Received: 9 June 2021 | Revised: 24 July 2021 | Accepted: 6 August 2021

DOI: 10.1111/pai.13620

SUPPLEMENT ARTICLE

WILEY

Endotyping allergic rhinitis in children: A machine learning approach

Velia Malizia[1] | Giovanna Cilluffo[1] | Salvatore Fasola[1] | Giuliana Ferrante[2] | Massimo Landi[1,3] | Laura Montalbano[1] | Amelia Licari[4,5] | Stefania La Grutta[1]

[1]Institute for Biomedical Research and Innovation, National Research Council, Palermo, Italy

[2]Department of Health Promotion, Mother and Child Care, Internal Medicine and Medical Specialties, University of Palermo, Palermo, Italy

[3]Pediatric National Healthcare System, Turin, Italy

[4]Pediatric Clinic, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy

[5]Department of Clinical, Chirurgical, Diagnostic and Pediatric Science, University of Pavia, Pavia, Italy

Correspondence
Giovanna Cilluffo, Institute for Biomedical Research and Innovation, National Research Council, Via Ugo La Malfa 153, 90146 Palermo, Italy.
Email: giovanna.cilluffo@irib.cnr.it

Editor: Elena Chiappini

Abstract

The diversity of allergic rhinitis (AR) phenotypes is particularly evident in childhood, suggesting the need to analyze and identify new approaches to capture such clinical heterogeneity. Nasal cytology (NC) is a very useful diagnostic tool for identifying and quantifying nasal inflammation. Data-driven approaches such as latent class analysis (LCA) assign subjects to classes based on their characteristics. We hypothesized that LCA based on NC, including the assessment of neutrophils, eosinophils, and mast cells, may be helpful for identifying AR endotypes in children. A total of 168 children were enrolled. Sociodemographic characteristics and detailed medical history were obtained from their parents. All children performed NC and skin prick tests. LCA was applied for identifying AR endotypes based on NC, using the R package *poLCA*. All the statistical analyses were performed using R 4.0.5 software. Statistical significance was set at $p \leq .05$. LCA identified two classes: Class 1 ($n = 126$, 75%): higher frequency of children with moderate/large number of neutrophils (31.45%); almost all the children in this class had no mast cells (91.27%) and Class 2 ($n = 42$, 25%): higher frequency of children with moderate/large number of eosinophils (45.24%) and moderate/large number of mast cells (50%). The present study used a machine learning approach for endotyping childhood AR, which may contribute to improve the diagnostic accuracy and to deliver personalized health care in the context of precision medicine.

KEYWORDS
allergic rhinitis, children, endotypes, latent class analysis, nasal cytology

Allergic rhinitis (AR) is a very common medical condition in children.[1] Allergic inflammation is characterized by the activation of a complex cellular network. Previous studies attempted to characterize AR based on immune-histological findings in the nasal mucosa.[2] Nasal cytology (NC) is a very useful diagnostic tool for identifying and quantifying nasal inflammation.[3] The complexity and variability of AR phenotypes are particularly evident in childhood,[4] suggesting the need to analyze and identify new approaches to capture the clinical heterogeneity of AR. However, childhood AR has been not extensively evaluated in such innovative fashion.

Data-driven approaches such as latent class analysis (LCA) could offer an advantage for identifying latent class membership among participants with multivariate categorical data.[5] Indeed, through LCA, subjects are assigned to classes based on their characteristics, rather than being arbitrarily assigned by the researchers.[6] We hypothesized that LCA based on NC, including the assessment of

neutrophils, eosinophils, and mast cells, may be helpful for identifying AR endotypes in children.

The inclusion criteria for this cross-sectional study were as follows: (1) age 6–14 years; (2) sensitization to inhalant allergens; and (3) clinical history of AR in the last 12 months, according to ARIA guidelines. The exclusion criteria were as follows: (1) asthma diagnosis; (2) upper or lower respiratory tract infections (having taken antibacterial therapy in the 4 weeks before the study entry); (3) use of leukotriene antagonists, systemic/topical antibiotics, or corticosteroids in the last 2 weeks; (4) ongoing allergen-specific immunotherapy; and (5) active smoking.

A total of 190 children (subjects without asthma enrolled in the CHASER study; identifier: NCT02433275) were assessed for eligibility at the outpatient clinic of the Institute for Biomedical Research and Innovation of the National Research Council of Palermo, and 168 (88%; 21 children were excluded due to violation of entry criteria and 1 due to consent withdrawn) were enrolled. The study was approved by the local ethics committee and conducted in accordance with the Declaration of Helsinki and Good Clinical Practice guidelines. All participants/caregivers provided written informed consent.

Sociodemographic characteristics and detailed medical history (disease duration and severity, atopy, environmental exposures) were obtained from the parents of all the participants through

Key Message

A machine learning approach applied to nasal cytology measurements could help to catch the heterogeneity of allergic rhinitis in children, in order to deliver a tailored solution in the context of precision medicine.

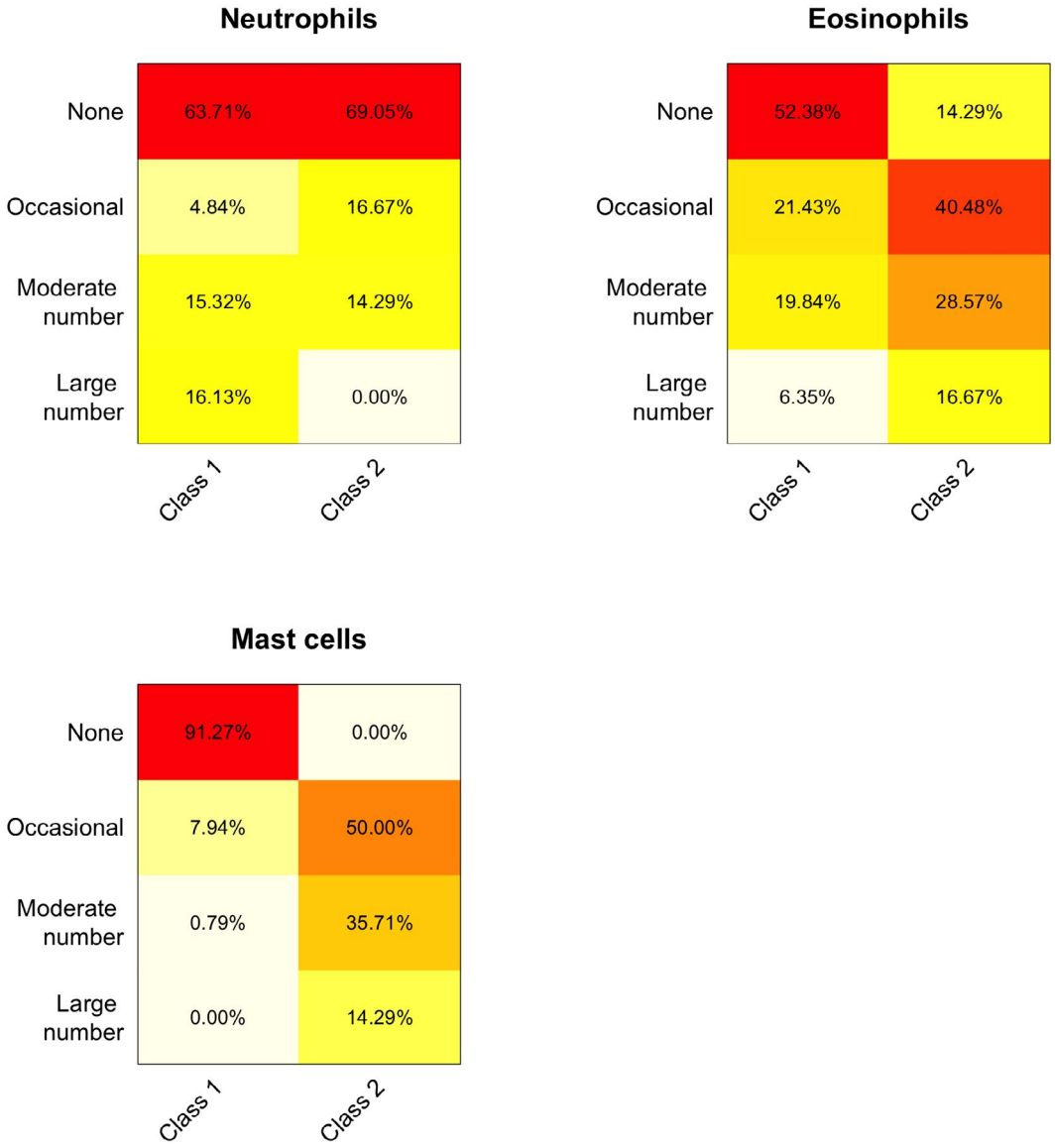

FIGURE 1 Allergic rhinitis classes identified by latent class analysis. The proportion of each variable in each class is represented through a color scale spanning from white (0%) to red (100%)

TABLE 1 Demographic and clinical characteristics of children by LCA class

	Whole sample	Class 1	Class 2	
	n = 168 (100%)	*n* = 126 (75%)	*n* = 42 (25%)	*p*-value
Gender: Female	53 (31.55)	41 (32.54)	12 (28.57)	.774
Age, years	9.81 (2.35)	9.60 (2.34)	10.43 (2.28)	**.040**
Height, cm	133.97 (16.31)	134.11 (16.48)	133.57 (15.97)	.934
Weight, kg	35.54 (14.84)	35.64 (14.78)	35.24 (15.20)	.776
Disease duration, years	4.73 (2.93)	4.53 (2.83)	5.27 (3.18)	.276
ARIA, Persistent	93 (55.36)	72 (57.14)	21 (50.00)	.272
Current pet exposure	38 (22.75)	25 (20.00)	13 (30.95)	.211
Current smoke exposure	55 (33.13)	40 (32.26)	15 (35.71)	.825
Current mold exposure	35 (20.96)	26 (20.80)	9 (21.43)	1.00
Sensitizations				
Indoor[a]	126 (82.35)	93 (81.58)	33 (84.62)	.852
Outdoor[b]	113 (76.35)	80 (72.07)	33 (89.19)	.058
Indoor + Outdoor	92 (62.16)	64 (57.66)	28 (75.68)	.078
Number of sensitizations	9.28 (1.41)	9.14 (1.40)	9.66 (1.36)	**.027**
T5SS total score	8.39 (3.14)	8.68 (3.07)	7.45 (3.22)	**.050**
Nasal itching	2.56 (1.03)	2.60 (1.00)	2.42 (1.12)	.398
Ocular itching	2.27 (1.07)	2.31 (1.10)	2.13 (0.99)	.404
Rhinorrhea	2.57 (1.08)	2.64 (1.06)	2.35 (1.11)	.201
Sneezing	3.02 (0.93)	3.05 (0.95)	2.90 (0.87)	.447
Nasal obstruction	2.99 (1.01)	3.10 (0.92)	2.65 (1.23)	**.028**
PRQLQ total score	2.87 (1.25)	2.88 (1.26)	2.84 (1.25)	.890
PSQI total score	5.88 (2.38)	5.97 (2.52)	5.57 (1.81)	.416

[a]Positive to at least one among Cat dander, Dog dander, Alternaria, and dermatophagoides.

[b]Positive to at least one among Grass mix, Olea, Parietaria, Artemisia, and Cupressus.

medical interviews (GF, SLG, and VM). Information about self-reported symptoms was collected through the Total 5 Symptom Score (T5SS), the Pittsburgh Sleep Quality Index (PSQI), and the Paediatric Rhinoconjunctivitis Quality of Life Questionnaire (PRQLQ). T5SS was used to assess symptom severity during the last week. It is based on 5 symptom domains: rhinorrhea, nasal obstruction, nasal itching, sneezing, and eye itching. Each symptom is scored from 0 to 3, so that the total score ranges from 0 (non-severe) to 15 (extremely severe). PSQI was used to measure sleep quality during the previous month. It includes 19 items, and the total score ranges from 0 (good sleeper) to 21 (poor sleeper). PRQLQ was used to assess the quality of life in the previous week. It includes 23 items and the total score scored ranges from 0 (not troubled) to 6 (extremely troubled). The cytologic sampling consisted in the collection of surface cells scraped from the middle portion of the inferior turbinate. The cellular material was spread on a glass slide, fixed by air-drying, and then stained by the method of May-Grunwald-Giemsa. Slides were read using a common optical microscope, at 10,009 in oil immersion, equipped with a digital camera. The analysis of rhinocytograms involved the reading of not less than 50 fields. Neutrophils, eosinophils, and mast cells were graded as follows: 0, none; 1, occasional; 2, moderate number; and 3, large number.[3] In children sensitized to seasonal aeroallergens, the sampling was performed outside the pollen season. Skin-prick tests (SPTs) were performed according to EAACI recommendations with a standard panel of inhalant allergens including a positive (histamine 1%) and a negative (saline) control (ALK-Abellò, Milan, Italy). Allergic sensitization was defined upon a positive skin response after 15 min (ie, a wheal ≥3 mm larger than the negative control test).

Data were presented as absolute and percentage frequencies (categorical variables), or mean and standard deviation (quantitative variables). Categorical variables were compared using the chi-squared test, and quantitative variables were compared using the Kruskal-Wallis test to avoid distributional assumptions. LCA was applied for identifying AR classes based on NC, using the R package *poLCA*. The Akaike information criterion (AIC) was used to determine the optimal number of classes (the smaller the better). All the statistical analyses were performed using R 4.0.5 software. Statistical significance was set at $p \leq .05$.

LCA identified two classes (Figure 1): Class 1 (n = 126, 75%): higher frequency of children with moderate/large number of neutrophils (31.45%); almost all the children in this class had no mast cells (91.27%) and Class 2 (n = 42, 25%): higher frequency of children

with moderate/large number of eosinophils (45.24%) and moderate/large number of mast cells (50%). Children in Class 1 reported higher scores of nasal obstruction and T5SS (Table 1). Children in Class 2 were older and had higher number of positive SPTs than children in Class 1. Detailed information on sensitization type, PRQLQ, and PSQI domains are reported in Table S1. Within classes, no differences in NC distribution were found for AR duration and severity according to ARIA classification (Table S2).

Compared to children in Class 2, those in Class 1 showed a higher frequency of neutrophils in the nasal mucosa, suggesting a condition of minimal persistent inflammation. Moreover, they reported higher score of nasal obstruction. In line with our findings, Ciprandi et al. also found that neutrophil infiltration in the nasal mucosa of adult subjects with AR is strongly associated with mild to moderate nasal obstruction when compared to no obstruction.[7] However, no association between neutrophil infiltration and AR duration and severity according to ARIA was found, in agreement with a previous finding on adults with AR.[8] Children in Class 2 showed more frequently moderate/large number of eosinophils and mast cells in the nasal mucosa. Furthermore, they showed a higher number of positive SPTs, regardless of type (indoor *vs* outdoor) of sensitization, and were older than children in Class 1. These findings are in line with previous data reporting that percentages of eosinophils and mast cells in the nasal mucosa increase with age[9] as well as that the point prevalence of aeroallergen sensitization increases significantly during childhood.[10] The finding of a higher number of mast cells in this class is not surprising, given the concomitant high infiltration of eosinophils in the nasal mucosa.

The strength of the current study was the identification of data-driven AR endotypes and investigation of the associated cytologic patterns. Nonetheless, we recognize some limitations. First, we did not test peripheral specific IgE; however, comparable levels of eosinophil inflammation in the airways have been previously reported in patients with AR. Moreover, we analyzed our children at a one-time point; therefore variability of the cytologic endotypes over time may limit the current findings. Further longitudinal studies are required for evaluating the stability of our results over time.

In conclusion, this study used a machine learning approach for endotyping childhood AR, which may contribute to improve the diagnostic accuracy and to deliver personalized health care in the context of precision medicine.

CONFLICT OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

AUTHOR CONTRIBUTIONS

All authors were involved in investigation and gave constructive criticism of the study manuscript. VM, LM, and ML participated in the data collection and database construction. GC performed the statistical analysis and data interpretation. GC, GF, SLG, and VM wrote the manuscript. SF and AL contributed to the interpretation of the results and reviewed the manuscript. All authors reviewed and approved the final version of the manuscript.

ORCID

Velia Malizia https://orcid.org/0000-0002-1754-9353
Giovanna Cilluffo https://orcid.org/0000-0003-4305-3981
Salvatore Fasola https://orcid.org/0000-0002-5378-5505
Giuliana Ferrante https://orcid.org/0000-0001-9917-2387
Massimo Landi https://orcid.org/0000-0001-7587-4800
Laura Montalbano https://orcid.org/0000-0002-3540-4801
Amelia Licari https://orcid.org/0000-0002-1773-6482
Stefania La Grutta https://orcid.org/0000-0001-8026-0715

REFERENCES

1. Cibella F, Ferrante G, Cuttitta G, et al. The burden of rhinitis and rhinoconjunctivitis in adolescents. *Allergy Asthma Immunol Res*. 2015;7(1):44-50.
2. KleinJan A, Willart M, van Rijt LS, et al. An essential role for dendritic cells in human and experimental allergic rhinitis. *J Allergy Clin Immunol*. 2006;118(5):1117-1125.
3. Heffler E, Landi M, Caruso C, et al. Nasal cytology: methodology with application to clinical practice and research. *Clin Exp Allergy*. 2018;48(9):1092-1106.
4. Papadopoulos NG, Aggelides X, Stamataki S, Prokopakis E, Katotomichelakis M, Xepapadaki P. New concepts in pediatric rhinitis. *Pediatr Allergy Immunol*. 2021;32(4):635-646.
5. Nosetti L, Paglietti M, Brunetti L, et al. A survey around the Italian pediatric units on current clinical practice for Sleep Disordered Breathing (SDB). *Ital J Pediatr*. 2019;45(1):75.
6. Ferrante G, Licari A, Fasola S, Marseglia GL, La Grutta S. Artificial intelligence in the diagnosis of pediatric allergic diseases. *Pediatr Allergy Immunol*. 2021;32(3):405-413.
7. Ciprandi G, Cirillo C, Klersy C, Marseglia GL, Caimmi D, Vizzaccaro A. Nasal obstruction is the key symptom in hay fever patients. *Otolaryngol–Head Neck Surg*. 2005;133(3):429-435.
8. Gelardi M, Incorvaia C, Passalacqua G, Quaranta N, Frati F. The classification of allergic rhinitis and its cytological correlate. *Allergy*. 2011;66(12):1624-1625.
9. Otsuka H, Otsuka K, Matsune S, Okubo K. Assessing the onset of allergic rhinitis by nasal cytology and immunoglobulin E antibody levels in children. *Am J Rhinol Allergy*. 2018;32(1):16-22.
10. Schoos A-M, Chawes B, Følsgaard N, Samandari N, Bønnelykke K, Bisgaard H. Disagreement between skin prick test and specific IgE in young children. *Allergy*. 2015;70(1):41-48.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Malizia V, Cilluffo G, Fasola S, et al. Endotyping allergic rhinitis in children: A machine learning approach. *Pediatr Allergy Immunol*. 2022;33(Suppl. 27):18–21. https://doi.org/10.1111/pai.13620