**Evolution of eczema, wheeze and rhinitis from infancy to early adulthood: analysis of four birth cohort studies**

Sadia Haider1, Sara Fontanella1, Anhar Ullah1, Stephen Turner2,3, Angela Simpson4, Graham Roberts5,6,7, Clare S Murray4, John W. Holloway5,6, John A Curtin4, Paul Cullinan1, Syed Hasan Arshad6,7,8, Guillem Hurault9, Raquel Granell10, and Adnan Custovic1

on behalf of STELAR/UNICORN11 investigators

1National Heart and Lung Institute, Imperial College London, UK

2 Royal Aberdeen Children's Hospital NHS Grampian Aberdeen, AB25 2ZG, UK

3Child Health, University of Aberdeen, Aberdeen, UK

4Division of Infection, Immunity and Respiratory Medicine, School of Biological Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Science Centre, NIHR Manchester Biomedical Research Unit, Manchester University NHS Foundation Trust, Manchester, UK

5Human Development and Health, Faculty of Medicine, University of Southampton, Southampton, UK

6NIHR Southampton Biomedical Research Centre, University Hospitals Southampton NHS Foundation Trust, Southampton, UK

7David Hide Asthma and Allergy Research Centre, Isle of Wight, UK

8Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK

9 Faculty of Engineering, Department of Bioengineering, Imperial College London, UK

10MRC Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol, UK

**Correspondence to:**Adnan Custovic MD PhD FMedSci, Imperial College London, a.custovic@imperial.ac.uk

**Contributions:** AC, AS, SH and SF conceived and planned the study, and wrote the manuscript. SH, SF and GH analysed the data. All authors contributed to the interpretation of the results. All authors provided critical feedback and helped shape the research, analysis and manuscript.

**Funding:** UNICORN consortium is funded by the UK Medical Research Council (MRC) Programme Grant MR/S025340/1; STELAR consortium was funded through MRC grants G0601361 and MR/K002449/1. RG is funded through Wellcome Trust Strategic Award 108818/15/Z. MAAS was supported by the Asthma UK Grants No 301 (1995-1998), No 362 (1998-2001), No 01/012 (2001-2004), No 04/014 (2004-2007), BMA James Trust (2005) and the JP Moulton Charitable Foundation (2004-2016), The North West Lung Centre Charity (1997-current) and the Medical Research Council (MRC) grant MR/L012693/1 (2014-2018). IOW was supported by Isle of Wight Health Authority, the National Asthma Campaign, UK (Grant No 364) and NIH R01 HL082925-01, R01 AI091905 and R01 AI121226.

**Abstract word count:** 250

**Word count:** 3498

**ABSTRACT**

**Background:** The relationship between atopic diseases (eczema, wheeze/asthma and rhinitis) is complex, and epidemiology and mechanisms of their comorbidities is unclear.

**Objective:** To investigate within-individual patterns of morbidity of eczema, wheeze and rhinitis from birth to adolescence/early adulthood.

**Methods:** We investigated onset/progression/resolution of eczema, wheeze and rhinitis using sequence mining and Latent Markov modelling (LMM) in four population-based birth cohorts. We used logistic regression to ascertain if early-life eczema or wheeze, or genetic factors (*filaggrin* mutations and 17q21 variants), increase the risk of multimorbidity.

**Results:** Single conditions, although the most prevalent, were observed significantly less frequently than by chance. There was considerable variation in the timing of onset/remission/persistence/intermittence. Multimorbidity of eczema+wheeze+rhinitis was rare, but significantly over-represented (3-6 times more often than by chance). Although infantile eczema was associated with subsequent multimorbidity, most children with eczema (75.4%) did not progress to any multimorbidity pattern. *FLG* mutations and rs7216389 were not associated with persistence of eczema/wheeze as single conditions, but both increased the risk of multimorbidity development (*FLG* by 2-3-fold, rs7216389 risk variants by 1.4-1.7-fold). LMM revealed 5 latent states (No disease/low risk; Mainly eczema; Mainly Wheeze; Mainly rhinitis; Multimorbidity). The most likely transition to Multimorbidity was from Eczema state (0.21). However, although this was one of the highest transition probabilities, only 1/5 of those with eczema transitioned to multimorbidity.

**Conclusions:** Atopic diseases fit a multimorbidity framework, with no evidence for sequential "atopic march” progression. The highest transition to multimorbidity was from eczema, but most children with eczema (~80%) had no comorbidities.

**INTRODUCTION**

Childhood eczema, wheezing/asthma and rhinitis are often collectively referred to as atopic diseases (2, 3). The clinical presentation encompasses multiple phenotypes, and some patients have symptoms affecting a single organ, while others have symptoms of varying severity affecting several organs (4, 5). The pathophysiological mechanisms which underpin this heterogeneity are largely unknown.

The relationship between atopic diseases is complex, and there is an ongoing controversy over the epidemiology and mechanisms of comorbidity (6). One paradigm is Atopic march, which, as originally proposed, described the progression of atopic disease in an individual as a sequential development starting with eczema in infancy and progressing to wheezing/asthma, and then rhinitis, in later childhood (8, 9). A specific sequence is implicit by the use of the term march (3). This framework is extended to the recommendation that primary care physicians “should inform parents that children with eczema may later develop asthma” (10), and has underpinned clinical trials specifically aiming to prevent wheezing/asthma in children with early-life eczema (11, 12). However, some studies have shown a substantial heterogeneity between patients in the chronology of symptom development (13-15), questioning a specific sequence of atopic march (16). Application of Bayesian machine learning to model the development of eczema, wheeze and rhinitis from birth to school-age in two population-based birth cohorts revealed eight latent profiles of atopic diseases development, each with different temporal patterns of symptoms co-manifestation (17), and distinct genetic associates (18). Thus, the evidence to date is convincing that atopic diseases coexist (2, 7, 19, 20), and although there is increasing acknowledgement of different trajectories (20, 21), a comprehensive analysis of their long-term evolution within individuals is lacking, and the mechanisms of their coexistence remain unclear (6).

Atopic comorbidities may occur due to the effects of an index disease (as in atopic march in which eczema, as the index disease, impacts upon the future risk of wheeze/asthma and rhinitis (9)), or in a multimorbidity framework, in which no single condition holds priority over any of the co-occurring conditions (22), via a common underlying pathogenic mechanism (e.g. impaired skin barrier leading to allergic sensitisation (23)). However, co-occurrence can also occur by chance; for example, if the population prevalence of eczema is 25%, and wheeze 30%, by chance alone, we would expect 7.5% of individuals (0.25\*0.3=0.075) to have both. To capture the spectrum of morbidity of atopic disease from birth to adulthood, we investigated patterns of onset, remission and persistence of eczema, wheeze and rhinitis using data from four population-based birth cohorts, and used sequence mining techniques to disaggregate and describe within-individual patterns. To ascertain whether there is evidence for shared genetic architecture across different patterns of co-occurring diseases, we took a candidate gene approach by investigating associations with *Filaggrin* loss-of-function mutations and a representative variant from 17q21 locus.

**METHODS**

**Study design, setting, participants and data sources**

Methods are described in detail in the Supplementary appendix. Briefly, we used data from four UK population−based birth cohorts in the STELAR consortium: Ashford (25), Isle of Wight (IOW) (26), Manchester Asthma and Allergy Study (MAAS) (27) and Aberdeen cohort (SEATON) (28). All studies recruited pregnant women who gave birth to 642, 1456, 1184 and 1924 children respectively, between 1989 and 1999. All studies were approved by research ethics committees. Informed consent was obtained from parents, and participants gave their assent/consent when applicable. Data were integrated in a web−based knowledge management platform to facilitate joint analyses (24).

Information on symptoms was collected using validated questionnaires administered on multiple occasions from infancy to adolescence/early adulthood (7 in ASHFORD over 14 years, 6 in MAAS over 16 years, 6 in SEATON over 14 years, and 6 in IOW over 26 years). The cohort-specific follow-up time-points, the questions used to define variables, and sample sizes are shown in Table E1.

**Definition of outcomes**

We ascertained current eczema, wheeze and rhinitis at each follow-up. For each individual at each time point we derived a variable summarising the coexistence of individual diseases, comprising 8 categories: (1) No disease; (2-4) Single disease: only eczema (E); only wheeze (W); only rhinitis (R); (5-7) Combinations of two diseases: eczema+wheeze (E+W), eczema+rhinitis (E+R), wheeze+rhinitis (W+R); (8) atopic triad: eczema+wheeze+rhinitis (E+W+R).

Definitions of all variables are presented in Supplementary Methods and Table E2.

**Genotyping**

Genotyping and quality control in each cohort are described in Supplementary Appendix. Briefly, *FLG* was genotyped using TaqMan based allelic discrimination assay for R501X and S3247X loss-of-function mutations, and a fluorescent-labeled PCR for 2282del4 (29). Data was analyzed as combined carriage of a *FLG* null allele, i.e. children carrying one or more of the three genetic variations were considered as having a *FLG* loss-of-function mutation. For 17q21 locus, we used the SNP rs7216389 in the *GSDMB*, which was coded for its risk allele (T); an additive (dosage) model was used.

**Statistical analysis**

Cross-sectional analyses focused on estimating the prevalence of single and co-occurring conditions at each time-point. Based on the point prevalence of eczema, wheeze and rhinitis at each time in each cohort, we calculated the probabilities of different symptoms coexistence in the same individual being observed by chance. We then compared observed and expected probabilities across populations and time points to ascertain which co-occurrence patterns were observed more frequently than by chance using the exact binomial test with Benjamini-Hochberg procedure to account for multiple comparisons.

We used multinomial logistic regression models to ascertain if early-life eczema or wheeze as index diseases, and rs7216389 and *FLG* (including their interaction) increased the risk of multi-morbidity thereafter; results are reported as relative risk ratios (RRR) with 95% confidence intervals (CI).

Longitudinal analyses among subjects with complete information on all 3 symptoms/diseases at all follow-ups comprised of two approaches: sequence analysis and multivariate Latent Markov modelling (LMM). The former described and visualized trajectories and transitions, while LMM was used for measuring the dynamics of change between successive time-points (30, 31). The optimal number of states was identified using the Bayesian Information Criterion (BIC) index in conjunction with interpretation of the conditional response probabilities. Finally, we explored associations between derived latent states and allergic sensitisation, and ascertained their genetic associates. All analyses were conducted in R using the *LMest* (32) and *TraMineR* (33) packages.

**RESULTS**

Descriptive characteristics of study populations and comparisons between included and excluded subjects are shown in Table E3. Maternal smoking was significantly less common among included participants in all cohorts. Table 1 shows data on prevalence of eczema, wheeze and rhinitis and their co-occurrence at each time-point across cohorts. Having a single disease was much more common than co-occurrence at all time-points and in all cohorts, with approximately one-third of study participants experiencing a single disease compared to 7-14% with two (Table E4). Atopic triad (E+W+R) multimorbidity was relatively rare throughout the observation period (~2-4% by the final time-point) and increased gradually from infancy to age 4-5 years, with little change thereafter (Tables 1 and E4).

**Co-occurrence patterns**

Figure 1 and Table E5 show the deviation of observed from expected probabilities of symptoms co-occurrence at each time point. Across all cohorts, single conditions, although the most prevalent cross-sectionally, were observed significantly less frequently than by chance at all follow-ups. In general, two-disease combinations tended to co-occur as often as would be expected by chance. Atopic triad, although rare, was significantly over-represented in all cohorts and time points (on average, 3-6 times more often than by chance).

**Longitudinal sequence analysis**

We carried out longitudinal analyses among 1898 participants with complete data at all follow-ups. Figure 2 shows individual-level sequences of symptoms across time. There was no typical trajectory, but considerable heterogeneity in the onset, remission, and persistence of symptoms. The number of person-unique sequences ranged from 220 to 351 across cohorts. The most common sequence was a single record of late-onset rhinitis. Figure E1 shows sequence frequency plots for 20 most common trajectories, which accounted for only ~26-32% of all sequences. Among children with eczema (Figure E2) or wheeze (Figure E3) in the first 3 years, transition to no disease was the most common sequence. All three symptoms were reported (including non-contemporaneously) by 374/1898 (19.6%), and 166 (8.7%) reported coincident E+W+R at least once.

***E+W+R multimorbidity:*** We carried out further analyses exploring symptom development among 166/1898 (8.7%) participants who experienced E+W+R at least once (Table E6). Of those, 157 (95%) had E+W+R in the school-age/adolescence/early adulthood, and 9 (5%) in early-life only.

Among 157 participants with E+W+R multimorbidity from school-age through adolescence, the majority (n=87, 55.4%) had eczema in the first year of life (Table E6). However, 41 (26.1%) did not have any symptoms in the first year, and 29 (18.5%) had wheeze only. Although early eczema was clearly associated with subsequent E+W+R multimorbidity, most children with eczema in the first year of life (267/354, 75.4%), as a single disease of comorbid condition, did not have E+W+R to adolescence/early adulthood.

**Early-life eczema and wheeze as “index” diseases**

We further investigated the relationship between eczema and wheeze in the first 3 years of life as index conditions with subsequent persistence, or development of different comorbidity patterns, to pre-school, mid-school and adolescence using multivariable logistic regression analyses of joint data at harmonised time-points (early life: 0-3 years; pre-school: 4-5 years; mid-childhood: 8-10 years; adolescence: 14-18 years). Early-life eczema only was associated with an increased risk of all profiles containing eczema through to adolescence (Table 2); the risk of eczema persistence as a single disease decreased significantly with increasing age, but there was no change in the magnitude of risk for co-morbid E+W or E+W+R. Early-life wheeze only was associated with persistence of wheeze, and a 3-fold increase in W+E and W+R at pre-school age, with no consistent comorbidity associations thereafter. Finally, E+W in the first 3 years was associated with substantially higher risk of all comorbidity patterns throughout childhood, with ~18-fold increase in E+W+R multimorbidity and ~14-21.5-fold higher risk of the persistence of W+E throughout childhood. In all three time-periods, early E+W increased the risk of all conditions more than single index diseases.

We found no significant associations between *FLG* mutations or rs7216389 with persistence of eczema or wheeze as single conditions. However, both were associated with the development of E+W+R multimorbidity. In all 3 models, *FLG* mutations were associated with a 2- to 3-fold higher risk of E+W+R, and RRRs for rs7216389 were smaller (1.4-1.7). rs7216389, but not *FLG,* was associated with W+R from mid-childhood (Table 2). We tested for an interaction effect of *FLG*\*rs7216389, however, this was not significant.

**Dynamics of change over time: Latent Markov modelling**

We applied LMM in a joint model to data from 2079 subjects with complete information on eczema, wheeze and rhinitis at 5 harmonised time-points (Table E7): Infancy (Age 1); Early life (age 2-3); Preschool (age 4-5; Mid-school (age 8-10); Adolescence (age 14-18 years). The optimal solution was a time-homogeneous model with five latent states (Table E8). There was a spectrum of co-morbidity risk in each latent state (conditional response probabilities, Table 3). We labelled the states based on the probability of dominant symptom as: (1) No disease/low risk; (2) Mainly eczema; (3) Mainly Wheeze; (4) Mainly rhinitis; (5) Multimorbidity.

Figure 3a shows predicted latent Markov states across all follow ups for each individual participant. The initial probabilities of state membership, and the probabilities of transitioning to different states are shown in Table 3; Figure 3b shows the relative size of transitions between latent states. The probability of starting in the Eczema and Wheeze states was similar (0.17 and 0.15) and was close to zero for Rhinitis and Multimorbidity states (0.03 and 0.02). Children in Eczema and Wheeze states were most likely to stay in these states (0.62 and 0.59). Children in Wheeze state were more likely to transition to Low risk than those in Eczema state (0.28 and 0.12), and the probability of transitioning from Eczema to Wheeze was very low (0.01). The most likely transition to Multimorbidity state was from Eczema state (0.21). However, whilst this was one of the highest transition probabilities, only 1 in 5 children transitioned from Eczema to Multimorbidity state (Figure 3b). For participants in the Multimorbidity state there was a high probability of persisting in this state (0.78). Figure 3c shows the individual-level transitions between the states at each time-point.

**Genetic associations of multimorbidity persistence**

To investigate whether *FLG* mutations and rs721389 were associated with Multimorbidity state persistence, we ran multinomial logistic regression analyses using the number of time periods in the Multimorbidity state (0, 1, 2-5) as the outcome (Table 4). Eczema and Wheeze states in early life were included as predictors. Neither *FLG* mutations nor rs721389 were significantly associated with having Multimorbidity once, but both significantly increased the risk of Persistent multimorbidity. In the model controlling for Eczema and Wheeze states in early life and sex, *FLG* mutations significantly increased the risk of Multimorbidity persistence (OR 1.75, 95% CI 1.05-2.92, p=0.032), and rs721389 was associated with ~50% increase in risk (OR 1.49, 95% CI 1.15-1.94, p=0.003). There was no significant interaction between *FLG* and rs721389.

**Associations of multimorbidity persistence with allergic sensitisation**

Table E9 shows associations between multimorbidity and sensitisation in pre-school and adolescence. Children in the Multimorbidity state were more likely to be sensitised, and sensitisation prevalence was consistently higher in the group with persistent multimorbidity (2-5 time-points). A similar trend is evident for poly-sensitisation. However, more than half of subjects with persistent multimorbidity were not sensitized at age 5, and ~30% were not sensitized in adolescence. Characteristics of children with persistent multimorbidity stratified by sensitisation status in childhood (age 5) and adolescence (age 14-18) is shown in Table E10. “Atopic multimorbidity” at both ages was associated with male sex. Maternal eczema was more common in those with “non-atopic multimorbidity” in school age, but paternal hay-fever was associated with a greater risk of “atopic multimorbidity”. There was a trend towards higher proportion of maternal smoking in “non-atopic multimorbidity”, however, the difference was not significant.

**DISCUSSION**

We used different temporal frameworks and different methodologies (descriptive statistics, frequentist methods and stochastic modelling) to investigate the sequence of the development of eczema, wheeze and rhinitis from infancy to early adulthood. Figure E4 provides a schematic overview of the results. Across all cohorts and time points, single conditions were considerably more prevalent than any co-occurrence. The combination of two diseases in the same individual occurred as frequently as expected by chance (apart from wheeze+rhinitis which occurred more frequently from mid-childhood onwards). Although the prevalence of E+W+R multimorbidity was low (2-4% by adolescence), a consistent finding was that this pattern was more prevalent in all study populations than by chance, and was stable from early school- age (e.g., in the IoW cohort in which data collection spanned to age 26 years, the proportion of participants with E+W+R multimorbidity remained at ~3% from age 4 years to adulthood).

We identified considerable variation in the timing of onset and remission, persistence and intermittence of symptoms. All methods led to similar conclusions, including the observation that most children with early-life eczema did not develop wheeze and/or rhinitis, and of those who experienced all three symptoms in the observation period, very few followed a sequence described as the “atopic march”. Sequence mining of individual trajectories highlighted the vast heterogeneity in individual-level symptom development, and no single pattern dominated, with different trajectories leading to multimorbidity. Whilst children with early-life eczema had a higher risk of developing multimorbidity than those with early wheeze, the attributable risk for an individual child with early-life eczema was small. This dynamic of change was confirmed by LMM, in that children had higher risk of transitioning to the Multimorbidity state from Eczema than from Wheeze state, but those in Eczema state were more likely to remain in the same state than to transition to Multimorbidity. Our results suggest that the relationship between atopic diseases fits a multimorbidity framework in which no single disease holds priority over any of the co-occurring conditions (34).

There may be a genetic predisposition for developing multimorbidity, and *FLG* may be important locus. *FLG* was not associated with early-life transient eczema, or with eczema persistence as a single disease. However, we showed a consistent association of *FLG* with persistent multimorbidity (i.e., all patterns leading to coexistence of all 3 symptoms in the same individual), which is consistent with 2 previous studies (18, 35). It is tempting to speculate that genotyping patients with early-life eczema (particularly those with co-occurring wheeze) for *FLG* mutations could help identify children who may benefit from interventions targeted at prevention of multimorbidity.

Our study has several limitations. There were differences in question wording between cohorts, and different definitions can impact upon prevalence estimates and associated risk factors (36, 37). However, we chose variables to be as consistent as possible. A further limitation relevant for interpretation is that we used symptom-based classifications by questionnaire-based definitions, and from these definitions we could not ascertain whether the severity of eczema (or wheeze) is associated with multimorbidity (12). We could not discern whether observations of the “same” symptoms in different children (or in the same child at different time points) may have arisen through different mechanisms (for example, whether eczema among children with eczema-only has the same underlying mechanism as eczema in patients with comorbidities).

*FLG* mutations which we used in this study play an important role in individuals of Caucasian ancestry, but their associations with clinical outcomes differ significantly by race (38). Our results are therefore not directly transferable to all ethnic groups.

We used wheeze instead of asthma. Whilst not all wheeze is asthma, wheeze is the most common manifestation of asthma. Furthermore, there is no formal operational definition of asthma, and pre-school age children are rarely diagnosed as asthmatics.

Food allergy might be involved in the transitions to multimorbidity. However, very few population-based birth cohorts have oral food challenge (OFC)-confirmed data on food allergy. In MAAS, we carried out OFCs to confirm peanut allergy (39-41), and have shown that the risk is markedly higher amongst children with persistent eczema (42), and those with co-morbid persistent eczema and wheeze, but not with transient phenotypes (43). In the exploratory single-cohort analysis in the current study, MAAS participants with multimorbidity persistence were 5-times more likely to be peanut allergic than those without multimorbidity (10% vs. 2%; data available on request), suggesting a link between food allergy and multimorbidity. However, we cannot quantify this confidently given the relatively small sample, and this warrants further investigation.

One strength of our approach is that we used data from four birth cohorts with detailed longitudinal phenotyping, which were harmonised to allow joint analyses. Further strength includes the application of various methodologies, with all findings pointing in the same directions, providing evidence of not only replication, but also triangulation, thereby strengthening confidence in our findings (44).

Rather than applying latent class (LC) models, which have been extensively used to study wheeze and eczema (45-51), we used LMM. A key difference is that in the LC models every subject remains in the same latent class across time, whilst in LMM subjects can transition between latent states, thereby allowing for phenotypic instability over time. An advantage of this approach is that it allows the time dependency between successive multivariate observations to be estimated. More specifically, we could observe whether the presence of one disorder increases the probability of developing (or transitioning) to others. Our results were obtained under the first-order Markov assumption, which states that the future state is independent of the historical events given the current state. This assumption could be relaxed by adopting a higher-order Markov chain, thereby allowing the conditional independence to include more time lags. However, over-parametrizing the transition probabilities increases the complexity, and affects the interpretability of the final model

The observation of co-occurrence does not imply any specific causal relationship (in particular in relation to sensitisation, as almost one third of individuals with E+W+R multimorbidity were not sensitised). Association of “non-atopic multimorbidity” with maternal eczema, and a trend towards higher frequency of maternal smoking, suggest the potential importance of skin barrier and specific environmental exposures in “non-atopic triad”. However, caution is required when interpreting these findings, since in the stratified analysis, the sample size was relatively low. The relationship between multimorbidity and sensitisation warrants further investigation.

In conclusion, our findings confirm that eczema, wheeze and rhinitis are not independent from each other, but there is no specific or typical sequence of symptoms development that characterises atopic multimorbidity. Overall, ~50% of children have at least one of these symptoms, but only ~4-6% of children with symptoms have multimorbidity that does not arise as a chance co-occurrence We found no evidence of a sequential “atopic march” progression. The early comorbidities increase the risk of future persistent multimorbidity, hence, early-life diseases should be examined (both clinically and epidemiologically) in the context of the co-occurrence of other conditions. We suggest that physicians should enquire about different atopic disorders if a child presents with one, but should not make recommendations about ways to prevent atopic march, or inform parents that children with eczema may later develop asthma. The term atopic march should not be used to describe atopic multimorbidity, and we should reform the taxonomy of atopic diseases from traditional symptom-based criteria towards a mechanism-based framework. However, for this change to be meaningful, the current symptom-based diagnoses will have to be surpassed by understanding of disease mechanisms.

**REFERENCES**

1. Dharmage SC, Perret JL, Custovic A. Epidemiology of Asthma in Children and Adults. *Frontiers in Pediatrics* 2019; 7: 246.

2. Moreno MA. JAMA Pediatrics Patient Page. Atopic Diseases in Children. *JAMA Pediatr* 2016; 170: 96.

3. Custovic A, Custovic D, Kljaic Bukvic B, Fontanella S, Haider S. Atopic phenotypes and their implication in the atopic march. *Expert Rev Clin Immunol* 2020: 1-9.

4. Custovic A, Henderson J, Simpson A. Does understanding endotypes translate to better asthma management options for all? *J Allergy Clin Immunol* 2019; 144: 25-33.

5. Akar-Ghibril N, Casale T, Custovic A, Phipatanakul W. Allergic Endotypes and Phenotypes of Asthma. *J Allergy Clin Immunol Pract* 2020; 8: 429-440.

6. Silverberg JI. Comorbidities and the impact of atopic dermatitis. *Ann Allergy Asthma Immunol* 2019; 123: 144-151.

7. Pinart M, Benet M, Annesi-Maesano I, von Berg A, Berdel D, Carlsen KC, Carlsen KH, Bindslev-Jensen C, Eller E, Fantini MP, Lenzi J, Gehring U, Heinrich J, Hohmann C, Just J, Keil T, Kerkhof M, Kogevinas M, Koletzko S, Koppelman GH, Kull I, Lau S, Melen E, Momas I, Porta D, Postma DS, Ranciere F, Smit HA, Stein RT, Tischer CG, Torrent M, Wickman M, Wijga AH, Bousquet J, Sunyer J, Basagana X, Guerra S, Garcia-Aymerich J, Anto JM. 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: 131-140.

8. Spergel JM. From atopic dermatitis to asthma: the atopic march. *Ann Allergy Asthma Immunol* 2010; 105: 99-106; quiz 107-109, 117.

9. Bantz SK, Zhu Z, Zheng T. The Atopic March: Progression from Atopic Dermatitis to Allergic Rhinitis and Asthma. *J Clin Cell Immunol* 2014; 5.

10. Munidasa D, Lloyd-Lavery A, Burge S, McPherson T. What Should General Practice Trainees Learn about Atopic Eczema? *J Clin Med* 2015; 4: 360-368.

11. Warner JO, Child ESGETotA. A double-blinded, randomized, placebo-controlled trial of cetirizine in preventing the onset of asthma in children with atopic dermatitis: 18 months' treatment and 18 months' posttreatment follow-up. *J Allergy Clin Immunol* 2001; 108: 929-937.

12. Schneider L, Hanifin J, Boguniewicz M, Eichenfield LF, Spergel JM, Dakovic R, Paller AS. Study of the Atopic March: Development of Atopic Comorbidities. *Pediatr Dermatol* 2016; 33: 388-398.

13. Williams H, Flohr C. How epidemiology has challenged 3 prevailing concepts about atopic dermatitis. *J Allergy Clin Immunol* 2006; 118: 209-213.

14. Barberio G, Pajno GB, Vita D, Caminiti L, Canonica GW, Passalacqua G. Does a 'reverse' atopic march exist? *Allergy* 2008; 63: 1630-1632.

15. Illi S, von Mutius E, Lau S, Nickel R, Gruber C, Niggemann B, Wahn U. The natural course of atopic dermatitis from birth to age 7 years and the association with asthma. *J Allergy Clin Immunol* 2004; 113: 925-931.

16. Hopper JL, Bui QM, Erbas B, Matheson MC, Gurrin LC, Burgess JA, Lowe AJ, Jenkins MA, Abramson MJ, Walters EH, Giles GG, Dharmage SC. Does eczema in infancy cause hay fever, asthma, or both in childhood? Insights from a novel regression model of sibling data. *J Allergy Clin Immunol* 2012; 130: 1117-1122.

17. Belgrave DC, Granell R, Simpson A, Guiver J, Bishop C, Buchan I, Henderson AJ, Custovic A. Developmental profiles of eczema, wheeze, and rhinitis: two population-based birth cohort studies. *PLoS medicine* 2014; 11: e1001748.

18. Clark H, Granell R, Curtin JA, Belgrave D, Simpson A, Murray C, Henderson AJ, Custovic A, Paternoster L. Differential associations of allergic disease genetic variants with developmental profiles of eczema, wheeze and rhinitis. *Clin Exp Allergy* 2019; 49: 1475-1486.

19. Ballardini N, Kull I, Lind T, Hallner E, Almqvist C, Ostblom E, Melen E, Pershagen G, Lilja G, Bergstrom A, Wickman M. Development and comorbidity of eczema, asthma and rhinitis to age 12: data from the BAMSE birth cohort. *Allergy* 2012; 67: 537-544.

20. Spergel JM. The atopic march: Where we are going? Can we change it? *Ann Allergy Asthma Immunol* 2021; 127: 283-284.

21. Paller AS, Spergel JM, Mina-Osorio P, Irvine AD. The atopic march and atopic multimorbidity: Many trajectories, many pathways. *J Allergy Clin Immunol* 2019; 143: 46-55.

22. Melen E, Himes BE, Brehm JM, Boutaoui N, Klanderman BJ, Sylvia JS, Lasky-Su J. Analyses of shared genetic factors between asthma and obesity in children. *J Allergy Clin Immunol* 2010; 126: 631-637 e631-638.

23. Chan A, Terry W, Zhang H, Karmaus W, Ewart S, Holloway JW, Roberts G, Kurukulaaratchy R, Arshad SH. Filaggrin mutations increase allergic airway disease in childhood and adolescence through interactions with eczema and aeroallergen sensitization. *Clin Exp Allergy* 2018; 48: 147-155.

24. Custovic A, Ainsworth J, Arshad H, Bishop C, Buchan I, Cullinan P, Devereux G, Henderson J, Holloway J, Roberts G. The Study Team for Early Life Asthma Research (STELAR) consortium ‘Asthma e-lab’: team science bringing data, methods and investigators together. *Thorax* 2015: thoraxjnl-2015-206781.

25. Cullinan P, MacNeill SJ, Harris JM, Moffat S, White C, Mills P, Newman Taylor AJ. Early allergen exposure, skin prick responses, and atopic wheeze at age 5 in English children: a cohort study. *Thorax* 2004; 59: 855-861.

26. Arshad SH, Holloway JW, Karmaus W, Zhang H, Ewart S, Mansfield L, Matthews S, Hodgekiss C, Roberts G, Kurukulaaratchy R. Cohort Profile: The Isle Of Wight Whole Population Birth Cohort (IOWBC). *Int J Epidemiol* 2018; 47: 1043-1044i.

27. Custovic A, Simpson BM, Murray CS, Lowe L, Woodcock A, Asthma NACM, Allergy Study G. The National Asthma Campaign Manchester Asthma and Allergy Study. *Pediatr Allergy Immunol* 2002; 13: 32-37.

28. Martindale S, McNeill G, Devereux G, Campbell D, Russell G, Seaton A. Antioxidant intake in pregnancy in relation to wheeze and eczema in the first two years of life. *Am J Resp Crit Care* 2005; 171: 121-128.

29. Palmer CN, Irvine AD, Terron-Kwiatkowski A, Zhao Y, Liao H, Lee SP, Goudie DR, Sandilands A, Campbell LE, Smith FJ, O'Regan GM, Watson RM, Cecil JE, Bale SJ, Compton JG, DiGiovanna JJ, Fleckman P, Lewis-Jones S, Arseculeratne G, Sergeant A, Munro CS, El Houate B, McElreavey K, Halkjaer LB, Bisgaard H, Mukhopadhyay S, McLean WH. Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis. *Nat Genet* 2006; 38: 441-446.

30. Bartolucci F, Farcomeni A, Pennoni F. Latent Markov models for longitudinal data. Boca Raton, FL: CRC Press, Taylor & Francis Group; 2013.

31. Langeheine R, Van de Pol F. A unifying framework for Markov modeling in discrete space and discrete time. *Sociological Methods & Research* 1990; 18: 416-441.

32. Bartolucci F, Pandolfi S, Pennoni F. LMest: An R Package for Latent Markov Models for Longitudinal Categorical Data. *2017* 2017; 81: 38.

33. Gabadinho A, Ritschard G, Müller NS, Studer M. Analyzing and Visualizing State Sequences in R with TraMineR. *2011* 2011; 40: 37.

34. Nicholson K, Makovski TT, Griffith LE, Raina P, Stranges S, van den Akker M. Multimorbidity and comorbidity revisited: refining the concepts for international health research. *J Clin Epidemiol* 2019; 105: 142-146.

35. Marenholz I, Esparza-Gordillo J, Ruschendorf F, Bauerfeind A, Strachan DP, Spycher BD, Baurecht H, Margaritte-Jeannin P, Saaf A, Kerkhof M, Ege M, Baltic S, Matheson MC, Li J, Michel S, Ang WQ, McArdle W, Arnold A, Homuth G, Demenais F, Bouzigon E, Soderhall C, Pershagen G, de Jongste JC, Postma DS, Braun-Fahrlander C, Horak E, Ogorodova LM, Puzyrev VP, Bragina EY, Hudson TJ, Morin C, Duffy DL, Marks GB, Robertson CF, Montgomery GW, Musk B, Thompson PJ, Martin NG, James A, Sleiman P, Toskala E, Rodriguez E, Folster-Holst R, Franke A, Lieb W, Gieger C, Heinzmann A, Rietschel E, Keil T, Cichon S, Nothen MM, Pennell CE, Sly PD, Schmidt CO, Matanovic A, Schneider V, Heinig M, Hubner N, Holt PG, Lau S, Kabesch M, Weidinger S, Hakonarson H, Ferreira MAR, Laprise C, Freidin MB, Genuneit J, Koppelman GH, Melen E, Dizier MH, Henderson AJ, Lee YA. Meta-analysis identifies seven susceptibility loci involved in the atopic march. *Nat Commun* 2015; 6: 8804.

36. Nakamura T, Haider S, Colicino S, Murray CS, Holloway J, Simpson A, Cullinan P, Custovic A, investigators S. Different definitions of atopic dermatitis: impact on prevalence estimates and associated risk factors. *Br J Dermatol* 2019; 181: 1272-1279.

37. Frainay C, Pitarch Y, Filippi S, Evangelou M, Custovic A. Atopic dermatitis or Eczema? Consequences of ambiguity in disease name for biomedical literature mining. *Clin Exp Allergy* 2021.

38. Margolis DJ, Mitra N, Wubbenhorst B, D'Andrea K, Kraya AA, Hoffstad O, Shah S, Nathanson KL. Association of Filaggrin Loss-of-Function Variants With Race in Children With Atopic Dermatitis. *JAMA Dermatol* 2019; 155: 1269-1276.

39. Nicolaou N, Poorafshar M, Murray C, Simpson A, Winell H, Kerry G, Härlin A, Woodcock A, Ahlstedt S, Custovic A. Allergy or tolerance in children sensitized to peanut: prevalence and differentiation using component-resolved diagnostics. *The Journal of allergy and clinical immunology* 2010; 125: 191-197.e191-113.

40. Nicolaou N, Murray C, Belgrave D, Poorafshar M, Simpson A, Custovic A. Quantification of specific IgE to whole peanut extract and peanut components in prediction of peanut allergy. *The Journal of allergy and clinical immunology* 2011; 127: 684-685.

41. Brough HA, Simpson A, Makinson K, Hankinson J, Brown S, Douiri A, Belgrave DC, Penagos M, Stephens AC, McLean WH, Turcanu V, Nicolaou N, Custovic A, Lack G. Peanut allergy: effect of environmental peanut exposure in children with filaggrin loss-of-function mutations. *The Journal of allergy and clinical immunology* 2014; 134: 867-875 e861.

42. Nakamura T, Haider S, Fontanella S, Murray CS, Simpson A, Custovic A. Modelling trajectories of parentally reported and physician-confirmed atopic dermatitis in a birth cohort study. *Br J Dermatol* 2022; 186: 274-284.

43. Kotsapas C, Nicolaou N, Haider S, Kerry G, Turner PJ, Murray CS, Simpson A, Custovic A. Early-life predictors and risk factors of peanut allergy, and its association with asthma in later-life: Population-based birth cohort study. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology* 2022.

44. Lawlor DA, Tilling K, Davey Smith G. Triangulation in aetiological epidemiology. *Int J Epidemiol* 2016; 45: 1866-1886.

45. Mulick AR, Mansfield KE, Silverwood RJ, Budu-Aggrey A, Roberts A, Custovic A, Pearce N, Irvine AD, Smeeth L, Abuabara K, Langan SM. Four childhood atopic dermatitis subtypes identified from trajectory and severity of disease and internally validated in a large UK birth cohort. *Br J Dermatol* 2021; 185: 526-536.

46. Oksel C, Granell R, Haider S, Fontanella S, Simpson A, Turner S, Devereux G, Arshad SH, Murray CS, Roberts G, Holloway JW, Cullinan P, Henderson J, Custovic A, Stelar investigators bTi. Distinguishing Wheezing Phenotypes from Infancy to Adolescence. A Pooled Analysis of Five Birth Cohorts. *Ann Am Thorac Soc* 2019; 16: 868-876.

47. Oksel C, Granell R, Mahmoud O, Custovic A, Henderson AJ, Stelar, Breathing Together i. Causes of variability in latent phenotypes of childhood wheeze. *J Allergy Clin Immunol* 2019; 143: 1783-1790 e1711.

48. Paternoster L, Savenije OEM, Heron J, Evans DM, Vonk JM, Brunekreef B, Wijga AH, Henderson AJ, Koppelman GH, Brown SJ. Identification of atopic dermatitis subgroups in children from 2 longitudinal birth cohorts. *J Allergy Clin Immunol* 2018; 141: 964-971.

49. Oksel C, Granell R, Henderson J, Custovic A. Distinguishing wheezing phenotypes in childhood: a pooled analysis of five birth cohorts. *Eur Respir J* 2018; 52.

50. Deliu M, Belgrave D, Sperrin M, Buchan I, Custovic A. Asthma phenotypes in childhood. *Expert Rev Clin Immunol* 2017; 13: 705-713.

51. Roduit C, Frei R, Depner M, Karvonen AM, Renz H, Braun-Fahrlander C, Schmausser-Hechfellner E, Pekkanen J, Riedler J, Dalphin JC, von Mutius E, Lauener RP, the Psg, Hyvarinen A, Kirjavainen P, Remes S, Roponen M, Dalphin ML, Kaulek V, Ege M, Genuneit J, Illi S, Kabesch M, Schaub B, Pfefferle PI, Doekes G. Phenotypes of Atopic Dermatitis Depending on the Timing of Onset and Progression in Childhood. *JAMA Pediatr* 2017; 171: 655-662.

**11STELAR/UNICORN investigators:**

Professor John Ainsworth, School of Health Sciences, The University of Manchester

Dr Andrew Boyd, University of Bristol

Professor Andrew Bush, NHLI, Imperial College London

Dr Philip Couch, School of Health Sciences, The University of Manchester

Professor Graham Devereux, Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool

Dr Ibrahim Emam, Department of Computing, Imperial College London

Professor Yi-ke Guo, Department of Computing, Imperial College London

Professor Sejal Saglani, NHLI, Imperial College London

Professor Ashley Woodcock, University of Manchester

**LEGENDS FOR FIGURES**

**Figure 1.** Trends in the deviation between observed and expected probabilities for each disease category over time (expressed as per cent point difference). Negative numbers show that observed probabilities were lower than expected probabilities, for example, single diseases were observed less frequently than expected in the population, and Eczema+Wheeze+Rhinitis was observed more than expected.

**Figure 2.** Individual longitudinal sequences of disease development. Each row is coloured by the disease state at each time-point and displays the duration an individual spent in each state. The number of person-unique sequences: 220 SEATON, 259 Ashford, 295 IoW, 351 MAAS.

**Figure 3**. Dynamics of change in eczema, wheeze and rhinitis over time: Latent Markov modelling in the joint cohort model (2079 children with complete observations on eczema, wheeze, and rhinitis at five time-points). Data were harmonised at overlapping time-points to represent five stages of development (infancy: age 1; early childhood: ages 2-3; pre-school: ages 4-5; mid-childhood: ages 8-10; adolescence: 14-18).

a) Predicted latent Markov states from joint modelling of all four cohorts; each row represents the individual-level latent states across time.

b) Alluvial plot to show relative size of transitions between latent states between t and t+1 (based on time-homogeneous transition probabilities displayed in Table 4).

Children from the Eczema (E) state are more likely to persist in the same state. Although relatively small, they are more likely to transition to Multimorbidity (MM) than children from other states. Children in the Wheeze (W) state are more likely to transition to Low risk than to any other state.

c) Dynamics of change in eczema, wheeze and rhinitis over time: Latent Markov modelling: Alluvial plot to show individual-level transitions between predicted latent Markov states at each time point.

**Table 1.** Prevalence of morbidity at each cross-sectional time-point.

Colour gradation tending towards red indicates higher prevalence; green indicates lowest prevalence.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Cohort** | **N** | **Eczema only** | **Wheeze only** | **Rhinitis only**  | **Wheeze + Eczema** | **Wheeze + Rhinitis**  | **Eczema + Rhinitis**  | **Eczema + wheeze + rhinitis**  | **No disease** |
| **MAAS** |  |  |  |  |  |  |  |  |  |
| **1 year** | ***935*** | 225 (24.1%) | 130 (13.9%) | 3 (0.3%) | 90 (9.6%) | 2 (0.2%) | 3 (0.3%) | (0.0%) | 482 (51.6%) |
| **3 years** | ***1049*** | 228 (21.7%) | 111 (10.6%) | 12 (1.1%) | 96 (9.2%) | 10 (1.0%) | 8 (0.8%) | 16 (1.5%) | 568 (54.2%) |
| **5 years** | ***1034*** | 172 (16.6%) | 82 (7.9%) | 111 (10.7%) | 40 (3.9%) | 54 (5.2%) | 69 (6.7%) | 54 (5.2%) | 452 (43.7%) |
| **8 years** | ***1020*** | 125 (12.3%) | 56 (5.5%) | 124 (12.2%) | 32 (3.1%) | 45 (4.4%) | 76 (7.5%) | 51 (5.0%) | 511 (50.1%) |
| **11 years** | ***912*** | 95 (10.4%) | 47 (5.2%) | 155 (17.0%) | 22 (2.4%) | 65 (7.1%) | 57 (6.3%) | 39 (4.3%) | 432 (47.4%) |
| **16 years** | ***734*** | 46 (6.3%) | 29 (4.0%) | 193 (26.3%) | 13 (1.8%) | 51 (7.0%) | 51 (7.0%) | 31 (4.2%) | 320 (43.6%) |
| **Ashford** |  |  |  |  |  |  |  |  |  |
| **1 year** | **454** | 22 (4.9%) | 141 (31.1%) | 10 (2.2%) | 24 (5.3%) | 14 (3.1%) | 1 (0.2%) | 4 (0.9%) | 238 (52.4%) |
| **2 years** | **615** | 53 (8.6%) | 129 (21.0%) | 22 (3.6%) | 26 (4.2%) | 23 (3.7%) | 7 (1.1%) | 10 (1.6%) | 345 (56.1%) |
| **3 years** | **615** | 62 (10.1%) | 99 (16.1%) | 33 (5.4%) | 36 (5.9%) | 23 (3.7%) | 10 (1.6%) | 13 (2.1%) | 339 (55.1%) |
| **4 years** | **611** | 54 (8.8%) | 73 (12.0%) | 36 (5.9%) | 18 (3.0%) | 28 (5.6%) | 10 (1.6%) | 13 (2.1%) | 379 (61.0%) |
| **5 years** | **604** | 47 (7.8%) | 48 (8.0%) | 51 (8.4%) | 13 (2.2%) | 27 (4.5%) | 9 (1.5%) | 20 (3.3%) | 389 (64.4%) |
| **8 years** | **593** | 40 (6.8%) | 30 (5.1%) | 74 (12.5%) | 12 (2.0%) | 26 (4.4%) | 17 (2.9%) | 11 (1.9%) | 383 (64.6%) |
| **14 years** | **499** | 20 (4.0%) | 21 (4.2%) | 110 (22.0%) | 3 (0.6%) | 33 (6.6%) | 28 (5.6%) | 15 (3.0%) | 269 (53.9%) |
| **IOW** |  |  |  |  |  |  |  |  |  |
| **1 year** | **1247** | 87 (7.0%) | 38 (3.1%) | 64 (5.1%) | 21 (1.7%) | 52 (4.2%) | 21 (1.7%) | 18 (1.4%) | 946 (75.9%) |
| **2 years**  | **1157** | 139 (12.0%) | 66 (5.7%) | 34 (2.9%) | 29 (2.5%) | 68 (5.9%) | 34 (2.9%) | 18 (1.6%) | 769 (66.5%) |
| **4 years** | **1157** | 151 (13.1%) | 90 (7.8%) | 73 (6.3%) | 40 (3.5%) | 40 (3.5%) | 34 (2.9%) | 32 (2.8%) | 697 (60.2%) |
| **10 years** | **1347** | 88 (6.5%) | 121 (9.0%) | 173 (12.8%) | 17 (1.3%) | 76 (5.6%) | 20 (1.5%) | 38 (2.8%) | 814 (60.4%) |
| **18 years** | **1080** | 37 (3.4%) | 78 (7.2%) | 211 (19.5%) | 6 (0.6%) | 108 (10.0%) | 24 (2.2%) | 26 (2.4%) | 590 (54.6%) |
| **26 years** | **1028** | 36 (3.5%) | 76 (7.4%) | 253 (24.6%) | 9 (0.9%) | 123 (12.0%) | 29 (2.8%) | 29 (2.8%) | 473 (46.0%) |
| **SEATON** | **N** |  |  |  |  |  |  |  |  |
| **6m** | **1585** | 151 (9.5%) | 188 (11.9%) | 171 (10.8%) | 30 (1.9%) | 64 (4.0%) | 12 (0.8%) | 24 (1.5%) | 945 (59.6%) |
| **1 year** | **1507** | 128 (8.5%) | 132 (8.8%) | 110 (7.3%) | 36 (2.4%) | 54 (3.6%) | 27 (1.8%) | 10 (0.7%) | 1010 (67.0%) |
| **2 years** | **1372** | 176 (12.8%) | 108 (7.9%) | 76 (5.5%) | 34 (2.5%) | 48 (3.5%) | 25 (1.8%) | 19 (1.4%) | 886 (64.6%) |
| **5 years** | **1175** | 174 (14.8%) | 79 (6.7%) | 16 (1.4%) | 48 (4.1%) | 11 (0.9%) | 8 (0.7%) | 17 (1.5%) | 822 (70.0%) |
| **10 years** | **883** | 53 (6.0%) | 36 (4.1%) | 128 (14.5%) | 5 (0.6%) | 39 (4.4%) | 40 (4.5%) | 26 (2.9%) | 556 (63.0%) |
| **15 years** | **703** | 48 (6.8%) | 19 (2.7%) | 163 (23.2%) | 8 (1.1%) | 35 (5.0%) | 42 (6.0%) | 16 (2.3%) | 372 (52.9%) |

**Table 2.** The association between Eczema only, Wheeze only, Eczema+wheeze in first three years as index diseases with subsequent persistence or development of different patterns of eczema, wheeze and rhinitis at pre-school, mid-school age and adolescence. Results are derived from jointly modelling the cohorts by harmonising time points (early life – age 0-3 years; pre-school - age 4-5; mid-childhood - age 8-10; adolescence – age 14-18). The model was adjusted by including a predictor for cohort to control for inter-cohort differences. Sex, FLG and rs7216389 were included as covariates. Results are presented as adjusted RRRs with 95% confidence intervals. ‘No disease’ is the reference category. E=eczema; W=wheeze; R=rhinitis**.**



**Table 3.** Estimated conditional responses and transition probabilities between latent states from latent Markov model with 5 optimal states and assuming time-homogeneous transitions. The transition matrix shows the probability of transitioning between latent state between time *t* to *t+1* assuming time-homogenous probabilities. Colour gradation tending towards red indicates highest probabilities, and green indicates lowest probabilities in the overall table.

|  |  |  |
| --- | --- | --- |
|   |   | **Conditional response probabilities of observed symptoms for each latent state** |
|   |   | **Low Risk** | **Eczema** | **Wheeze** | **Rhinitis** | **Multimorbidity** |
| *Observed symptoms*  | **Eczema** | 0.05 | 0.857 | 0.061 | 0.061 | 0.459 |
| **Wheeze** | 0.013 | 0.176 | 0.734 | 0.123 | 0.705 |
| **Rhinitis** | 0.001 | 0.138 | 0.092 | 0.562 | 0.845 |
|   |   | **Initial probabilities of starting in each latent state** |
|  |  | **Low Risk** | **Eczema** | **Wheeze** | **Rhinitis** | **Multimorbidity** |
|   |   | 0.627 | 0.166 | 0.154 | 0.031 | 0.022 |
|  |  | **Matrix of transition probabilities** |
|  |  | ***t+1*** |
|  |  | **Low Risk** | **Eczema** | **Wheeze** | **Rhinitis** | **Multimorbidity** |
| ***t*** | **Low Risk** | 0.798 | 0.028 | 0.031 | 0.139 | 0.003 |
| **Eczema** | 0.116 | 0.619 | 0.011 | 0.048 | 0.207 |
| **Wheeze** | 0.278 | 0.021 | 0.591 | 0.083 | 0.028 |
| **Rhinitis** | 0.086 | 0.006 | 0.022 | 0.873 | 0.014 |
| **Multimorbidity** | 0.044 | 0.054 | 0.062 | 0.064 | 0.777 |

**Table 4:** Multinomial regression analyses to investigate genetic associations with Multimorbidity state persistence.For rs7216389, an additive (dosage) model was used, where the number of risk alleles was treated as a continuous variable in the regression analysis, where 0=CC, 1=CT, 2=TT. Outcome is 0: No Multimorbidity (MM), 1: MM at 1 time-point (TP), 2:MM at 2-5 time-points (TP). No multimorbidity is the omitted category. Results are expressed as relative risk ratios (RRR) with 95% CI.

|  |  |  |
| --- | --- | --- |
|  | Model 1 (n=1463) | Model 2 (n=1463) |
|   | MM at 1 TP (n=84)  | MM at 2-5 TPs (n=205) | MM at 1 TP (n=84)  | MM at 2-5 TPs (n=205) |
|   | RRR/95% CI | p-value | RRR/95% CI | p-value | RRR/95% CI | p-value | RRR/95% CI | p-value |
| Wheeze state in early-life  | 3.00 | **0.008** | 1.00 | 1.000 | 3.00 | **0.008** | 1.00 | 1.000 |
|   | [1.33,6.76] |  | [0.51,1.95] |   | [1.33,6.77] |  | [0.51,1.95] |   |
| Eczema state in early-life  | 39.65 | **<0.001** | 19.72 | **<0.001** | 39.60 | **<0.001** | 19.73 | **<0.001** |
|   | [20.58,76.39] |  | [12.64,30.77] |   | [20.54,76.37] |  | [12.64,30.79] |   |
| *Filaggrin* loss-of-function mutation  | 0.88 | 0.771 | 1.75 | **0.032** | 0.40 | 0.298 | 1.53 | 0.399 |
|   | [0.37,2.10] |  | [1.05,2.92] |   | [0.07,2.23] |  | [0.57,4.15] |   |
| rs7216389  | 1.03 | 0.881 | 1.49 | **0.003** | 0.95 | 0.814 | 1.48 | **0.007** |
|   | [0.71,1.49] |  | [1.15,1.94] |   | [0.64,1.42] |  | [1.11,1.96] |   |
| *Filaggrin*\*rs7216389 |   |  |  |   | 2.01 | 0.260 | 1.13 | 0.743 |
|   |   |  |  |   | [0.60,6.80] |  | [0.54,2.39] |   |
| Male  | 0.96 | 0.892 | 1.06 | 0.753 | 0.96 | 0.882 | 1.06 | 0.753 |
|   | [0.57,1.63] |  | [0.73,1.53] |   | [0.57,1.62] |  | [0.73,1.53] |   |

**Figure 1.** Trends in the deviation between observed and expected probabilities for each disease category over time (expressed as per cent point difference). Negative numbers show that observed probabilities were lower than expected probabilities, for example, single diseases were observed less frequently than expected in the population, and Eczema+Wheeze+Rhinitis was observed more than expected.



**Figure 2.** Individual longitudinal sequences of disease development over time. Each row, which represents a participant, is colored by the disease state at each time-point, and displays the number of time-points spent in each state. The number of person-unique sequences: 220 SEATON, 259 Ashford, 295 IoW, 351 MAAS.



**Figure 3**. Dynamics of change in eczema, wheeze and rhinitis over time: Latent Markov modelling

a) Predicted latent Markov states from joint modelling of all four cohorts; each row represents the individual-level latent states across time. Data were harmonised at overlapping time-points to represent five stages of development (infancy: age 1; early childhood: ages 2-3; pre-school: ages 4-5; mid-childhood: ages 8-10; adolescence: 14-18). The sample comprised 2079 children with complete observations on eczema, wheeze, and rhinitis at five time-points.

**a)**

****

**b)** Alluvial plot to show relative size of transitions between latent states between t and t+1 (based on time-homogeneous transition probabilities displayed in Table 4).

Children from the Eczema (E) state are more likely to persist in the same state. Although relatively small, they are more likely to transition to Multimorbidity (MM) than children from other states. Children in the Wheeze (W) state are more likely to transition to Low risk than to any other state.



**c)** Dynamics of change in eczema, wheeze and rhinitis over time: Latent Markov modelling: Alluvial plot to show individual-level transitions between predicted latent Markov states at each time point

