

Differential factors associated with challenge-proven food allergy phenotypes in a population cohort of infants: a latent class analysis

R. L. Peters^{1,2}, K. J. Allen^{1,2,3,4}, S. C. Dharmage^{1,5}, C. J. Lodge^{1,5}, J. J. Koplin^{1,5}, A.-L. Ponsonby^{1,5}, M. Wake^{1,2}, A. J. Lowe^{1,5}, M. L. K. Tang^{1,2,3}, M. C. Matheson⁵ and L. C. Gurrin^{1,5} for the HealthNuts study

¹Murdoch Childrens Research Institute, Parkville, Vic., Australia, ²Department of Paediatrics, University of Melbourne, Parkville, Vic., Australia,

³Department of Allergy and Immunology, Royal Children's Hospital, Parkville, Vic., Australia, ⁴School of Inflammation and Repair, The University of

Manchester, Manchester, UK and ⁵Melbourne School of Population and Global Health, Centre for Epidemiology and Biostatistics, University of Melbourne, Parkville, Vic., Australia

Clinical & Experimental Allergy

Summary

Background Food allergy, eczema and wheeze are early manifestations of allergic disease and commonly co-occur in infancy although their interrelationship is not well understood. Data from population studies are essential to determine whether there are differential drivers of multi-allergy phenotypes. We aimed to define phenotypes and risk factors of allergic disease using latent class analysis (LCA).

Methods The HealthNuts study is a prospective, population-based cohort of 5276 12-month-old infants in Melbourne, Australia. LCA was performed using the following baseline data collected at age 12 months: food sensitization (skin prick test ≥ 2 mm) and allergy (oral food challenge) to egg, peanut and sesame; early (< 4 months) and late-onset eczema; and wheeze in the first year of life. Risk factors were modelled using multinomial logistic regression.

Results Five distinct phenotypes were identified: no allergic disease (70%), non-food-sensitized eczema (16%), single egg allergy (9%), multiple food allergies (predominantly peanut) (3%) and multiple food allergies (predominantly egg) (2%). Compared to the baseline group of no allergic disease, shared risk factors for all allergic phenotypes were parents born overseas (particularly Asia), delayed introduction of egg, male gender (except for single egg allergy) and family history of allergic disease, whilst exposure to pet dogs was protective for all phenotypes. Other factors including filaggrin mutations, vitamin D and the presence of older siblings differed by phenotype.

Conclusions and clinical relevance Multiple outcomes in infancy can be used to determine five distinct allergy phenotypes at the population level, which have both shared and separate risk factors suggesting differential mechanisms of disease.

Keywords eczema, food allergy, latent class analysis, phenotypes, population study

Submitted 28 August 2014; revised 22 October 2014; accepted 21 November 2014

Correspondence:

Professor Katrina J. Allen, Murdoch Childrens Research Institute, Royal Children's Hospital, Flemington Road, Parkville 3052, Vic., Australia.

E-mail: katie.allen@rch.org.au

Cite this as: R. L. Peters, K. J. Allen, S. C. Dharmage, C. J. Lodge, J. J. Koplin, A.-L. Ponsonby, M. Wake, A. J. Lowe, M. L. K. Tang, M. C. Matheson, L. C. Gurrin and for the HealthNuts study, *Clinical & Experimental Allergy*, 2015 (45) 953–963.

Introduction

Children with eczema and food allergy in infancy are at higher risk of developing asthma and rhinitis in childhood, a progression commonly referred to as the atopic march [1–3]. The expression of allergic disease in childhood varies and evidence suggests that these variations are associated with different risks of other allergic diseases later in life [4–7]. The interrelationships between food allergy, eczema and wheeze in infancy are not well understood, and the phenotypic heterogeneity of

allergic diseases in early life has not been well characterized.

Previous attempts to identify distinct allergic phenotypes have been based on essentially arbitrary classifications, such as the presence of specific sensitizations, or timing of onset and remission [4, 5, 8]. An alternative is to use a statistical modelling approach to identify phenotypes of allergy, such as latent class analysis (LCA). An LCA detects the presence of latent or unobserved classes at the population level based on patterns of association between multiple outcomes. Methods

based on LCA have been used to identify multiple phenotypes of sensitizations and allergic disease, particularly wheeze and asthma [9–14]. Not only has this method identified novel phenotypes of wheeze, differential risk factors for the phenotypes suggest that the underlying mechanism for developing wheezing diseases differs [9, 13–15].

Only two studies using LCA have focused on the expression of allergic disease in the first 2 years of life, however, neither considered food allergy [11, 12]. Inclusion of food allergy in the determination of phenotypes is important because it is one of the first objective expressions of allergic disease. In this study, we used LCA to synthesize multiple allergic conditions into distinct phenotypes to characterize the expression of allergic disease in infancy, with an emphasis on challenge-proven food allergy outcomes. In addition, we aimed to identify parental, infant and environmental risk factors for these phenotypes to explore mechanisms of disease development.

Methods

Study design and recruitment

The methods of 'HealthNuts', a population-based, longitudinal food allergy study in Melbourne, Australia, have been described previously [16]. Briefly, 5276 infants aged 11–15 months were recruited through 131 council-run immunization sessions from September 2007 to August 2011. Parental questionnaires captured information on demographic factors, environmental exposures and common childhood illnesses in the first year of life. Infants underwent skin prick testing (SPT) to four common food allergens – peanut, hen's egg, sesame and either cow's milk or shrimp (ALK, Spain), along with positive (10 mg/mL histamine) and negative (saline) controls. Infants with a detectable SPT wheal (≥ 1 mm) were invited for an oral food challenge (OFC), along with a random sample of 200 infants with negative SPT (0 mm) (OFC controls) using methods previously described [16, 17]. Genotyping to identify filaggrin gene (FLG) loss-of-function mutations was performed with the Sequenom MassARRAY system (Sequenom, San Diego, CA, USA) [18]. Serum 25(OH)D3 level was measured using liquid chromatography–tandem mass spectrometry at the Royal Melbourne Institute of Technology [19]. The online repository contains further detail on study methods.

Definitions for variables used in the latent class analysis model

Sensitization: SPT ≥ 2 mm using the recruitment SPT.

IgE-mediated Food allergy: positive OFC (at least one of the following objective reactions within 2 h of the last challenge dose: three or more concurrent, non-contact urticarial lesions lasting ≥ 5 min, vomiting, angioedema or anaphylaxis) in the presence of positive test of sensitization at clinic attendance (SPT ≥ 2 mm or sIgE ≥ 0.35 kua/L (ImmunoCAP FEIA)).

Food tolerance: negative OFC or negative SPT at community recruitment.

Eczema: parent report of a doctor diagnosis of eczema; early onset (< 4 months) or late onset (≥ 4 months).

Wheeze: parent report of a wheezing episode ever in the first year of life.

Potential risk factors

Risk factors were identified from review of the literature and from our own previously published work and collected from the parent questionnaire at recruitment [18–22]. They were divided into three categories: parental, infant and environmental. Genetic polymorphisms (FLG) and vitamin D insufficiency (≤ 50 nmol/L) were only available on the subgroup of infants who attended the HealthNuts study clinic for OFC and consented to a blood test ($n = 835$ and $n = 707$, respectively.)

Ethical conduct in human research

Ethics approval was obtained from the Victorian State Government (Office for Children, (CDF/07/492) and Department of Human Services (10/07)) and the Royal Children's Hospital (27047).

Statistical methods

We allowed the definition of the latent classes to depend on multiple outcomes and estimated the parameters of the underlying statistical model using patterns among the observed outcome data at the individual level. The LCA model can be used to generate estimates of (i) the population prevalence of each of the classes, (ii) the *item response probability* which can be interpreted as the class-specific prevalence of each outcome variable and (iii) participant-specific *posterior probabilities* of belonging to each class. To determine the number of classes, we fitted the LCA model multiple times with 1–7 classes. The optimal number of classes was determined by considering the following criteria: convergence of iterative routine for parameter estimation and goodness-of-fit statistics (Akaike's information criteria (AIC), the corrected AIC, the Bayesian information criteria (BIC) and the adjusted BIC).

Summary data for categorical risk factors for each class are presented as proportions and compared using the Wald test; classes are weighted for each individual participant's posterior probability of membership for that class. Multinomial logistic regression (MLR) was used to determine risk factors for each class, also weighted for posterior probabilities of class membership. Three separate multivariable models were fitted for the three categories of risk factors (parental, infant and environmental). Risk factors within each category were included simultaneously in the multivariable model if the *P*-value was < 0.1 for the association between the risk factor and any class from a univariable MLR analysis. After adjustment, any significant parental risk factors were retained as potential confounders of the association between infant and environmental risk factors, and the classes. Likewise, any infant risk factors that remained significantly associated with an allergy phenotype in the multivariable model were included as potential confounders in the environmental risk factor model. FLG and vitamin D levels were only available for a subgroup of infants. The MLR was fitted separately for both of these variables. Other risk factors were considered as potential confounding variables and retained in the model if they altered the crude OR by $> 10\%$. Finally, we tested for an interaction between vitamin D levels and parents country of birth, because we have previously found that the association between vitamin D insufficiency and food allergy is modified by parent's country of birth [19]. Stata 12 and the LCA Stata plug-in were used for statistical analyses [23].

Results

Study participation and determination of allergic phenotypes

Of the 7134 eligible infants, 5276 infants participated (74% response rate). Of these, 5142 underwent SPT to egg, peanut or sesame and 1089 infants were eligible for hospital assessment of whom 908 participated in OFC (83% participation). Data from all 5276 infants were included in the LCA model. The following variables were included in the LCA model: sensitization to egg, peanut or sesame, IgE-mediated food allergy to egg, peanut or sesame, early- and late-onset eczema and wheeze (Table 1). The 5-class model was selected (Table S1 online repository). Item response probabilities for the five phenotypes of allergic disease based on the five classes are shown in Fig. 1 and Table S2 (online repository). Two phenotypes had no evidence of food sensitization, whilst the remaining phenotypes were differentiated by the probability of food allergy.

Table 1. Variables used in the latent class analysis model

Characteristic	Prevalence
Egg sensitized	14% (696/5127)
Egg allergy	9% (471/4982)
Peanut sensitized	8% (390/5129)
Peanut allergy	3% (156/5080)
Sesame sensitized	2% (96/5074)
Sesame allergy	0.7% (33/5097)
No eczema	74% (3657/4931)
Early-onset eczema (< 4 months old)	12% (592/4931)
Late-onset eczema (≥ 4 months old)	14% (682/4931)
Wheeze ever	18% (821/4603)

Class 1: No allergic disease (70% of the cohort). The no allergic disease class had the highest estimated population prevalence and was characterized by the absence of food sensitization, food allergy and eczema. The estimated prevalence of an episode of wheeze was 17%.

Class 2: Non-food-sensitized eczema (16% of the cohort). The estimated prevalence of eczema was 93%, although it was predominately late-onset (≥ 4 months) eczema (60%) compared to early-onset (33%). Food sensitizations were notably absent (all below 6%), whilst the estimated prevalence of wheeze was 24%.

Class 3: Single egg allergy (9% of the cohort). The estimated prevalence of allergy to egg alone was 95%, with little allergy or sensitization to other foods. Half the infants had eczema (27% early-onset and 23% late-onset (≥ 4 months)) and 15% had wheeze.

Class 4: Multiple food allergies (predominantly peanut) (3% of the cohort). The estimated prevalence of peanut sensitization and peanut allergy were 100% and 89%, respectively. Approximately half (56%) were also allergic to egg, but none were allergic to sesame. The estimated prevalence of eczema was 72%, with early-onset eczema being more common (42%), whilst 23% had a history of wheezing.

Class 5: Multiple food allergies (predominantly egg) (2% of the cohort). Infants in this class were likely to have multiple food allergies: egg (estimated prevalence 91%), sesame (50%) and peanut (39%). Early-onset eczema had the highest estimated prevalence among all classes (58%), and an additional 31% were estimated to have reported late-onset (≥ 4 months) eczema. The estimated prevalence of the history of wheeze was 14%.

Parental risk factors for allergy phenotypes

Demographic characteristics and the prevalence of risk factors considered are presented for each allergy

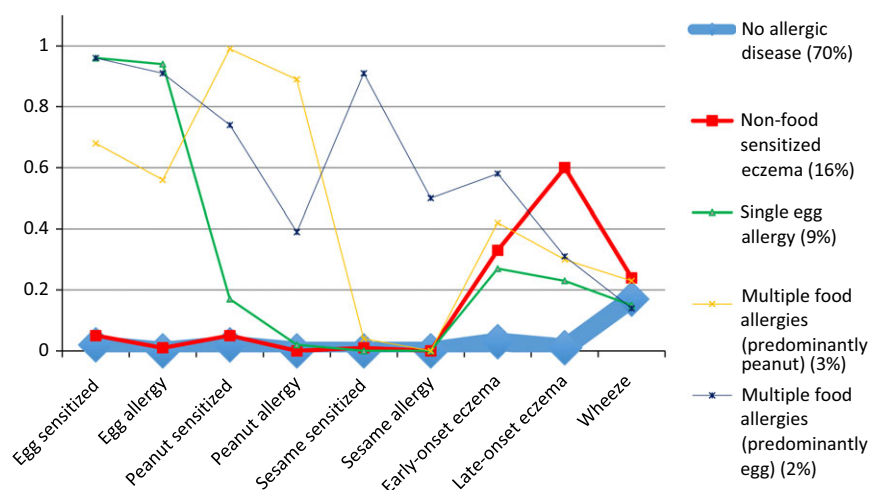


Fig. 1. Item response probabilities for each phenotype identified in the 5-class latent class analysis model. (Percentage in parentheses and the thickness of lines represent the relative size of the classes based on class membership probability).

phenotype in Table 2. The MLR model for risk factors for allergy phenotypes is shown in Table 3. The 'no allergic disease' class was selected as the baseline because it had the largest estimated population prevalence and is characterized by an absence of disease features.

After adjustment for potential confounders in MLR, parents' country of birth and family history of allergy remained significant risk factors for all the allergy phenotypes (classes 2–5). Infants with one or both parents born in Asia, compared to infants with both parents born in Australia, were at an increased risk of all three food allergy phenotypes (class 3–5) and marginally increased risk of belonging to the non-food-sensitized eczema phenotype. Maternal age and maternal smoking in pregnancy were not associated with any allergy phenotypes.

Infant risk factors for allergy phenotypes

Compared to the baseline class of 'no allergic disease', males were at increased risk of belonging to the non-food-sensitized eczema and multiple food allergy phenotypes (classes 2, 4 and 5) but not single egg allergy, (Table 3B). Premature birth reduced the risk of membership to multiple food allergies (predominantly peanut), whilst both colic and bronchiolitis increased the risk of membership to the non-food-sensitized eczema phenotype. Reflux, mode of delivery and season of birth were not associated with any allergy phenotype.

Environmental risk factors for allergy phenotypes

Environmental risk factors differed between the phenotypes. Compared to the baseline class of no allergic

disease, exposure to pet dogs allowed inside the home was associated with a reduced risk of membership for the non-food-sensitized eczema and multiple food allergies (predominantly peanut) phenotypes, and although not statistically significant, there was consistent evidence of reduction in risk of membership to the single egg allergy and multiple food allergy phenotypes (predominantly egg) (Table 3C). Presence of siblings reduced the risk of single egg allergy and non-food-sensitized eczema, but not multiple food allergies. For each month increase in breastfeeding (up to 12 months of age), there was a slight increase in risk of membership of the multiple food allergies (predominantly egg) phenotype. Delayed age of egg introduction (> 12 months) was a risk factor for all the allergy phenotypes (classes 2–5). Infants who were introduced peanut at 11–12 months compared to 4–10 months were less likely to belong to the multiple food allergy phenotypes (predominantly egg); age of peanut introduction was not associated with other phenotypes. Childcare attendance, cat ownership and exposure to tobacco smoke in the home were not associated with any allergy phenotypes.

Filaggrin and vitamin D studies

In the subgroup who attended OFC clinic, mutations in the filaggrin gene were associated with a higher risk for the non-food-sensitized eczema and egg allergy phenotypes, (aOR 2.37 (95% CI 1.27–4.43, $P = 0.01$) and 1.78 (95% CI 0.99–3.19, $P = 0.05$)), but it was not associated with either multiple food allergy phenotypes. Testing for interactions revealed that the association between vitamin D insufficiency and allergy phenotypes was modified by parent's country of birth (Table 4). Vitamin

Table 2. Demographic characteristics of the classes (weighted for probability of class membership)

	No allergic disease <i>N</i> = 3693, %	Non-food-sensitized eczema <i>N</i> = 844, %	Single egg allergy <i>N</i> = 475, %	Multiple food allergies (peanut) <i>N</i> = 158, %	Multiple food allergies (egg) <i>N</i> = 105, %	<i>P</i> -value
Parental characteristics						
Socioeconomic status Quintile 1*	21	20	18	17	19	0.28
Quintile 2	20	19	19	27	19	
Quintile 3	21	20	20	16	21	
Quintile 4	19	20	21	19	17	
Quintile 5	19	20	22	20	24	
Parent's ethnicity†						
Both Australia	61	62	49	41	41	< 0.001
One or both Asian	13	17	27	33	30	
One or both other	26	21	25	27	29	
FHx any allergy	67	77	76	76	73	< 0.001
Maternal smoking in pregnancy	5	5	3	5	1	0.11
Maternal age at birth						
18–24 year	6	5	5	3	5	0.85
25–29 year	19	19	20	21	26	
30–34 year	41	42	44	41	37	
35–39 year	28	27	26	29	24	
40+	7	7	5	6	7	
Infant characteristics						
Male	49	56	52	63	66	< 0.001
Preterm	6	7	4	5	5	0.22
Caesarean delivery	33	34	33	35	25	0.42
Season of birth						
Summer	24	23	22	21	13	0.19
Autumn	23	25	28	22	26	
Winter	26	24	25	29	32	
Spring	27	27	25	29	28	
Colic	13	17	14	11	15	0.03
Reflux	22	25	24	25	25	0.08
Bronchiolitis	14	18	14	11	8	0.001
Filaggrin mutation (<i>n</i> = 835)	8 (<i>n</i> = 275)	16 (<i>n</i> = 134)	11 (<i>n</i> = 259)	8 (<i>n</i> = 117)	7 (<i>n</i> = 50)	0.04
Environmental exposures						
Childcare attendance	24	27	24	22	26	0.42
No dogs	68	71	75	83	82	< 0.001
Dogs inside	22	19	14	9	13	
Dogs outside only	10	10	11	8	5	
No cats	82	83	87	88	88	0.06
Cats inside	14	14	10	9	10	
Cats outside only	3	3	3	3	3	
No siblings	48	50	61	50	50	< 0.001
1 sibling	34	35	27	33	36	
2 siblings	14	11	9	14	11	
3+ siblings	5	4	3	4	3	
Current smoking in the household	22	22	20	18	19	0.45
Duration of breastfeeding						
< 1 month	14	14	10	10	2	< 0.001
1–3 month	15	15	11	11	11	
4–6 month	15	15	14	21	18	
7–9 month	13	11	15	9	15	
10–12 month	14	14	15	11	10	
> 12 month	30	32	35	38	44	
Age egg introduction						
4–6 month	25	24	20	19	28	< 0.001

(continued)

Table 2 (continued)

	No allergic disease N = 3693, %	Non-food-sensitized eczema N = 844, %	Single egg allergy N = 475, %	Multiple food allergies (peanut) N = 158, %	Multiple food allergies (egg) N = 105, %	P-value
7–9 month	41	38	35	40	24	
10–12 month	30	31	34	30	28	
> 12 month	3	6	9	11	20	
< 4 month	1	1	2	0.1	0.1	
Age peanut introduction [‡]						
4–10 month	17	16	11	13	10	< 0.00
11–12 month	13	11	9	9	3	
Not yet given	70	73	80	78	88	
Vitamin D insufficiency (<50nM) (n = 707)		(n = 219)	(n = 106)	(n = 233)	(n = 106)	(n = 43)
	16	21	25	19	27	0.09

*Socio economic status as defined by Index of Relative Socio-economic Disadvantage (SEIFA) [27].

[†]Parent's ethnicity was based on reported country of birth and defined as follows: both Australian – both parents born in Australia; Asian – one of both parents born in Asia; Other – one or both parents born in other countries (Europe, United Kingdom, America, Africa, Middle East and Oceania.)

[‡]Only seven infants were introduced peanut before 4 months of age; because of small numbers, this group was excluded.

D insufficiency was a risk factor for the three food allergy phenotypes (classes 3–5), but only among infants with both parents born in Australia.

Discussion

This is the first study to use latent class analysis to describe phenotypes of allergic disease using challenge-proven food allergy outcomes. Using data from multiple allergic outcomes, we were able to synthesize these into five distinct phenotypic classes of allergic disease in infancy: no allergic disease, non-food-sensitized eczema, single egg allergy, multiple food allergies (predominantly peanut) and multiple food allergies (predominantly egg). Shared and differential risk factors for membership to one or more of the classes were identified.

The five allergic phenotypes identified are clearly interpretable and relevant to patient presentations in clinical practice. The two multiple food allergy phenotypes were the smallest, collectively representing 5% of the population, whilst single egg allergy accounted for a further 9%. Late-onset eczema was a prominent feature in the non-food-sensitized eczema phenotype, whilst the prevalence of early-onset eczema was greater in the three food allergy phenotypes. This supports the dual-allergen exposure hypothesis for the development of food allergy as proposed by Lack, that food sensitization occurs through cutaneous exposure early in life, facilitated by increased skin permeability in eczema [24]. Infants who were food-sensitized but tolerant were spread across the phenotypes, rather than being identi-

fied in a separate phenotype, which suggests that asymptomatic food sensitization is not a distinguishing feature of allergic disease in infancy. An interesting finding is that wheeze in infancy was non-discriminatory across the allergy phenotypes. This may be because not all wheeze in infancy is allergic in origin.

We found that several risk factors were consistently associated with the food allergy phenotypes, including parents' country of birth, family history of allergy, delayed introduction of egg, pet ownership and vitamin D insufficiency. However, several risk factors were differentially associated with the food allergy phenotypes. For example, FLG and the presence of siblings were only associated with non-food-sensitized eczema and single egg allergy, and neither were associated with the two multiple food allergy phenotypes of which peanut allergy was a prominent feature. This concurs with previous reports from this cohort that FLG was associated with food sensitization but not peanut allergy [18]. Conversely, male sex was strongly associated with the two multiple food allergy phenotypes, but not single egg allergy, and preterm birth was only associated with the multiple food allergy (predominantly peanut) phenotype. These differential risk factors suggest that the mechanisms that drive the development of allergic disease differ depending on food allergy phenotype.

In the present study, we found that delayed introduction of egg (> 12 months) was a risk factor for all the allergy phenotypes, of which egg allergy was a marked feature of three. This supports previous findings from this cohort [21]. Delayed introduction of egg was also a risk factor for non-food-sensitized eczema, a

Table 3. Multinomial logistic regression to identify risk factors for allergy phenotypes, compared to the baseline class of "no allergic disease"

A. Parent characteristics								
	Non-food sensitised eczema N = 844		Single egg allergy N = 475		Multiple food allergies (peanut) N = 158		Multiple food allergies (egg) N = 105	
	Adjusted OR*	P-value	Adjusted OR*	P-value	Adjusted OR*	P-value	Adjusted OR*	P-value
Parents country of birth								
Both Australia	1.0		1.0		1.0		1.0	
One or both Asian	1.35 (1.12–1.63)	0.002	2.73 (2.16–3.44)	< 0.001	4.24 (2.99–6.01)	< 0.001	3.39 (2.12–5.41)	< 0.001
Other	0.85 (0.72–1.00)	0.06	1.24 (0.98–1.56)	0.04	1.57 (1.09–2.26)	0.02	1.69 (1.06–2.71)	0.03
Socio economic status [†]								
Quintile 1 (most disadvantaged)	1.0	P trend	1.0	P trend	1.0	P trend	1.0	P trend
Quintile 2	0.96 (0.78–1.19)	0.29	1.02 (0.75–1.38)	0.07	1.59 (1.03–2.46)	0.65	1.02 (0.54–1.92)	0.40
Quintile 3	0.95 (0.76–1.18)		1.04 (0.77–1.41)		0.99 (0.62–1.58)		1.10 (0.59–2.05)	
Quintile 4	1.03 (0.83–1.27)		1.23 (0.91–1.66)		1.28 (0.80–2.03)		0.87 (0.44–1.70)	
Quintile 5	1.10 (0.89–1.36)		1.25 (0.92–1.69)		1.31 (0.82–2.10)		1.43 (0.77–2.64)	
Family history of allergy [‡]	1.66 (1.43–1.94)	< 0.001	1.80 (1.45–2.24)	< 0.001	1.93 (1.36–2.73)	< 0.001	1.64 (1.06–2.55)	0.03
Maternal smoking in pregnancy	1.08 (0.78–1.49)	0.64	0.72 (0.41–1.25)	0.24	1.40 (0.71–2.79)	0.33	0.39 (0.08–1.75)	0.22
B. Infant characteristics								
	Non-food sensitised eczema N = 844		Single egg allergy N = 475		Multiple food allergies (peanut) N = 158		Multiple food allergies (egg) N = 105	
	Adjusted OR [§]	P-value	Adjusted OR [§]	P-value	Adjusted OR [§]	P-value	Adjusted OR [§]	P-value
Sex (male)	1.33 (1.14–1.56)	< 0.001	1.11 (0.89–1.39)	0.34	2.00 (1.38–2.90)	< 0.001	2.27 (1.37–3.71)	0.001
Premature birth [¶]	0.99 (0.72–1.37)	0.95	0.68 (0.40–1.15)	0.16	0.34 (0.13–0.88)	0.03	0.83 (0.27–2.56)	0.75
Mode of delivery ^{**}	1.02 (0.86–1.20)	0.85	1.02 (0.81–1.29)	0.87	1.24 (0.86–1.78)	0.25	0.93 (0.56–1.60)	0.79
Season on birth – Summer	1.0		1.0		1.0		1.0	
Autumn	1.08 (0.86–1.35)	0.52	1.22 (0.89–1.67)	0.22	0.77 (0.46–1.34)	0.36	1.55 (0.73–3.23)	0.25
Winter	0.93 (0.74–1.16)	0.53	0.91 (0.66–1.25)	0.56	1.12 (0.68–1.82)	0.66	1.71 (0.84–3.45)	0.14
Spring	1.05 (0.86–1.32)	0.57	1.01 (0.74–1.38)	0.96	1.22 (0.75–1.98)	0.43	1.17 (0.55–2.49)	0.69
Colic	1.29 (1.03–1.62)	0.03	1.02 (0.72–1.44)	0.92	0.85 (0.49–1.48)	0.57	1.32 (0.67–2.60)	0.42
Reflux	1.11 (0.91–1.35)	0.29	1.11 (0.84–1.47)	0.47	1.07 (0.69–1.64)	0.75	0.92 (0.51–1.66)	0.78
Bronchiolitis	1.21 (1.02–1.42)	0.03	0.92 (0.72–1.19)	0.54	0.87 (0.56–1.35)	0.53	0.51 (0.25–1.01)	0.06
C. Environmental characteristics								
	Non-food sensitised eczema N = 844		Single egg allergy N = 475		Multiple food allergies (peanut) N = 158		Multiple food allergies (egg) N = 105	
	Adjusted OR ^{††}	P-value	Adjusted OR ^{††}	P-value	Adjusted OR ^{††}	P-value	Adjusted OR ^{††}	P-value
Childcare attendance	1.11 (0.92–1.34)	0.28	0.91 (0.69–1.21)	0.52	1.05 (0.68–1.63)	0.82	1.62 (0.89–2.91)	0.11
No dogs	1.0		1.0		1.0		1.0	
Dogs allowed inside the home	0.78 (0.63–0.96)	0.02	0.76 (0.56–1.05)	0.10	0.40 (0.21–0.73)	0.003	0.59 (0.26–1.34)	0.21
Dogs outside only	0.99 (0.76–1.29)	0.93	1.56 (1.10–2.21)	0.01	0.82 (0.44–1.54)	0.54	0.39 (0.13–1.18)	0.09

Table 3. continued

C. Environmental characteristics		Non-food sensitised eczema N = 844		Single egg allergy N = 475		Multiple food allergies (peanut) N = 158		Multiple food allergies (egg) N = 105	
		Adjusted OR ^{††}	P-value	Adjusted OR ^{††}	P-value	Adjusted OR ^{††}	P-value	Adjusted OR ^{††}	P-value
Pet cats		0.98 (0.79–1.22)	0.89	0.80 (0.57–1.12)	0.19	0.83 (0.47–1.47)	0.53	0.86 (0.38–1.91)	0.71
No siblings		1.0		1.0		1.0		1.0	
1 sibling		0.91 (0.76–1.10)	P trend	0.59 (0.45–0.77)	P trend	0.96 (0.64–1.44)	P trend	1.34 (0.76–2.35)	P trend
2 siblings		0.68 (0.52–0.88)	0.004	0.47 (0.32–0.70)	< 0.001	0.88 (0.49–1.59)	0.58	1.09 (0.47–2.55)	0.61
3+ siblings		0.73 (0.49–1.09)		0.39 (0.20–0.75)		0.81 (0.33–1.99)		1.19 (0.33–4.21)	
Duration of breastfeeding ^{††}		0.99 (0.98–1.01)	0.98	1.02 (0.99–1.04)	0.29	1.00 (0.96–1.05)	0.99	1.17 (1.09–1.24)	< 0.001
Age any egg introduced									
< 4 months		1.29 (0.52–3.24)	0.58	2.10 (0.75–6.10)	0.15	0.28 (0.05–1.61)	0.15	0.27 (0.05–1.44)	0.13
4–6 months		1.0		1.0		1.0		1.0	
7–9 months		0.97 (0.79–1.21)	0.80	1.04 (0.75–1.46)	0.80	1.68 (0.99–2.83)	0.05	0.55 (0.27–1.12)	0.10
10–12 months		1.05 (0.84–1.31)	0.71	1.51 (1.08–2.11)	0.02	1.68 (0.97–2.93)	0.06	0.82 (0.39–1.71)	0.56
> 12 months		2.02 (1.36–3.01)	< 0.001	3.46 (2.14–5.57)	< 0.001	4.29 (2.11–8.73)	< 0.001	3.14 (1.23–7.99)	0.02
Age peanut introduced									
4–10 months		1.0		1.0		1.0		1.0	
11–12 months		0.98 (0.72–1.32)	0.89	1.06 (0.64–1.73)	0.83	1.15 (0.54–2.45)	0.71	0.08 (0.03–0.21)	< 0.001
Not yet given		0.98 (0.78–1.23)	0.88	1.29 (0.89–1.86)	0.18	1.16 (0.67–2.02)	0.59	1.68 (0.70–4.03)	0.25

*The presented parental risk factors were associated ($P < 0.1$) with at least one phenotype on univariable analysis and included in the multivariable multinomial logistic regression (MLR) model simultaneously. We also considered adjusting for maternal age at birth, however it did not materially alter the adjusted OR presented and was not included in the final model.

[†]Socioeconomic status was measured by using the Index of Relative Socio-economic Disadvantage, a general socioeconomic index that summarizes a range of information about the economic and social conditions of households within an area; these data were obtained from the Australian Census (2006) and matched by postcode.

[‡]History of either asthma, eczema, hay fever or food allergy in an immediate family member.

[§]The presented infant risk factors were associated ($P < 0.1$) with at least one phenotype on univariable analysis and are fitted in the model simultaneously. The model was also adjusted for the following parental risk factors: parent's country of birth, family history of allergy and socioeconomic status. Parent's country of birth confounded the association between season of birth and class membership. Other parental risk factors did not alter the association between sex, premature birth, mode of delivery, breastfeeding, colic or bronchiolitis and class membership.

[¶]< 37 weeks gestation.

**Caesarean compared to vaginal birth.

^{††}The presented environmental risk factors were associated ($P < 0.1$) with at least one phenotype on univariable MLR analysis and are fitted in the model simultaneously. The model was also adjusted for the following parental risk factors: parent's country of birth, family history of allergy and socioeconomic status, and the following infant risk factors: sex, premature birth, duration of breastfeeding, colic and bronchiolitis. The following parent or infant risk factors confounded the association between the stated environmental risk factor and class membership: parent's country of birth and number of siblings; parent's country of birth and dog/cat ownership; bronchiolitis and childcare attendance; bronchiolitis and age of egg/peanut introduction; colic and age of egg/peanut introduction; pre-term birth and age of peanut introduction.

^{‡‡}One month increase in breastfeeding up to 12 months of age.

Values in bold indicate those with a P-value < 0.05.

Table 4. Association between allergy phenotypes and vitamin D insufficiency, compared to the baseline class of 'no allergic disease' ($n = 219$)

	Non-food-sensitized eczema ($n = 106$)		Single egg allergy ($n = 233$)		Multiple food allergies (peanut) ($n = 106$)		Multiple food allergies (egg) ($n = 43$)	
	Adjusted OR*	P	Adjusted OR*	P	Adjusted OR*	P	Adjusted OR*	P
Vitamin D insufficiency	1.26 (0.71–2.23)	0.42	1.67 (1.02–2.74)	0.04	1.36 (0.72–2.57)	0.34	1.65 (0.70–3.87)	0.25
P-value for interaction with parents country of birth		0.49		0.004		< 0.001		0.74
Vitamin D insufficiency stratified by parents country of birth								
Both parents born in Australia	1.59 (0.69–3.69)	0.27	3.24 (1.58–6.65)	0.001	4.40 (1.77–10.92)	0.001	3.76 (1.23–11.45)	0.02
One or both Asian	1.34 (0.35–5.12)	0.67	1.32 (0.44–3.95)	0.62	1.58 (0.45–5.52)	0.47	0.60 (0.09–4.25)	0.61
Other	0.97 (0.35–2.67)	0.95	0.58 (0.23–1.51)	0.27	0.03 (0.007–0.10)	< 0.001	0.78 (0.19–3.15)	0.73

*Adjusted for parents country of birth, duration of breastfeeding, premature birth and age of egg introduction.

phenotype that was characterized by a high prevalence of eczema, but no egg sensitization or allergy. This may be reflective of reverse causation, that is, once the infant develops eczema, parents may have chosen to delay the introduction of potentially allergenic foods. We did not find strong evidence of association between timing of peanut introduction and the multiple food allergies (predominantly peanut) phenotype. This suggests that different phenotypes of allergy may respond differently to preventative measures, including timing of introduction of allergenic foods, because they are driven by other mechanism such as genetics and/or environment.

The strengths of this study include the large sample size and population-based sampling frame. In addition, this study objectively ascertained sensitization status on over 5000 infants and utilized the gold standard OFC; this is therefore the largest study identifying phenotypes on the basis of objective measures of allergy. Using an LCA model, the phenotypes identified were inferred entirely from the actual data, rather than relying on arbitrary definitions, and the resulting output produced interpretable phenotypes.

A potential limitation of this study is that although objective food allergy testing was conducted, we relied on parent-reported measures of eczema (parent report of a doctor diagnosis) and wheeze. This study focused on sensitization and food challenge data on 3 common food allergens; milk and shrimp SPT were only measured on half the cohort but were considered in a supplementary analysis. The prevalence of milk and shrimp sensitization within the LCA defined phenotypes were, however, consistent with the patterns observed with other food sensitizations. For example, milk and shrimp sensitization were very low in the no allergic disease and non-food-sensitized eczema phenotype and highest in the multiple food allergy phenotypes (data not shown). We repeated the LCA analysis using half of the

cohort who underwent milk SPT. The model produced was remarkably similar to that currently presented in the study, and no new phenotypes were observed. This indicates that our model is robust and that food allergy phenotypes in infancy are primarily driven by sensitization to the three common food allergens for which OFC-confirmed allergy is most common. We did not have sensitization data on inhalant allergens at the population level; however, sensitization to inhalant allergens increases later in childhood and is not common in early infancy, the focus of this study [25, 26].

Two recent studies have also identified allergy phenotypes using LCA from cross-sectional data of a similar age group to this cohort. Havstad et al. [11] described phenotypes of atopy in a subgroup of 594 2-year-old children, from a larger birth cohort of 1258. Four phenotypes were identified on the basis of serum-specific IgE measurements to 10 food and inhalant allergens: 'no to low sensitization', 'highly sensitized', 'milk and egg dominated' and 'peanut and inhalant'. Although it is difficult to directly compare these phenotypes to ours because they did not include common allergic manifestations such as food allergy and eczema, and sIgE was serologically measured, they also found that race and maternal atopy were risk factors for the allergy phenotypes. A birth cohort of 1831 18-month-old infants used LCA to identify four phenotypes on the basis of parent report of respiratory and allergic outcomes and utilization of healthcare resources [12]. The analysis was based primarily on respiratory data and did not include data specific to food sensitization or allergy.

Several longitudinal studies following children from early childhood to adolescence have employed LCA to explore primarily wheeze phenotypes or patterns of sensitization, but not food allergy directly [6, 9, 10, 14]. Collectively, these studies highlight that allergic disease should be considered beyond a dichotomous definition

of asthma or atopy as various subphenotypes have been identified that may have different clinical courses as well as risk factors, pathogenesis and prognosis. The present study builds upon previous knowledge by including challenge-proven food allergy outcomes. Our findings support that as with other allergic diseases, food allergy should not be considered as a single condition as multiple subphenotypes exist.

In conclusion, we defined five distinct allergy phenotypes which were synthesized from common multiple allergic outcomes in infancy. Differences in risk factors suggest that the mechanisms that drive the development of allergic disease in infancy and beyond may differ. The HealthNuts cohort age 4- and 6-year follow-up, that is in progress, will enable us to investigate how these phenotypes progress along the atopic march and whether the development or resolution allergic diseases in childhood is differential.

Acknowledgements

We thank the children and parents who participated in the HealthNuts study as well as the staff of Melbourne's Local Government Areas for providing access to community immunization clinics. We thank ALK-Abello SA, Madrid, Spain, for supplying the skin prick testing reagents. We acknowledge the contributions of the HealthNuts safety committee – Associate Professor Noel Cranswick (Australian Paediatric Pharmacology Research Unit/Murdoch Childrens Research Institute), Dr Jo Smart (Department of Allergy and Immunology, Royal Children's Hospital, Melbourne, Australia) and Associate Professor Jo Douglass (Head of Allergy, Alfred Hospital, Melbourne, Australia). The

HealthNuts study Investigators also include Melanie Matheson, Melissa Wake, Adrian Lowe, Thanh Dang, Pamela Martin, Tina Tan, Nick Osborne, Giovanni Zurzolo, Leone Thiele, Helen Czech, Deborah Anderson, Jana Eckert, Nadine Bertalli and Jeeva Sanjeevan.

Author contributions

K. A., L. G., S. D., M. W. and M. T. contributed to conception and design of the HealthNuts study. R. P. conducted the data analysis with the assistance of L. G. and was responsible for the first draft of the article. All authors contributed to interpretation of the data, revision of the draft manuscript and approved the final version to be published.

Funding

This work was supported by funding from the National Health & Medical Research Council (NHMRC) of Australia, Ilhan Food Allergy Foundation, AnaphylaxiStop, the Charles and Sylvia Viertel Medical Research Foundation, the Victorian Government's Operational Infrastructure Support Program and the NHMRC Centre for Food and Allergy Research. K.J.A. is a Viertel senior medical research fellow. R. P. is an Australian Postgraduate Award scholar. L. G., M. M., A. L., A.-L. P., J. K. and S. D. hold NHMRC awards.

Conflict of interest

The authors have no relevant conflict of interests.

References

- Burgess JA, Dharmage SC, Byrnes GB *et al*. Childhood eczema and asthma incidence and persistence: a cohort study from childhood to middle age. *J Allergy Clin Immunol* 2008; **122**: 280–5.
- Spergel JM. Epidemiology of atopic dermatitis and atopic march in children. *Immunol Allergy Clin North Am* 2010; **30**:269–80.
- Penard-Morand C, Raherison C, Kopferschmitt C *et al*. Prevalence of food allergy and its relationship to asthma and allergic rhinitis in schoolchildren. *Allergy* 2005; **60**:1165–71.
- Illi S, von Mutius E, Lau S *et al*. The pattern of atopic sensitization is associated with the development of asthma in childhood. *J Allergy Clin Immunol* 2001; **108**:709–14.
- Kjaer HF, Eller E, Andersen KE, Host A, Bindslev-Jensen C. The association between early sensitization patterns and subsequent allergic disease. The DARC birth cohort study. *Pediatr Allergy Immunol* 2009; **20**:726–34.
- Lodge CJ, Lowe AJ, Gurrin LC *et al*. House dust mite sensitization in toddlers predicts current wheeze at age 12 years. *J Allergy Clin Immunol* 2011; **128**:782–8 e9.
- Martin PE, Matheson MC, Gurrin L *et al*. Childhood eczema and rhinitis predict atopic but not nonatopic adult asthma: a prospective cohort study over 4 decades. *J Allergy Clin Immunol* 2011; **127**:1473–9 e1.
- Kurukulaaratchy RJ, Matthews S, Arshad SH. Defining childhood atopic phenotypes to investigate the association of atopic sensitization with allergic disease. *Allergy* 2005; **60**:1280–6.
- Caudri D, Savenije OE, Smit HA *et al*. Perinatal risk factors for wheezing phenotypes in the first 8 years of life. *Clin Exp Allergy* 2013; **43**:1395–405.
- Garden FL, Simpson JM, Marks GB, Investigators C. Atopy phenotypes in the Childhood Asthma Prevention Study (CAPS) cohort and the relationship with allergic disease: clinical mechanisms in allergic disease. *Clin Exp Allergy* 2013; **43**:633–41.
- Havstad S, Johnson CC, Kim H *et al*. Atopic phenotypes identified with latent class analyses at age 2 years. *J*

- Allergy Clin Immunol* 2014; 134 (3):722–7 e2.
- 12 Herr M, Just J, Nikasinovic L *et al.* Risk factors and characteristics of respiratory and allergic phenotypes in early childhood. *J Allergy Clin Immunol* 2012; 130:389–96 e4.
 - 13 Lodge CJ, Zaloumis S, Lowe AJ *et al.* Early-life risk factors for childhood wheeze phenotypes in a high-risk birth cohort. *J Pediatr* 2013; 164(2):289–94 e1–2.
 - 14 Spycher BD, Silverman M, Brooke AM, Minder CE, Kuehni CE. Distinguishing phenotypes of childhood wheeze and cough using latent class analysis. *Eur Respir J* 2008; 31:974–81.
 - 15 Henderson J, Granell R, Heron J *et al.* Associations of wheezing phenotypes in the first 6 years of life with atopy, lung function and airway responsiveness in mid-childhood. *Thorax* 2008; 63:974–80.
 - 16 Osborne NJ, Koplin JJ, Martin PE *et al.* The HealthNuts population-based study of paediatric food allergy: validity, safety and acceptability. *Clin Exp Allergy* 2010; 40:1516–22.
 - 17 Koplin JJ, Tang ML, Martin PE *et al.* Predetermined challenge eligibility and cessation criteria for oral food challenges in the HealthNuts population-based study of infants. *J Allergy Clin Immunol* 2012; 129:1145–7.
 - 18 Tan HT, Ellis JA, Koplin JJ *et al.* Filaggrin loss-of-function mutations do not predict food allergy over and above the risk of food sensitization among infants. *J Allergy Clin Immunol* 2012; 130:1211–3 e3.
 - 19 Allen KJ, Koplin JJ, Ponsonby AL *et al.* Vitamin D insufficiency is associated with challenge-proven food allergy in infants. *J Allergy Clin Immunol* 2013; 131:1109–16, 16 e1–6.
 - 20 Koplin JJ, Dharmage SC, Ponsonby AL *et al.* Environmental and demographic risk factors for egg allergy in a population-based study of infants. *Allergy* 2012; 67:1415–22.
 - 21 Koplin JJ, Osborne NJ, Wake M *et al.* Can early introduction of egg prevent egg allergy in infants? A population-based study. *J Allergy Clin Immunol* 2010; 126:807–13.
 - 22 Lack G. Update on risk factors for food allergy. *J Allergy Clin Immunol* 2012; 129:1187–97.
 - 23 University Park: The Methodology Center PS. LCA Stata Plugin (Version 1.1). 2014. p. [Software].
 - 24 Lack G. Epidemiologic risks for food allergy. *J Allergy Clin Immunol* 2008; 121:1331–6.
 - 25 Kulig M, Bergmann R, Klettke U, Wahn V, Tacke U, Wahn U. Natural course of sensitization to food and inhalant allergens during the first 6 years of life. *J Allergy Clin Immunol* 1999; 103:1173–9.
 - 26 Nissen SP, Kjaer HF, Host A, Nielsen J, Halken S. The natural course of sensitization and allergic diseases from childhood to adulthood. *Pediatr Allergy Immunol* 2013; 24:549–55.
 - 27 The Index Of Relative Socio-economic Disadvantage. Australian Bureau of Statistics. Available from: www.abs.gov.au/ausstats/abs@.nsf/Lookup/2033.0.55.001 (Last accessed June 6, 2013)

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Model fit statistics for 1–7 classes.

Table S2. Description of classes in the 5 class model: item response probabilities.

Data S1. Methods.

Supplementary files:

Table E1: Model fit statistics for 1-7 classes:

No. of classes	Log-likelihood	AIC	Corrected AIC	BIC	Adjusted BIC
1	-12202.56	4473.20	4559.33	4550.33	4521.73
2	-10383.51	873.10	1016.93	997.93	937.55
3	-10176.25	478.57	698.10	669.10	576.94
4	-10153.12	452.32	747.54	708.54	584.61
5	-10087.85	341.79	712.71	663.71	508.00
6	-10052.16	290.41	737.03	678.03	490.55
7	-10044.96	296.00	818.32	749.32	530.06

AIC: Akaike's Information criteria; BIC: Bayesian Information Criterion. Values in bold indicate the optimal number of classes based on the minimal value of the test statistic.

Both the 5 and 6 class models were indicated as the best models by the fit statistics. The 5 class model was selected because (i) the classes were more easily interpretable than the 6 class model; and (ii) in the 6 class model the estimated population prevalences of three classes were very small (2-3%).

Table E2: Description of classes in the 5 class model: item response probabilities.

	No allergic disease (N=3693)	Non-food sensitised eczema (N=844)	Single egg allergy (N=475)	Multiple food allergies (peanut) (N=158)	Multiple food allergies (egg) (N=106)
Membership probability	0.70	0.16	0.09	0.03	0.02
Egg sensitised	0.02	0.05	0.96	0.68	0.96
Egg allergy	0.00	0.01	0.94	0.56	0.91
Peanut sensitised	0.02	0.05	0.17	0.99	0.74
Peanut allergy	0.00	0.00	0.02	0.89	0.39
Sesame sensitised	0.00	0.01	0.00	0.04	0.91
Sesame allergy	0.00	0.00	0.00	0.00	0.50
No eczema	0.96	0.07	0.50	0.28	0.11
Early-onset eczema	0.03	0.33	0.27	0.42	0.58
Late-onset eczema (<u>≥ 4 months</u>)	0.01	0.60	0.23	0.30	0.31
Wheeze	0.17	0.24	0.15	0.23	0.14

Methods:

SPT to shrimp was conducted on the first half of the cohort and changed to cow's milk for the second half. Initially we were reluctant to conduct cow's milk SPT in the community because of concerns that a high prevalence of sensitization would be observed among infants already tolerating cow's milk in their diet, leading to unnecessary anxiety among parents, and possibly extending the already very long Royal Children's Hospital Allergy department waiting lists beyond 12 months. A pilot study subsequently showed the prevalence of sensitization to cow's milk was relatively low in the community. Unfortunately, we were unable to proceed with food challenges to cow's milk due to insufficient funding. Because SPT to shrimp and cow's milk were only available for half the cohort, they could not be included in the LCA model.

Any infant who showed objective evidence of sensitisation (SPT >1mm) was invited to the royal Childrens hospital for OFC. We chose this low threshold for community screening to ensure that we did not miss any potential cases of food allergy. SPT and sIgE were repeated at clinic attendance and were used for the classification of food allergy. We chose to use a threshold of 2mm to define sensitisation because of the young age of our cohort (12 months). Evidence suggests that skin reactivity is lower in infancy and we felt that this threshold was more appropriate to the young age of this cohort. [\(1,2\)](#) We assumed that infants with negative SPT in the community (0mm) did not have IgE-mediated food allergy. To test this assumption a random sample of 200 negative control (non-sensitized) infants were invited for OFC, of which all were food challenge negative. We therefore believe we have strong evidence to support that these non-sensitized infants do not have IgE-mediated food allergy

1. [Menardo JL, Bousquet J, Rodiere M, Astruc J, Michel FB. Skin test reactivity in infancy. The Journal of allergy and clinical immunology. 1985;75\(6\):646-51.](#)

2. [Skassa-Brociek W, Manderscheid JC, Michel FB, Bousquet J. Skin test reactivity to histamine from infancy to old age. The Journal of allergy and clinical immunology. 1987;80\(5\):711-6.](#)