics, Technical University of Munich, Munich, Germany

[11]Research Institute, Department of Pediatrics, Marien-Hospital Wesel, Wesel, Germany

[12]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

[13]Centre for Occupational and Environmental Medicine, Stockholm, Sweden

[14]Evangelisches Krankenhaus Düsseldorf, Children's Hospital, Düsseldorf, Germany

[15]Allergy and Lung Health Unit, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Australia

[16]Department of Environmental Immunology, Helmholtz Centre for Environmental Research-UFZ, Leipzig, Germany

[17]Department of Pediatrics, Dr. von Hauner Children's Hospital, University Hospital, LMU Munich, Munich, Germany

[18]Department of Pediatrics, Gastroenterology and Nutrition, School of Medicine Collegium Medicum, University of Warmia and Mazury, Olsztyn, Poland

[19]Södersjukhuset, Karolinska Institutet, Stockholm, Sweden

[20]Sachs Children's and Youth Hospital, Stockholm, Sweden

[21]Chair of Epidemiology, Ludwig-Maximilians University, Munich, Germany

Abbreviations: BAMSE, Barn (=Children) Allergy Milieu Stockholm Epidemiology; FEV_1, Forced Expiratory Volume in one second; FLG, Filaggrin; FVC, Forced Vital Capacity; GINIplus, German Infant Study on the Influence of Nutrition Intervention plus Air pollution and Genetics on Allergy Development; GLI, Global Lung Initiative; GWAS, Genome-wide association studies; HRC, Haplotype Reference Consortium; IgE, Immunoglobulin E; ISAAC, International Study of Asthma and Allergies in Childhood; LISA, Influence of Life-style factors on Development of the Immune System and Allergies in East and West Germany study; LOCF, Last observation carried forward; PRS, Polygenic Risk Score; RRR, Relative Risk Ratio; SNP, Single Nucleotide Polymorphism.

Correspondence
Marie Standl, Helmholtz Zentrum München - Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH), Institute of Epidemiology, Ingolstädter Landstraße 1, D-85764 Neuherberg, Germany.
Email: marie.standl@helmholtz-muenchen.de

Funding information
Bundesministerium für Bildung und Forschung; Commission of the European Communities, the 7th Framework Program; Federal Ministry for Environment (IUF Düsseldorf), Grant/Award Number: FKZ 20462296; Fondation Nestlé; H2020 European Research Council; Helmholtz Zentrum München; Helmholtz-Zentrum für Umweltforschung; Hjärt-Lungfonden; Karolinska Institutet SFO Epidemiology and Region Stockholm; Leibniz Research-Institute for Environmental Medicine at the University of Düsseldorf; Ludwig-Maximilians-Universität München; Marien-Hospital Wesel; Mead Johnson Nutrition; Pediatric Practice, Bad Honnef; Swedish Research Council; Technische Universität München

Abstract

Background: Allergic diseases often develop jointly during early childhood but differ in timing of onset, remission, and progression. Their disease course over time is often difficult to predict and determinants are not well understood.

Objectives: We aimed to identify trajectories of allergic diseases up to adolescence and to investigate their association with early-life and genetic determinants and clinical characteristics.

Methods: Longitudinal k-means clustering was used to derive trajectories of allergic diseases (asthma, atopic dermatitis, and rhinitis) in two German birth cohorts (GINIplus/LISA). Associations with early-life determinants, polygenic risk scores, food and aeroallergen sensitization, and lung function were estimated by multinomial models. The results were replicated in the independent Swedish BAMSE cohort.

Results: Seven allergic disease trajectories were identified: "Intermittently allergic," "rhinitis," "early-resolving dermatitis," "mid-persisting dermatitis," "multimorbid," "persisting dermatitis plus rhinitis," and "early-transient asthma." Family history of allergies was more prevalent in all allergic disease trajectories compared the non-allergic controls with stronger effect sizes for clusters comprising more than one allergic disease (e.g., RRR = 5.0, 95% CI = [3.1–8.0] in the multimorbid versus 1.8 [1.4–2.4] in the mild intermittently allergic cluster). Specific polygenic risk scores for single allergic diseases were significantly associated with their relevant trajectories. The derived trajectories and their association with genetic effects and clinical characteristics showed similar results in BAMSE.

Conclusion: Seven robust allergic clusters were identified and showed associations with early life and genetic factors as well as clinical characteristics.

KEYWORDS
allergic diseases, epidemiology, longitudinal clustering, polygenic risk score, trajectories

1 | INTRODUCTION

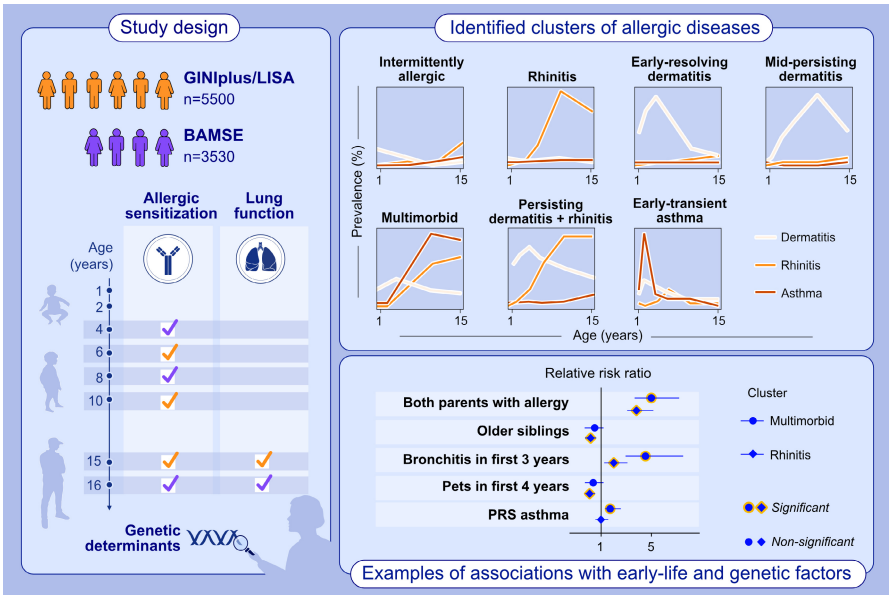
Allergic diseases, such as asthma, atopic dermatitis, and rhinitis, impose high impact on quality of life.[1] While atopic dermatitis often develops in early infancy with high remission rate up to adolescence, asthma, and rhinitis usually occur later in childhood.[2] In general, allergic diseases often develop jointly, in temporal succession and differing severities, highlighting the role of heritability, joint mechanisms, and genetic susceptibility.[2–5]

Prediction of onset, progression or remission of allergic diseases are often difficult to obtain. It is therefore of high importance, to characterize the patterns of joint disease development and their relating factors.

Based on this need, a pioneer investigation of clinical phenotypes taking age and time into account, reported classes of childhood wheezing in 1995.[6] With increasing sample sizes and long-lasting birth cohorts, researchers have now turned to data-driven approaches to identify more detailed developmental patterns of allergic diseases,[7] for example, building on the asthma trajectories in the Millennium Cohort Study up to seven years[8] and in BAMSE up to young adulthood.[9] The combination of the allergic disease trajectories, asthma, rhinitis, and dermatitis, together, was previously described until school-age[10–13] and from school age to adulthood.[14]

Genetic susceptibility to allergic diseases was investigated in several genome-wide association studies (GWAS).[5,15–18] Risk variants, single nucleotide polymorphisms (SNPs), identified in these analyses can be used to calculate individual predisposition. These aggregated polygenic risk scores (PRS) were applied in the context of various diseases,[19] including allergic ones[20] and their derived latent classes.[21] In addition, specific gene expression signatures have been identified for multimorbid asthma, rhinitis, and dermatitis, as compared with single-disease allergic phenotypes,[22] highlighting the current focus on genetic origins of allergic diseases.

However, given the rise of allergic disease prevalence in the recent decades,[23] the increase cannot be explained by genetic factors alone. Several environmental factors have been investigated in the



GRAPHICAL ABSTRACT

We identified seven allergic disease trajectories up to adolescence, which are corresponding to clinical observation in the German GINIplus and LISA studies and replicated the results in the Swedish BAMSE cohort. The clusters can be characterized using polygenic risk scores and early-life determinants, which support the hygiene hypothesis. The clusters also pose clinical implications for allergic sensitization, increasing with number of present allergic diseases, and lung function.

context of the hygiene hypothesis[24,25] and are commonly analyzed as determinants of allergic diseases, for example, in Hu et al.,[26] stating poor distinction of dermatitis phenotypes for example by breastfeeding or pet ownership alone.

Summarizing the need for a more specific distinction of phenotypes and their temporal patterns is apparent. Therefore, the aim of this study is to identify joint trajectories of allergic diseases using independent birth cohorts. Following objectives are to assess their association with (1) early-life determinants, (2) GWAS-derived PRS, and (3) clinical characteristics, such as allergic sensitization and spirometry.

2 | METHODS

2.1 | Study population

Data were obtained from two prospective, population-based German birth cohorts with a focus on the development of allergic diseases, the German Infant Study on the Influence of Nutrition Intervention plus Air pollution and Genetics on Allergy Development (GINIplus) and the Influence of Life-style factors on Development of the Immune System and Allergies in East and West Germany (LISA) study.For GINIplus, 5991 full-term, healthy newborns were recruited in Munich and Wesel between 1995 and 1998. The LISA study included 3094 healthy, full-term newborns from the study centers Munich, Wesel, Bad Honnef and Leipzig, born between 1997 and 1999. Ethics approval by the respective ethics committees, and written consent from all participating families was obtained. Both studies had comparable follow-up time points, which were harmonized to time points 1, 2, 4, 6, 10, and 15 years. More details on both studies can be found elsewhere.[27]

2.2 | Allergic diseases

Allergic diseases were defined as parent-reported doctor diagnosis of asthma, atopic dermatitis, hay fever or allergic rhinitis for each year of life, which were aggregated to cover the time periods between follow-ups (Table S1).

2.3 | Early-life determinants

We investigated early-life determinants, previously reported to be associated with allergic diseases. These include sex, parental education level, family history of allergic diseases, Caesarean section, exclusive breastfeeding, presence of older siblings, maternal smoking during pregnancy, second-hand tobacco smoke exposure, pet exposure, urbanicity at birth residency, and early bronchitis infections (Table S2).

2.4 | Aeroallergen and food sensitization

Specific immunoglobulin E (IgE) was measured at 6, 10, and 15 years using the CAP-RAST FEIA system (Pharmacia Diagnostics, Freiburg, Germany) according to the manufacturer's instructions. Sensitization

against common aero- (SX1 mix) and food allergens (FX5 mix) was defined with a cut-off for allergen specific IgE of >0.35kU/l in the screening test (Table S1).

2.5 | Spirometry

Lung function was measured using spirometry at the 15-year follow-up investigation according to guidelines from the American Thoracic Society and the European Respiratory Society[28] and has been described previously.[29] The present study investigates forced expiratory volume in one second (FEV_1), forced vital capacity (FVC), and the Tiffeneau-Index (FEV_1/FVC), all standardized according to the Global Lung Initiatives (GLI) formula to control for non-linear effects of age, height, and sex.[30]

2.6 | Polygenic risk scores

Combining the joint effect of genetic variation, we calculated four different PRS based on current GWAS summary statistics for any allergic disease,[5,15] asthma,[17] dermatitis,[16] and rhinitis.[18] SNPs reported as significant in the GWAS meta-analyses were extracted from existing genome-wide data (Methods S1). After quality control, four PRS were calculated for each participant, weighting the allele dosage with the effect size reported in the GWAS. Finally included SNPs and annotated genes[31] can be found in Table S3a–d and a more detailed description of the calculation in Methods S2. Additionally, the two most common Filaggrin (FLG) loss-of-function mutations (R501X, 2282del4) were genotyped.[32] FLG mutation carrier status is defined as at least one mutation compared with no mutation.

2.7 | Replication study

The prospective, population-based BAMSE (Swedish abbreviation for Children, Allergy, Milieu, Stockholm, Epidemiology) project recruited 4093 newborns in the area of Stockholm between 1994 and 1996. Participants were followed-up at 1, 2, 4, 8, 12, and 16 years, and ethics approval was given by the Regional Ethics Board (EPN). Allergic disease diagnoses were defined based on parent-reported symptoms[33] according to the questionnaire used in the International Study of Asthma and Allergies in Childhood (ISAAC).[34] Further details on the study, including genetic data can be found elsewhere[35] and in the supplementary information (Methods S1 and S3 and Table S1).

The BAMSE cohort was used for replication of clusters, pre-defined early-life determinants, genetic variation, and clinical characteristics.

2.8 | Statistical analysis

All participants with available data on at least three out of six time-points per allergic disease were included in the analysis. Remaining missing observations were filled in carrying the last observation forward (LOCF).

To identify joint trajectories of allergic disease development, longitudinal k-means clustering using the kml3d package[36] in R[37] was employed, which enables to cluster longitudinal data with multiple trajectories. To ensure the focus on allergic development, a case-only approach including only those who reported any allergic disease for at least one time point was used. The non-allergic cluster, comprising all remaining participants, was added afterward. Differing numbers of clusters from two to nine clusters were tested and the optimal number of clusters was chosen based on the quality criterion of Davies-Bouldin[38] in GINIplus/LISA. The cluster number in BAMSE was fixed according to this for comparability and independent clustering results are provided in the supplement. We specified the algorithm to run 40 times with random starting conditions and used standard Euclidean distance. Sensitivity analyses for missingness were conducted using only complete observations as well as performing analyses based on bagging tree imputation of diagnoses in GINIplus/LISA to test the assumptions of LOCF (further described in methods S4) and latent class modeling.[9] Here, the optimal partitioning was based on the Bayesian information criterion and log-likelihood. As it is not possible to consider both dimensions, time and multiple allergic diseases, in this approach, the longitudinal clustering was the preferred approach. We also added symptom-based clustering for GINIplus/LISA to check for differences between the disease definitions within the used studies. K-means clusters are always presented for their optimal solution and pre-set for seven allergic disease clusters to ease comparison.

The total population and the allergic and non-allergic sub-populations were descriptively assessed regarding the distribution of early-life determinants and differences between the sub-populations were analyzed with a Chi-squared test.

Early-life determinants were assessed using post hoc Tukey-tests to identify differences between clusters and compact letter design was used to indicate similarities between the trajectories as well as a multiple-testing adjusted p-value[39] from the respective analysis of variance (ANOVA) (Table S4). Significant factors (FDR-adjusted ANOVA p-value <0.05) were selected for a joint multinomial regression model. Influences of genetic factors on clusters were assessed using a joint multinomial model followed by a fixed-effect meta-analysis between GINIplus/LISA and BAMSE using the meta package.[40] The association with allergic sensitization and spirometric z-scores was also assessed using multinomial regression models for each time point and measurement. All models were additionally adjusted for sex and further for study and study center in GINIplus/LISA. Relative Risk Ratios (RRR) and corresponding 95% confidence intervals are presented. Further sensitivity analyses were performed testing the robustness of genetic association for additional adjustment for (a) family history of allergic diseases, (b) early-life determinants, and (c) interaction effects with sex in GINIplus/LISA.

All analyses were conducted using R[37] version 4.0.3 (GINIplus/LISA) and version 4.0.4 (BAMSE) and code is available on request.

3 | RESULTS

Starting with descriptive assessments, the main analysis comprised 5550 participants of the GINIplus/LISA studies (Figure S1). Population characteristics, cumulative allergic disease prevalences, and investigated determinants are presented in Table 1 for GINIplus/LISA and Table S5 for pre-defined BAMSE determinants. Allergic disease development in GINIplus/LISA shows early dermatitis as most common allergic disease in infancy with decreasing prevalence after four years and increasing rhinitis and asthma prevalence during childhood and adolescence (Figure S2).

3.1 | Joint trajectories of allergic diseases

Within GINIplus/LISA, seven allergic disease cluster plus one non-allergic cluster were identified (Figure 1A) with a comparable cluster allocation observed in BAMSE (Figure 1B): Intermittently allergic (GINIplus/LISA:17.5%, BAMSE:15.5%), rhinitis (7.5%, 5.3%), early-resolving dermatitis (6.3%, 11.5%), mid-persisting dermatitis (4.1%, 12.0%), multimorbid (4.0%, 8.2%), persisting dermatitis plus rhinitis (2.2%, 6.6%), early-transient asthma (0.5%, 10.2%), and non-allergic (57.8%, 30.6%). Names for the clusters were chosen descriptively and do not impose strict clinical definitions. The independent clustering in BAMSE yielded six allergic disease clusters (Figure S3) and only using complete observations in GINIplus/LISA led to nine optimal allergic disease clusters (Figure S4a; see Figure S4b for pre-defined 7 + 1 cluster solution in complete observations only).

The sensitivity analysis using latent class modeling, yielded eight optimal classes, including a non-allergic class in GINIplus/LISA and seven in BAMSE (Figure S5). The identified classes are highly similar and differ only in a split of the multimorbid cluster. Here, we observe, instead of only one cluster featuring all three allergic diseases now one with low level dermatitis and one with high level dermatitis, supporting the original differentiation. GINIplus/LISA looses the early-transient wheeze cluster and BAMSE the intermittent and early-resolving dermatitis ones.

3.2 | Early-life determinants

To assess which factors potentially influence disease development, Table 2 presents results of the multinomial model containing early-life determinants, which were identified using Tukey-Tests (Table S4) in GINIplus/LISA. Male sex is associated with the clusters comprising a high prevalence of rhinitis and inversely with the persisting dermatitis cluster. Parental history of allergies increases the risk for all identified clusters compared with the non-allergic cluster, with intermittently allergic showing the weakest association, significantly differing from the clusters comprising rhinitis and early-transient asthmatics. Pet exposure and presence of older siblings are inversely associated with the rhinitis cluster in GINIplus/LISA. In BAMSE associations with family history show similar patterns, although not all effects reach significance and are mostly visible for only one parent with allergies. Nonetheless, an even stronger inverse relationship can be seen for early bronchitis infections.

3.3 | Genetic factors

Further, we investigated genetic influences on allergic disease trajectories to identify the effects of genetic predisposition on disease development. For this, we used a fixed-effect meta-analysis of the calculated PRS and FLG mutation carrier status on the derived clusters (Table 3; Table S6 for individual studies). In GINIplus/LISA all specific PRS are significantly associated with their respective trajectories and therefore more specific than the PRS for any allergic disease, which is only associated with the rhinitis cluster. Further, the PRS for atopic dermatitis shows a significant association with early-onset or transient but not mid-onset or persisting atopic dermatitis.

Within BAMSE, the PRS for any allergic disease shows significant associations with the intermittently allergic, early dermatitis, multimorbid, and dermatitis plus rhinitis cluster, whereas the specific PRS are less clear associated.

Comparing the independent samples, we observe similar emerging patterns, even though clusters differ in their specifics. Despite differences in significance, effect directions are mostly similar between GINIplus/LISA and BAMSE and the meta-analysis shows significant associations between all allergic clusters and the PRS for any allergic disease, except the mid-persisting dermatitis and early-transient asthma cluster. The multimorbid cluster is specified by additional significant associations with the PRS for asthma and rhinitis, differentiating it from the rhinitis cluster. The FLG-mutation is significantly associated with the multimorbid and persisting dermatitis plus rhinitis cluster.

Our sensitivity analyses showed that genetic associations seen between PRS and clusters are rather robust to adjustment for family history of allergic diseases and early-life determinants (Tables S7 and S8). Investigating genetic effects and their modification by sex yielded no significant interaction (Table S9).

3.4 | Aeroallergen and food sensitization

For the clinical characterization of trajectories, we investigated associations of aero- and food allergen sensitization measured at several time points with the allergic disease trajectories, as shown in Table 4. All clusters, except the early-transient asthma and the mid-persisting dermatitis clusters, show significant associations compared to the non-allergic control in all cohorts. However, the magnitude of effects is distinctly higher in rhinitis containing trajectories. Especially the clusters with more than one allergy are robustly associated and are significantly higher compared with all

TABLE 1 Population characteristics and determinants in GINIplus/LISA

	Total ($N = 5550$)	Allergic disease cases ($N = 2342$)	No allergic diseases ($N = 3208$)	p-value (Chi-squared)
Sex				
Male	50.8% (2818)[0]	52.7% (1235)[0]	49.3% (1583)[0]	0.014
Study				
GINI intervention	24.9% (1381)[0]	31.4% (735)[0]	20.1% (646)[0]	<0.001
GINI control	35.5% (1971)[0]	32.4% (759)[0]	37.8% (1212)[0]	<0.001
LISA	39.6% (2198)[0]	36.2% (848)[0]	42.1% (1350)[0]	<0.001
Study center				
Munich	51.5% (2856)[0]	52.9% (1238)[0]	50.4% (1618)[0]	0.079
Leipzig	10.2% (564)[0]	10.5% (247)[0]	9.9% (317)[0]	0.444
Bad Honnef	4.4% (244)[0]	3.8% (90)[0]	4.8% (154)[0]	0.099
Wesel	34.0% (1886)[0]	32.7% (767)[0]	34.9% (1119)[0]	0.104
Cumulative prevalence of allergic disease up to 15 years				
Any allergic disease	42.2% (2342)[0]	100.0% (2342)[0]	0.0% (0)[0]	
Asthma	9.2% (513)[0]	21.9% (513)[0]	0.0% (0)[0]	
Dermatitis	27.9% (1549)[0]	66.1% (1549)[0]	0.0% (0)[0]	
Rhinitis	21.3% (1183)[0]	50.5% (1183)[0]	0.0% (0)[0]	
Family history of allergic diseases				
No parent	43.0% (2272)[263]	33.7% (758)[94]	49.8% (1514)[169]	<0.001
One parent	40.2% (2126)[263]	43.1% (970)[94]	38.0% (1156)[169]	<0.001
Both parents	16.8% (889)[263]	23.1% (520)[94]	12.1% (369)[169]	<0.001
Parental education				
Low	6.6% (366)[23]	6.6% (154)[10]	6.6% (212)[13]	1.000
Medium	27.8% (1536)[23]	27.3% (637)[10]	28.1% (899)[13]	0.520
High	65.6% (3625)[23]	66.1% (1541)[10]	65.2% (2084)[13]	0.528
Caesarean section				
Yes	19.8% (1037)[325]	20.8% (466)[99]	19.1% (571)[226]	0.154
Breastfeeding during first 4 months				
Yes	62.6% (3417)[92]	61.5% (1417) [39]	63.4% (2000)[53]	0.169
Presence of older siblings				
Yes	46.6% (2582)[11]	43.5% (1015)[7]	48.9% (1567)[4]	<0.001
Maternal smoking during 2nd or 3rd trimester of pregnancy				
Yes	9.6% (521)[141]	8.8% (202)[46]	10.2% (319)[95]	0.082
Environmental tobacco smoke exposure in first 4 years				
Yes	37.4% (2012)[172]	36.3% (825)[72]	38.2% (1187)[100]	0.176
Bronchitis infection in first 3 years				
Yes	39.3% (2111)[180]	44.5% (1007)[77]	35.6% (1104)[103]	<0.001
Cat or dog in first 4 years				
Yes	27.1% (1253)[924]	24.3% (480)[368]	29.1% (773)[556]	<0.001
Urbanicity at birth				
City	47.0% (2472)[285]	47.2% (1054)[107]	46.8% (1418)[178]	0.817
Town or suburb	40.6% (2139)[285]	41.4% (925)[107]	40.1% (1214)[178]	0.349
Rural area	12.4% (654)[285]	11.5% (256)[107]	13.1% (398)[178]	0.074

Note: Presented are prevalence, total case numbers and number of missing values for non-allergic participants, allergic participants, and the total sample.

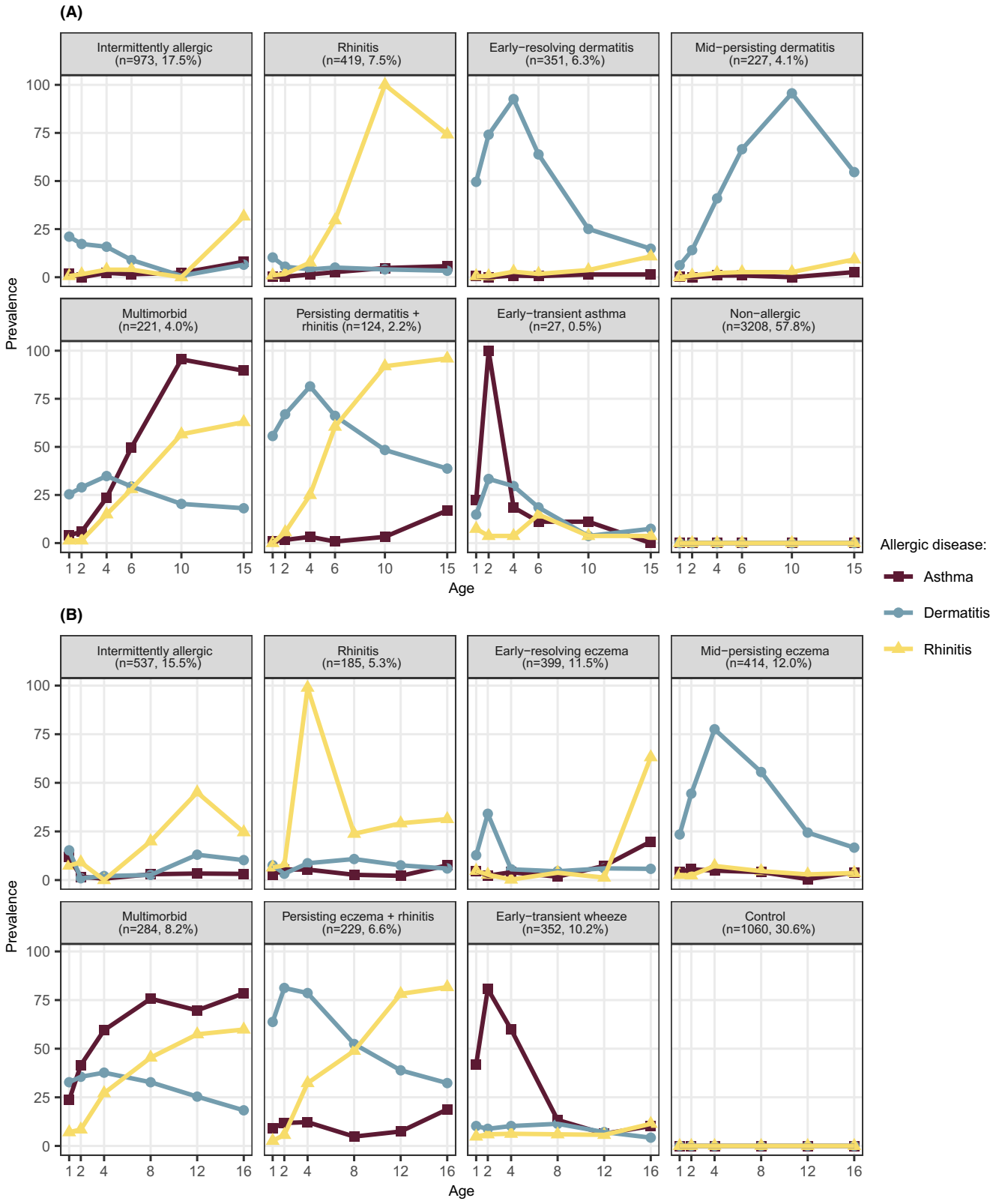


FIGURE 1 Clusters of allergic diseases in (A) GINIplus/LISA and (B) BAMSE. The allergic disease clusters were derived using longitudinal k-means for allergic patients, the non-allergic cluster was added afterward. The best number of clusters was defined in GINIplus/LISA following the quality criterion of Davies-Bouldin.

TABLE 2 Results from multinomial regression of clusters on early-life determinants including all early-life factors, which showed an association in the univariate analysis, at once

	Intermittently allergic	Rhinitis	Early-resolving dermatitis	Mid-persisting dermatitis	Multimorbid	Persisting dermatitis + rhinitis	Early-transient asthma
GINIplus & LISA							
Male sex	1.1 [0.9–1.3]	**1.3* [1–1.6]**	1.0 [0.8–1.3]	**0.6* [0.4–0.8]**	**1.8* [1.3–2.5]**	**1.8* [1.2–2.8]**	2.0 [0.8–4.7]
One parent with allergies	**1.3* [1–1.6]**	**2.3* [1.7–3.1]**	**1.9* [1.3–2.6]**	1.1 [0.7–1.7]	**1.7* [1.1–2.6]**	**3.5* [1.9–6.7]**	**3.7* [1.1–12.1]**
Both parents with allergies	**1.8* [1.4–2.4]**	**3.6* [2.5–5.1]**	**3.4* [2.3–5.1]**	**2.8* [1.7–4.5]**	**5.0* [3.1–8]**	**7.9* [3.9–15.9]**	**15.9* [4.5–56.3]**
Older siblings	**0.8* [0.7–0.9]**	**0.7* [0.5–0.9]**	1.0 [0.8–1.3]	1.1 [0.8–1.5]	0.8 [0.6–1.1]	0.8 [0.5–1.3]	2.7 [1–7.6]
Bronchitis infection in first 3 years	**1.6* [1.3–1.9]**	**1.6* [1.2–2]**	1.2 [0.9–1.6]	1.1 [0.8–1.6]	**4.8* [3.4–6.8]**	**2.0* [1.3–3.1]**	**6.0* [2.4–15.0]**
Cat or dog in first 4 years	1.0 [0.8–1.2]	**0.6* [0.5–0.8]**	0.8 [0.6–1.0]	1.2 [0.8–1.7]	0.7 [0.5–1.1]	0.6 [0.4–1.1]	1.8 [0.7–4.2]
BAMSE							
Male sex	1.2 [1.0–1.5]	1.2 [0.9–1.7]	1.0 [0.8–1.2]	**0.8* [0.6–0.9]**	**1.4* [1.1–1.9]**	1.3 [1.0–1.7]	**1.6* [1.2–2.0]**
One parent with allergies	**1.6* [1.2–2.2]**	1.2 [0.7–1.9]	**1.5* [1.1–2.0]**	**1.7* [1.2–2.3]**	**4.6* [3.4–6.2]**	**2.7* [1.9–3.8]**	**2.7* [2.0–3.6]**
Both parents with allergies	1.2 [0.3–4.5]	2.3 [0.5–11.1]	2.0 [0.6–7.0]	1.0 [0.2–4.8]	**9.8* [3.7–26.5]**	2.1 [0.4–10.0]	**5.3* [1.9–14.9]**
Older siblings	0.9 [0.7–1.1]	0.7 [0.5–1.0]	0.8 [0.7–1.1]	0.9 [0.7–1.1]	1.1 [0.8–1.4]	1.0 [0.7–1.3]	1.3 [1.0–1.6]
Bronchitis infection in first 3 years	**1.5* [1.0–2.3]**	1.7 [0.9–3.1]	1.4 [0.9–2.2]	1.5 [0.9–2.3]	**2.5* [1.6–3.9]**	1.3 [0.7–2.3]	**3.9* [2.7–5.7]**
Cat or dog in first 4 years	**0.7* [0.5–0.8]**	**0.6* [0.4–0.9]**	0.7* [0.6–0.9]	0.8 [0.7–1.1]	**0.5* [0.4–0.8]**	**0.6* [0.4–0.9]**	0.9 [0.7–1.2]

Note: Models in GINIplus/LISA are additionally adjusted for study and study center. Values show RRR and corresponding 95% confidence intervals in comparison to the control cluster. Significance is indicated by bold font and asterisks.

TABLE 3 Results from fixed-effect meta-analysis on multinomial regression of clusters on PRS and FLG mutation status, controlled for sex, study and study center in GINIplus/LISA and sex only in BAMSE

Meta-analysis (fixed effect)	Intermittently allergic	Rhinitis	Early-resolving dermatitis	Mid-persisting dermatitis	Multimorbid	Persisting dermatitis + rhinitis	Early-transient asthma
PRS any allergic disease	**1.1* [1.0–1.3]**	**1.3* [1.1–1.6]**	**1.2* [1.1–1.4]**	1.1 [0.9–1.3]	**1.3* [1.1–1.6]**	**1.6* [1.3–1.9]**	1.1 [0.9–1.3]
PRS asthma	1.1 [0.9–1.2]	1.0 [0.9–1.2]	**1.2* [1.0–1.3]**	1.0 [0.9–1.2]	**1.2* [1.0–1.4]**	1.0 [0.8–1.2]	1.1 [0.9–1.3]
PRS dermatitis	1.1 [1.0–1.2]	1.1 [0.9–1.2]	1.1 [0.9–1.2]	1.1 [1.0–1.3]	1.0 [0.9–1.2]	1.1 [1.0–1.4]	1.0 [0.9–1.2]
PRS rhinitis	1.1 [0.9–1.2]	1.1 [1.0–1.3]	1.0 [0.9–1.2]	1.0 [0.9–1.2]	**1.4* [1.2–1.6]**	1.1 [0.9–1.4]	**1.2* [1.0–1.5]**
Any FLG-mutation	0.9 [0.6–1.3]	0.8 [0.5–1.6]	0.9 [0.5–1.5]	1.6 [1.0–2.5]	**1.7* [1.0–2.8]**	**2.0* [1.2–3.4]**	1.4 [0.8–2.7]

Note: Table shows RRR and corresponding 95% confidence intervals. Significance is indicated by bold script and asterisks. A RRR higher than 1 indicates a positive association between increasing PRS and likelihood to be allocated to this trajectory. Tables for single cohorts are presented in Table S6.

other clusters in BAMSE. Additionally, we see significant association of sensitization already before the peak of rhinitis prevalence. In food sensitization, this might enable differentiation between the early-resolving and mid-persisting dermatitis cases.

3.5 | Spirometry

To assess the impact of trajectories on lung function in adolescence and thus long-term impairment of spirometric function, Table 5 displays the association of spirometry measures z-scores at age 15 years with trajectory allocation. Significant results are observed for FEV_1 and the Tiffeneau-Index in the multimorbid (all cohorts) and early-transient asthma clusters (only BAMSE). Thus, reduced spirometric z-scores can be seen in the clusters comprising a high percentage of asthmatic participants. Borderline significant associations are seen for the intermittently allergic and early-resolving dermatitis cluster.

4 | DISCUSSION

In the present study, we identified seven allergic disease trajectories, using an unsupervised, longitudinal clustering approach in population-based birth cohorts, and one non-allergic control group. The role of early-life determinants was investigated and significant differences in relation to allergic sensitization and spirometric indices between the trajectories were observed. Furthermore, the specific GWAS-derived PRS showed significant associations with their disease-associated trajectories. Successful replication of trajectories, their association with genetic factors, lung function, and sensitization was achieved in the independent Swedish BAMSE cohort, and a comparable classification was obtained when using latent class modeling, further strengthening the robustness of our findings.

Comparing our results with the available evidence, our clusters could confirm results from two of the papers previously characterizing allergic disease trajectories identifying five clusters of allergic diseases up to four years[10] and six up to nine years in the PARIS birth cohort[13] as well as eight classes up to eleven years in UK cohorts.[12] We extend the covered time period up to adolescence, showing the continued developments indicated by the classification found in UK cohorts[12] and being able to distinguish two different dermatitis trajectories, not reported in the previous clustering approaches. Furthermore, the PRS based on one GWAS on any allergic disease[5] was significantly associated with all identified classes in the UK cohorts.[21] Our application of PRS for all single disease entities allows further distinction of sub-phenotypes.

Another publications tracking allergic disease development from nine to 34 years of age identified six classes of allergic diseases ("No symptoms," "Rhinoconjunctivits only," "Late-onset wheeze," "Rhinoconjunctivitis+Wheeze," and "Eczema+Rhinoconjunctivitis+Wheeze") using latent classes.[14] Identified clusters are comparable. The two identified multimorbid clusters resemble our multimorbid and persisting dermatitis plus

TABLE 4 Results from single multinomial regression of clusters on aeroallergen sensitization (SX1 mix) and food allergen sensitization (FX5 mix) at different time points, controlled for sex, study, and study center in GINIplus/LISA and BAMSE

	Intermittently allergic	Rhinitis	Early-resolving dermatitis	Mid-persisting dermatitis	Multimorbid	Persisting dermatitis + rhinitis	Early-transient asthma
Aeroallergen sensitization							
GINIplus & LISA							
Aeroallergen sensitization at 6 years	**2.2* [1.8–2.8]**	**23.0* [16.6–31.8]**	**4.4* [3.2–6.2]**	1.6 [1–2.5]	**18.9* [12.7–28.1]**	**55.3* [27.2–112.4]**	0.7 [0.1–5.1]
Aeroallergen sensitization at 10 years	**3.0* [2.4–3.7]**	**23.0* [15.7–33.9]**	**2.4* [1.7–3.4]**	**2.1* [1.4–3.1]**	**11.9* [7.8–18.1]**	**40.5* [17.4–94.4]**	2.1 [0.5–9.2]
Aeroallergen sensitization at 15 years	**3.1* [2.5–3.8]**	**13.2* [9–19.2]**	**2.8* [2–3.8]**	1.4 [0.9–2.1]	**12.2* [7.5–19.7]**	**22.7* [9.7–52.9]**	1.5 [0.5–4.8]
BAMSE							
Aeroallergen sensitization at 4 years	**4.1* [2.5–6.7]**	**4.3* [2.3–8.0]**	**3.5* [2.0–5.9]**	**2.6* [1.5–4.6]**	**28.7* [17.9–46.1]**	**27.1* [16.5–44.4]**	**2.2* [1.2–4.0]**
Aeroallergen sensitization at 8 years	**5.8* [4.0–8.3]**	**4.2* [2.5–6.9]**	**4.3* [2.9–6.4]**	**2.1* [1.3–3.3]**	**27.0* [18.0–40.6]**	**32.1* [20.8–49.7]**	**1.9* [1.2–3.1]**
Aeroallergen sensitization at 16 years	**3.1* [2.4–4.1]**	**2.2* [1.5–3.2]**	**3.6* [2.7–4.7]**	**1.8* [1.3–2.4]**	**13.5* [9.2–19.9]**	**14.1* [9.3–21.4]**	**1.7* [1.2–2.3]**
Food allergen sensitization							
GINIplus & LISA							
Food allergen sensitization at 6 years	**2.1* [1.5–2.8]**	**4.0* [2.8–5.6]**	**2.5* [1.6–3.9]**	1.0 [0.5–2.0]	**4.9* [3.2–7.4]**	**6.2* [3.8–10.2]**	1.6 [0.2–12.8]
Food allergen sensitization at 10 years	**1.8* [1.4–2.3]**	**5.0* [3.7–6.7]**	**1.8* [1.2–2.7]**	1.3 [0.8–2.2]	**4.2* [2.9–6.2]**	**5.9* [3.7–9.5]**	2.5 [0.5–12.8]
Food allergen sensitization at 15 years	**2.0* [1.5–2.8]**	**4.0* [2.7–5.8]**	**2.4* [1.5–3.8]**	0.9 [0.4–2.0]	**5.1* [3.2–7.9]**	**6.2* [3.6–10.8]**	2.8 [0.6–13.2]
BAMSE							
Food allergen sensitization at 4 years	1.2 [0.8–1.8]	1.2 [0.7–2.1]	1.3 [0.9–2.0]	1.2 [0.8–1.8]	**5.6* [3.9–8.0]**	**3.9* [2.6–5.8]**	0.9 [0.5–1.4]
Food allergen sensitization at 8 years	**1.8* [1.3–2.6]**	**1.9* [1.2–3.1]**	**1.6* [1.1–2.4]**	1.1 [0.8–1.7]	**6.4* [4.5–9.2]**	**6.0* [4.1–8.8]**	1.3 [0.9–2.0]
Food allergen sensitization at 16 years	**1.6* [1.0–2.4]**	1.3 [0.7–2.5]	**1.7* [1.1–2.6]**	1.0 [0.6–1.7]	**7.3* [4.9–10.8]**	**7.0* [4.6–10.6]**	**1.7* [1.0–2.7]**

Note: Table shows RRR and corresponding 95% confidence intervals. Significance is indicated by bold script and asterisks. A RRR lower than 1 indicates a lower sensitization value in this trajectory.

TABLE 5 Results from single multinomial regression of clusters on spirometry measurements, controlled for study and study center in GINIplus/LISA and BAMSE

	Intermittently allergic	Rhinitis	Early-resolving dermatitis	Mid-persisting dermatitis	Multimorbid	Persisting dermatitis + rhinitis	Early-transient asthma
GINIplus & LISA							
FEV1 (GLI)	**0.9* [0.8–1.0]**	0.9 [0.8–1.1]	1.1 [1–1.4]	1.0 [0.8–1.3]	**0.7* [0.6–0.8]**	0.9 [0.7–1.2]	0.5 [0.3–1.0]
FVC (GLI)	0.9 [0.8–1.0]	0.9 [0.8–1.1]	1.1 [0.9–1.3]	0.9 [0.7–1.1]	1.0 [0.8–1.2]	0.9 [0.7–1.2]	0.6 [0.3–1.1]
Tiffeneau-Index (GLI)	0.9 [0.9–1.0]	1.0 [0.8–1.1]	1.1 [0.9–1.3]	1.2 [1.0–1.4]	**0.6* [0.5–0.7]**	1.0 [0.8–1.3]	0.9 [0.5–1.5]
BAMSE							
FEV1 (GLI)	0.9 [0.8–1.1]	1.0 [0.8–1.3]	0.8* [0.7–1.0]	0.9 [0.8–1.1]	**0.6* [0.5–0.7]**	0.9 [0.8–1.1]	**0.7* [0.6–0.8]**
FVC (GLI)	1.0 [0.8–1.1]	1.0 [0.8–1.2]	0.9 [0.7–1.0]	1.0 [0.8–1.1]	0.9 [0.7–1.0]	0.9 [0.7–1.1]	0.9 [0.8–1.1]
Tiffeneau-Index (GLI)	0.9 [0.8–1.1]	1.0 [0.8–1.3]	0.9 [0.8–1.1]	1.0 [0.8–1.1]	**0.6* [0.5–0.7]**	1.1 [0.9–1.3]	**0.6* [0.5–0.8]**

Note: Table shows RRR and 95% corresponding confidence intervals. Significance is indicated by bold script and asterisks. A RRR lower than 1 indicates a lower lung function value in this trajectory.

rhinitis clusters, with lower prevalence of dermatitis and higher prevalence of wheeze, respectively. They further found that all allergic trajectories were already established by adolescence, not changing much after the age of 18 supporting our approach to include participants from infancy to adolescence in our longitudinal clustering, still ensuring comparability to previously reported approaches covering early life windows.

Within the first intermittently allergic cluster it is likely that the allocated participants have mild or transient forms of allergic diseases, as they all reported an allergic disease for at least one time point but prevalences were generally low. Nonetheless, they have a significant risk associated with family history and parents might be more attentive for symptoms of dermatitis and rhinitis.

The rhinitis cluster was inversely associated with environmental factors such as pet exposure and presence of older siblings, which have been discussed in the context of the hygiene hypothesis.[25] This cluster was further significantly associated with the PRS for any allergic diseases and rhinitis, underlining the genetic contribution to this allergic phenotype. It further indicates that within the SNPs identified to be associated with any allergic disease,[5,15] rhinitis is the most common allergic disease and GWAS are commonly not adjusted for comorbid allergic diseases.

As previously reported, early-onset dermatitis is, compared to late-onset dermatitis, associated with sensitization against aeroallergens and higher genetic susceptibility.[41] This is also seen in GINIplus/LISA, with the early dermatitis cluster being significantly associated with the respective PRS and higher aeroallergen sensitization rates at six years. In BAMSE, instead the mid-persisting trajectory is significantly associated with the PRS for dermatitis, but the trajectories prevalence also peaks earlier, potentially blurring the differentiation between the two trajectories, with the early-transient cluster being detected rather in the independent BAMSE clustering solution with six allergic disease trajectories only (Figure S3). Another aspect to consider is the structurally higher percentage of missing diagnoses in the mid-persisting trajectory (Figure S7a), which might lead to falsely carried-forward dermatitis cases in GINIplus/LISA, while non-missing cases might be still following a persisting disease course but on a lower level than assumed here. Further differentiation between the dermatitis clusters is possible in regard to food sensitization, which was observed in the early-resolving but not the mid-persisting cluster, as supported by the literature,[42] which might indicate earlier remission in those children.

The multimorbid cluster is significantly characterized by a higher percentage of male participants, following the previous results reporting an increased risk for allergic multimorbidity in males with family history of allergies.[43] There are also significant associations of the multimorbid cluster with the PRS for asthma and rhinitis in GINIplus/LISA, as well as any allergy and rhinitis in BAMSE.

Both, the multimorbid and rhinitis clusters, show similar rhinitis prevalences and a high sensitization rate. However, one group of participants develops additional asthma, leading to a higher burden of disease and worse lung function. One possible explanation might be higher genetic susceptibility as indicated by the effect of the

asthma PRS; another one could be the higher percentage of bronchitis infections in infancy, negatively affecting spirometry measures[29] as also seen in our results. This development from early bronchitis to asthma and reduced spirometry differentiates the multimorbid from the persisting dermatitis plus rhinitis cluster, where asthma prevalence remains low.

Other early-life factors, such as maternal smoking during pregnancy or second-hand tobacco smoke exposure during early childhood, both risk factors for reduced lung function, did not differ between these clusters. We also did not see significant differences for further common determinants, such as Caesarean section or urbanicity, potentially due to low variance in our sample or low case numbers.

The persisting dermatitis plus rhinitis cluster was previously described in UK studies as "persistent eczema with later-onset rhinitis."[12] It seems to be very robust, as it was found within five different birth cohorts (ALSPAC & MAAS[12]; GINIplus, LISA, and BAMSE). It has the strongest associations with aeroallergen sensitization, nearly double compared with the multimorbid cluster in GINIplus/LISA, but none of the PRS is significantly associated with the trajectory in GINIplus/LISA and only the PRS for any allergic disease in BAMSE.

Nonetheless, FLG mutation, known to be the strongest genetic risk factor for dermatitis,[44] was significantly associated with the persisting dermatitis plus rhinitis cluster in BAMSE. Higher effect sizes, although not significant, can also be seen for the other clusters comprising relevant proportions of dermatitis patients (early-resolving (only GINIplus/LISA), mid-onset and multimorbid). However, the association with FLG seems higher in persisting forms of dermatitis, in comparison with the early-onset resolving cluster.

The last cluster is weakly powered in GINIplus/LISA with only 27 participants, all of which reported an asthma diagnosis in infancy but with a prevalence below 20% later in life. This early-transient asthma sub-phenotype is often identified in clustering approaches as "early-transient wheeze"[9,10] and much more common in English-speaking countries where a higher prevalence of asthma has been reported.[43] This was also observed during our replication in the BAMSE cohort using a symptom-based phenotype definition. Mostly missing sensitization indicates that this cluster does not depict an allergic disease trajectory but rather an early respiratory infection trajectory with potentially ongoing impairment[6] as underlined by the significant association with early bronchitis infections. Previous studies report on the connection between wheezing inducing airway infections and future asthma development[45,46] and although these participants here do not develop further asthma until 15 years of age, they present reduced lung function outcomes in adolescence. The presence of early-asthma or wheezing symptoms is much higher in our symptom-based clustering for GINIplus/LISA (Figure S8) but these cases may just be early respiratory infections instead of asthma, which is difficult to diagnose in infancy.

Similar results in lung function reduction as seen for our multimorbid and early-transient asthma clusters were also reported for cohorts covering nine to 34 and seven to 53 years of age, respectively.[14,47] Forster et al. reported lower Tiffeneau-indices in young adulthood for classes involving wheeze[14] and Bui et al. showing that early-onset persisting asthma trajectories had stronger associations with lower lung function at 53 years than later appearing asthma cases.[47]

One limitation is the difference in allergic disease definitions based on clinical practices in Germany and Sweden. Although all cohorts employ similar time points of follow-up, GINIplus/LISA analyzed parent-reported doctor diagnoses, while in BAMSE, the analysis was based on parent-reported symptoms and prescribed asthma medication use. Remarkably, we were still able to derive the same cluster solution, showing the robustness of the results, further supported by our sensitivity analysis clustering symptom-diagnoses in GINIplus/LISA (Figure S8). Nevertheless, all included studies are observational cohort studies and cannot investigate causality and are impacted by missing validated food allergy assessments, non-measured sensitization in infancy and missing indicators of severity or quality of life. Also, the issue with missing diagnoses due to loss to follow-up needs to be considered in the context of prospective birth cohorts, which collect valuable data but loose participants with each further time point, leading to higher number of missing values at the later time points of this study. Possible solutions for addressing this are different imputation methods but here no clear standard has been established yet. Allergic diseases are chronic but still every imputation might lead to false assumptions or might omit especially transient or late-onset trajectories. Of note, percentage of missing diagnoses further differs not only related to the time points but also between the clusters, especially the mid-persisting dermatitis. Potential reasons include a higher loss-to-follow-up in non-allergic parents (compared to those with positive family history), which are also overrepresented in the mid-persisting dermatitis cluster (Table S4). Furthermore, the early-transient wheeze cluster displays a higher proportion of missing diagnoses at earlier time-points. However, the generally low sample size in this cluster (n = 27) and thus missingness might simply occur by chance in this cluster.

Further, our PRS were calculated purely on genome-wide significant SNP sets and not as proposed by newer publications on complete GWAS results, not restricted by a significance threshold.[48] Nevertheless, in this paper, we aimed to show associations of disease-specific scores, which we assume to be represented by the respective large-scale GWAS results. This approach also enables direct comparison to other publications using a similar PRS for any allergic disease.[21]

The strengths of our study include the long follow-up period, covering both childhood and adolescence and the ability to cover many discussed early-life determinants. The successful replication in an independent cohort further underlines robustness of results. Furthermore, this study demonstrates the association of disease-specific PRS with allergic disease trajectories.

Using the results from this paper, we want to further drive clinical prediction of allergic trajectories, enabling pediatricians to forecast future allergic developments and initiate prevention strategies. Early identification of patients at risk, might help them to mitigate

further risk factors or concretely prevent asthma exacerbations with life-threatening potential.

In conclusion, we aimed to classify allergic disease development to add knowledge about the characteristics and determinants of the derived seven allergic disease trajectories from birth to adolescence. The derived trajectories allow a clearer classification of common allergic disease courses in times of increasing prevalence and burden of disease, which might be further facilitated to improve prediction in the future.

AUTHOR CONTRIBUTIONS

Anna Kilanowski involved in formal analysis, methodology, visualization, writing—original draft preparation, and writing—review and editing. Elisabeth Thiering contributed to conceptualization, methodology, supervision, and writing—review and editing. Gang Wang involved in formal analysis, validation, writing—review and editing. Ashish Kumar and Sara Kress contributed to validation and writing—review and editing. Claudia Flexeder performed writing—review and editing. Carl-Peter Bauer and Dietrich Berdel involved in investigation, resources, and writing—review and editing. Andrea von Berg contributed to funding acquisition, investigation, resources, and writing—review and editing. Anna Bergström performed investigation, resources, and writing—review and editing. Monika Gappa contributed to investigation and writing—review and editing. Joachim Heinrich performed data curation, funding acquisition, investigation, resources, and writing—review and editing. Gunda Herberth contributed to investigation, resources, writing—review and editing. Sibylle Koletzko involved in funding acquisition, investigation, resources, and writing—review and editing. Inger Kull involved in investigation, resources, and writing—review and editing. Erik Melén involved in data curation, funding acquisition, investigation, resources, and writing—review and editing. Tamara Schikowski performed data curation, investigation, resources, and writing—review and editing. Annette Peters involved in supervision and writing—review and editing. Marie Standl involved in conceptualization, data curation, methodology, project administration, resources, supervision, and writing—review and editing.

ACKNOWLEDGMENTS

GINIplus: The authors thank all the families for their participation in the GINIplus study. Furthermore, we thank all the members of the GINIplus Study Group for their excellent work. The GINIplus Study group consists of the following: Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg (Heinrich J, Brüske I, Schulz H, Flexeder C, Zeller C, Standl M, Schnappinger M, Ferland M, Thiering E, Tiesler C); Department of Pediatrics, Marien-Hospital, Wesel (Berdel D, von Berg A); Ludwig-Maximilians-University of Munich, Dr von Hauner Children's Hospital (Koletzko S); Child and Adolescent Medicine, University Hospital rechts der Isar of the Technical University Munich (Bauer CP, Hoffmann U); IUF-Environmental Health Research Institute, Düsseldorf (Schikowski T, Link E, Klümper C, Krämer U, Sugiri D). LISA: The authors thank all the families for their participation in the LISA study. Furthermore, we thank all members of the LISA Study Group for their excellent work. The LISA Study group consists of the following: Helmholtz Zentrum München, German Research Center for Environmental Health, Institute of Epidemiology, Munich (Heinrich J, Schnappinger M, Brüske I, Ferland M, Schulz H, Zeller C, Standl M, Thiering E, Tiesler C, Flexeder C); Department of Pediatrics, Municipal Hospital "St. Georg," Leipzig (Borte M, Diez U, Dorn C, Braun E); Marien Hospital Wesel, Department of Pediatrics, Wesel (von Berg A, Berdel D, Stiers G, Maas B); Pediatric Practice, Bad Honnef (Schaaf B); Helmholtz Centre of Environmental Research—UFZ, Department of Environmental Immunology/Core Facility Studies, Leipzig (Lehmann I, Bauer M, Röder S, Schilde M, Nowak M, Herberth G, Müller J); Technical University Munich, Department of Pediatrics, Munich (Hoffmann U, Paschke M, Marra S); Clinical Research Group Molecular Dermatology, Department of Dermatology and Allergy, Technische Universität München (TUM), Munich (Ollert M, J. Grosch). BAMSE: We thank all the participants and parents participating in the BAMSE cohort and the staff involved in the study through the years. Open Access funding enabled and organized by Projekt DEAL.

FUNDING INFORMATION

GINIplus: The GINIplus study was mainly supported for the first 3 years of the Federal Ministry for Education, Science, Research, and Technology (interventional arm) and Helmholtz Zentrum Munich (former GSF) (observational arm). The 4 year, 6 year, 10 year, and 15 year follow-up examinations of the GINIplus study were covered from the respective budgets of the 5 study centres (Helmholtz Zentrum Munich (former GSF), Research Institute at Marien-Hospital Wesel, LMU Munich, TU Munich and from 6 years onwards also from IUF—Leibniz Research-Institute for Environmental Medicine at the University of Düsseldorf) and a grant from the Federal Ministry for Environment (IUF Düsseldorf, FKZ 20462296). Further, the 15 year follow-up examination of the GINIplus study was supported by the Commission of the European Communities, the 7th Framework Program: MeDALL project, and as well by the companies Mead Johnson and Nestlé. LISA: The LISA study was mainly supported by grants from the Federal Ministry for Education, Science, Research, and Technology and in addition from Helmholtz Zentrum Munich (former GSF), Helmholtz Centre for Environmental Research—UFZ, Leipzig, Research Institute at Marien-Hospital Wesel, Pediatric Practice, Bad Honnef for the first 2 years. The 4 year, 6 year, 10 year, and 15 year follow-up examinations of the LISA study were covered from the respective budgets of the involved partners (Helmholtz Zentrum Munich (former GSF), Helmholtz Centre for Environmental Research—UFZ, Leipzig, Research Institute at Marien-Hospital Wesel, Pediatric Practice, Bad Honnef, IUF—Leibniz-Research Institute for Environmental Medicine at the University of Düsseldorf) and in addition by a grant from the Federal Ministry for Environment (IUF Düsseldorf, FKZ 20462296). Further, the 15-year follow-up examination of the LISA study was supported by the Commission of the European Communities, the 7th Framework Program: MeDALL project. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research

and innovation programme (grant agreement No. 949906). BAMSE: The BAMSE study is supported by grants from the Swedish Research Council, the Swedish Heart-Lung Foundation, Karolinska Institutet SFO Epidemiology and Region Stockholm (ALF project, and for cohort and database maintenance).

CONFLICT OF INTEREST

MG reports grants from Nestlé Vevey, Switzerland, during the conduct of the study and personal fees from Aimmune, ALK, AstraZeneca, Boehringer, GSK, Nestle, Novartis, and Sanofi outside the submitted work. SKoletzko reports grants from Mead Johnson Company during the conduct of the study; personal fees from Nestle, Danone, Shire, AbbVie, ThermoFisher, Janssen, Pfizer, Takeda, Mead Johnson, grants from BioGaia outside the submitted work. EM has received lecture and/or advisory board fees from ALK, AstraZeneca, Chiesi, Novartis and Sanofi outside the submitted work. All other authors have no interests to declare.

ORCID

Anna Kilanowski https://orcid.org/0000-0001-5328-6274
Elisabeth Thiering https://orcid.org/0000-0002-5429-9584
Joachim Heinrich https://orcid.org/0000-0002-9620-1629
Gunda Herberth https://orcid.org/0000-0003-0212-3509
Sibylle Koletzko https://orcid.org/0000-0003-2374-8778
Inger Kull https://orcid.org/0000-0001-6096-3771
Annette Peters https://orcid.org/0000-0001-6645-0985
Marie Standl https://orcid.org/0000-0002-5345-2049

REFERENCES

1. Vos T, Lim SS, Abbafati C, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global burden of disease study 2019. *Lancet*. 2020;396(10258):1204-1222. doi:10.1016/S0140-6736(20)30925-9
2. Nissen SP, Kjær HF, Høst A, Nielsen J, Halken S. The natural course of sensitization and allergic diseases from childhood to adulthood. *Pediatr Allergy Immunol*. 2013;24(6):549-555. doi:10.1111/pai.12108
3. Pinart M, Benet M, Annesi-Maesano I, et al. Comorbidity of eczema, rhinitis, and asthma in IgE-sensitised and non-IgE-sensitised children in MeDALL: a population-based cohort study. *Lancet Respir Med*. 2014;2(2):131-140. doi:10.1016/S2213-2600(13)70277-7
4. Fuertes E, Standl M, Von Berg A, et al. Parental allergic disease before and after child birth poses similar risk for childhood allergies. *Allergy*. 2015;70(7):873-876. doi:10.1111/all.12609
5. Ferreira MA, Vonk JM, Baurecht H, et al. Shared genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. *Nat Genet*. 2017;49(12):1752-1757. doi:10.1038/ng.3985
6. Martinez FD, Wright AL, Taussig LM, Holberg CJ, Halonen M, Morgan WJ. Asthma and wheezing in the first six years of life. The group health medical associates. *N Engl J Med*. 1995;332(3):133-138. doi:10.1056/NEJM199501193320301
7. Lee E, Hong SJ. Phenotypes of allergic diseases in children and their application in clinical situations. *Korean J Pediatr*. 2019;62(9):325-333. doi:10.3345/kjp.2018.07395
8. Panico L, Stuart B, Bartley M, Kelly Y. Asthma trajectories in early childhood: identifying modifiable factors. *PLOS One*. 2014;9(11):e111922. doi:10.1371/journal.pone.0111922
9. Ödling M, Wang G, Andersson N, et al. Characterization of asthma trajectories from infancy to young adulthood. *J Allergy Clin Immunol Pract*. 2021;16:2368-2376.e3. doi:10.1016/j.jaip.2021.02.007
10. 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*. 2013;68(9):1158-1167. doi:10.1111/all.12208
11. Garcia-Aymerich J, Benet M, Saeys Y, et al. Phenotyping asthma, rhinitis and eczema in MeDALL population-based birth cohorts: an allergic comorbidity cluster. *Allergy*. 2015;70(8):973-984. doi:10.1111/all.12640
12. Belgrave DCM, Granell R, Simpson A, et al. Developmental profiles of eczema, wheeze, and rhinitis: two population-based birth cohort studies. *PLoS Med*. 2014;11(10):e1001748. doi:10.1371/journal.pmed.1001748
13. Bougas N, Just J, Beydon N, et al. Unsupervised trajectories of respiratory/allergic symptoms throughout childhood in the PARIS cohort. *Pediatr Allergy Immunol*. 2019;30(3):315-324. doi:10.1111/pai.13027
14. Forster F, Ege MJ, Gerlich J, et al. Trajectories of asthma and allergy symptoms from childhood to adulthood. *Allergy*. 2021;2:1192-1203. doi:10.1111/all.15075
15. Ferreira MAR, Vonk JM, Baurecht H, et al. Eleven loci with new reproducible genetic associations with allergic disease risk. *J Allergy Clin Immunol*. 2019;143(2):691-699. doi:10.1016/j.jaci.2018.03.012
16. Paternoster L, Standl M, Waage J, et al. Multi-ancestry genome-wide association study of 21,000 cases and 95,000 controls identifies new risk loci for atopic dermatitis. *Nat Genet*. 2015;47(12):1449-1456. doi:10.1038/ng.3424
17. El-Husseini ZW, Gosens R, Dekker F, Koppelman GH. The genetics of asthma and the promise of genomics-guided drug target discovery. *Lancet Respir Med*. 2020;8(10):1045-1056. doi:10.1016/S2213-2600(20)30363-5
18. Waage J, Standl M, Curtin JA, et al. Genome-wide association and HLA fine-mapping studies identify risk loci and genetic pathways underlying allergic rhinitis. *Nat Genet*. 2018;50(8):1072-1080. doi:10.1038/s41588-018-0157-1
19. Khera AV, Chaffin M, Aragam KG, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet*. 2018;50(9):1219-1224. doi:10.1038/s41588-018-0183-z
20. Arabkhazaeli A, Ahmadizar F, Leusink M, et al. The association between a genetic risk score for allergy and the risk of developing allergies in childhood—results of the WHISTLER cohort. *Pediatr Allergy Immunol*. 2018;29(1):72-77. doi:10.1111/pai.12824
21. Clark H, Granell R, Curtin JA, et al. Differential associations of allergic disease genetic variants with developmental profiles of eczema, wheeze and rhinitis. *Clin Exp Allergy*. 2019;49(11):1475-1486. doi:10.1111/cea.13485
22. Lemonnier N, Melén E, Jiang Y, et al. A novel whole blood gene expression signature for asthma, dermatitis, and rhinitis multimorbidity in children and adolescents. *Allergy*. 2020;75(12):3248-3260. doi:10.1111/all.14314
23. Pawankar R, Holgate ST, Canonica GW, Lockey RF, Blaiss MS. WAO white book on allergy 2013 update. 2013.
24. Strachan DP. Hay fever, hygiene, and household size. *Br Med J*. 1989;299(6710):1259-1260. doi:10.1136/bmj.299.6710.1259
25. Roduit C, Frei R, von Mutius E, Lauener R. The hygiene hypothesis. In: Esser C, ed. *Environmental Influences on the Immune System*. Springer; 2016:77-96. doi:10.1007/978-3-7091-1890-0_4
26. Hu C, Duijts L, Erler NS, et al. Most associations of early-life environmental exposures and genetic risk factors poorly differentiate between eczema phenotypes: the generation R study. *Br J Dermatol*. 2019;181(6):1190-1197. doi:10.1111/bjd.17879

27. Heinrich J, Brüske I, Cramer C, et al. GINIplus and LISAplus. Design and selected results of two German birth cohorts about natural course of atopic diseases and its determinants. *Allergol Sel*. 2017;1(1):85-95. doi:10.5414/ALX01455E
28. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. *Eur Respir J*. 2005;26(2):319-338. doi:10.1183/09031936.05.00034805
29. Luzak A, Fuertes E, Flexeder C, et al. Which early life events or current environmental and lifestyle factors influence lung function in adolescents? – results from the GINIplus & LISAplus studies. *Respir Res*. 2017;18, 138. doi:10.1186/s12931-017-0619-5
30. Quanjer PH, Stanojevic S, Cole TJ, et al. Multi-ethnic REFERENCE values for spirometry for the 3–95 year age range: the GLOBAL lung function 2012 equations. *Eur Respir J*. 2012;40(6):1324-1343. doi:10.1183/09031936.00080312
31. McLaren W, Gil L, Hunt SE, et al. The Ensembl variant effect predictor. *Genome Biol*. 2016;17(1):122. doi:10.1186/s13059-016-0974-4
32. Weidinger S, Illig T, Baurecht H, et al. Loss-of-function variations within the filaggrin gene predispose for atopic dermatitis with allergic sensitizations. *J Allergy Clin Immunol*. 2006;118(1):214-219. doi:10.1016/j.jaci.2006.05.004
33. Ballardini N, Bergström A, Wahlgren CF, et al. IgE antibodies in relation to prevalence and multimorbidity of eczema, asthma, and rhinitis from birth to adolescence. *Allergy*. 2016;71(3):342-349. doi:10.1111/all.12798
34. Asher MI, Keil U, Anderson HR, et al. International study of asthma and allergies in childhood (ISAAC): rationale and methods. *Eur Respir J*. 1995;8(3):483-491. doi:10.1183/09031936.95.08030483
35. Gref A, Merid SK, Gruzieva O, et al. Genome-wide interaction analysis of air pollution exposure and childhood asthma with functional follow-up. *Am J Respir Crit Care Med*. 2017;195(10):1373-1383. doi:10.1164/rccm.201605-1026OC
36. Genolini C, Alacoque X, Sentenac M, Arnaud C. Kml and kml3d: R packages to cluster longitudinal data. *J Stat Software*. 2015;65(1):1-34. doi:10.18637/jss.v065.i04
37. R Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing; 2019. https://www.R-project.org/
38. Davies DL, Bouldin DW. A cluster separation measure. *IEEE Trans Pattern Anal Mach Intell*. 1979;PAMI-1(2):224-227. doi:10.1109/TPAMI.1979.4766909
39. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol*. 1995;57(1):289-300.
40. Schwarzer G, Carpenter JR, Rücker G. Fixed effect and random effects meta-analysis. In: Schwarzer G, Carpenter JR, Rücker G, eds. *Meta-Analysis with R*. Use R! Springer International Publishing; 2015:21-53. doi:10.1007/978-3-319-21416-0_2
41. Gustafsson D, Sjöberg O, Foucard T. Sensitization to food and airborne allergens in children with atopic dermatitis followed up to 7 years of age. *Pediatr Allergy Immunol*. 2003;14(6):448-452. doi:10.1046/j.0905-6157.2003.00093.x
42. Martin PE, Eckert JK, Koplin JJ, et al. Which infants with eczema are at risk of food allergy? Results from a population-based cohort. *Clin Exp Allergy*. 2015;45(1):255-264. doi:10.1111/cea.12406
43. Gough H, Grabenhenrich L, Reich A, et al. Allergic multimorbidity of asthma, rhinitis and eczema over 20 years in the German birth cohort MAS. *Pediatr Allergy Immunol*. 2015;26(5):431-437. doi:10.1111/pai.12410
44. Margolis DJ, Apter AJ, Gupta J, et al. The persistence of atopic dermatitis and filaggrin (FLG) mutations in a US longitudinal cohort. *J Allergy Clin Immunol*. 2012;130(4):912-917. doi:10.1016/j.jaci.2012.07.008
45. Mikhail I, Grayson MH. Asthma and viral infections: an intricate relationship. *Ann Allergy Asthma Immunol*. 2019;123(4):352-358. doi:10.1016/j.anai.2019.06.020
46. Rossi GA, Colin AA. Infantile respiratory syncytial virus and human rhinovirus infections: respective role in inception and persistence of wheezing. *Eur Respir J*. 2015;45(3):774-789. doi:10.1183/09031936.00062714
47. Bui DS, Lodge CJ, Perret JL, et al. Trajectories of asthma and allergies from 7 years to 53 years and associations with lung function and extrapulmonary comorbidity profiles: a prospective cohort study. *Lancet Respir Med*. 2021;9(4):387-396. doi:10.1016/S2213-2600(20)30413-6
48. Choi SW, Mak TSH, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. *Nat Protoc*. 2020;15(9):2759-2772. doi:10.1038/s41596-020-0353-1

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Kilanowski A, Thiering E, Wang G, et al. Allergic disease trajectories up to adolescence: Characteristics, early-life, and genetic determinants. *Allergy*. 2022;00:1-15. doi: 10.1111/all.15511