# Title Page

 **Atopic dermatitis trajectories to age 8 years in the GUSTO cohort**

**Running title: Childhood atopic dermatitis trajectories identified using machine learning approach**

**Author list & Affiliations**

Noor H. A. Suaini1, Gaik Chin Yap2, Bui Do Phuong Tung3, Evelyn Xiu Ling Loo1,2, Anne Eng Neo Goh4, Oon Hoe Teoh5, Kok Hian Tan6, Keith M. Godfrey7,8, Bee Wah Lee2, Lynette Pei-chi Shek2,9, Hugo Van Bever2,9, Yap Seng Chong1,10,Elizabeth Huiwen Tham1,2,9,11

**Affiliations:**

1Singapore Institute for Clinical Sciences (SICS), Agency for Science, Technology and Research (A\*STAR), Singapore

2 Department of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore (NUS), Singapore

3Department of Architecture, School of Design and Environment, National University of Singapore (NUS), Singapore

4Allergy service, Department of Paediatrics, KK Women’s and Children’s Hospital (KKH), Singapore

5Respiratory Service, Department of Paediatrics, KK Women’s and Children’s Hospital (KKH), Singapore

6Department of Maternal Fetal Medicine, KK Women’s and Children’s Hospital (KKH), Singapore

7NIHR Southampton Biomedical Research Centre, University of Southampton and University Hospital Southampton NHS Foundation Trust, SO16 6YD, Southampton, United Kingdom

8Medical Research Council Lifecourse Epidemiology Unit, SO16 6YD, Southampton, United Kingdom

9Khoo Teck Puat-National University Children’s Medical Institute, National University Health System (NUHS), Singapore

10Department of Obstetrics and Gynaecology, Yong Loo Lin School of Medicine, National University of Singapore (NUS), National University Health System (NUHS), Singapore

11Human Potential Translational Research Programme, Yong Loo Lin School of Medicine, National University of Singapore

# Statement of contribution

Noor H. A. Suaini contributed to the design of the study, carried out data cleaning and data analysis, was involved in interpretation of the data, provided intellectual input, drafted and revised the manuscript.

Gaik Chin Yap and Bui Do Phuong Tung contributed to the design of the study, carried out data cleaning and data analysis, were involved in interpretation of the data and provided intellectual input in the draft and revised manuscripts.

Evelyn Xiu Ling Loo, Anne Eng Neo Goh, Oon Hoe Teoh, Kok Hian Tan, Keith M. Godfrey, Bee Wah Lee, Lynette Pei-chi Shek, Hugo Van Bever, Yap Seng Chong provided intellectual input, contributed to the interpretation of data and critically revised the manuscript.

Elizabeth Huiwen Tham led the design and conception of the study and data analysis, provided intellectual input, contributed to the interpretation of data and contributed to the draft and revised manuscripts.

All authors reviewed the manuscript and approved the final version to be published.

**Corresponding Author**

**Name:** Dr Elizabeth Huiwen Tham

**Address:** Division of Allergy & Immunology

Department of Paediatrics

Yong Loo Lin School of Medicine

National University of Singapore

1E Kent Ridge Road

Level 12 NUHS Tower Block

Singapore 119228

**Telephone Number:** +65 – 6772 3394

**Email:** elizabeth\_tham@nuhs.edu.sg

**Word Count for abstract:** 220 words

**Word Count for text :** 3433 words

**Number of Tables:** 3 main tables and 2 supplementary tables

**Number of Figures:** 3 main figures and 4 supplementary figures

**Conflict of Interest Statement:** Chong YS has received reimbursement for speaking at conferences sponsored by Abbott Nutrition, Nestle, and Danone. Godfrey KM has received reimbursement for speaking at conferences sponsored by Nestle and Shek LP has received reimbursement for speaking at conferences sponsored by Danone and Nestle and consulting for Mead Johnson and Nestle. Godfrey KM, Chong YS are part of an academic consortium that has received research funding from Abbott Nutrition, Nestle and Danone. Shek LP has received research funding from Danone.

**Declaration of funding source:** This research is supported by the Singapore National Research Foundation under its Translational and Clinical Research (TCR) Flagship Program and administered by the Singapore Ministry of Health’s National Medical Research Council (NMRC), Singapore—NMRC/TCR/004-NUS/2008; NMRC/TCR/012-NUHS/ 2014. Additional funding is provided by the Singapore Institute for Clinical Sciences, Agency for Science Technology and Research (A\*STAR), Singapore. E.H. Tham is supported by the National Medical Research Council (NMRC) Transition Award grant [MOH-TA18nov-003] from NMRC, Singapore. K.M. Godfrey is supported by the UK Medical Research Council (MC\_UU\_12011/4), the National Institute for Health Research (NIHR Senior Investigator (NF-SI-0515-10042) and the NIHR Southampton Biomedical Research Centre) and by the European Union’s Erasmus+ Programme (ImpENSA 598488-EPP-1-2018-1-DE-EPPKA2-CBHE-JP).

# Acknowledgements

The authors acknowledge the contribution of the rest of the GUSTO study group which includes Lee Yung Seng, Wei Wei Pang, Pratibha Agarwal, Dennis Bier, Arijit Biswas, Shirong Cai, Jerry Kok Yen Chan, Cornelia Yin Ing Chee, Helen Y. H Chen, Audrey Chia, Amutha Chinnadurai, Chai Kiat Chng, Shang Chee Chong, Mei Chien Chua, Chun Ming Ding, Eric Andrew Finkelstein, Doris Fok, Marielle Fortier, Yam Thiam Daniel Goh, Joshua J. Gooley, Wee Meng Han, Mark Hanson, Christiani Jeyakumar Henry, Joanna D Holbrook,Chin-Ying Hsu, Hazel Inskip, Jeevesh Kapur, Birit Leutscher-Broekman, Sok Bee Lim, Seong Feei Loh, Yen-Ling Low, Iliana Magiati, Lourdes Mary Daniel, Michael Meaney, Susan Morton, Cheryl Ngo, Krishnamoorthy Niduvaje, Anqi Qiu, Boon Long Quah, Victor Samuel Rajadurai, Mary Rauff, Jenny L. Richmond, Anne Rifkin-Graboi, Allan Sheppard, Borys Shuter, Leher Singh, Wing Chee So, Walter Stunkel, Lin Lin Su, Soek Hui Tan, Rob, M. van Dam, Sudhakar K. Venkatesh, Inez Bik Yun Wong, P. C. Wong, George Seow Heong Yeo. We would also like to thank the GUSTO Executive committee, research staff and all parents and children who participated in the GUSTO study.

**Key messages**

1. We used machine learning to characterise atopic dermatitis phenotypes in a birth cohort study.
2. Three AD trajectories were identified – early transient, early persistent and late persistent AD.
3. Molecular and immunological work is needed to understand the origins of these distinct phenotypes.

# Abstract

**Background:** The heterogeneity of childhood atopic dermatitis (AD) underscores the need to understand latent phenotypes that may inform risk stratification and disease prognostication.

**Objective:** To identify AD trajectories across the first 8 years of life, investigate risk factors associated with each trajectory and their relationships with other comorbidities.

**Methods:**  Data were collected prospectively from 1152 mother-offspring dyads in the Growing Up in Singapore Towards healthy Outcomes (GUSTO) cohort from ages 3 months to 8 years. AD was defined based on parent-reported doctor’s diagnosis. An unsupervised machine learning technique was used to determine AD trajectories.

**Results:** Three AD trajectories were identified: early onset transient (6.3%), late onset persistent (6.3%) and early onset persistent (2.1%), alongside a no AD/reference group (85.2%). Early onset transient AD was positively associated with male gender, family history of atopy, house dust mite sensitization and some measures of wheezing. Early onset persistent AD was associated with antenatal/intrapartum antibiotic use, food sensitization and some measures of wheezing. Late onset persistent AD was associated with a family history of atopy, some measures of house dust mite sensitization and some measures of allergic rhinitis and wheezing.

**Conclusion and Clinical Relevance:**  Three AD trajectories were identified in this birth cohort, with different risk factors and prognostic implications. Further work is needed to understand the molecular and immunological origins of these phenotypes.

**Keywords:** atopic dermatitis, trajectories, wheezing, rhinitis, machine learning

**Abbreviations**

AD: Atopic Dermatitis

APT: Atopy Patch Test

CRP: C-Reactive Protein

EM: Expectation Maximization

EOT: Early Onset Transient AD

EOP: Early Onset Persistent AD

GUSTO: Growing Up in Singapore Towards healthy Outcomes (GUSTO)

HDM: House Dust Mite

ISAAC: The International Study of Asthma and Allergies in Childhood

LOP: Late Onset Persistent AD

SCORAD: Scoring of Atopic Dermatitis

SPT: Skin Prick Test

TEWL: Trans-Epidermal Water Loss

**Data Request Statement**

The data that support the findings of this study are available from the corresponding author upon reasonable request.

# Introduction

Atopic dermatitis (AD) is a chronic inflammatory condition characterized by itchy, dry and red skin which often develops early in life and can have debilitating effects. AD in infancy is a precursor of food allergy, asthma and allergic rhinitis later in childhood1. Mechanistically, defects in the skin epithelial barrier are postulated to promote allergen entry through the skin, inducing epicutaneous sensitization and food allergy, followed by allergic airway disorders such as allergic rhinitis and asthma 2, 3.

This progression of atopic manifestations is often termed the ‘allergic march’ or ‘atopic march’. At an individual level, this pathway has been observed in about 3.1% of children, but more than 90% of children with symptoms do not follow the classical atopic march trajectory 4. Current literature has also shown that these allergic diseases share some common genetic risk markers and pathogenic pathways 5-7. Nonetheless, the atopic march may be an oversimplification of more complex and multifactorial processes involved in allergic disease progression 8.

Given its heterogeneity and complex etiology, 9, 10 there are likely to be several endotypes of AD which may not be captured solely based on classical diagnostic tools measuring severity and epidermal barrier function, such as Scoring of Atopic Dermatitis (SCORAD) and trans-epidermal water loss (TEWL) 11. There has been a recent shift towards the use of machine learning approaches to classify latent classes of AD and other allergic diseases. Machine learning techniques represent an unbiased and hypothesis-free approach in explaining clinical outcomes by generating clusters of phenotypes based on a set of variables added to the model 12, 13. Since they are not based on a pre-determined hypothesis, results from data-driven analyses may better reflect underlying biological mechanisms of action14. Although hypothesis-driven techniques are preferred in investigating an existing phenomenon/theory, they aim to pursue and address a different set of research questions, which is not the main aim of the present study.

While there have been several studies exploring latent AD phenotypes, these were predominantly in Caucasian populations and not in ethnically Asian populations. Fewer studies have investigated the development of wheeze and allergic rhinitis in relation to AD phenotypes over time. This study was therefore designed to identify latent AD trajectories in the first 8 years of life, elucidate risk factors and characterize their relationships with other comorbidities such as wheezing, allergic sensitization and rhinitis in a multi-ethnic Asian population.

# Methods

## Study cohort

This study was conducted in the Growing Up in Singapore Towards healthy Outcomes (GUSTO) mother-offspring cohort. The GUSTO cohort has previously been described in detail15. In brief, GUSTO recruited women of Chinese, Malay and Indian ethnicity in their first trimester of pregnancy. Detailed information on demographic characteristics, family history of allergy and lifestyle factors were collected prospectively. Data on allergic outcomes were obtained from questionnaires administered at 3, 6, 12, 15, 18, 24 months and 3, 4, 5, 6, 7 and 8 years , based on the International Study of Asthma and Allergies in Childhood (ISAAC) modified questionnaire which has been extensively validated for epidemiological studies worldwide. 16-19 All mothers gave written informed consent and ethics approval was obtained from the Domain Specific Review Board of Singapore National Healthcare Group (D/2009/021; 26/02/2009) and the Centralised Institutional Review Board of SingHealth (2018/2767; 02/03/2009).

## Definition of main outcome

*Atopic dermatitis:* From months 3 to 18 and at month 36, this was defined asan affirmative response to the question *“*Has your child ever been diagnosed with eczema?”.From month 24 onwards, this was defined as an affirmative response to the question *“*Has your child ever been diagnosed with eczema since the (previous timepoint)/last visit?”. All other definitions of allergic diseases can be found in the Supporting Information.

## Statistical analysis

### Unsupervised machine learning

The AD trajectories analysis was performed on doctor diagnosed AD data from 1152 participants over twelve total time points from 3 months to 8 years of age (at 3, 6, 12, 15, 18, 24 months, 3, 4, 5, 6, 7, and 8 years). The expectation maximization (EM) method in conjunction with 10-fold cross validation was used to identify and determine the appropriate number of AD trajectories within the dataset. Only the AD variables over the 12 timepoints were included in the EM method. The EM method also accounts for missing data in its calculations and therefore all participants remained in the model. Detailed information on the EM method is provided in the Supporting Information.

### Association analyses

Statistical comparisons for the demographic characteristics and allergic outcomes between the trajectories were obtained using Pearson’s chi square or Fisher’s exact test with Bonferroni correction. Multinomial logistic regressions were used to examine associations between trajectories and three sets of variables: prenatal risk factors, environmental risk factors and comorbidities which we hypothesized to modulate AD pathogenesis. Odds ratios and 95% confidence intervals with statistical significance of p<0.05 were used.

The EM unsupervised machine learning method was carried out in Weka, version 3.8.4 (Machine Learning Group, University of Waikato, Hamilton, New Zealand20), while other analyses were carried out using IBM SPSS version 24 (Armonk, NK).

# Results

## Study participants

Of the 1237 GUSTO mothers recruited, 1152 children conceived naturally were included in the analysis. The point prevalence of AD steadily increased from 3.7% at 3 months to 9.3% at 12 months. Thereafter, the prevalence at each time-point fluctuated from 8.7% at 15 months to 7.3% at 8 years (12.5% at 18 months, 8.3% at 24 months, 15.1% at 3 years, 8.3% at 4 years, 4.8% at 5 years, 6.8% at 6 years and 5.2% at 7 years), with 22.8% of our study population having AD at any time-point (n=258/1152). Figure 1 shows the flow diagram of participants with AD at each timepoint. Complete data at all time points were available on 448 children and their trajectories shown in Figures S1 and S2 of Supporting Information.

Three AD trajectories were identified – which we termed early onset transient (EOT; n=73, 6.3%), early onset persistent (EOP; n=24, 2.1%) and late onset persistent (LOP; n=73, 6.3%) based on the trajectory patterns over time (Figure 2). The group of children with low calculated probability of AD throughout all the time-points were designated as the reference group (n=982, 85.2%). The demographics of participants in each cluster are also provided in Table 1.

## Early onset transient (EOT) AD

This trajectory (n=73, 6.3%) was characterized by a sharp increase in the probability of AD in the first 12 months of life, which subsequently decreased until year 8 (Figure 2). More children in this trajectory were male (65.8%), had a family history of atopy (69%) and owned pets in first 12 months of life (14.3%) compared to the reference group (Table 1). A higher proportion of children in this trajectory had current allergic rhinitis, persistent allergic rhinitis with or without sensitization, current wheezing with or without sensitization and any allergen sensitization compared to the reference group (Table 2).

We also carried out multinomial regression analyses on three models. In the first model of prenatal risk factors, male gender [aOR 1.98 (95% CI 1.16-3.39) p=0.012] and a family history of atopy [aOR 1.85 (95% CI 1.08-3.18) p=0.025] were associated with increased odds of EOT AD. Additionally, being of Indian ancestry was protective for EOT AD [aOR 0.38 (95% CI 0.14-0.98) p=0.046] (Figure 3).

EOT AD was also associated with HDM sensitization at 3 years and wheezing at several time-points (Table 3). Of note, this trajectory was not associated with any risk of food sensitization or rhinitis.

## Early onset persistent (EOP) AD

The EOP AD trajectory (n=24, 2.1%) was marked by a steady increase in the probability of AD from month 3 to month 18 and remained high up till year 8 (Figure 2). More children in this trajectory were from households earning > $4000 USD (47.8%) compared to the reference group (25.4%) (Table 1).

Antenatal and intrapartum antibiotic use [aOR 2.71 (95% CI 1.12-6.58) p=0.027] and breastfeeding were associated with increased odds of EOP AD [aOR 5.14 (95% CI 1.04-25.4) p=0.045] (Figure 3).

EOP AD was the only AD phenotype associated with food sensitization. Apart from current wheezing at 8 years, there were no associations with any other comorbidities (Table 3). Food allergy prevalence is known to be very low in the GUSTO cohort,21 thus inclusion of food allergy outcomes in the multinomial analyses was not possible. Of note, the use of topical steroids, as a proxy of eczema severity, was the highest in this group compared to other trajectories (55.6% at 18 months, 40.9% at 5 years and 38.1% at 8 years).

## Late onset persistent (LOP) AD

In this trajectory (n=73, 6.3%), the probability of AD was lowest among the three AD trajectories in the first year of life; increased only after the 15-month time-point and peaked at 3 years before stabilizing thereafter up till year 8, but remained lower than the EOP trajectory (Figure 2). The majority of the children with LOP AD were from atopic families (72.1%) (Table 1). Frequencies of the various comorbidities observed in this trajectory are presented in Table 2.

In multinomial analyses, family history of atopy was associated with increased odds of LOP AD [aOR 2.56 (95% CI 1.45-4.52) p=0.001] while breastfeeding was associated with reduced odds of LOP AD [aOR 0.21 (95% CI 0.04-0.97) p=0.046] (Figure 3). LOP AD was also associated with increased odds of almost all of the comorbidities assessed at 8 years as well as HDM sensitization at 3 years and 5 years, and wheezing at 3 years (Table 3).

## Sensitivity analyses

To demonstrate reproducibility and robustness of our results, we performed sensitivity analyses based on a different definition of AD, AD symptoms of itchy rash with topical steroid use, as well as parent reported doctor’s diagnosis in the full cohort.

 The AD trajectories generated from the two sensitivity analyses were similar to the original trajectories (Figures S3 and S4 in Supporting Information). Risk factors identified in the sensitivity analyses remained the same in both sensitivity analyses. In the multinomial analyses for associations with allergic co-morbidities, both analyses using the definition of a doctor’s diagnosis of AD demonstrated similar findings (Table S1 Supporting Information), while the analyses using the “itchy rash with steroid use” definition were similar, but not statistically significant for HDM sensitization and current allergic rhinitis in late onset persistent AD (Table S2, Supporting Information).

 We also assessed the sensitivity and specificity of both AD definitions against SCORAD, which is the standard for measurement of AD severity. SCORAD was carried out at 18 months in a subset of participants, by a trained physician who diagnosed AD on physical examination. The “doctor’s diagnosis” definition had a higher sensitivity (sensitivity 93.3%, specificity 90.5%), while itchy rash with use of topical steroids had a high specificity but low sensitivity (sensitivity 50%, specificity 96.7%). The higher sensitivity of the “doctor diagnosis of AD” definition indicated that it performed better than AD symptoms in accurately identifying participants with the disease, and thus this definition was retained in the main analyses.

# Discussion

The unsupervised machine learning approach carried out in this study identified three AD trajectories during the first 8 years of life – early onset transient (EOT), early onset persistent (EOP) and late onset persistent (LOP) AD. The EOT trajectory was linked to non-modifiable risk factors, notably male gender, ethnicity and family history of atopy, and these children tended to develop wheezing later in childhood. The EOP AD trajectory was associated with maternal antibiotic use and breastfeeding. Children in this trajectory also had increased odds of developing food sensitization and wheezing disorders. Of all the trajectories, the LOP AD trajectory showed the strongest links to other atopic risk and allergic outcomes at age 8 years.

Of the more recent studies on AD trajectories, one carried out in an Asian population described four phenotypes: Early-onset AD with high eosinophils and food sensitization; Early-onset, non-allergic AD; Early-onset AD with high C-reactive protein (CRP); and middle-onset AD with inhalant sensitization11. This study provided further evidence to the heterogeneity of AD in early childhood, whereby six variables such as age of onset, age of diagnosis, white blood cell count and eosinophil count, were sufficient to accurately classify 95.5% of Korean children under 3 years of age. However, despite having high accuracy, its utility in the clinical setting is impractical as CRP and eosinophil counts are not routinely performed as part of allergy evaluation.

Our study builds on the current literature by exploring an extended number of demographic and risk factors. In addition, Roduit et al identified four clusters - early onset, transient; early-onset, persistent; late-onset (after age 2); and never/infrequent in a European cohort 22. The highest risk of asthma was again observed in the early persistent AD phenotype, however, the study was restricted to a white European population and was unable to reflect observations in a non-white population. Our present study therefore bridges the gap of existing studies by using a machine learning method to identify latent AD trajectories and assess associations between a wide range of potential risk factors in an Asian population group. Nonetheless, direct comparisons between the various studies should be interpreted cautiously given the different methodologies used.

## Early onset transient AD

In this study we have found that EOT AD was associated with wheezing at multiple time points, including allergic wheeze at age 8 years. The association of early transient AD with persistent wheeze throughout childhood to 8 years with HDM sensitization is unique to our study. Other cohorts, such as the ALSPAC23, 24, Tucson25, ORCA26, PASTURE22 cohorts, which have described early AD phenotypes in relation to allergic outcomes, have linked persistent wheezing/asthma with EOP AD and multiple sensitization. Animal studies suggest that although mild and transient AD may not induce persistent or severe skin inflammation, it is still associated with systemic immune priming and triggering distal airway inflammation 27, 28.

## Early onset persistent AD

Antenatal and intrapartum antibiotic use was associated with increased odds of EOP AD in this study. A possible mechanism for this is through alterations of maternal and fetal microbiota, 29 which in turn modulate fetal immune programming30, 31. Current literature on the links between antibiotic use and AD is contradictory. A recent systematic review evaluating antibiotic use and allergic outcomes found only three publications meeting inclusion criteria 32. One reported a significant relationship with AD 33: another reported a significant relationship only when the intra-partum antibiotic exposure lasted more than 24 hours34 while the other reported a significant relationship only when the mother was atopic and had taken antibiotics throughout pregnancy 35. Intrapartum antibiotics have been associated with dysbiosis, particularly richness and diversity, of intestinal microbiota in full-term newborns 36, however, breastfeeding appears to be able to minimize these deleterious effects37.

Findings on breastfeeding however are more conflicting. We found that breastfeeding was associated with increased odds of EOP AD, and postulate that reverse causation may be a key reason for this observation, whereby pre-existing AD causes a change in breastfeeding behaviour. This is consistent with findings by Paternoster et al24 comparing two cohorts - the ALSPAC and PIAMA cohort. In the ALSPAC cohort, breastfeeding was associated with early onset persistent and late resolving AD trajectories while there was a lack of evidence for an association in the PIAMA cohort. Some studies however, have shown support for breastfeeding as a protective factor for AD38 as is observed in the LOP AD group in our study. These findings suggest that the presence of early infantile AD could have potentially modified breastfeeding practices (i.e. mothers breastfeeding for a longer duration) in the hope that this might reduce AD severity or the risk of other allergic diseases.

Food sensitization was strongly associated with EOP AD. This is consistent with the dual allergen hypothesis where epicutaneous sensitization occurs due to a disrupted skin barrier in early life, and predisposes a child to food sensitization and food allergy 39, 40. Although AD severity assessments were not available for the majority of participants in our cohort, the persistent nature of this AD phenotype suggests more chronic skin inflammation which would increase the risks of epicutaneous food sensitization. Topical steroid usage, a proxy of AD severity, was more prevalent in the EOP than the EOT trajectory, which may account for the association observed with food sensitization in the former but not the latter trajectory. However, detailed analyses for associations were not performed due to the small sample size.

Furthermore, in this study EOP AD does not appear to be a risk factor for other allergic sequelae, contrary to the classical atopic march hypothesis. This runs contrary to other studies where EOP AD has been shown to be associated with the greatest risk of allergic morbidity in later childhood9, 22, 41, 42. Longitudinal cohort studies also now suggest that only a small proportion of children actually follow the typical atopic march trajectory 24, 43, lending support to the notion that multiple allergic phenotypes exist.

## Late onset persistent AD

We found that the LOP AD trajectory was the phenotype most closely associated with allergic airway disorders from 3 years onwards, particularly in children with a family history of atopy. Children in this trajectory were more likely to be HDM sensitized, and this was associated with current/persistent allergic rhinitis and atopic wheezing up to age 8 years. It might be postulated that environmental exposure to HDM from infancy, together with a defective epithelial barrier in the presence of AD, predisposes to acquisition of HDM sensitization that occurs a little later in life, alongside the later onset of AD. This is supported by our prior findings that overall HDM sensitization in this cohort doubled from 11% at 18 months to 22% at 3 years44; increasing to 34.4% by age 5 years and 53.1% by age 8 years.,.

HDM aeroallergen sensitization plays a major role in the pathogenesis of allergic airway disorders in tropical Singapore, primarily due to the high environmental HDM exposures in early life compared to other countries where HDM exposures are lower45, 46. Furthermore, an assessment of TEWL in children up to 12 months showed increased risk of aeroallergen sensitization among those with AD47. 89% of AD children had a positive atopy patch test (APT) result, with a higher mean TEWL observed among those with two or more positive APT. Murine studies similarly suggest that HDM sensitization through the impaired skin barrier can induce systemic and airway inflammation leading to the preferential development of allergic airway disorders 48-52.

## Strengths and limitations

 A key strength of this study is our large cohort which allows subgroups to be defined on the basis of their longitudinal disease trajectories. The unsupervised machine learning approach in our study design maintained prediction accuracy without the need for separate training and validation datasets, while allowing data-driven discovery of latent AD trajectories unbiased by pre-selected investigator-determined definitions. Limitations of this study were mainly the parental-reported outcomes which may be subject to recall bias. However, this was mitigated by the very close intervals between each follow-up time-point, use of ISAAC questionnaires, which have been extensively validated as reliable tools for assessing allergic disease outcomes, and the comprehensive clinical and sensitization profiling in this cohort.

# Conclusion

Collectively, disparities in the longitudinal progression of early, late, transient and persistent AD support the notion that there are distinct subtypes of AD53. We have shown unique characteristics of each AD trajectory which differ from that shown in cohorts in other parts of the world. Unlike the typical atopic march, in which early onset persistent AD precedes allergic morbidity, LOP AD appears to be the strongest phenotype linked to allergic disorders in our cohort, with HDM sensitization appearing to play a major role. Nonetheless, it is yet to be seen whether the trajectories differ at the molecular and immunological level and if these differences translate to differential treatment outcomes53.

This emphasizes the importance of a personalized treatment approach in AD. The routine use of prophylactic emollients from the neonatal period as a proactive AD prevention strategy has been suggested to reduce the risk of AD 54, and this can be made possible through early identification of at-risk infants. This has important clinical implications as evidence suggests that early diagnosis and intervention improves the quality of life for AD patients at high risk of developing comorbidities55, 56. Understanding the various AD phenotypes and endotypes can also guide diagnosis, risk-stratification, treatment and the development of preventive strategies, leading the way towards a phenotype-driven, precision medicine approach to the management of childhood AD. More refined and holistic endotyping predictive models can also be employed with the addition of biomarkers from microbiome, blood and saliva/buccal samples 57. In doing so, earlier identification of at-risk endophenotypes can occur earlier in life during the critical window for intervention.

# References

1. Tsakok T, Marrs T, Mohsin M, et al. Does atopic dermatitis cause food allergy? A systematic review. *J Allergy Clin Immunol* 2016; 137: 1071-1078. 2016/02/22. DOI: 10.1016/j.jaci.2015.10.049.

2. Hudson TJ. Skin barrier function and allergic risk. *Nature Genetics* 2006; 38: 399. DOI: 10.1038/ng0406-399.

3. Baurecht H, Irvine AD, Novak N, et al. Toward a major risk factor for atopic eczema: Meta-analysis of filaggrin polymorphism data. *Journal of Allergy and Clinical Immunology* 2007; 120: 1406-1412. DOI: <https://doi.org/10.1016/j.jaci.2007.08.067>.

4. Belgrave DC, Simpson A, Buchan IE, et al. Atopic Dermatitis and Respiratory Allergy: What is the Link. *Curr Dermatol Rep* 2015; 4: 221-227. 2015/11/14. DOI: 10.1007/s13671-015-0121-6.

5. Ferreira MA, Vonk JM, Baurecht H, et al. Shared genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. *Nat Genet* 2017; 49: 1752-1757. 2017/10/31. DOI: 10.1038/ng.3985.

6. Marenholz I, Esparza-Gordillo J, Ruschendorf F, et al. Meta-analysis identifies seven susceptibility loci involved in the atopic march. *Nature communications* 2015; 6: 8804. 2015/11/07. DOI: 10.1038/ncomms9804.

7. Gupta J, Johansson E, Bernstein JA, et al. Resolving the etiology of atopic disorders by using genetic analysis of racial ancestry. *Journal of Allergy and Clinical Immunology* 2016; 138: 676-699. Review. DOI: 10.1016/j.jaci.2016.02.045.

8. Busse WW. The atopic march: Fact or folklore? *Ann Allergy Asthma Immunol* 2018; 120: 116-118. 2018/02/08. DOI: 10.1016/j.anai.2017.10.029.

9. Kim JP, Chao LX, Simpson EL, et al. Persistence of atopic dermatitis (AD): A systematic review and meta-analysis. *Journal of the American Academy of Dermatology* 2016; 75: 681-687.e611. DOI: <https://doi.org/10.1016/j.jaad.2016.05.028>.

10. Mortz CG, Andersen KE, Dellgren C, et al. Atopic dermatitis from adolescence to adulthood in the TOACS cohort: prevalence, persistence and comorbidities. *Allergy* 2015; 70: 836-845. DOI: 10.1111/all.12619.

11. Seo E, Yoon J, Jung S, et al. Phenotypes of atopic dermatitis identified by cluster analysis in early childhood. *J Dermatol* 2019; 46: 117-123. 2018/12/07. DOI: 10.1111/1346-8138.14714.

12. Eyerich K, Brown SJ, Perez White BE, et al. Human and computational models of atopic dermatitis: A review and perspectives by an expert panel of the International Eczema Council. *J Allergy Clin Immunol* 2019; 143: 36-45. 2018/11/11. DOI: 10.1016/j.jaci.2018.10.033.

13. Brew BK, Chiesa F, Lundholm C, et al. A modern approach to identifying and characterizing child asthma and wheeze phenotypes based on clinical data. *PLoS One* 2019; 14: e0227091. 2019/12/31. DOI: 10.1371/journal.pone.0227091.

14. Spycher BD, Silverman M and Kuehni CE. Phenotypes of childhood asthma: are they real? *Clin Exp Allergy* 2010; 40: 1130-1141. 2010/06/16. DOI: 10.1111/j.1365-2222.2010.03541.x.

15. Soh SE, Tint MT, Gluckman PD, et al. Cohort profile: Growing Up in Singapore Towards healthy Outcomes (GUSTO) birth cohort study. *Int J Epidemiol* 2014; 43: 1401-1409. DOI: 10.1093/ije/dyt125.

16. Bunyavanich S, Rifas-Shiman SL, Platts-Mills TA, et al. Peanut, milk, and wheat intake during pregnancy is associated with reduced allergy and asthma in children. *J Allergy Clin Immunol* 2014; 133: 1373-1382. DOI: 10.1016/j.jaci.2013.11.040.

17. Hardjojo A, Goh A, Shek LP, et al. Rhinitis in the first 18 months of life: exploring the role of respiratory viruses. *Pediatr Allergy Immunol* 2015; 26: 25-33. 2015/01/06. DOI: 10.1111/pai.12330.

18. Bousquet J, Khaltaev N, Cruz AA, et al. Allergic Rhinitis and its Impact on Asthma (ARIA) 2008 update (in collaboration with the World Health Organization, GA(2)LEN and AllerGen). *Allergy* 2008; 63 Suppl 86: 8-160. 2008/03/26. DOI: 10.1111/j.1398-9995.2007.01620.x.

19. Brozek JL, Bousquet J, Agache I, et al. Allergic Rhinitis and its Impact on Asthma (ARIA) guidelines-2016 revision. *J Allergy Clin Immunol* 2017; 140: 950-958. 2017/06/13. DOI: 10.1016/j.jaci.2017.03.050.

20. Witten IH, Frank E, Hall MA, et al. *Data Mining, Fourth Edition: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann Publishers Inc., 2016.

21. Tham EH, Lee BW, Chan YH, et al. Low Food Allergy Prevalence Despite Delayed Introduction of Allergenic Foods-Data from the GUSTO Cohort. *J Allergy Clin Immunol Pract* 2018; 6: 466-475.e461. 2017/07/25. DOI: 10.1016/j.jaip.2017.06.001.

22. Roduit C, Frei R, Depner M, et al. Phenotypes of Atopic Dermatitis Depending on the Timing of Onset and Progression in Childhood. *JAMA Pediatr* 2017; 171: 655-662. 2017/05/23. DOI: 10.1001/jamapediatrics.2017.0556.

23. Belgrave DC, Granell R, Simpson A, et al. Developmental profiles of eczema, wheeze, and rhinitis: two population-based birth cohort studies. *PLoS Med* 2014; 11: e1001748. 2014/10/22. DOI: 10.1371/journal.pmed.1001748.

24. Paternoster L, Savenije OEM, Heron J, et al. Identification of atopic dermatitis subgroups in children from 2 longitudinal birth cohorts. *J Allergy Clin Immunol* 2018; 141: 964-971. 2017/11/14. DOI: 10.1016/j.jaci.2017.09.044.

25. Taussig LM, Wright AL, Holberg CJ, et al. Tucson Children's Respiratory Study: 1980 to present. *J Allergy Clin Immunol* 2003; 111: 661-675; quiz 676. 2003/04/22. DOI: 10.1067/mai.2003.162.

26. Amat F, Saint-Pierre P, Bourrat E, et al. Early-onset atopic dermatitis in children: which are the phenotypes at risk of asthma? Results from the ORCA cohort. *PLoS One* 2015; 10: e0131369. 2015/06/25. DOI: 10.1371/journal.pone.0131369.

27. Ubags ND, Trompette A, Pernot J, et al. Microbiome-induced antigen-presenting cell recruitment coordinates skin and lung allergic inflammation. *J Allergy Clin Immunol* 2020 2020/07/18. DOI: 10.1016/j.jaci.2020.06.030.

28. Leyva-Castillo JM, Yoon J and Geha RS. IL-22 promotes allergic airway inflammation in epicutaneously sensitized mice. *J Allergy Clin Immunol* 2019; 143: 619-630 e617. 2018/06/20. DOI: 10.1016/j.jaci.2018.05.032.

29. Rautava S, Luoto R, Salminen S, et al. Microbial contact during pregnancy, intestinal colonization and human disease. *Nature reviews Gastroenterology & hepatology* 2012; 9: 565-576. 2012/08/15. DOI: 10.1038/nrgastro.2012.144.

30. Jenmalm MC. The mother-offspring dyad: microbial transmission, immune interactions and allergy development. *J Intern Med* 2017; 282: 484-495. 2017/07/21. DOI: 10.1111/joim.12652.

31. Gomez de Agüero M, Ganal-Vonarburg SC, Fuhrer T, et al. The maternal microbiota drives early postnatal innate immune development. *Science* 2016; 351: 1296. DOI: 10.1126/science.aad2571.

32. Baron R, Taye M, der Vaart IB, et al. The relationship of prenatal antibiotic exposure and infant antibiotic administration with childhood allergies: a systematic review. *BMC pediatrics* 2020; 20: 312. 2020/07/01. DOI: 10.1186/s12887-020-02042-8.

33. Dom S, Droste JH, Sariachvili MA, et al. Pre- and post-natal exposure to antibiotics and the development of eczema, recurrent wheezing and atopic sensitization in children up to the age of 4 years. *Clin Exp Allergy* 2010; 40: 1378-1387. 2010/06/16. DOI: 10.1111/j.1365-2222.2010.03538.x.

34. Wohl DL, Curry WJ, Mauger D, et al. Intrapartum antibiotics and childhood atopic dermatitis. *J Am Board Fam Med* 2015; 28: 82-89. 2015/01/09. DOI: 10.3122/jabfm.2015.01.140017.

35. Timm S, Schlünssen V, Olsen J, et al. Prenatal antibiotics and atopic dermatitis among 18-month-old children in the Danish National Birth Cohort. *Clin Exp Allergy* 2017; 47: 929-936. 2017/03/10. DOI: 10.1111/cea.12916.

36. Gasparrini AJ, Crofts TS, Gibson MK, et al. Antibiotic perturbation of the preterm infant gut microbiome and resistome. *Gut Microbes* 2016; 7: 443-449. 2016/07/30. DOI: 10.1080/19490976.2016.1218584.

37. Azad MB, Konya T, Persaud RR, et al. Impact of maternal intrapartum antibiotics, method of birth and breastfeeding on gut microbiota during the first year of life: a prospective cohort study. *Bjog* 2016; 123: 983-993. 2015/09/29. DOI: 10.1111/1471-0528.13601.

38. Ip S, Chung M, Raman G, et al. Breastfeeding and maternal and infant health outcomes in developed countries. *Evidence report/technology assessment* 2007: 1-186. 2007/09/04.

39. Flohr C, Perkin M, Logan K, et al. Atopic dermatitis and disease severity are the main risk factors for food sensitization in exclusively breastfed infants. *J Invest Dermatol* 2014; 134: 345-350. 2013/07/23. DOI: 10.1038/jid.2013.298.

40. Lack G. Epidemiologic risks for food allergy. *Journal of Allergy and Clinical Immunology* 2008; 121: 1331-1336. DOI: 10.1016/j.jaci.2008.04.032.

41. Hu C, Nijsten T, van Meel ER, et al. Eczema phenotypes and risk of allergic and respiratory conditions in school age children. *Clin Transl Allergy* 2020; 10: 7. 2020/02/27. DOI: 10.1186/s13601-020-0310-7.

42. Lowe AJ, Angelica B, Su J, et al. Age at onset and persistence of eczema are related to subsequent risk of asthma and hay fever from birth to 18 years of age. *Pediatr Allergy Immunol* 2017; 28: 384-390. 2017/03/17. DOI: 10.1111/pai.12714.

43. Belgrave DCM, Granell R, Simpson A, et al. Developmental Profiles of Eczema, Wheeze, and Rhinitis: Two Population-Based Birth Cohort Studies. *PLOS Medicine* 2014; 11: e1001748. DOI: 10.1371/journal.pmed.1001748.

44. Loo EX, Sim JZ, Goh A, et al. Predictors of allergen sensitization in Singapore children from birth to 3 years. *Allergy, asthma, and clinical immunology : official journal of the Canadian Society of Allergy and Clinical Immunology* 2016; 12: 56. 2016/11/02. DOI: 10.1186/s13223-016-0161-x.

45. Andiappan AK, Puan KJ, Lee B, et al. Allergic airway diseases in a tropical urban environment are driven by dominant mono-specific sensitization against house dust mites. *Allergy* 2014; 69: 501-509. 2014/01/25. DOI: 10.1111/all.12364.

46. Chew FT, Zhang L, Ho TM, et al. House dust mite fauna of tropical Singapore. *Clin Exp Allergy* 1999; 29: 201-206. 1999/03/03. DOI: 10.1046/j.1365-2222.1999.00493.x.

47. Boralevi F, Hubiche T, Leaute-Labreze C, et al. Epicutaneous aeroallergen sensitization in atopic dermatitis infants - determining the role of epidermal barrier impairment. *Allergy* 2008; 63: 205-210. 2008/01/12. DOI: 10.1111/j.1398-9995.2007.01556.x.

48. Stremnitzer C, Manzano-Szalai K, Starkl P, et al. Epicutaneously applied Der p 2 induces a strong TH 2-biased antibody response in C57BL/6 mice, independent of functional TLR4. *Allergy* 2014; 69: 741-751. 2014/04/17. DOI: 10.1111/all.12399.

49. Shimura S, Takai T, Iida H, et al. Epicutaneous Allergic Sensitization by Cooperation between Allergen Protease Activity and Mechanical Skin Barrier Damage in Mice. *J Invest Dermatol* 2016; 136: 1408-1417. 2016/03/19. DOI: 10.1016/j.jid.2016.02.810.

50. Coquet JM, Schuijs MJ, Smyth MJ, et al. Interleukin-21-Producing CD4(+) T Cells Promote Type 2 Immunity to House Dust Mites. *Immunity* 2015; 43: 318-330. 2015/08/20. DOI: 10.1016/j.immuni.2015.07.015.

51. Deckers J, Sichien D, Plantinga M, et al. Epicutaneous sensitization to house dust mite allergen requires interferon regulatory factor 4-dependent dermal dendritic cells. *J Allergy Clin Immunol* 2017; 140: 1364-1377 e1362. 2017/02/13. DOI: 10.1016/j.jaci.2016.12.970.

52. Lee HJ, Lee NR, Jung M, et al. Atopic March from Atopic Dermatitis to Asthma-Like Lesions in NC/Nga Mice Is Accelerated or Aggravated by Neutralization of Stratum Corneum but Partially Inhibited by Acidification. *J Invest Dermatol* 2015; 135: 3025-3033. 2015/09/25. DOI: 10.1038/jid.2015.333.

53. Wan J, Mitra N, Hoffstad OJ, et al. Longitudinal atopic dermatitis control and persistence vary with timing of disease onset in children: A cohort study. *J Am Acad Dermatol* 2019; 81: 1292-1299. 2019/05/16. DOI: 10.1016/j.jaad.2019.05.016.

54. Lowe AJ, Leung DYM, Tang MLK, et al. The skin as a target for prevention of the atopic march. *Ann Allergy Asthma Immunol* 2018; 120: 145-151. 2018/02/08. DOI: 10.1016/j.anai.2017.11.023.

55. Irvine AD and Mina-Osorio P. Disease trajectories in childhood atopic dermatitis: an update and practitioner's guide. *Br J Dermatol* 2019; 181: 895-906. 2019/02/14. DOI: 10.1111/bjd.17766.

56. Boguniewicz M, Nicol N, Kelsay K, et al. A multidisciplinary approach to evaluation and treatment of atopic dermatitis. *Seminars in cutaneous medicine and surgery* 2008; 27: 115-127. 2008/07/16. DOI: 10.1016/j.sder.2008.05.001.

57. Paller AS, Spergel JM, Mina-Osorio P, et al. The atopic march and atopic multimorbidity: Many trajectories, many pathways. *J Allergy Clin Immunol* 2019; 143: 46-55. 2018/11/21. DOI: 10.1016/j.jaci.2018.11.006.

# Tables

**Table 1** Demographics of study participants in each trajectory.

|  | Reference group(n=982)N (%) | Early onset transient(n=73)N (%) | Early onset persistent(n=24)N (%) | Late onset persistent(n=73)N (%) |
| --- | --- | --- | --- | --- |
| Male | 468 (51) | ***48 (65.8)*** | 15 (62.5) | 40 (54.8) |
| Family history of atopy  | 410 (51.3) | ***49 (69)*** | 17 (70.8) | ***49 (72.1)*** |
| Household income > USD 4000  |  232 (25.4) | 26 (36.6) | 11 (47.8) | 21 (30) |
| Ethnicity |  |  |  |  |
| *Chinese* | 520 (53) | 43 (58.9) | 16 (66.7) | 47 (64.4) |
| *Malay* | 267 (27.2) | 24 (32.9) | 7 (29.2) | 16 (21.9) |
| *Indian* | 193 (19.7) | 6 (8.2) | 1 (4.2) | 10 (13.7) |
| *Others* | 2 (0.2) | 0 (0) | 0 (0) | 0 (0) |
| Presence of siblings | 533 (58.1) | 37 (50.7) | 12 (50) | 38 (52.1) |
| Antenatal & intrapartum antibiotics  | 404 (44) | 33 (45.8) | 16 (66.7) | 30 (41.1) |
| Caesarean delivery | 270 (29.4) | 26 (36.1) | 4 (16.7) | 22 (30.1) |
| Breastfeeding level |  |  |  |  |
| *Low* | 381 (47.7) | 26 (36.6) | 5 (21.7) | 27 (39.7) |
| *Mid* | 319 (40) | 36 (50.7) | 14 (60.9) | 39 (57.4) |
| *High* | 98 (12.3) | 9 (12.7) | 4 (17.4) | 2 (2.9) |
| Pet ownership in first year of life | 47 (6.1) | ***10 (14.3)*** | 4 (16.7) | 5 (7.5) |
| Household exposure to tobacco smoke | 226 (38.7) | 27 (42.9) | 3 (15.8) | 20 (35.1) |
| Childcare attendance in first year of life | 53 (8.3) | 11 (16.9) | 5 (25) | 3 (5.1) |
| Infant antibiotic use in first year of life | 287 (46.1) | 33 (50) | 12 (57.1) | 19 (34.5) |

Cells in bold and italics indicate significance at p value ≤0.05 compared to reference group in univariate analysis after Bonferroni correction.

**Table 2** Prevalence of comorbidities at 8 years in each AD trajectory.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|   | **Reference group** | **Early onset transient** | **Early onset persistent** | **Late onset persistent** |
| **(n=982)** | **(n=73)** | **(n=24)** | **(n=73)** |
| N | **n (%)** | N | **n (%)** | N | **n (%)** | N | **n (%)** |
| **AR diagnosis (n=744)** | 604 | 27 (4.5) | 59 | 4 (6.8) | 24 | ***5 (20.8)*** | 57 | ***9 (15.8)*** |
| **Current AR (n=744)** | 604 | 101 (16.7) | 59 | ***18 (30.5)*** | 24 | ***9 (37.5)*** | 57 | ***22 (38.6)*** |
| **Current AR + sensitization (n=734)** | 599 | 73 (12.2) | 58 | 11 (19) | 22 | 6 (27.3) | 55 | ***18 (32.7)*** |
| **Persistent AR (n=483)** | 385 | 32 (8.3) | 49 | ***10 (20.4)*** | 17 | 4 (23.5) | 32 | ***11 (34.4)*** |
| **Persistent AR + sensitization (n=483)** | 385 | 19 (4.9) | 49 | ***7 (14.3)*** | 17 | 2 (11.8) | 32 | ***8 (25)*** |
| **Current wheezing (n=722)** | 586 | 24 (4.1) | 59 | ***10 (16.9)*** | 24 | ***8 (33.3)*** | 53 | 6 (11.3) |
| **Current wheezing + sensitization (n=350)** | 301 | 15 (5) | 21 | ***9 (42.9)*** | 8 | ***5 (62.5)*** | 20 | 4 (20) |
| **Food allergic (n=736)** | 600 | 3 (0.5) | 58 | 0 (0) | 22 | 1 (4.5) | 56 | 0 (0) |
| **Food sensitized tolerant (n=736)** | 600 | 13 (2.2) | 58 | 5 (8.6) | 22 | 3 (13.6) | 56 | 4 (7.1) |
| **Any sensitization (n=716)** | 586 | 287 (49.1) | 57 | ***44 (77.2)*** | 20 | ***16 (80)*** | 53 | ***36 (67.9)*** |
| **Any HDM sensitization (n=716)** | 586 | 284 (48.5) | 57 | ***44 (77.2)*** | 20 | ***16 (80)*** | 53 | ***36 (67.9)*** |
| **Any food sensitization (n=716)** | 585 | 36 (6.2) | 57 | ***10 (17.5)*** | 20 | 4 (20) | 54 | 6 (11.1) |

AR: Allergic Rhinitis; HDM: House Dust Mite

Any sensitization: positive skin prick test result to any of the house dust mite allergens (*Dermatophagoides pteronyssinus*, *Dermatophagoides farinae, Blomia tropicalis*), food allergens (egg, peanut, cow’s milk, shrimp, crab) or aeroallergens (*Blatella germanica and Periplanata Americana).*

Cells in bold and italics indicate significance at p value ≤0.05 compared to reference group in univariate analysis after Bonferroni correction.

**Table 3** Multinomial logistic regression of comorbidities at 18 months, 3 years, 5 years and 8 years with AD trajectoriesa

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Early onset transient** | **Early onset persistent** | **Late onset persistent** |
|  | **OR (95% CI)** | **P value** | **OR (95% CI)** | **P value** | **OR (95% CI)** | **P value** |
| **18 months (n=548)** |
| **HDM sensitization**  | 3.15 (1.4-7.06) | **0.005** | 3.13 (0.76-12.97) | 0.115 | 1.6 (0.6-4.23) | 0.346 |
| **Food sensitization**  | 2.94 (0.83-10.45) | 0.096 | 16.94 (3.14-91.35) | **0.001** | 1.67 (0.32-8.79) | 0.546 |
| **AR** | 1.1 (0.4-3.05) | 0.853 | 1.41 (0.23-8.57) | 0.709 | 1.57 (0.62-4.01) | 0.343 |
| **Wheezing**  | 1.04 (0.25-4.3) | 0.961 | 0 (0-.c) | 0.995 | 2.02 (0.56-7.33) | 0.285 |
|  **36 months (n=524)** |
| **HDM sensitization**  | 4.34 (2.25-8.4) | **<0.001** | 1.55 (0.43-5.59) | 0.506 | 1.95 (0.96-3.96) | 0.066 |
| **Food sensitization**  | NA | NA | NA | NA | NA | NA |
| **AR** | 2.11 (0.98-4.57) | 0.057 | 0.58 (0.1-3.36) | 0.54 | 2.55 (1.2-5.42) | **0.015** |
| **Wheezing**  | 2.65 (1.09-6.47) | **0.032** | 2.41 (0.49-11.77) | 0.276 | 3.84 (1.64-8.99) | **0.002** |
| **60 months (n=482)** |
| **HDM sensitization** | 2.28 (1.1-4.73) | **0.027** | 1.61 (0.39-6.67) | 0.514 | 4.77 (2.2-10.35) | **<0.001** |
| **Food sensitization** | 1.36 (0.32-5.88) | 0.679 | 24.6 (3.97-152.43) | **0.001** | 1.37 (0.35-5.41) | 0.655 |
| **AR** | 1.48 (0.46-4.76) | 0.511 | 0.25 (0.02-3.47) | 0.301 | 1.27 (0.38-4.28) | 0.697 |
| **Wheezing** | 2.82 (1.09-7.29) | **0.033** | 1.69 (0.29-9.94) | 0.562 | 0.33 (0.07-1.57) | 0.165 |
| **8 yearsb** |
| **Current AR (n=462)** | 1.19 (0.56-2.52) | 0.656 | 0.75 (0.17-3.34) | 0.71 | 2.52 (1.14-5.57) | **0.022** |
| **Current wheezing (n=462)** | 3.9 (1.44-10.61) | **0.008** | 27.27 (4.61-161.42) | **<0.001** | 1.39 (0.34-5.71) | 0.645 |
| **Any sensitization (n=462)** | 2.06 (0.99-4.28) | 0.053 | 2.35 (0.55-9.97) | 0.248 | 2.27 (1-5.16) | 0.051 |
| **Any HDM sensitization (n=462)** | 2.09 (1-4.35) | **0.05** | 2.38 (0.56-10.14) | 0.241 | 2.28 (1-5.19) | **0.05** |
| **Any food sensitization (n=463)** | 2.44 (0.92-6.46) | 0.072 | 4.66 (0.84-25.72) | 0.077 | 2.09 (0.67-6.51) | 0.205 |
| **Current AR + sensitization (n=478)** | 0.91 (0.38-2.18) | 0.831 | 0.8 (0.19-3.46) | 0.768 | 3.53 (1.58-7.88) | **0.002** |
| **Current wheezing + sensitization (n=223)** | 9.35 (1.87-46.91) | **0.007** | 696.56 (1.99-243432.57) | **0.028** | 9.1 (1.24-66.84) | **0.03** |
| **Persistent AR (n=365)** | 1.54 (0.58-4.14) | 0.389 | 0.63 (0.03-12.66) | 0.761 | 5.35 (1.87-15.35) | **0.002** |
| **Persistent AR + sensitization (n=376)** | 1.63 (0.51-5.2) | 0.408 | 1.11 (0.06-20.23) | 0.946 | 4.94 (1.53-15.91) | **0.007** |

a All analyses were adjusted for prenatal and environmental factors shown in Figure 3, including pet ownership, household smoking, childcare attendance and infant’s use of antibiotics during the first year. All comparisons were made with the non-AD group as the reference group.

b For this model, multinomial logistic regressions were carried out stepwise adding each non-overlapping variable to the model separately. This was repeated until all the variables were included in a model. As a result, the numbers included in each analysis vary as shown in the table.

AR: Allergic Rhinitis; HDM: House dust mite; NA: Not applicable

Any sensitization: positive skin prick test result to any of the house dust mite allergens (*Dermatophagoides pteronyssinus*, *Dermatophagoides farinae, Blomia tropicalis*), food allergens (egg, peanut, cow’s milk, shrimp, crab) or aeroallergens (*Blatella germanica and Periplanata Americana).*

# Figure Legend

**Figure 1** Flow diagram of participants with and without AD at each timepoint.

**Figure 2** Three atopic dermatitis (AD) and one reference trajectory generated from the unsupervised machine learning approach.

Figure legend: The unsupervised machine learning method identified four atopic dermatitis trajectories. Early onset transient (EOT) was made up of children with high probability of reporting AD in the first 12 months of life and this subsequently decreased up till year 8. The early onset persistent (EOP) AD trajectory was marked by a gradual increase in probability of AD from month 3 to month 18 and this remained high throughout till year 8. In the late onset persistent (LOP) trajectory, probability of AD was low in the first year of life, peaking at 3 years and stabilized thereafter up till year 8. Those with low probability of reporting atopic dermatitis were classified in the reference group.

**Figure 3** Forest plot representing multinomial regression analyses for prenatal and environmental factors.

Figure legend: Multinomial logistic regressions were used to examine associations between prenatal and environmental risk factors with each trajectory. ♦ indicate significant factors, p<0.05. x-axes are on a logarithmic scale. All comparisons were made with the non-AD group as the reference group. Reference groups for the categorical variables are as follows in the order presented in the figure: female, no family history of atopy, household income < 4000 USD, Chinese race, no siblings, no antibiotic use during labour and pregnancy, vaginal delivery and low breastfeeding level.

**Supporting Information Figures**

**Figure S1 AD trajectories among those with complete data at all twelve timepoints (n=448).**

Figure legend: Apart from the late onset persistent trajectory, all other trajectories were consistent with the trajectories of the main analyses. This is possibly due to the smaller sample size (n=448), which renders the model less stable and resulted in loss of power to perform downstream analyses.

**Figure S2 AD trajectories in those with A) missing data at 1 timepoint, B) missing data at 4 timepoints, C) missing data at 7 timepoints and D) missing data at 11 timepoints.**

Figure legend: Trajectories of participants with varying numbers of missing values (those with complete data on 11 or less timepoints, out of a maximum of 12 time-points), were all consistent with the main analysis.

**Figure S3 AD trajectories based on the definition of doctor diagnosis of AD in full cohort**

Figure legend: As a measure of comparison, analyses based on the definition of doctor diagnosis of AD in the full cohort was also performed (n=1237). Similarly, three AD trajectories were obtained.

**Figure S4 AD trajectories based on the definition of itchy rash and use of topical steroids in full cohort.**

Figure legend: Due to the small numbers in some of the exposure variables, analysis based on definition of itchy rash and use of topical steroids was carried out on the full cohort (n=1237). Similar to the main analysis, the machine learning method identified three AD trajectories.