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

Sadia Haider[1], Sara Fontanella[1], Anhar Ullah[1], Stephen Turner[2,3], Angela Simpson[4], Graham Roberts[5,6,7], Clare S Murray[4], John W. Holloway[5,6], John A Curtin[4], Paul Cullinan[1], Syed Hasan Arshad[6,7,8], Guillem Hurault[9], Raquel Granell[10], and Adnan Custovic[1]

on behalf of STELAR/UNICORN[11] investigators

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

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

[3]Child Health, University of Aberdeen, Aberdeen, UK

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

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

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

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

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

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

[10]MRC 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: 3496

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 (1, 2). 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 (3, 4). 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 (5). 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 (6, 7). A specific sequence is implicit by the use of the term march (2). This framework is extended to the recommendation that primary care physicians "should inform parents that children with eczema may later develop asthma" (8), and has underpinned clinical trials specifically aiming to prevent wheezing/asthma in children with early-life eczema (9, 10). However, some studies have shown a substantial heterogeneity between patients in the chronology of symptom development (11-13), questioning a specific sequence of atopic march (14). 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 (15), and distinct genetic associates (16). Thus, the evidence to date is convincing that atopic diseases coexist (1, 17-19), and although there is increasing acknowledgement of different trajectories (19, 20), a comprehensive analysis of their long-term evolution within individuals is lacking, and the mechanisms of their coexistence remain unclear (5).

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 (7)), or in a multimorbidity framework, in which no single condition holds priority over any of the co-occurring conditions (21), via a common underlying pathogenic mechanism (e.g. impaired skin barrier leading to allergic sensitisation (22)). 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 (23), Isle of Wight (IOW) (24), Manchester Asthma and Allergy Study (MAAS) (25) and Aberdeen cohort (SEATON) (26). 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 (27).

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 (28). 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 (29, 30). 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* (31) and *TraMineR* (32) 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 E+W 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. 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 then 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 (33).

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 (16, 34). 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 (35, 36). 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 (10). 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 (37). Our results are therefore not directly transferable to all ethnic groups.

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 (38-40), and have shown that the risk is markedly higher amongst children with persistent eczema (41), and those with co-morbid persistent eczema and wheeze, but not with transient phenotypes (42). 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 (43).

Rather than applying latent class (LC) models, which have been extensively used to study wheeze and eczema (44-50), 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. Moreno MA. JAMA Pediatrics Patient Page. Atopic Diseases in Children. *JAMA Pediatr* 2016; 170: 96.
2. 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.
3. Custovic A, Henderson J, Simpson A. Does understanding endotypes translate to better asthma management options for all? *The Journal of allergy and clinical immunology* 2019; 144: 25-33.
4. Akar-Ghibril N, Casale T, Custovic A, Phipatanakul W. Allergic Endotypes and Phenotypes of Asthma. *J Allergy Clin Immunol Pract* 2020; 8: 429-440.
5. Silverberg JI. Comorbidities and the impact of atopic dermatitis. *Ann Allergy Asthma Immunol* 2019; 123: 144-151.
6. Spergel JM. From atopic dermatitis to asthma: the atopic march. *Ann Allergy Asthma Immunol* 2010; 105: 99-106; quiz 107-109, 117.
7. Bantz SK, Zhu Z, Zheng T. The Atopic March: Progression from Atopic Dermatitis to Allergic Rhinitis and Asthma. *J Clin Cell Immunol* 2014; 5.
8. 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.
9. 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.
10. 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.
11. Williams H, Flohr C. How epidemiology has challenged 3 prevailing concepts about atopic dermatitis. *J Allergy Clin Immunol* 2006; 118: 209-213.
12. Barberio G, Pajno GB, Vita D, Caminiti L, Canonica GW, Passalacqua G. Does a 'reverse' atopic march exist? *Allergy* 2008; 63: 1630-1632.
13. 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.
14. 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. *The Journal of allergy and clinical immunology* 2012; 130: 1117-1122.
15. 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 Med* 2014; 11: e1001748.
16. 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.
17. 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.

18. 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.
19. Spergel JM. The atopic march: Where we are going? Can we change it? *Ann Allergy Asthma Immunol* 2021; 127: 283-284.
20. 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.
21. 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.
22. 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.
23. 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.
24. 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.
25. 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.
26. 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.
27. 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.
28. 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.
29. Bartolucci F, Farcomeni A, Pennoni F. Latent Markov models for longitudinal data. Boca Raton, FL: CRC Press, Taylor & Francis Group; 2013.
30. 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.
31. Bartolucci F, Pandolfi S, Pennoni F. LMest: An R Package for Latent Markov Models for Longitudinal Categorical Data. *2017* 2017; 81: 38.
32. Gabadinho A, Ritschard G, Müller NS, Studer M. Analyzing and Visualizing State Sequences in R with TraMineR. *2011* 2011; 40: 37.
33. 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.
34. 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.

35. 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.
36. 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.
37. 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.
38. 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.
39. 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.
40. 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.
41. 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.
42. 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.
43. Lawlor DA, Tilling K, Davey Smith G. Triangulation in aetiological epidemiology. *Int J Epidemiol* 2016; 45: 1866-1886.
44. 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.
45. 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.
46. 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.

47. 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.
48. 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.
49. Deliu M, Belgrave D, Sperrin M, Buchan I, Custovic A. Asthma phenotypes in childhood. *Expert Rev Clin Immunol* 2017; 13: 705-713.
50. 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.

[11]STELAR/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. Index plots of individual longitudinal sequences of disease development. Each row is coloured by the disease state at each time-point and displays the duration 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**.**

	E		**W**		**R**		**E+W**		**W+R**		**E+R**		**E+W+R**	
	RRR/95% CI	p	RRR/95% CI	p	RRR/95% CI	p	RRR/95% CI	p	RRR/95% CI	p	RRR/95% CI	p	RRR/95% CI	p
Predictors	**Outcomes at pre-school N=2314**													
Early eczema only	8.32	**<0.001**	1.21	0.550	0.58	0.106	3.04	**0.002**	0.13	**0.044**	6.98	**<0.001**	4.64	**<0.001**
	[6.20,11.17]		[0.65,2.23]		[0.30,1.12]		[1.52,6.11]		[0.02,0.95]		[4.10,11.88]		[2.25,9.58]	
Early wheeze only	0.51	0.094	6.03	**<0.001**	1.05	0.882	2.67	**0.018**	2.96	**<0.001**	0.95	0.933	2.01	0.211
	[0.23,1.12]		[4.04,8.99]		[0.57,1.91]		[1.18,6.00]		[1.66,5.28]		[0.28,3.18]		[0.67,6.03]	
Early eczema & wheeze	6.62	**<0.001**	8.20	**<0.001**	1.29	0.653	37.20	**<0.001**	6.41	**<0.001**	8.07	**<0.001**	58.65	**<0.001**
	[3.38,12.97]		[3.94,17.06]		[0.43,3.90]		[17.92,77.25]		[2.65,15.49]		[2.96,22.01]		[27.39,125.62]	
Filaggrin loss-of-function mutation	1.11	0.625	0.67	0.241	1.47	0.164	1.99	**0.039**	0.84	0.690	1.49	0.281	2.53	**0.006**
	[0.73,1.70]		[0.34,1.31]		[0.86,2.51]		[1.04,3.81]		[0.35,1.99]		[0.72,3.07]		[1.30,4.94]	
rs7216389	1.11	0.276	1.12	0.346	0.97	0.805	1.75	**0.002**	1.31	0.099	0.98	0.904	1.69	**0.007**
	[0.92,1.33]		[0.88,1.42]		[0.76,1.23]		[1.23,2.49]		[0.95,1.80]		[0.70,1.37]		[1.15,2.47]	
Sex (male)	0.81	0.116	1.36	0.066	1.53	**0.014**	1.44	0.143	1.43	0.119	1.05	0.845	1.37	0.251
	[0.63,1.05]		[0.98,1.89]		[1.09,2.14]		[0.88,2.35]		[0.91,2.24]		[0.65,1.70]		[0.80,2.34]	
	Outcomes at mid-childhood N=2409													
Early eczema only	3.86	**<0.001**	1.76	**0.038**	1.19	0.367	3.41	**0.007**	1.38	0.265	5.65	**<0.001**	5.49	**<0.001**
	[2.73,5.47]		[1.03,3.01]		[0.82,1.72]		[1.41,8.25]		[0.78,2.43]		[3.58,8.92]		[3.11,9.69]	
Early wheeze only	0.82	0.571	4.79	**<0.001**	0.88	0.613	1.14	0.868	1.79	0.055	1.71	0.161	1.71	0.282
	[0.42,1.62]		[3.04,7.54]		[0.55,1.42]		[0.25,5.08]		[0.99,3.24]		[0.81,3.61]		[0.64,4.54]	
Early eczema & wheeze	4.35	**<0.001**	6.88	**<0.001**	1.90	0.085	40.10	**<0.001**	6.01	**<0.001**	6.00	**<0.001**	24.82	**<0.001**
	[2.14,8.81]		[3.21,14.72]		[0.92,3.93]		[17.52,91.80]		[2.81,12.83]		[2.53,14.22]		[12.01,51.32]	
Filaggrin loss-of-function mutation	1.22	0.426	1.07	0.832	1.10	0.666	1.41	0.472	1.50	0.163	1.28	0.452	3.09	**<0.001**
	[0.75,2.01]		[0.58,1.97]		[0.72,1.68]		[0.55,3.57]		[0.85,2.64]		[0.67,2.46]		[1.74,5.48]	
rs7216389	1.00	0.983	1.20	0.147	1.24	**0.017**	1.65	**0.030**	1.43	**0.008**	0.89	0.439	1.41	**0.041**
	[0.81,1.24]		[0.94,1.54]		[1.04,1.47]		[1.05,2.60]		[1.10,1.86]		[0.67,1.19]		[1.01,1.97]	
Sex (male)	0.72	**0.032**	1.40	0.060	1.23	0.100	0.95	0.866	1.40	0.078	0.78	0.240	0.75	0.243
	[0.53,0.97]		[0.99,2.00]		[0.96,1.57]		[0.50,1.80]		[0.96,2.03]		[0.52,1.18]		[0.47,1.21]	
	Outcomes at adolescence N=1978													
Early eczema only	2.22	**0.003**	1.29	0.485	1.28	0.136	9.54	**0.001**	0.93	0.821	6.95	**<0.001**	3.43	**<0.001**
	[1.31,3.77]		[0.63,2.62]		[0.92,1.78]		[2.60,34.96]		[0.51,1.70]		[4.36,11.09]		[1.78,6.62]	
Early wheeze only	1.16	0.714	4.55	**<0.001**	0.90	0.605	12.48	**<0.001**	1.56	0.127	0.88	0.798	1.32	0.616
	[0.53,2.52]		[2.61,7.94]		[0.59,1.36]		[3.22,48.42]		[0.88,2.76]		[0.34,2.30]		[0.45,3.89]	
Early eczema & wheeze	3.98	**0.003**	6.58	**<0.001**	1.53	0.209	58.80	**<0.001**	3.78	**0.001**	2.57	0.101	17.63	**<0.001**
	[1.59,9.94]		[2.71,16.02]		[0.79,2.95]		[14.47,239.01]		[1.70,8.39]		[0.83,7.92]		[7.91,39.30]	
Filaggrin loss-of-function mutation	0.85	0.686	1.03	0.942	1.28	0.207	0.49	0.493	1.24	0.476	2.10	**0.015**	2.31	**0.019**
	[0.38,1.90]		[0.48,2.23]		[0.87,1.87]		[0.06,3.81]		[0.69,2.23]		[1.15,3.81]		[1.15,4.66]	
rs7216389	0.88	0.396	1.22	0.196	0.96	0.598	1.54	0.177	1.32	**0.028**	1.20	0.246	1.66	**0.008**
	[0.65,1.18]		[0.90,1.66]		[0.82,1.12]		[0.82,2.88]		[1.03,1.70]		[0.88,1.62]		[1.14,2.42]	
Sex (male)	0.49	**0.001**	0.89	0.596	1.24	0.057	0.45	0.098	1.05	0.780	1.14	0.548	0.48	**0.010**
	[0.32,0.76]		[0.58,1.36]		[0.99,1.55]		[0.18,1.16]		[0.74,1.49]		[0.74,1.75]		[0.28,0.84]	

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				
Observed symptoms		**Low Risk**	**Eczema**	**Wheeze**	**Rhinitis**	**Multimorbidity**
	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 [1.33,6.76]	**0.008**	1.00 [0.51,1.95]	1.000	3.00 [1.33,6.77]	**0.008**	1.00 [0.51,1.95]	1.000
Eczema state in early-life	39.65 [20.58,76.39]	**<0.001**	19.72 [12.64,30.77]	**<0.001**	39.60 [20.54,76.37]	**<0.001**	19.73 [12.64,30.79]	**<0.001**
Filaggrin loss-of-function mutation	0.88 [0.37,2.10]	0.771	1.75 [1.05,2.92]	**0.032**	0.40 [0.07,2.23]	0.298	1.53 [0.57,4.15]	0.399
rs7216389	1.03 [0.71,1.49]	0.881	1.49 [1.15,1.94]	**0.003**	0.95 [0.64,1.42]	0.814	1.48 [1.11,1.96]	**0.007**
*Filaggrin**rs7216389					2.01 [0.60,6.80]	0.260	1.13 [0.54,2.39]	0.743
Male	0.96 [0.57,1.63]	0.892	1.06 [0.73,1.53]	0.753	0.96 [0.57,1.62]	0.882	1.06 [0.73,1.53]	0.753

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.

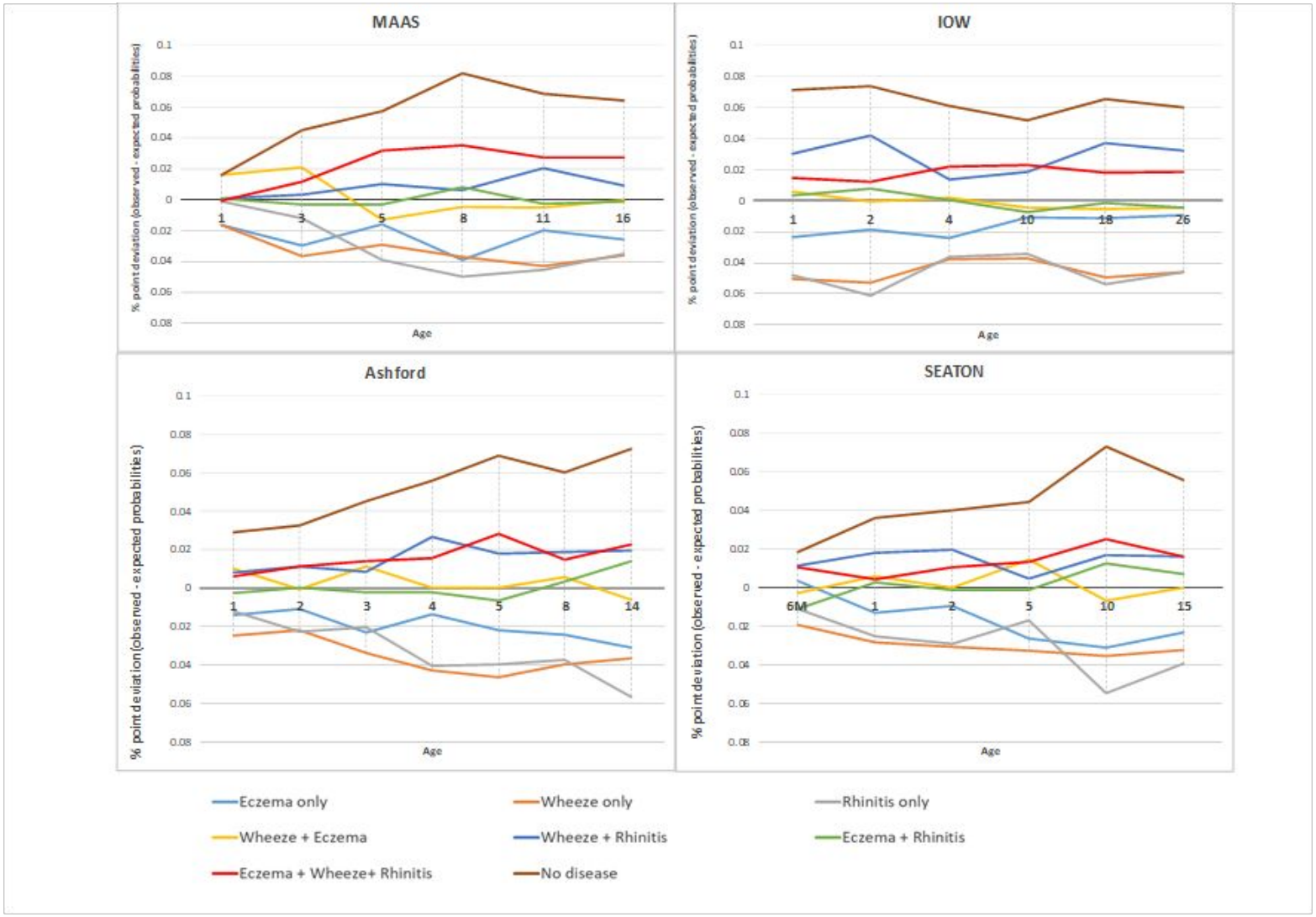


Figure 2. Index plots of individual longitudinal sequences of disease development. Each row is coloured by the disease state at each time-point and displays the duration spent in each state. The number of person-unique sequences: 220 SEATON, 259 Ashford, 295 IoW, 351 MAAS)

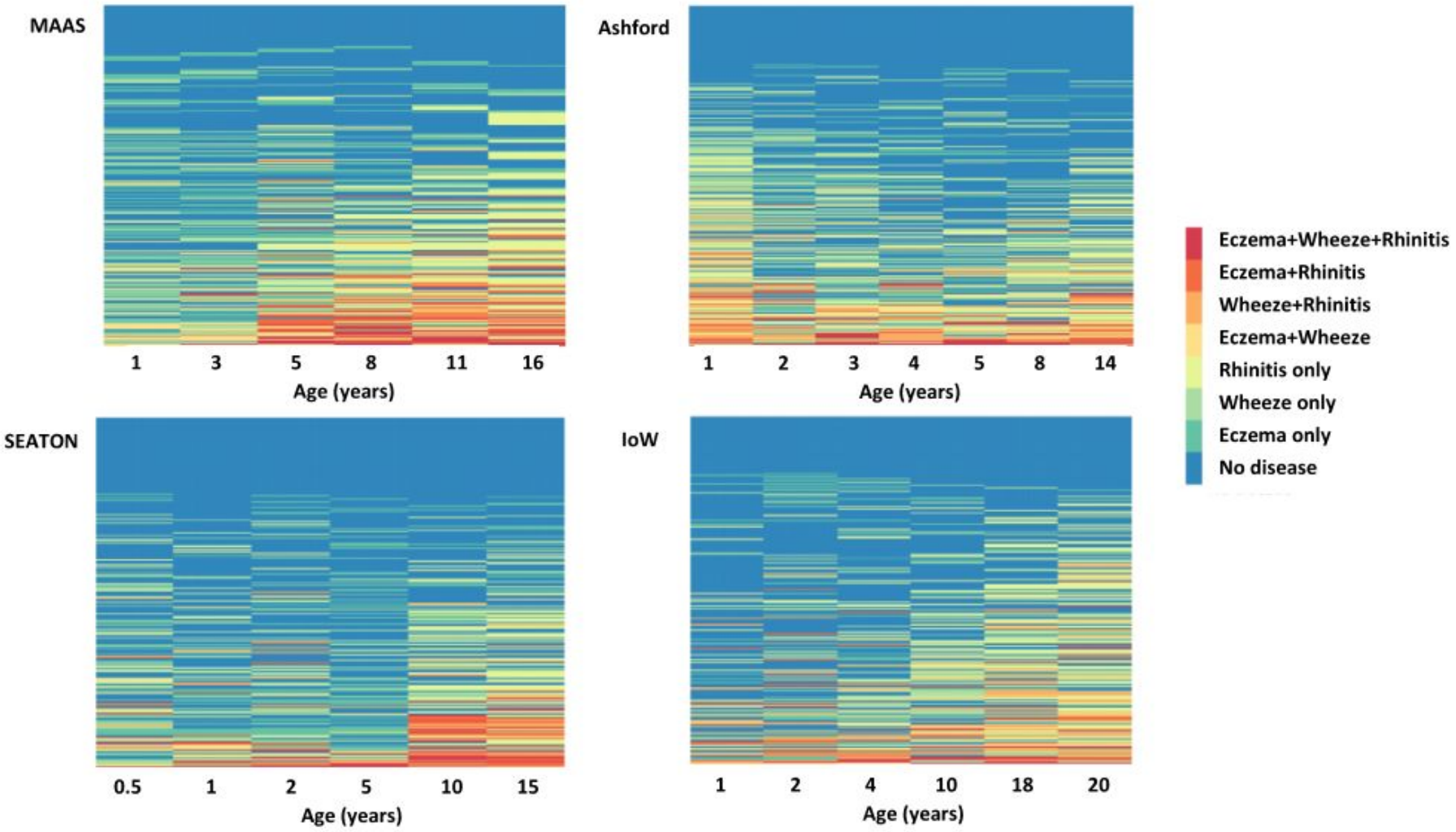
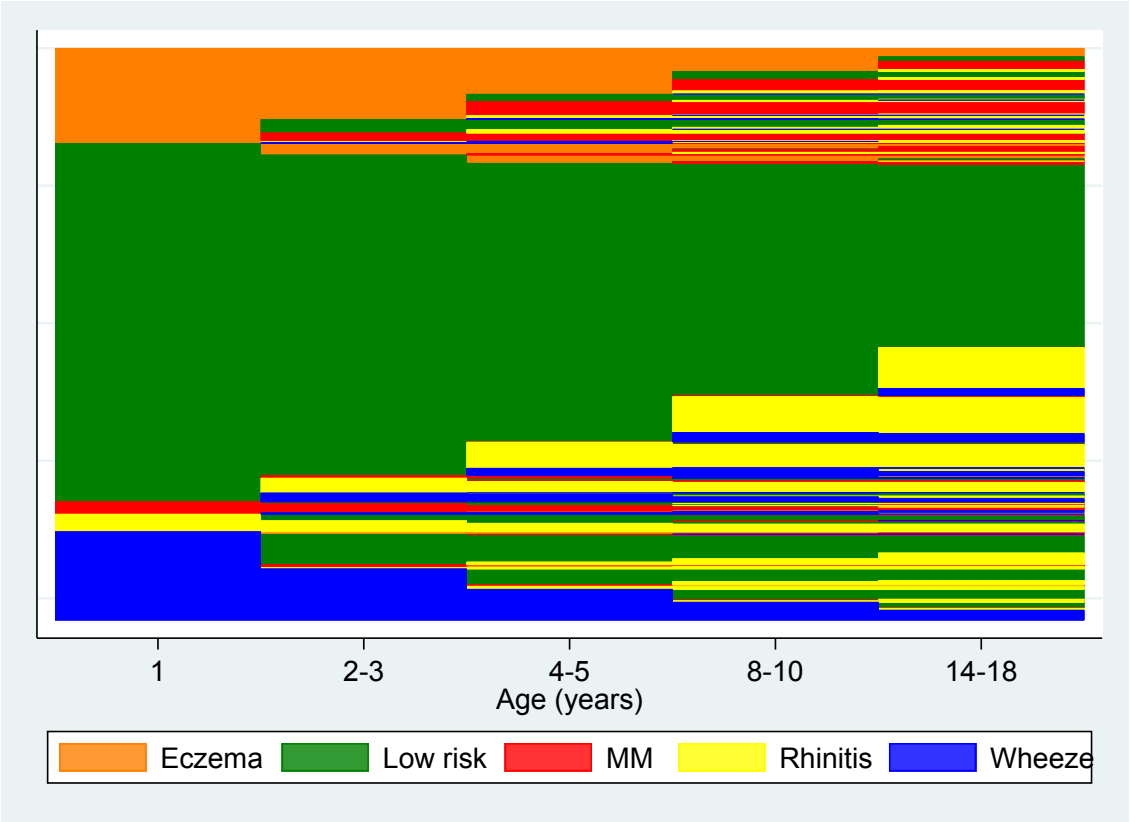


Figure 3. Dynamics of change in eczema, wheeze and rhinitis over time: Latent Markov modelling

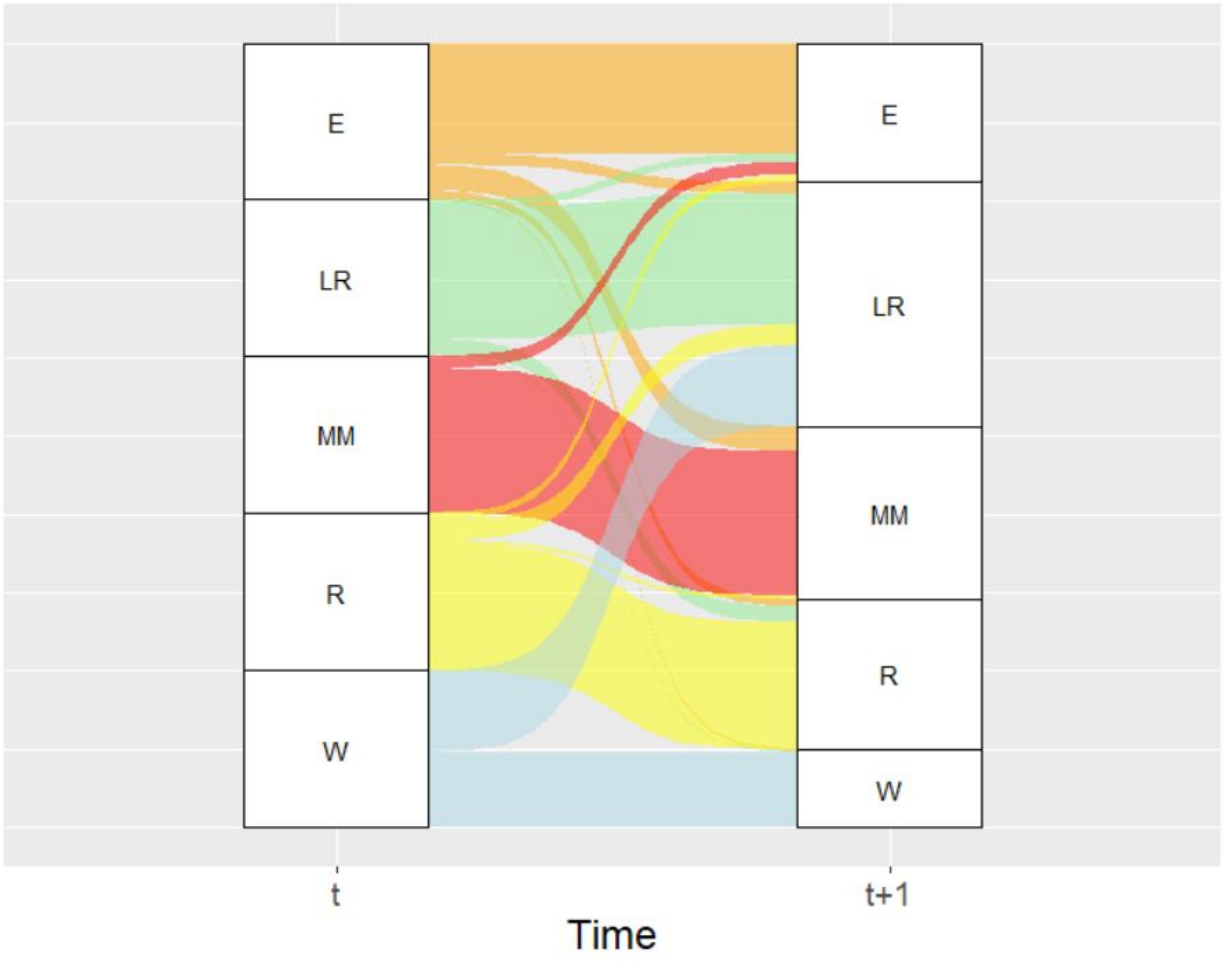
a) Predicted latent Markov states from joint modelling of all four cohorts; each row represents the individual-level latent states across time. Data were harmonised at overlapping time-points to represent five stages of development (infancy: age 1; early childhood: ages 2-3; pre-school: ages 4-5; mid-childhood: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

a)

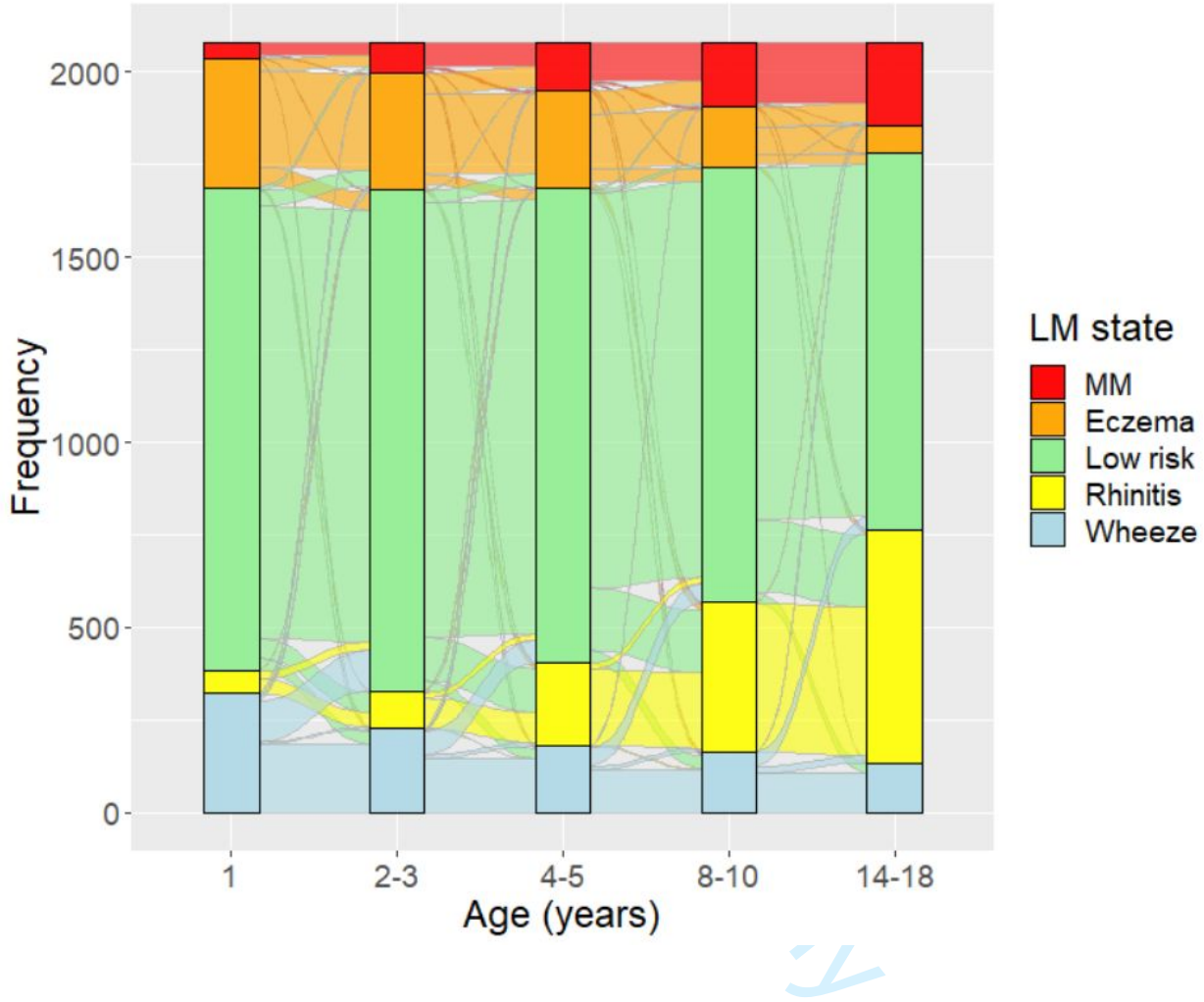


b) Alluvial plot to show relative size of transitions between latent states between t and t+1 (based on time-homogeneous transition probabilities displayed in Table 4).

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



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

Sadia Haider[1], Sara Fontanella[1], Anhar Ullah[1], Stephen Turner[2,3], Angela Simpson[4], Graham Roberts[5,6,7], Clare S Murray[4], John W. Holloway[5,6], John A Curtin[4], Paul Cullinan[1], Syed Hasan Arshad[6,7,8], Guillem Hurault[9], Raquel Granell[10], and Adnan Custovic[1]

on behalf of STELAR/UNICORN[11] investigators

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

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

[3]Child Health, University of Aberdeen, Aberdeen, UK

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

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

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

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

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

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

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

SUPPLEMENTARY APPENDIX

SUPPLEMENTARY METHODS

Data sources: Description of cohorts

MAAS

MAAS is an unselected birth cohort study established in 1995 in Manchester, UK.[1] It consists of a mixed urban-rural population within 50 square miles of South Manchester and Cheshire, United Kingdom located within the maternity catchment area of Wythenshawe and Stepping Hill Hospitals. All pregnant women were screened for eligibility at antenatal visits (8-10th week of pregnancy). Of the 1499 couples who met the inclusion criteria (≤10 weeks of pregnancy, maternal age ≥18 years, and questionnaire and skin prick data test available for both parents), 288 declined to take part in the study and 27 were lost to follow-up between recruitment and the birth of a child. A total of 1184 children were born into the study between February 1996 and April 1998. They were followed prospectively for 19 years to date and attended follow-up clinics for assessments, which included lung function measurements, skin prick testing, biological samples (serum, plasma and urine), and questionnaire data collection. The study was approved by the North West – Greater Manchester East Research Ethics Committee.

ASHFORD

The Ashford study is an unselected birth cohort study established in 1991 in Ashford, UK.[2] It included 642 children born between 1992 and 1993. Participants were recruited prenatally and followed to age 14 years. Detailed standardised questionnaires were administered at each follow-up to collect information on the natural history of asthma and other allergic diseases. Lung function measurements and SPT was carried out at 5, 8 and 14 years of age. In 2015, the study children aged 20 were sent a self-completion questionnaire, which was returned by 60% of the participants.

IOW

The Isle of Wight (IOW) is an unselected birth cohort study established in 1989 on the Isle of Wight.[3,4 3-5] After the exclusion of adoptions, perinatal deaths, and refusal for follow-up, written informed consent was obtained from parents to enrol 1,456 newborns born between 1st January 1989 and 28th February 1990. Follow-up-up assessments were conducted to 26 years of age to prospectively study the development of asthma and allergic diseases. At each follow-up, validated questionnaires were completed by the parents. Additionally, the Skin Prick Test (SPT) was performed on 980, 1036 and 853 participants at 4, 10 and 18 years of age to check allergic reactions to common allergens. At 10, 18, and 26 years, spirometry and methacholine challenge tests were performed to diagnose lung problems. Ethics approvals were obtained from the Isle of Wight Local Research Ethics Committee (now named the National Research Ethics Service, NRES Committee South Central – Southampton B) at recruitment and for the subsequent follow-ups.

SEATON

The Study of Eczema and Asthma to Observe the influence of Nutrition (SEATON) is an unselected birth cohort study established in 1997 in Aberdeen, UK, which was designed to explore the relationship between antenatal dietary exposures and asthma outcomes in childhood. 2000 healthy pregnant women attending an antenatal clinic, at median 12 weeks gestation, were recruited. An interviewer administered a questionnaire to the women and atopic status was ascertained by skin prick test (SPT). The cohort included 1924 children born between April 1998 and December 1999. Participants were recruited prenatally and followed up by self−completion questionnaire to 15 years of age using postal questionnaires to record the presence of asthma and allergic diseases. Lung function measurements and SPT to common allergens was performed at 5, 10 and 15 years. The study was approved by the North of Scotland Research Ethics Committee.

Definition of variables

The definitions of variables were chosen to be as homogeneous as possible across cohorts and time-points. The presence of current eczema, wheeze and rhinitis was defined as a positive response to questions on parentally-reported symptoms in the past twelve months. The precise definitions are presented in Table E1. In general, the following definitions were used:

Current eczema: A positive response to "Has your child had an itchy rash (affecting creases) / eczema in the last 12 months?"

Current wheeze: A positive response to "Has your child had wheezing or whistling in the chest in the last 12 months?"

Current rhinitis: A positive response to "Has your child had a problem with sneezing, or a runny or blocked nose when he/she did not have a cold or the flu in the last 12 months?"

To examine the impact of early-life eczema on the subsequent morbidity risk, we created a variable which indicated whether a child had eczema in the first 3 years of life, and whether she/he subsequently developed wheeze and rhinitis. This produced 4 categories: (1) No disease; (2) Early-life eczema, no wheeze and/or rhinitis after age three; (3) Early-life eczema and concomitant wheeze and/or rhinitis after age three; (4) No early-life eczema, wheeze and/or rhinitis after age 3. The analysis was repeated using early wheeze as an index disease.

Genotyping and imputation

FLG genotyping in all four cohorts was performed using probes and primers as previously described. Genotyping for R501X, S3247X and R2447X loss-of-function mutations were performed using a TaqMan based allelic discrimination assay (Applied Biosystems, Cheshire, UK). Mutation 2282del4 was genotyped by sizing of a fluorescent-labelled PCR fragment on a 3100 or 3730 DNA sequencer. *FLG* mutations 3673delC and 3702delG were assessed by GeneScan analysis of fluorescently labelled polymerase chain reaction products.

Genome-wide genotyping platforms

MAAS: Study participants were genotyped using the Illumina 610 quad genome-wide SNP genotyping platform (Illumina Inc., San Diego, CA, USA). Prior to imputation samples were excluded on the basis of gender mismatches; minimal or excessive heterozygosity, genotyping call rates of < 97%. SNPS were excluded if they had call rates of < 95%, minor allele frequencies of < 0.5% and HWE p<3x10-8. Prior to imputation each chromosome was pre-phased using EAGLE2 (v2.0.5) as recommended by the sanger imputation server[6]. We then imputed with PBWT with the Haplotype Reference Consortium (release 1.1) of 32,470 reference genomes using the Sanger Imputation Server.

IOW, SEATON and ASHFORD: Participants were genotyped using the Illumina Infinium Omni2.5-8 v1.3 BeadChip genotyping platform (Illumina Inc., San Diego, CA, USA). Genotype QC and imputation was carried out as described for MAAS.

Choice of representative SNP in 17q21

The first GWAS of asthma reported in 2007 by Moffatt *et al.* identified multiple markers on chromosome 17q21 as associates of childhood-onset asthma. This remains the best replicated asthma locus to date. We specifically focused on rs7216389, a single nucleotide polymorphism (SNP) in the *ORMDL3* gene at the 17q21 locus, as it was strongly associated with childhood asthma as initially identified in this study. Furthermore, Bisgaard *et al's* longitudinal study of the Danish population found that rs7216389 was an independent risk factor for recurrent wheeze, asthma, asthma exacerbations, and bronchial

hyperresponsiveness from early infancy to school age, however, no increased risk was identified for eczema, rhinitis, or allergic sensitization. We wished to investigate whether the finding of no association between rs7216389 and allergic diseases could be replicated in the STELAR cohorts. An additive (dosage) model was used, where the number of risk alleles was treated as a continuous variable in the regression analysis. Alleles were coded as 0=CC, 1=CT, 2=TT.

Statistical analysis: Latent Markov Modelling

A multivariate latent Markov approach was used to model transitions of eczema, wheeze, and rhinitis longitudinally. Let Y_{ijt} denotes the binary outcome variable for the jth item (eczema, wheeze, rhinitis) recorded at the tth time-point for the ith subject, with $i = 1,\ldots,n$, $j = 1,\ldots,J$ and $t = 1,\ldots,T$. In the LM framework, the health status of the ith subject at t is represented by a discrete latent variable C_{it} with k levels ($k = 1,\ldots,k$). These report how observed eczema, wheeze, and rhinitis relate to the "true" disease states, which are assumed to follow a first-order Markov chain (that is, the latent state occupied in the current period depends on the latent state at the previous time-point). Each of these levels corresponds to a latent state with a specific conditional distribution of the observed variables. The response variables are assumed to be conditionally independent given the latent process.

Three key outputs from this model are the conditional response probabilities of Y_{ijt} given C_{it}, the initial probabilities of the latent process, and a $k \times k$ matrix of transition probabilities, which shows the dependencies between latent states. Each cell denotes the probability of belonging to a state at t given state membership at the previous time-point. Diagonal probabilities show the degree of persisting in the same state (with a probability of 1 showing no movement in to or out of a particular state), whereas the off-diagonals represent the degree of mobility for each pair of latent states. We tested test for stationarity of transition probabilities across time. This means that the transition matrix is constant throughout the evolution of disease development and that transitions among states have the same probability structure across time. The expectation maximization algorithm to estimate relevant parameters. For jointly modelling cohort data, a dummy variable for cohort were entered into the model to control for inter-cohort differences in initial and transition probabilities.

Methodological considerations

We chose LMM because in the situation when data are subject to measurement error, observed transitions are the product of true mobility and spurious change arising from measurement error. The LMM makes it possible to separate those through defining a structural part, which describes the true dynamics among latent states by means of first-order Markov chains, and a measurement part, which relates each latent state to its observed counterpart by a conditional response matrix. As such, the "true" rates of mobility or persistence between latent states can be estimated by controlling for measurement error. Latent class analysis (LCA) has been extensively applied to longitudinal data of single allergic diseases. Unlike LMM, this framework assumes that an individual's membership to a latent class is immutable once assigned (i.e., children cannot transition to different classes). LCA has been useful for characterising the risk of disease within each class.

Association of multimorbidity persistence with allergic sensitisation

We analysed associations of multimorbidity persistence (derived from latent Markov modelling) sensitisation, defined as a positive skin prick test to cat, grass, and house dust mite. Sensitisation was ascertained by skin prick tests (SPT) at pre-school age (4-5 years) and adolescence (14-18 years) for 3 allergens (cat, grass pollen, and house dust-mite). We defined sensitization as a mean wheal diameter 3 mm larger than that elicited by the negative control to at least 1 of the allergens tested. Tests of associations were carried out using and chi-squared test.

Table E1. Definition of variables for each cohort and cross-sectional and longitudinal sample sizes

Ashford							
Age (Years)	**1**	**2**	**3**	**4**	**5**	**8**	**14**
Eczema							
Has your child (daughter) had an itchy skin condition?	**x**	**x**	**x**				
In the past twelve months has your daughter had an itchy skin rash?				**x**	**x**	**x**	**x**
Has it ever affected the skin creases in the past?	**x**	**x**	**x**	**x**	**x**	**x**	**x**
Wheeze							
Has your child has wheezing or whistling in the chest in the last 12 months?	**x**	**x**	**x**	**x**	**x**	**x**	**x**
Which of the following best describes your child's (daughter's) wheeze over the past twelve months? (Response: 0,1-6,7+)	**x**	**x**	**x**	**x**	**x**	**x**	**x**
Rhinitis							
Has she ever suffered from hay fever	**x**	**x**	**x**	**x**			
In the last twelve months has she had a problem with sneezing or a runny or blocked nose?					**x**	**x**	**x**
N with complete data for all symptoms (%)	**454/623 (72.4%)**	**615/617 (99.7%)**	**615/615 (100%)**	**611/611 (100%)**	**604/604 (100%)**	**593/593 (100%)**	**499/499 (100%)**
N individuals with complete longitudinal trajectories = 418; N unique trajectories=259							

IOW						
Age (years)	**1**	**2**	**4**	**10**	**18**	**26**
Eczema						
Parent reported eczema (12 months)	**x**	**x**	**x**			
Itchy rash affected creases in past 12 months (derived from 3 linked questions)				**x**	**x**	**x**
Wheeze						
Has your child had any wheezing episodes?	**x**					
Has your child had any wheezing episodes since review at Year 1?		**x**				

Has your child had asthma/wheezing episodes since review at Year 2?			x			
Has your child had wheezing or whistling in the chest in the last 12 months?				x	x	x
Rhinitis						
Rhinitis nasal symptoms based on clinical examination at follow-ups (blockage, discharge, symptoms)	x	x	x			
In the past 12 months have you had a problem with sneezing, or a runny or blocked nose when you DID NOT have a cold or the flu?				x	x	x
N with complete data for all symptoms (%)	**1247/1369 (91.1%)**	**1157/1231 (94%)**	**1157/1218 (95%)**	**1347/1373 (98.1%)**	**1080/1312 (82.3%)**	**1028/1031 (99.7%)**
N individuals with complete longitudinal trajectories = 519; N unique trajectories=295						

MAAS						
Age (years)	**1**	**3**	**5**	**8**	**11**	**16**
Eczema						
Did your doctor ever tell you that your child had eczema?	x	x				
Has your child ever had an itchy rash which was coming and going for at least six months?			x	x	x	x
Has your child had this itchy rash at any time in the past 12 months?			x	x	x	x
Wheeze						
Has your child had wheezing or whistling in the chest in the last 12 months?	x	x	x	x	x	x
Rhinitis						
Has your doctor ever told you that your child has hay fever or allergic rhinitis?	x	x				
Has your child ever had a problem with sneezing or a runny blocked nose when she did not have a cold or the flu? (yes; no)			x	x	x	x
N with complete data for all symptoms (%)	**935/1093 (85.5%)**	**1049/1110 (94.5%)**	**1034/1059 (97.6%)**	**1020/1029 (99.1%)**	**912/928 (98.3%)**	**734/753 (97.5%)**
N individuals with complete longitudinal trajectories = 553; N unique trajectories=351						

SEATON						
Age (years)	**6m**	**1**	**2**	**5**	**10**	**15**
Eczema						
Has child ever had itchy skin	**x**			**x**		
Has child had itchy skin since 6months old		**x**				
Has child had itchy skin since 1 year			**x**			
Itchy skin in skin creases in last 12 months				**x**		
Has this affected the skin creases	**x**	**x**	**x**			
Child suffered from eczema in last 12 months					**x**	**x**
Wheeze						
Child ever had wheeze	**x**					
Child had wheeze since 6 months		**x**				
Child had wheeze since 1 year			**x**			
Child wheeze last 12 months				**x**	**x**	**x**
Rhinitis						
Has child been sneezy without cold	**x**					
Has child been sneezy without cold since six months		**x**				
Has child been sneezy without cold since 1 year			**x**			
Doctor confirmed hay fever ever				**x**		
Child suffered from hay fever last 12 months					**x**	**x**
N with complete data for all symptoms (%)	**1585/1637 (96.8%)**	**1507/1512 (99.7%)**	**1372/1374 (99.9%)**	**1173/1253 (93.6%)**	**883/934 (94.5%)**	**703/763 (92.1%)**
N individuals with complete longitudinal trajectories = 408; N unique trajectories=220						

Table E2. Definitions of variables (demographic, exposures and outcomes)

Variable	Definition
	Cohort: IOW
Mother–asthma	Do you or have you suffered from asthma or wheezing? (recruitment)
Mother–eczema	Do you or have you suffered from eczema? (recruitment)
Mother–hay fever	Do you or have you suffered from hay fever or allergic rhinitis? (recruitment)
Father–asthma	Does your partner or has your partner suffered from asthma or wheezing? (recruitment)
Father–eczema	Does your partner or has your partner suffered from eczema? (recruitment)
Father–hay fever	Does your partner or has your partner suffered from hay fever or allergic rhinitis? (recruitment)
Mother smoking	Do you smoke in the house? (recruitment)
Father smoking	Does your partner smoke in the house? (recruitment)
Pet	Do you have any pets in the house? (recruitment)
Allergic sensitisation (pre-school & adolescence)	Positive skin prick test to cat, house dust mite or grass at age 4 and age 18
	Cohort: ASHFORD
Mother–asthma	Do you have or have you ever been told you to have asthma? (recruitment)
Mother–eczema	Do you have or have you ever been told you to have eczema? (recruitment)
Mother–hay fever	Do you have or have you ever been told you to have hay fever? (recruitment)
Father–asthma	The father has had or has ever been told to have asthma? (recruitment)
Father–eczema	The father has had or has ever been told to have eczema? (recruitment)
Father–hay fever	The father has had or has ever been told to have hay fever? (recruitment)
Mother smoking	Do you smoke cigarettes? (recruitment)
Father smoking	Does your partner smoke cigarettes? (recruitment)
Pet	Do you have any pets in the home? (recruitment)
Allergic sensitisation (pre-school & adolescence)	Positive skin prick test to cat, house dust mite or grass at age 4 and age 14
	Cohort: MAAS
Mother–asthma	Has a doctor ever told you that you had asthma? (recruitment)
Mother–eczema	Has a doctor ever told you that you had eczema? (recruitment)
Mother–hay fever	Has a doctor ever told you that you had hay-fever? (recruitment)
Father–asthma	Has a doctor ever told you that you had asthma? (recruitment)
Father–eczema	Has a doctor ever told you that you had eczema? (recruitment)
Father–hay fever	Has a doctor ever told you that you had hay-fever? (recruitment)

Mother smoking	Do you smoke– mother (recruitment)
Father smoking	Do you smoke– father (recruitment)
Pet	Do you own a pet? (recruitment)
Allergic sensitisation (pre-school & adolescence)	Positive skin prick test to cat, house dust mite or grass at age 5 and age 16
Cohort: SEATON	
Wheezing	Has your child had wheezing in the chest in the last 12 months? (year 1, 2, 5, 10 and 15)
Mother–asthma	Do you suffer from asthma? (recruitment)
Mother–eczema	Do you suffer from eczema? (recruitment)
Mother–hay fever	Do you suffer from hay fever? (recruitment)
Father–asthma	Does your baby's father suffer from asthma? (recruitment)
Father–eczema	Does your baby's father suffer from eczema? (recruitment)
Father–hay fever	Does your baby's father suffer from hay fever? (recruitment)
Mother smoking	Which of the following best describes your smoking status? (recruitment)
Father smoking	Who smokes in home? (6 months)
Pet	Dog, cat and pet in the home combined (6 months)
Allergic sensitisation (pre-school & adolescence)	Positive skin prick test to cat, house dust mite or grass at age 5 and age 15

SUPPLEMENTARY RESULTS

Characteristics of study populations

Table E3 Descriptive characteristics of the study populations: Included participants are those with complete observations on eczema, wheeze, and rhinitis at all time points. Excluded participants are those with incomplete data. The p-values show whether there was a statistically significant difference in characteristics between included and excluded groups.

MAAS	Whole (N=1136)	Included (N=553)	Excluded (N=583)	P-value	IOW	Whole (N=1488)	Included (N=519)	Excluded (N=969)	P-value
	Mean (SD)					Mean (SD)			
Birth weight, kg	3.46 (0.51)	3.48 (0.50)	3.45 (0.52)	0.416	*Birth weight, kg*	3.41 (0.53)	3.42 (0.52)	3.40 (0.54)	0.467
Maternal age	30.5 (4.8)	30.9 (4.5)	29.9 (4.9)	<0.001	*Maternal age*	26.8 (5.4)	27.7 (5.2)	26.4 (5.4)	<0.001
	Frequency (%)					Frequency (%)			
Sex (Male)	617/1136 (54.3)	275/553 (49.7)	342/583 (58.7)	0.003	*Sex (Male)*	753/1488 (50.6)	236/519 (45.5)	517/969 (53.4)	0.004
Low birth weight (≤2.5 kg)	37/1095 (3.4)	21/548 (3.8)	16/547 (2.9)	0.406	*Low birth weight (≤2.5 kg)*	52/1454 (3.6)	16/506 (3.2)	36/948 (3.8)	0.534
Maternal smoking	118/1024 (11.5)	50/551 (9.1)	68/473 (14.4)	0.008	*Maternal smoking*	377/1468 (25.7)	92/516 (17.8)	285/952 (29.9)	<0.001
Pet ownership	371/1115 (33.3)	189/553 (34.2)	182/562 (32.4)	0.525	*Pet ownership*	833/1474 (56.5)	304/517 (58.8)	529/957 (55.3)	0.193
Maternal eczema	179/1131 (15.8)	92/553 (16.6)	87/578 (15.1)	0.465	*Maternal eczema*	180/1473 (12.2)	51/516 (9.9)	129/957 (13.5)	0.044
Maternal asthma	226/1135 (19.9)	121/553 (21.9)	105/582 (18.0)	0.105	*Maternal asthma*	160/1476 (10.8)	48/517 (9.3)	112/959 (11.7)	0.158
Maternal hay-fever	298/1131 (26.4)	144/553 (26.0)	154/578 (26.6)	0.818	*Maternal hay-fever*	296/1476 (20.1)	107/517 (20.7)	189/959 (19.7)	0.651
Paternal eczema	99/1131 (8.8)	50/553 (9.0)	49/578 (8.5)	0.737	*Paternal eczema*	96/1465 (6.4)	47/515 (9.1)	46/950 (4.8)	0.001
Paternal asthma	155/1133 (13.7)	89/553 (16.1)	66/580 (11.4)	0.021	*Paternal asthma*	147/1466 (10.0)	44/516 (8.5)	103/950 (10.8)	0.159
Paternal hay-fever	263/1131 (23.3)	126/553 (22.8)	137/578 (23.7)	0.715	*Paternal hay-fever*	216/1467 (14.7)	70/515 (13.6)	146/952 (15.3)	0.368
SEATON	**Whole (N=1731)**	**Included (N=408)**	**Excluded (N=1323)**	**P-value**	**Ashford**	**Whole (N=642)**	**Included (N=418)**	**Excluded (N=224)**	**P-value**
	Mean (SD)					Mean (SD)			
Birth weight, kg	3.45 (0.56)	3.50 (0.56)	3.43 (0.55)	0.030	*Birth weight, kg*	3.38 (0.59)	3.39 (0.57)	3.32 (0.64)	0.357
Maternal age	29.2 (5.5)	30.9 (4.5)	28.7 (5.6)	<0.001	*Maternal age*	27.4 (4.8)	27.6 (4.7)	26.9 (4.9)	0.115
	Frequency (%)					Frequency (%)			
Sex (Male)	876/1713 (51.1)	183/408 (44.9)	693/1305 (53.1)	0.004	*Sex (Male)*	327/611 (53.5)	218/418 (52.2)	109/193 (56.5)	0.319
Low birth weight (≤2.5 kg)	74/1639 (4.5)	14/390 (3.6)	60/1249 (4.8)	0.313	*Low birth weight (≤2.5 kg)*	39/621 (6.3)	28/418 (6.7)	11/203 (5.4)	0.537
Maternal smoking	466/1731 (26.9)	64/408 (15.7)	402/1323 (30.4)	<0.001	*Maternal smoking*	115/623 (18.5)	63/418 (15.1)	52/205 (25.4)	0.002
Pet ownership	580/1616 (35.9)	149/404 (36.9)	431/1212 (35.6)	0.632	*Pet ownership*	392/623 (62.9)	262/418 (62.7)	130/205 (63.4)	0.858
Maternal eczema	292/1731 (16.9)	73/408 (17.9)	219/1323 (16.6)	0.528	*Maternal eczema*	113/623 (18.1)	78/418 (18.7)	35/205 (17.1)	0.629
Maternal asthma	278/1731 (16.1)	54/408 (13.2)	224/1323 (16.9)	0.075	*Maternal asthma*	87/623 (14.0)	58/418 (13.9)	29/205 (14.2)	0.927
Maternal hay-fever	431/1731 (24.9)	99/408 (24.3)	332/1323 (25.1)	0.753	*Maternal hay-fever*	165/623 (26.5)	111/418 (26.6)	54/205 (26.3)	0.955
Paternal eczema	161/1731 (9.3)	43/408 (10.5)	118/1323 (8.9)	0.325	*Paternal eczema*	85/618 (13.8)	59/425 (14.2)	26/203 (12.8)	0.633
Paternal asthma	227/1731 (13.1)	61/408 (15.0)	166/1323 (12.6)	0.209	*Paternal asthma*	85/618 (13.8)	63/415 (15.2)	22/203 (10.8)	0.141
Paternal hay-fever	400/1731 (23.1)	107/408 (26.2)	293/1323 (22.2)	0.087	*Paternal hay-fever*	159/617 (25.8)	120/415 (28.9)	39/202 (19.3)	0.010

Table E4. Prevalence of the co-occurrence of atopic morbidity at each cross-sectional time-point

0 = No disease; 1 = any single disease (wheeze, eczema or rhinitis); 2 = co-occurrence of any 2 diseases; 3 = atopic triad

Cohort/age	N	0 No disease		1 Single disease		2 Any 2 diseases		3 Atopic triad		Total
MAAS		n	%	n	%	n	%	n	%	
1 year	**935**	482	51.6	358	38.28	95	10.2	0	0.0	100
3 years	**1049**	568	54.2	351	33.5	114	10.9	16	1.5	100
5 years	**1034**	452	43.7	365	35.3	163	15.8	54	5.2	100
8 years	**1020**	511	50.1	305	29.9	153	15.0	51	5.0	100
11 years	**912**	432	47.4	297	32.6	144	15.8	39	4.3	100
16 years	**734**	320	43.6	268	36.5	115	15.7	31	4.2	100
Mean			48.4		34.3		13.9		3.4	100
Ashford										
1 year	**454**	238	52.4	173	38.1	39	8.6	4	0.9	100
2 years	**615**	345	56.1	204	33.2	56	9.1	10	1.6	100
3 years	**615**	339	55.1	194	31.6	69	11.2	13	2.1	100
4 years	**611**	379	62.0	163	26.7	56	9.2	13	2.1	100
5 years	**604**	389	64.4	146	24.2	49	8.1	20	3.3	100
8 years	**593**	383	64.6	144	24.3	55	9.3	11	1.9	100
14 years	**499**	269	53.9	151	30.3	64	12.8	15	3.0	100
Mean			58.4		29.7		9.8		2.1	100
IOW										
1 year	**1247**	946	75.9	189	15.2	94	7.5	18	1.4	100
2 years	**1157**	769	66.5	239	20.7	131	11.3	18	1.6	100
4 years	**1157**	697	60.2	314	27.1	114	9.9	32	2.8	100
10 years	**1347**	814	60.4	382	28.4	113	8.4	38	2.8	100
18 years	**1080**	590	54.6	326	30.2	138	12.8	26	2.4	100
26 years	**1028**	473	46.0	365	35.5	161	15.7	29	2.8	100
Mean			60.6		26.2		10.9		2.3	
SEATON										
6m	**1585**	945	59.6	510	32.2	106	6.7	24	1.5	100
1 year	**1507**	1010	67.0	370	24.6	117	7.8	10	0.7	100
2 years	**1372**	886	64.6	360	26.2	107	7.8	19	1.4	100
5 years	**1175**	822	70.0	269	22.9	67	5.7	17	1.5	100
10 years	**883**	556	63.0	217	24.6	84	9.5	26	2.9	100
15 years	**703**	372	52.9	230	32.7	85	12.1	16	2.3	100
Mean			62.8		27.2		8.3		1.7	

Co-occurrence patterns of eczema, wheeze and rhinitis

Table E5. Comparison of cross-sectional observed and expected probabilities for each disease category. P-values derived using the exact binomial test to test the hypothesis that the observed probabilities do not differ from expected. Highlighted p-values denote significance against a Benjamini-Hochberg FDR corrected significance level to adjust for multiple comparisons; values less than this threshold denote that observed and expected probabilities differ significantly more than by chance.

MAAS (FDR significance level=0.0320)		Eczema only	Wheeze only	Rhinitis only	Wheeze + Eczema	Wheeze + Rhinitis	Eczema + Rhinitis	Eczema + Wheeze+ Rhinitis	No disease
	Observed	0.241	0.139	0.003	0.096	0.002	0.003	0.000	0.516
	Expected	0.257	0.155	0.004	0.080	0.001	0.002	0.001	0.500
Age 1	P-Value	0.015	0.013	0.164	0.011	0.363	0.271	0.547	0.017
	Observed	0.217	0.106	0.011	0.092	0.010	0.008	0.015	0.542
	Expected	0.247	0.142	0.023	0.071	0.007	0.011	0.003	0.496
Age 3	P-Value	0.002	<0.001	0.001	0.002	0.067	0.067	<0.001	<0.001
	Observed	0.166	0.079	0.107	0.039	0.052	0.067	0.052	0.437
	Expected	0.182	0.108	0.146	0.052	0.042	0.070	0.020	0.380
Age 5	P-Value	0.012	<0.001	<0.001	0.009	0.019	0.045	<0.001	<0.001
	Observed	0.123	0.055	0.122	0.031	0.044	0.075	0.050	0.501
	Expected	0.162	0.092	0.172	0.036	0.038	0.067	0.015	0.419
Age 8	P-Value	<0.001	<0.001	<0.001	0.044	0.045	0.032	<0.001	<0.001
	Observed	0.104	0.052	0.170	0.024	0.071	0.063	0.043	0.474
	Expected	0.124	0.095	0.215	0.029	0.051	0.066	0.015	0.405
Age 11	P-Value	0.007	<0.001	<0.001	0.047	0.002	0.050	<0.001	<0.001
	Observed	0.063	0.040	0.263	0.018	0.070	0.070	0.042	0.436
	Expected	0.089	0.076	0.298	0.018	0.061	0.071	0.015	0.372
Age16	P-Value	0.002	<0.001	0.004	0.110	0.037	0.057	<0.001	<0.001

Ashford (FDR significance level=0.0360)		Eczema only	Wheeze only	Rhinitis only	Wheeze + Eczema	Wheeze + Rhinitis	Eczema + Rhinitis	Eczema + Wheeze+ Rhinitis	No disease
	Observed	0.049	0.311	0.022	0.053	0.031	0.002	0.009	0.524
	Expected	0.063	0.335	0.034	0.043	0.023	0.004	0.003	0.495
Age 1	P-Value	0.036	0.021	0.027	0.049	0.061	0.141	0.034	0.019
	Observed	0.086	0.210	0.036	0.042	0.037	0.011	0.016	0.561
	Expected	0.097	0.232	0.059	0.043	0.026	0.011	0.005	0.529
Age 2	P-Value	0.033	0.017	0.003	0.079	0.029	0.157	0.002	0.009
	Observed	0.101	0.161	0.054	0.059	0.037	0.016	0.021	0.551
	Expected	0.124	0.195	0.074	0.048	0.029	0.018	0.007	0.506
Age 3	P-Value	0.011	0.004	0.009	0.033	0.051	0.105	0.001	0.003
	Observed	0.088	0.120	0.059	0.030	0.056	0.016	0.021	0.610
	Expected	0.102	0.163	0.099	0.030	0.029	0.018	0.005	0.554
Age 4	P-Value	0.026	0.001	<0.001	0.095	<0.001	0.106	<0.001	0.001
	Observed	0.078	0.080	0.084	0.022	0.045	0.015	0.033	0.644
	Expected	0.100	0.126	0.124	0.022	0.027	0.021	0.005	0.575
Year 5	P-Value	0.023	0.001	0.002	0.139	0.018	0.079	<0.001	0.001
	Observed	0.068	0.051	0.125	0.020	0.044	0.029	0.019	0.646
	Expected	0.092	0.091	0.162	0.014	0.025	0.026	0.004	0.586
Year 8	P-Value	0.006	<0.001	0.002	0.085	0.002	0.087	<0.001	<0.001
	Observed	0.040	0.042	0.220	0.006	0.066	0.056	0.030	0.539
	Expected	0.071	0.078	0.276	0.012	0.046	0.042	0.007	0.467
Year 14	P-Value	0.001	<0.001	0.001	0.045	0.015	0.034	<0.001	<0.001

IOW (FDR significance level=0.0350)		Eczema only	Wheeze only	Rhinitis only	Wheeze + Eczema	Wheeze + Rhinitis	Eczema + Rhinitis	Eczema + Wheeze+ Rhinitis	No disease
	Observed	0.070	0.031	0.051	0.017	0.042	0.017	0.014	0.759
	Expected	0.094	0.082	0.099	0.011	0.012	0.014	0.002	0.688
Age 1	P-Value	<0.001	<0.001	<0.001	0.018	<0.001	0.054	<0.001	<0.001
	Observed	0.120	0.057	0.030	0.025	0.059	0.029	0.016	0.665
	Expected	0.139	0.110	0.091	0.026	0.017	0.021	0.004	0.591
Age 2	P-Value	0.005	<0.001	<0.001	0.071	<0.001	0.020	<0.001	<0.001
	Observed	0.131	0.078	0.063	0.035	0.035	0.029	0.028	0.602
	Expected	0.155	0.116	0.099	0.033	0.021	0.028	0.006	0.541
Age 4	P-Value	0.002	<0.001	<0.001	0.062	0.001	0.070	<0.001	<0.001
	Observed	0.065	0.090	0.128	0.013	0.056	0.015	0.028	0.604
	Expected	0.076	0.127	0.162	0.017	0.037	0.022	0.005	0.552
Age 10	P-Value	0.012	<0.001	<0.001	0.035	<0.001	0.013	<0.001	<0.001
	Observed	0.034	0.072	0.195	0.006	0.100	0.022	0.024	0.546
	Expected	0.045	0.122	0.249	0.011	0.063	0.023	0.006	0.481
Age 18	P-Value	0.010	<0.001	<0.001	0.021	<0.001	0.075	<0.001	<0.001
	Observed	0.035	0.074	0.246	0.009	0.120	0.028	0.028	0.460
	Expected	0.044	0.120	0.292	0.013	0.088	0.032	0.010	0.400
Age 26	P-Value	0.017	<0.001	<0.001	0.052	<0.001	0.048	<0.001	<0.001

SEATON (FDR significance level=0.0310)		Eczema only	Wheeze only	Rhinitis only	Wheeze + Eczema	Wheeze + Rhinitis	Eczema + Rhinitis	Eczema + Wheeze+ Rhinitis	No disease
	Observed	0.095	0.119	0.108	0.019	0.040	0.008	0.015	0.596
	Expected	0.092	0.138	0.119	0.022	0.028	0.019	0.005	0.577
Age 6 months	P-Value	0.031	0.002	0.012	0.052	0.002	<0.001	<0.001	0.007
	Observed	0.085	0.088	0.073	0.024	0.036	0.018	0.007	0.670
	Expected	0.098	0.116	0.098	0.018	0.018	0.015	0.003	0.634
Age 1	P-Value	0.008	<0.001	<0.001	0.017	<0.001	0.054	0.007	<0.001
	Observed	0.128	0.079	0.055	0.025	0.035	0.018	0.014	0.646
	Expected	0.138	0.109	0.084	0.025	0.015	0.019	0.003	0.606
Age 2	P-Value	0.018	<0.001	<0.001	0.069	<0.001	0.074	<0.001	<0.001
	Observed	0.148	0.067	0.014	0.041	0.009	0.007	0.015	0.700
	Expected	0.174	0.100	0.030	0.027	0.005	0.008	0.001	0.655
Age 5	P-Value	0.002	<0.001	<0.001	0.001	0.013	0.104	<0.001	<0.001
	Observed	0.060	0.041	0.145	0.006	0.044	0.045	0.029	0.630
	Expected	0.091	0.076	0.200	0.012	0.027	0.033	0.004	0.557
Age 10	P-Value	<0.001	<0.001	<0.001	0.023	0.001	0.012	<0.001	<0.001
	Observed	0.068	0.027	0.232	0.011	0.050	0.060	0.023	0.529
	Expected	0.092	0.059	0.271	0.011	0.034	0.053	0.007	0.474
Age 15	P-Value	0.005	<0.001	0.002	0.140	0.006	0.051	<0.001	<0.001

Longitudinal sequence analysis

Figure E1. Sequence frequency plots for 20 most common individual patterns. The y-axis shows the cumulative percentage of these sequences and the bar widths are proportional to their frequencies. Data for children reporting at least one symptom on at least one time point. For MAAS, the top 20 most frequently occurring sequences accounted for 26% of all sequences, of which no disease from age1-11 was followed by rhinitis at age 16. This was also the most common sequence in SEATON & IOW, and the second most common in Ashford cohorts.

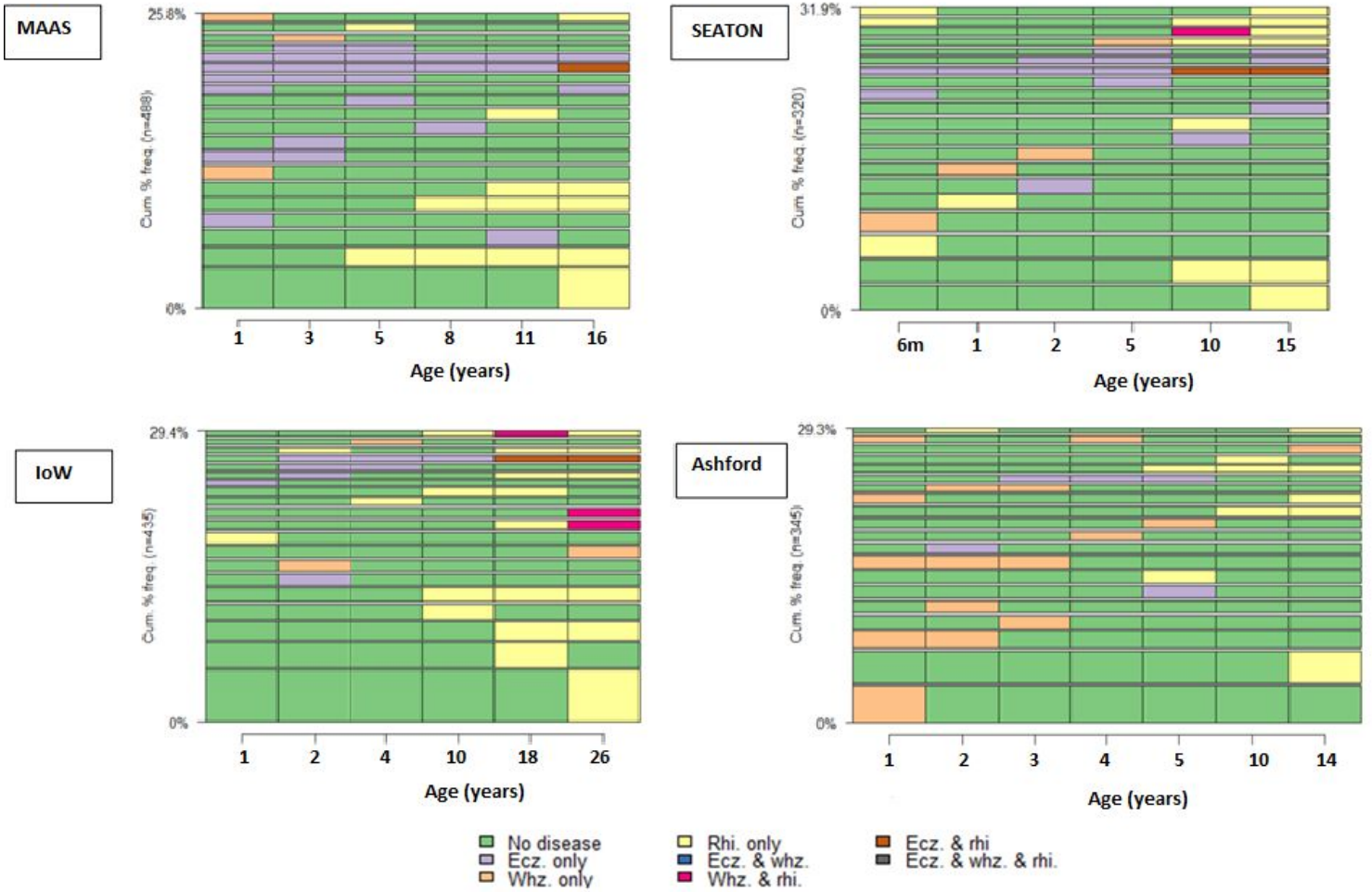


Figure E2. Sequence frequency plots for the 20 most common individual patterns among children with eczema as a single disease or eczema with comorbidities in the first three years of life. The y-axis represents the cumulative percentage of these sequences and the bar widths are proportional to their frequencies.

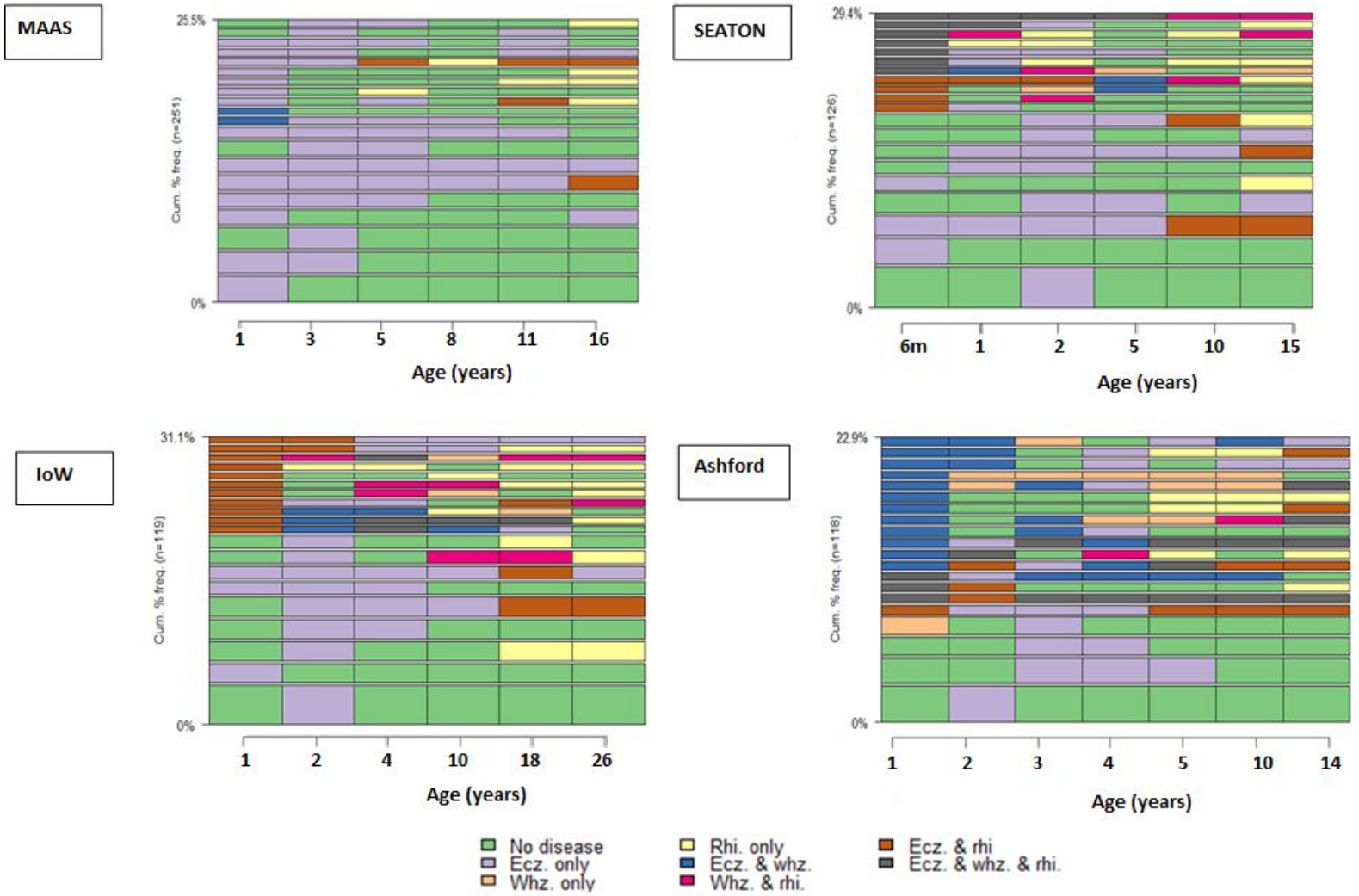


Figure E3. Sequence frequency plots for the twenty most common individual patterns among children with wheeze as a single disease or wheeze with comorbidities in the first three years of life. The y-axis represents the cumulative percentage of these sequences and the bar widths are proportional to their frequencies.

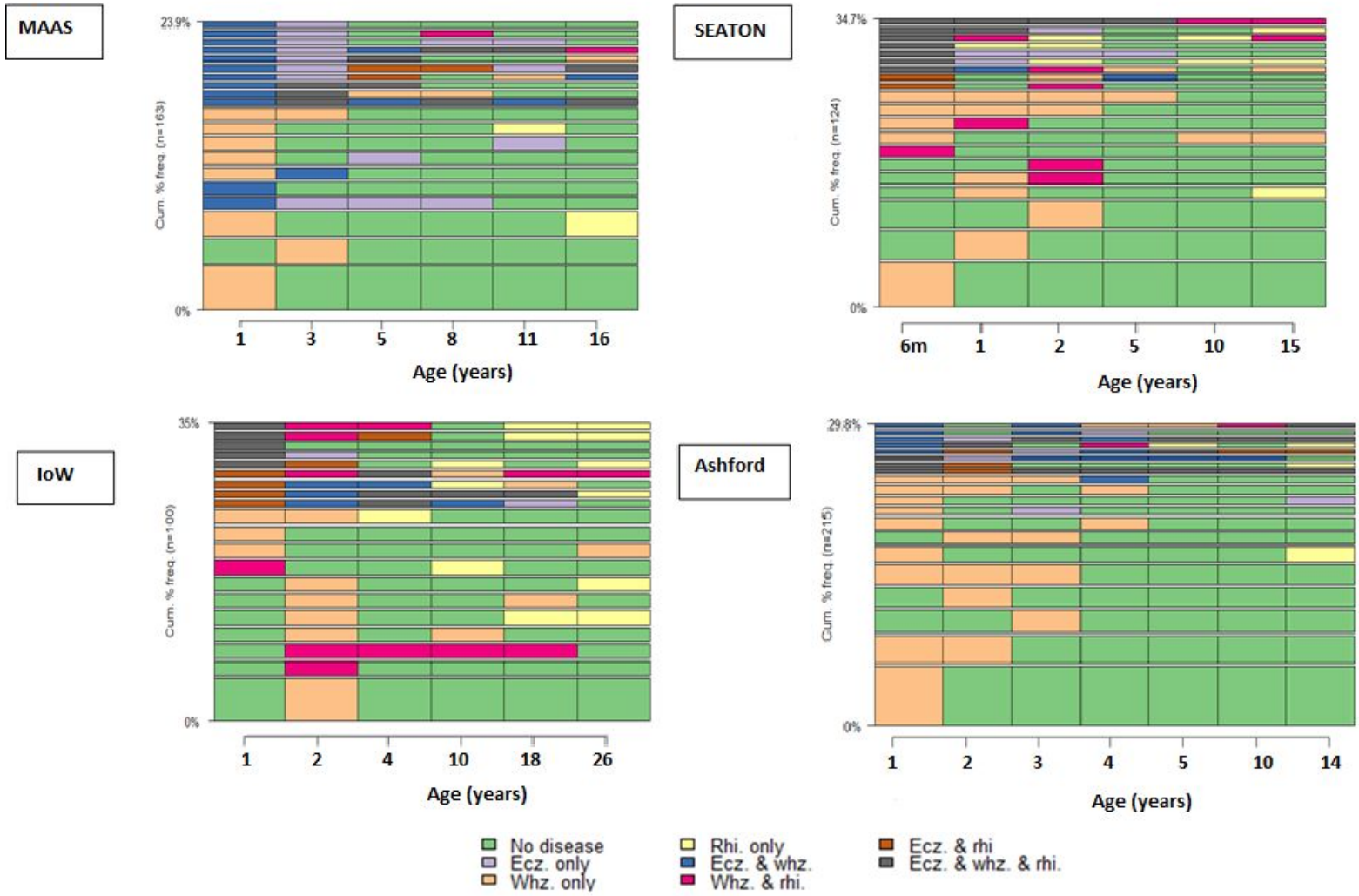


Table E6. The number (and proportion) of children with and without symptoms in the first 12 months of life amongst 157 participants who reported E+W+R multimorbidity at least once during the study period

	MAAS		Ashford		SEATON		IOW		TOTAL
	N	%	N	%	N	%	N	%	
Complete trajectories across all time points	553		418		408		519		1898
All three symptoms observed at least once (including non-contemporaneously)	143	26	79	19	70	17	82	16	**374**
E+W+R multimorbidity observed at least once	61	11	34	8	32	8	39	8	**166**
Reported symptoms in the first 12 months of life									
No early eczema or wheeze & no later E+W+R multimorbidity	283	51.2	235	56.2	288	70.6	409	78.8	1215
Early eczema only & later E+W+R multimorbidity	24	4.3	4	1.0	12	2.9	11	2.1	51
Early life eczema only & no later E+W+R multimorbidity	122	22.1	13	3.1	42	10.3	36	6.9	213
Early wheeze only & later E+W+R multimorbidity	7	1.3	16	3.8	2	0.5	4	0.8	29
Early life wheeze only & no later E+W+R multimorbidity	61	11.0	127	30.4	41	10.1	30	5.8	259
Early E+W & later E+W+R multimorbidity	20	3.6	6	1.4	4	1.0	1	0.2	31
Early E+W & no later E+W+R multimorbidity	26	4.7	9	2.2	5	1.2	5	1.0	45
No early eczema or wheeze & later E+W+R multimorbidity	10	1.8	5	1.2	9	2.2	17	3.3	41
Early triad & later E+W+R multimorbidity	0	0.0	1	0.2	3	0.7	1	0.2	5
Early triad & no later E+W+R multimorbidity	0	0.0	2	0.5	2	0.5	5	1.0	9

Dynamics of change over time: Latent Markov modelling

Table E7. Harmonised time-points (years) and sample sizes per cohort used for joint LMM analysis:
Infancy: Age 1 year; Early life: Age 2-3 years; Preschool: 4-5 years; Mid-school: Age 8-10 years; Adolescence: Age 14-18 years

	Infancy	Early life	Preschool	Mid-school	Adolescence	Observations	%
MAAS	1	3	5	8	16	574	28
Ashford	1	2	4	8	18	419	20
IOW	1	2	4	10	14	666	32
SEATON	1	2	5	10	15	420	20
Total						2079	100

Table E8. BIC index for optimal latent Markov model selection

Cohort	Time homogeneous	Time heterogeneous	Cohort	Time homogeneous	Time heterogeneous
	MAAS			**IOW**	
1	10512.65	10512.65	1	8510.24	8510.24
2	9608.12	9743.36	2	7783.22	7780.05
3	9180.28	9236.92	3	7546.09	7585.51
4	8953.91	9129.22	4	7451.48	7594.51
5	8860.01	9201.74	5	**7446.76**	7776.08
6	**8852.48**	9394.22	6	7466.41	8028.37
7	8874.77	9745.86	7	7529.46	8360.67
8	8934.07	10124.97	8	7611.51	8737.22
	SEATON			**Ashford**	
1	6476.22	6476.22	1	8112.82	8112.82
2	5937.33	5978.95	2	7359.22	7405.42
3	5837.72	5804.50	3	7082.69	7205.40
4	5767.09	5839.84	4	6871.42	7091.73
5	**5737.93**	5989.85	5	**6731.60**	7148.09
6	5823.20	6508.45	6	6761.91	7447.10
7	5772.94	6248.28	7	6799.76	7821.29
8	5894.74	6893.66	8	6861.08	8269.49
	Joint cohorts				
1	29552.35	29552.35			
2	27375.52	27418.75			
3	26418.38	26318.43			
4	25966.23	25917.76			
5	**25792.55**	25822.46			
6	25947.97	25959.86			
7	26036.09	26189.27			
8	26519.35	26705.17			

The BIC index indicated that a time-homogeneous model with 5 latent states was optimal for the joint cohorts and 3 individual cohorts (IOW, SEATON, Ashford); BIC indicated 6 states for MAAS, but the estimates of the conditional response probabilities suggested that the model with five latent states was more interpretable and consistent with other cohorts and the joint model.

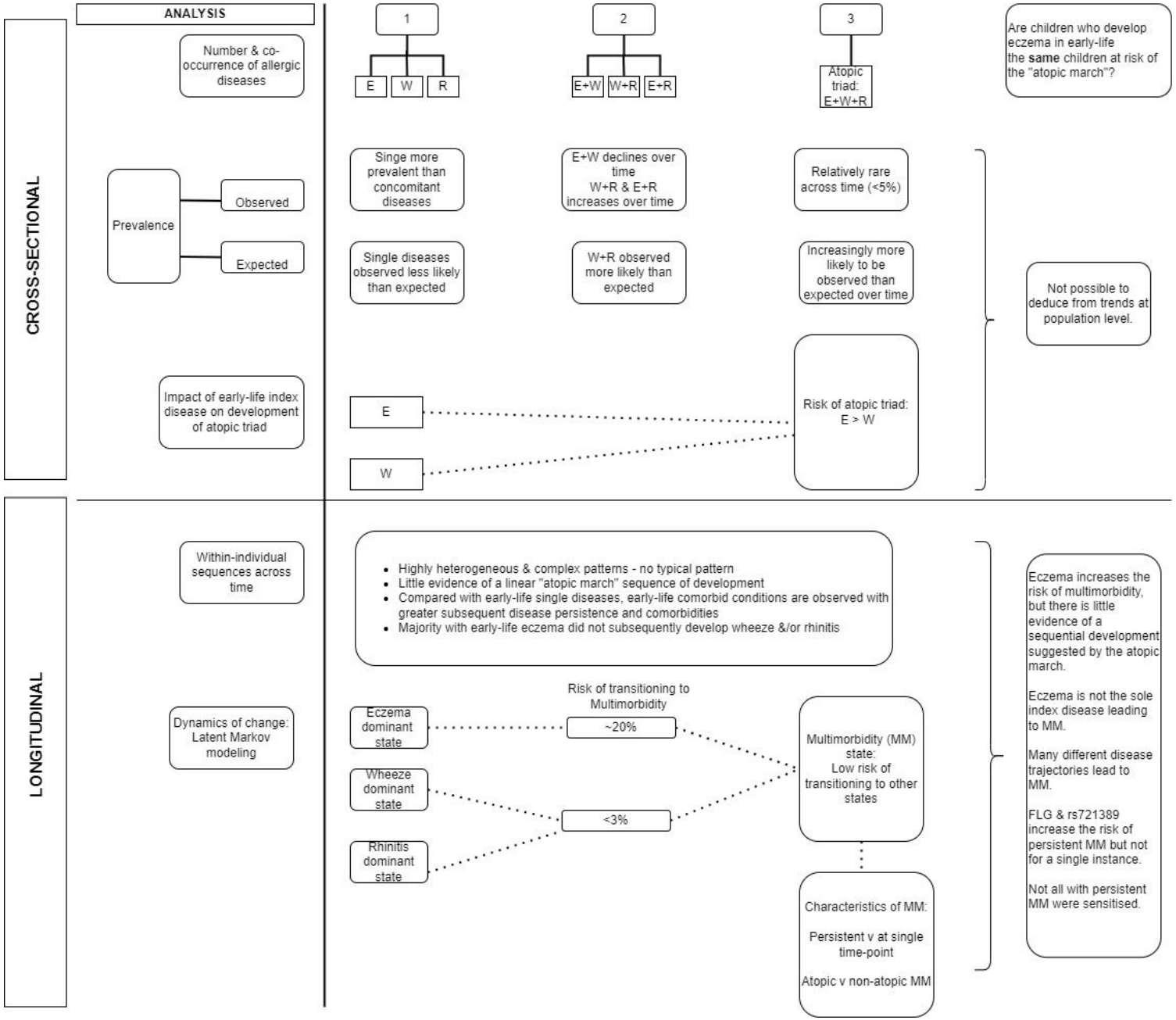
Table E9. Associations of multimorbidity persistence with allergic sensitisation. Tests of associations were carried out using ANOVA for continuous variables and chi-square test for categorical variables.

			Number of time-points in Multimorbidity state			
Variable	**Statistics**	**Level**	**0; N=1792**	**1; N=84**	**2-5; N=203**	**P-value**
Allergic sensitisation at age 5						
SPT: Cat	N (Col %)	Negative	1412 (94.7)	65 (91.55)	126 (76.36)	**<.001**
	N (Col %)	Positive	79 (5.3)	6 (8.45)	39 (23.64)	
SPT: Grass	N (Col %)	Negative	1361 (91.28)	58 (81.69)	114 (69.09)	**<.001**
	N (Col %)	Positive	130 (8.72)	13 (18.31)	51 (30.91)	
SPT: House dust-mite	N (Col %)	Negative	1360 (91.15)	55 (76.39)	113 (68.9)	**<.001**
	N (Col %)	Positive	132 (8.85)	17 (23.61)	51 (31.1)	
Sensitised (≥1 +ve SPT)	N (Col %)	No	1251 (83.9)	49 (69.01)	89 (53.94)	**<.001**
	N (Col %)	Yes	240 (16.1)	22 (30.99)	76 (46.06)	
N. of sensitisations	N (Col %)	None	1251 (83.9)	49 (69.01)	89 (53.94)	**<.001**
	N (Col %)	One	164 (11.0)	13 (18.31)	29 (17.58)	
	N (Col %)	Poly	76 (5.10)	9 (12.68)	47 (28.48)	
Allergic sensitisation at age 14-18						
SPT: Cat	N (Col %)	Negative	1253 (88.93)	46 (73.02)	99 (60.74)	**<.001**
	N (Col %)	Positive	156 (11.07)	17 (26.98)	64 (39.26)	
SPT: Grass	N (Col %)	Negative	1050 (74.36)	30 (48.39)	67 (41.61)	**<.001**
	N (Col %)	Positive	362 (25.64)	32 (51.61)	94 (58.39)	
SPT: House dust-mite	N (Col %)	Negative	1094 (77.59)	39 (61.9)	86 (52.44)	**<.001**
	N (Col %)	Positive	316 (22.41)	24 (38.1)	78 (47.56)	
Sensitised (≥1 +ve SPT)	N (Col %)	No	887 (63.13)	22 (35.48)	49 (29.88)	**<.001**
	N (Col %)	Yes	518 (36.87)	40 (64.52)	115 (70.12)	
N. of sensitisations	N (Col %)	None	887 (63.13)	22 (35.48)	49 (29.88)	**<.001**
	N (Col %)	One	282 (20.07)	17 (27.42)	37 (22.56)	
	N (Col %)	Poly	236 (16.80)	23 (37.10)	78 (47.56)	

Table E10: Characteristics of children with non-atopic and atopic multimorbidity persistence (as defined by MM at 2-5 time-points according to predicted latent state from LMM) stratified by sensitisation status in childhood (age 5) and adolescence (age 14-18).

	Sensitisation status (age 5)				**Sensitisation status (age 14-18)**			
	Not sensitised	Sensitised	Total	p-value	Not sensitised	Sensitised	Total	p-value
Male	37/89	44/76	81/165	**0.037**	15/49	59/115	74/164	**0.015**
	41.6%	57.9%	49.1%		30.1%	51.3%	45.1%	
Low birth weight	7/86	4/76	11/162	0.468	3/47	7/114	10/161	0.954
	8.1%	5.3%	6.8%		6.4%	6.1%	6.2%	
Environmental								
Maternal smoking	15/89	7/76	22/165	0.150	8/49	12/115	20/164	0.291
	16.9%	9.2%	13.3%		16.3%	10.4%	12.2%	
Pets in home	39/89	33/75	72/164	0.982	24/49	43/115	67/164	0.167
	43.8%	44%	43.9%		48.98%	37%	40.9%	
Parental morbidity								
Maternal asthma	17/89	19/76	36/165	0.360	8/49	27/115	35/164	0.306
	19.1%	25%	21.8%		13.6%	23.5%	21.3%	
Maternal eczema	28/89	13/76	41/165	**0.033**	12/49	27/115	39/164	0.889
	31.5%	17.1%	24.9%		24.5%	23.5%	23.8%	
Maternal hayfever	35/89	33/73	68/165	0.594	18/49	48/115	66/164	0.550
	39.3%	43.4%	41.2%		36.7%	41.7%	40.2%	
Paternal asthma	21/89	14/76	35/165	0.418	9/49	25/115	34/164	0.626
	23.6%	18.4%	21.2%		18.4%	21.7%	20.7%	
Paternal eczema	12/89	15/76	27/165	0.279	9/49	17/115	26/164	0.565
	13.5%	19.7%	16.4%		18.4%	14.8%	15.9%	
Paternal hayfever	23/89	28/76	51/165	0.128	10/49	44/115	54/164	**0.026**
	25.8%	36.8%	30.9%		20.4%	38.3%	32.9%	
Genetic								
FLG	13/78	13/63	26/141	0.546	8/45	18/95	26/140	0.868
	16.7%	20.6%	18.4%		17.8%	19.%	18.6%	
rs7216389 CC	16/80	18/70	34/150	0.672	8/47	27/103	35/150	0.434
	20.0%	25.7%	22.7%		17.0%	26.2%	23.3%	
CT	40/80	31/70	70/150		25/47	46/103	71/150	
	50.0%	44.3%	47.3%		53.2%	44.7%	47.3%	
TT	24/80	21/70	45/150		14/47	30/103	44/150	
	30.0%	30.0%	30.0%		29.8%	29.1%	29.3%	

Figure E5. Schematic overview of cross-sectional and longitudinal results (E=eczema, W=wheeze, R=rhinitis)



REFERENCES

1. Custovic A, Simpson BM, Murray CS, et al. The National Asthma Campaign Manchester Asthma and Allergy Study. *Pediatr Allergy Immunol.* 2002;13(s15):32-37.
2. Atkinson W, Harris J, Mills P, et al. Domestic aeroallergen exposures among infants in an English town. *Eur Respir J.* 1999;13(3):583-589.
3. Kurukulaaratchy RJ, Fenn M, Twiselton R, Matthews S, Arshad SH. The prevalence of asthma and wheezing illnesses amongst 10-year-old schoolchildren. *Respir Med.* 2002;96(3):163-169.
4. Kurukulaaratchy RJ, Fenn MH, Waterhouse LM, Matthews SM, Holgate ST, Arshad SH. Characterization of wheezing phenotypes in the first 10 years of life. *Clin Exp Allergy.* 2003;33(5):573-578.
5. Arshad SH, Holloway JW, Karmaus W, et al. Cohort Profile: The Isle Of Wight Whole Population Birth Cohort (IOWBC). *Int J Epidemiol.* 2018;47(4):1043-1044i.
6. McCarthy S, Das S, Kretzschmar W, et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet.* 2016;48(10):1279-1283.

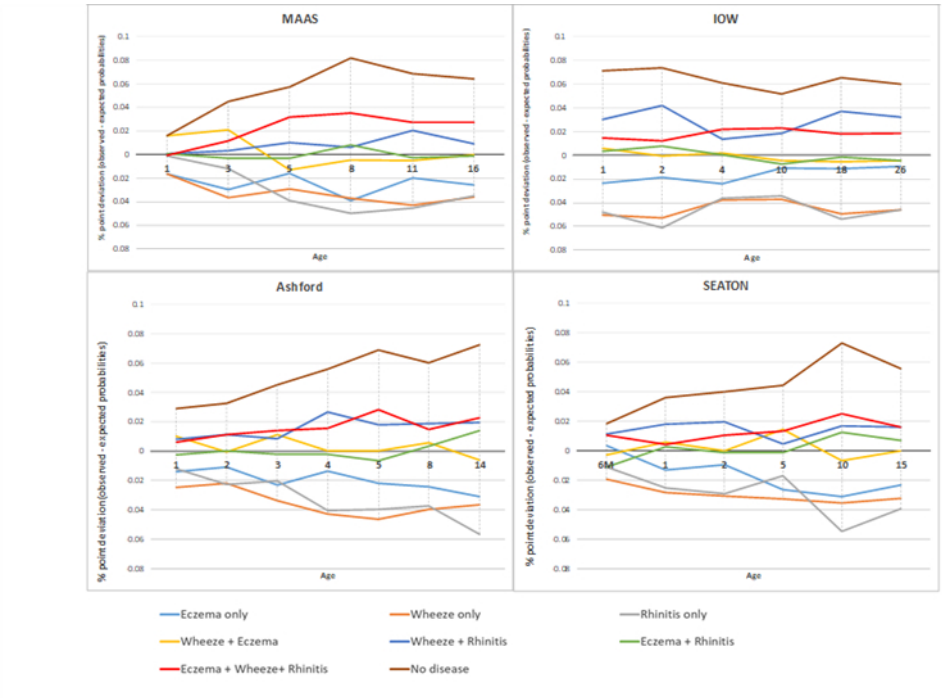


Figure 1. Trends in the deviation between observed and expected probabilities for each disease category over time (expressed as per cent point difference). Negative numbers show that observed probabilities were lower than expected probabilities, for example, single diseases were observed less frequently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

206x143mm (96 x 96 DPI)

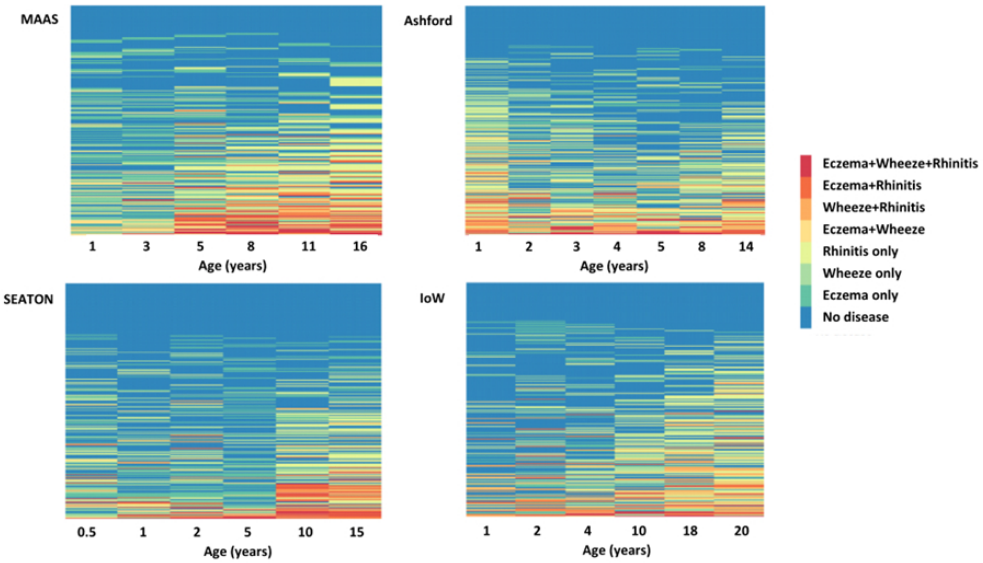


Figure 2. Index plots of individual longitudinal sequences of disease development. Each row is coloured by the disease state at each time-point and displays the duration spent in each state. The number of person-unique sequences: 220 SEATON, 259 Ashford, 295 IoW, 351 MAAS)

241x142mm (96 x 96 DPI)

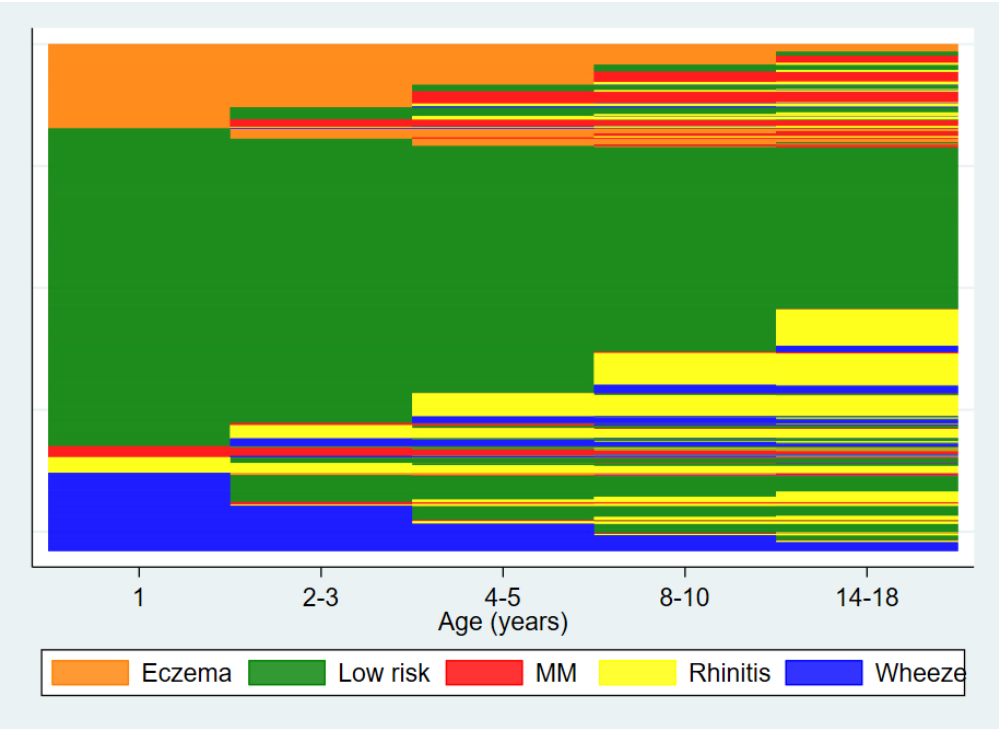
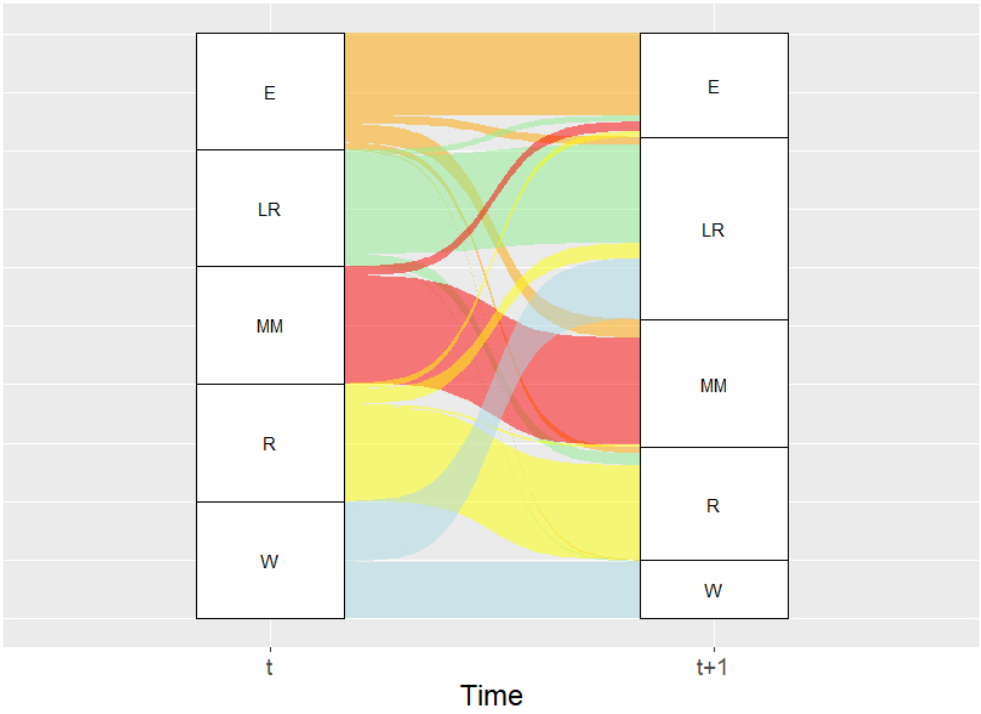


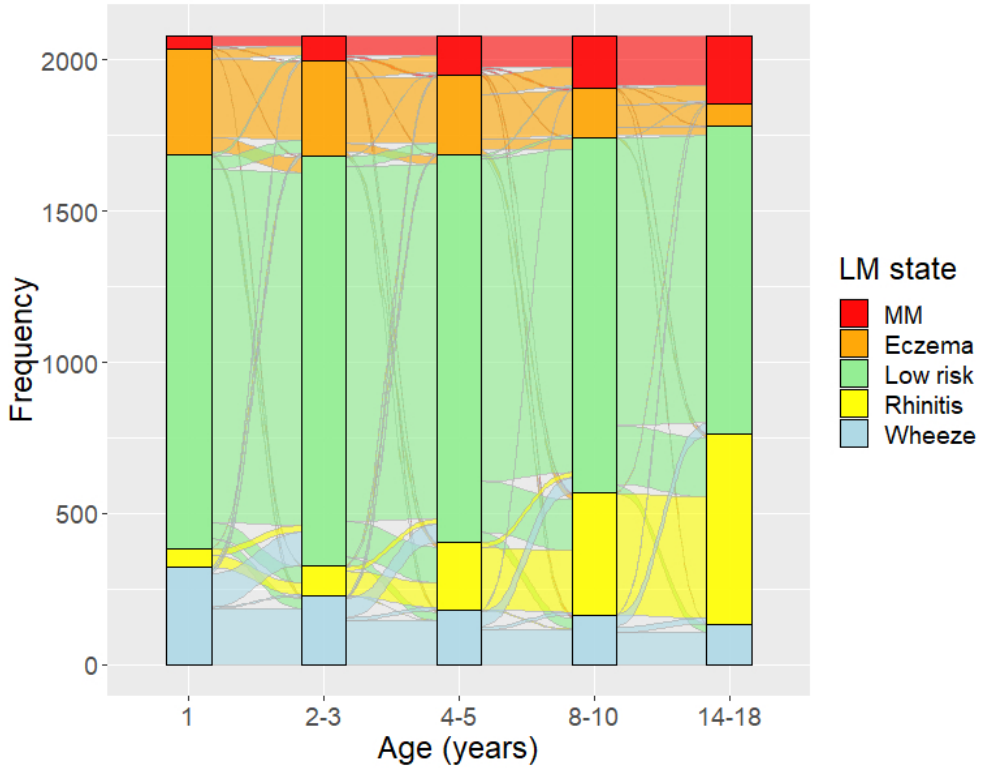
Figure 3. Dynamics of change in eczema, wheeze and rhinitis over time: Latent Markov modelling in the joint cohort model (2079 children with complete observations on eczema, wheeze, and rhinitis at five time-points). Data were harmonised at overlapping time-points to represent five stages of development (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.

322x234mm (72 x 72 DPI)



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.

217x157mm (96 x 96 DPI)



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.

202x157mm (96 x 96 DPI)