Pediatric and Adult Asthma Clinical Phenotypes: A Real World, Big Data Study Based on Acute Exacerbations

Jie Xu, PhD[1], **Jiang Bian, PhD**[1], **Jennifer N Fishe, MD**[2,3]

[1]Department of Health Outcomes and Bioinformatics, University of Florida, Gainesville, Florida, USA

[2]Center for Data Solutions, University of Florida College of Medicine – Jacksonville, Jacksonville, Florida, USA

[3]Department of Emergency Medicine, University of Florida College of Medicine – Jacksonville, Jacksonville, Florida, USA

Abstract

Introduction: Asthma is a heterogeneous disease with a range of observable phenotypes. To date, the characterization of asthma phenotypes is mostly limited to allergic versus non-allergic disease. Therefore, the aim of this big data study was to computationally derive asthma subtypes from the OneFlorida Clinical Research Consortium

Methods: We obtained data from 2012-2020 from the OneFlorida Clinical Research Consortium. Longitudinal data for patients greater than two years of age who met inclusion criteria for an asthma exacerbation based on International Classification of Diseases codes. We used matrix factorization to extract information and K-means clustering to derive subtypes. The distributions of demographics, comorbidities, and medications were compared using Chi-square statistics.

Results: A total of 39,807 pediatric patients and 23,883 adult patients met inclusion criteria. We identified five distinct pediatric subtypes and four distinct adult subtypes. Pediatric subtype P1 had the highest proportion of black patients, but the lowest use of inhaled corticosteroids and allergy medications. Subtype P2 had a predominance of patients with gastroesophageal reflux disease, whereas P3 had a predominance of patients with allergic disorders. Adult subtype A2 was the most severe and all patients were on biologic agents. Most of subtype A3 patients were not taking controller medications, whereas most patients (>90%) in subtypes A2 and A4 were taking corticosteroids and allergy medications.

*Corresponding Author: Jennifer Fishe, MD; 655 W 8th St., Jacksonville, FL, 32209, Telephone: 904-244-4046, Jennifer.Fishe@jax.ufl.edu.

Conclusion: We found five distinct pediatric asthma subtypes and four distinct adult asthma subtypes. Future work should externally validate these subtypes and characterize response to treatment by subtype to better guide clinical treatment of asthma.

Introduction

Asthma is one of the most prevalent diseases, affecting 339 million people worldwide and over 25 million people in the United States.[1,2] The diagnosis of asthma is made clinically by observing a constellation of symptoms such as wheezing, shortness of breath, chest tightness, and cough, in combination with expiratory airflow limitation.[3] However, rather than being a uniform condition, asthma is a term that describes a variety of observable phenotypes that result in chronic airway inflammation.[4,5] Therefore, patients with asthma have different etiologies of their respiratory disease, triggers, clinical presentation, and response to treatment.

Since asthma is a heterogeneous disease, it follows that characterizing asthma phenotypes would aid in improved primary preventative care and treatment of asthma exacerbations. However, to date, the characterization of asthma phenotypes is largely confined to allergic versus non-allergic asthma,[6,7] which ignores overlapping phenotypes, and does not account for how asthma varies between pediatric and adult patients.[4,8] Patients with allergic asthma are candidates for targeted biologic therapies which have helped decrease the number and severity of exacerbations.[6,7] Accordingly, further characterization of other asthma phenotypes could also lead to targeted individualized therapy that decreases morbidity. In particular, patients with severe asthma whose exacerbations are resistant or poorly responsive to corticosteroids are ideal candidates for personalized management based on better characterization of their asthma phenotype.[5,9]

Prior studies on asthma phenotypes have either contained small sample sizes, and/or were confined to outpatient data.[5,8,10] Yet accounting for a patient's presentation during an asthma exacerbation (i.e., severity and response to treatment) is arguably one of the most clinically relevant ways to define asthma phenotypes.[11] Management of asthma exacerbations currently follows a "one size fits all" approach, which although evidence-based and tailored for severity and pediatric versus adult patients, does not account for different asthma phenotypes.[12,13] Therefore, as a first step towards better characterization of both adult and pediatric asthma phenotypes in order to improve management of exacerbations and chronic symptoms, the aim of this big data study was to computationally derive asthma subtypes from the OneFlorida Clinical Research Consortium.

Methods

Data Sources and Population

We obtained individual-level patient data from 2012-2020 from the OneFlorida Clinical Research Consortium,[14] which contains robust longitudinal and linked patient-level real world data of ~15 million (>60%) Floridians, including data from Medicaid claims, cancer registries, vital statistics, and electronic health records (EHRs) from its clinical partners. As one of the Clinical Data Research Networks contributing to the national Patient-Centered Clinical Research Network (PCORnet), OneFlorida includes 12 healthcare organizations that provide care through 4,100 physicians, 914 clinical practices, and 22 hospitals, covering all 67 Florida counties. OneFlorida follows the PCORnet Common Data Model including patient demographics, enrollment status, vital signs, conditions, encounters, diagnoses, procedures, prescribing (i.e., provider orders for medications), dispensing (i.e., outpatient pharmacy dispensing), and laboratory testing results.[15] This study was approved by the study institution's Institutional Review Board.

Longitudinal EHR data of patients who met inclusion criteria for an asthma exacerbation based on International Classification of Diseases (ICD) codes from the OneFlorida network between 2012 and 2020 were analyzed. We identified the asthma patients in our data based on diagnostic codes (ICD codes 9th and 10th revision) and required at least one asthma event. Selected codes include (ICD-9: 493.01, 493.02, 493.11, 493.12, 493.21, 493.22, 493.91, 493.92; ICD-10: J45.901, J45.902, J45.21, J45.31, J45.41, J45.51, J45.22, J45.32, J45.42, J45.52). Patients younger than two years of age were excluded from our study to avoid confounding with bronchiolitis.

Feature Representation

We set the index date to the first ever encounter in which an asthma exacerbation event was assigned (index date). We refer to the observed time as two years before and two years after the index date as the baseline period and use the information collected during that time period to derive subtypes. We extracted features from both adult and pediatric cohorts including demographics (e.g., age, sex, race / ethnicity), diagnoses (i.e., ICD 9 / 10 – see Supplementary Table 1), and drug prescriptions (i.e., National Drug Codes / RXNorm). We calculated age based on the date of birth and date of the index asthma exacerbation. We grouped the selected drugs to drug classes (see Supplementary Table 2). By determining whether a patient had received a diagnosis or a drug in those categories or not, we assigned a one or zero to each feature (i.e., one-hot encoding). We concatenated all features to represent each patient as a binary matrix $X = [x_1, \ldots, x_n] \in \{0,1\}^{p \times n}$. $x_{i,j} = 1$ denotes the i-th patient has p-th feature (i.e., diagnosis or drug classes).

Clustering With Nonnegative Matrix Factorization

Non-negative matrix factorization (NMF) is a machine learning technique that has been demonstrated to extract meaningful information from high-dimensional data such as gene expression microarrays.[16,17] More specifically, NMF uses unsupervised learning, which is the analysis of unlabeled datasets using machine learning algorithms to discover patterns and groupings that otherwise may be hidden. After representing the patient data as a data

matrix $X \in \{0,1\}^{p \times n}$, NMF is applied to decompose the data matrix X into the product of two matrices W and H,

$$X \approx W H,$$

where $X \in \mathbb{R}^{p \times n}$, $W \in \mathbb{R}^{p \times r}$ is a binary matrix and $H \in \mathbb{R}^{r \times n}$ is a non-negative matrix. In our study, rank r is determined by taking consideration of both the final objective values and learning efficiency. Matrices W and H preserve the most important properties of the input matrix X. Matrix H was, then, used as our new feature representation. The K-means clustering algorithm was applied on H to derive subtypes, and the number of clusters was determined by Elbow Method, Davies-Bouldin Index and Gap Statistic.[18] After deriving subtypes, we calculated the frequency of each variable in each subtype. The distributions of demographics, comorbidities, and medications were compared using Chi-square statistics.

Results

A total of 63,690 patients met study inclusion criteria (Figure 1). There were 39,807 pediatric patients ages 2 to18 years and 23,883 adult patients. Salient characteristics of both adult and pediatric patients overall are displayed in Figure 2.

Pediatric Patient Subtypes

Overall, the mean age of pediatric patients (N=39,807) was 7.1 years, 41.2% of were female, 28.2% were White and 40.5% were Black. The most common comorbidities were allergic rhinitis (45.1%), sinusitis (27.6%), and eczema (24.6%). The most common drug classes were systemic corticosteroids (90.4%), inhaled corticosteroids (73.4%), and allergy medications (67.8%).

There were five subtypes identified in pediatric patients (P1-P5). Characteristics of pediatric patients and their five subtype groups (P1-P5) are displayed in Table 1 and Figure 3. Patients in subtype P4 were relatively younger (mean age is 6.3) than other four subtypes (mean age greater than 7 years). Subtype P1 had a higher proportion of Black patients (N=5913, 46.9%) and less White patients (N=3092, 24.5%), whereas Subtype P2 was the opposite, with 28.3% Black and 35.9% White patients.

Compared with the other four subtypes, obesity, allergic rhinitis, and gastroesophageal reflux disease (GERD) were more common in subtype P2. Figure 3 represents the connections between five subtypes (represented by different color) and several most common comorbidities or medications. Each comorbidity or medication is represented by a fragment on the outer part of the circular layout. The size of the different colors of the arc is proportional to the number of patients in each subtype. In particular, GERD was found in 74.6% (N=2342) of patients with subtype P2, and less than 10% of patients with other subtypes. Subtype P3 had a higher proportion of patients with sinusitis (N=6201, 85%) and allergic rhinitis (N=4513, 61.8%).

Almost all patients in subtypes P2, P3 and P5 were taking inhaled corticosteroids, but only 49% (N=6,188) patients in subtype P1 and 29.6% patients (N=1,561) in subtype P4 were

taking inhaled corticosteroids. Most patient in subtypes P2-P5 took allergy medications, while only 0.5% (N=67) patients took allergy medications in Subtype P1.

Adult Patient Subtypes

Overall, the mean age of adult patients (N=23,883) was 44.4 years. The adult study population had a female predominance (80.8%). In contrast with pediatric patients, there were more adult White patients (40.9%) than Black patients (33.7%). The most common comorbidities were hypertension (59.1%), obesity (47.0%), GERD (39.7%), depression (32.2%), and tobacco use disorder (30.5%). The most common drug classes were systemic corticosteroids (66.8%), allergy medications (47.6%), and inhaled corticosteroids (45.9%)

There were four subtypes identified in adult patients (A1-A4). Characteristics of adult patients and their four subtypes groups are displayed in Table 2 and Figure 4. Patients with subtype A1 were younger (mean age 36.3 years) and had less comorbidities than the other three subtypes. Patients with subtype A3 were the oldest (mean age 55.1 years), and most patients had hypertension (N=6543, 97.6%). Subtype A2 appeared to be the most severe, as patients had more comorbidities than other subtypes. Additionally, all patients with subtype A2 were taking biologic drugs (i.e., Omalizumab, Mepolizumab, Benralizumab, Reslizumab). Subtype A4 had the most patients with tobacco use order (38.3%). Most patients in subtype A2 (> 95%) and A4 (> 90%) were taking systemic corticosteroids, inhaled corticosteroids, and allergy medications, while most patients in subtype A3 were not taking those medications. Additionally, less than 30% of patients in subtype A1 and A3 were taking allergy drugs.

Discussion

In this big data study of the OneFlorida Clinical Research Consortium, we found five distinct pediatric asthma subtypes and four distinct adult asthma subtypes. To our knowledge, to date this is the largest sample size for a computationally derived asthma phenotyping study. Our results reinforce the traditional allergic versus non-allergic asthma subtype distinction,[6,7] but also further divide both allergic and non-allergic asthma into more distinct subtypes that account for severity of disease and response to treatment, as well as potentially asthma determinants such as race, ethnicity, and environmental triggers.

In the 39,807 included pediatric patients, we found five distinct subtypes (P1-P5). Although the overall pediatric cohort was 40.5% Black patients, Subtype P1 had the highest proportion of Black patients (46.9%) and lowest proportion of White patients (24.5%), while Subtype P2 had the lowest proportion of Black patients (28.3%) and the highest proportion of White patients (35.9%). Corresponding to that, subtype P2 had a significantly higher incidence of GERD, obesity, and obstructive sleep apnea comorbidities compared to P1 and the other subtypes with a larger proportion of Black patients. However, despite P1 having the highest proportion of Black patients, and the known higher morbidity and mortality suffered by Black children due to asthma,[19] patients in P1 had significantly lower rates of all asthma and allergy medication administrations and prescriptions, including inhaled corticosteroids (82.3% in P1 versus 99.6% in P2) and short-acting beta agonists (6.1% in P1 versus 26.4% in P2). That finding may represent a lack of access to care or other social determinants of

health for the more racially diverse P1 patients. It could also represent the known differences in pharmacogenomic response to asthma medications, particularly bronchodilators, among Black pediatric patients,[20, 21] which may lead to clinicians not prescribing those medications due to a perceived lack of effectiveness.

Based off the comorbidities of allergic rhinitis and eczema, subtypes P2, P3, and P5 appear to constitute what has been traditionally grouped as the allergic or atopic subtype of asthma. However, subtype P4, while not having higher rates of allergic rhinitis compared to other subtypes, did have a higher rate of patients with eczema and environmental allergies. Regardless, P2 is distinguished from the other subtypes with allergic characteristics by a statistically significant higher percentage of patients with a history of prematurity and low birth weight (8% in P2 versus around 1% in all other subtypes), higher rates of pulmonary hypertension that likely correspond to a history of prematurity (3% in P2 for near 0% for all other subtypes), and nearly one-third of P2 patients with a diagnosis of obesity. Corresponding to that, P2 had the most patients prescribed a biologic agent. However, P2 and subtype P5 both had nearly equivalent high rates of systemic and inhaled corticosteroid prescriptions (99.6% for both in P2 and 100% for both in P5). Thus, it appears subtypes P2 and P5 appear to be the most severe pediatric phenotypes. Further study, particularly of P2 and P5, but also of all the pediatric subtypes, is warranted to characterize in more detail patients' response to treatments (particularly inhaled and systemic corticosteroids, cornerstones of asthma controller and exacerbation therapy) and the burden of emergency department visits and intensive care admissions.

We found four distinct adult asthma phenotypes, with subtype A1 being the youngest (mean age 36.3 years), most male predominant (78.8% female), and least obese (16.4% obese versus > 60% obese in all other subtypes). Subtype A3 was the oldest and had significantly higher rates of both systemic hypertension (97.6% of A3 patients) and pulmonary hypertension (14.2% of A3 patients). Pulmonary hypertension could be related to obstructive sleep apnea, which is known to worsen chronic lung diseases such as asthma.[22] Further, poor control of systemic hypertension could also have a deleterious effect on asthma outcomes.[23]

However, despite those characteristics of subtype A3, it appears subtype A2 is the most severe phenotype, as 100% of A2 patients were on biologic agents. Subtype A2 also had the highest rates of obstructive sleep apnea (46.5%), allergic rhinitis (65.1%), and eczema (15.5%). Thus, it appears that subtype A2 falls into the classic allergic phenotype. However, it also appears that subtype A4 has allergic characteristics as well (36.3% of patients with allergic rhinitis and 10.5% of patients with eczema), thus perhaps A4 is also allergic but of a less severe phenotype. Further to this, A4 had the highest rates of tobacco use disorder, therefore perhaps the allergic characteristics of A4 of environmentally-induced and modifiable, in contrast to perhaps more a genetically driven allergic phenotype of A2.[7] As with the pediatric subtypes, further characterization of outcomes in terms of emergency department visits, intensive care admissions, response to treatment, and even pharmacogenetic data would help guide definitive preventative and emergency management of these distinct subtypes of patients.

Limitations

This study has limitations that merit consideration. First, although the OneFlorida Clinical Research Consortium is a large repository of real-world clinical data, it contains encounters from a distinct region of the United States, and thus the patient data may not be generalizable to the rest of the United States or more globally. Second, due to the unsupervised nature of the clustering methods we used, we ultimately decided the number of subtypes. Thus, there is some potential human error inherent in that approach, even though we used a variety of statistical methods to minimize this potential (i.e., the "Elbow" method and the Gap Statistic). Third, we performed our subtype characterization using a total of four years of clinical data per patient: two years before and two years after an index encounter. For older pediatric patients and for adult patients with mild asthma and / or infrequent encounters with the healthcare system, this may miss some patients or mischaracterize other patients. We also did not include more qualitative assessments of patients' asthma, such as quality of life or other validated asthma severity scores, which are difficult to obtain from the EHR (or not contained in the EHR). However, further study using natural language processing of free text note could help fill that knowledge gap. Further to this, we did not have access to granular treatment response data or genomic data which could add more breadth and explanation to the subtype characteristics.

Conclusion

We found five distinct pediatric asthma subtypes and four distinct adult asthma subtypes. These subtypes reinforce the traditional allergic versus non-allergic asthma distinction, but also add further characterization, and vary by race and comorbidities. Further work should be done to externally validate these subtypes and more definitively characterize their response to current therapies to better guide clinical treatment of asthma.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding:

Abbreviations:

EHR	electronic health record
GERD	gastroesophageal reflux disease
ICD	international classification of diseases
LABA	long-acting beta agonist
RWD	real world data
SABA	short-acting beta agonist

References

1. The Global Asthma Report. 2018. Auckland, New Zealand: Global Asthma Network, 2018. Available at: http://globalasthmareport.org/resources/Global_Asthma_Report_2018.pdf Accessed April 7, 2022.

2. National Health Interview Survey, National Center for Health Statistics, CDC Compiled 12/7/2020, available at: https://www.cdc.gov/asthma/most_recent_national_asthma_data.htm. Accessed April 7, 2022.

3. Pocket Guide for Asthma Management and Prevention: Adults and Children Older than 5 Years. Global Initiative for Asthma. Updated 2021. Available at: https://ginasthma.org/wp-content/uploads/2021/05/GINA-Pocket-Guide-2021-V2-WMS.pdf Accessed April 7, 2022.

4. Wenzel SE. Asthma: defining of the persistent adult phenotypes. Lancet. 2006;368:804–13. [PubMed: 16935691]

5. Wardlaw AJ, Silverman M, Siva R, Pavord ID, Green R. Multi-dimensional phenotyping: towards a new taxonomy for airway disease. Clin Exp Allergy. 2005;35:1254–1262. [PubMed: 16238783]

6. Kuruvilla ME, Lee FE, Lee GB. Understanding asthma phenotypes, endotypes, and mechanisms of disease. Clin Rev Allergy Immunol. 2019;56(2):219–233. [PubMed: 30206782]

7. Akar-Ghibril N, Casale T, Custovic A, Phipatanakul W. Allergic Endotypes and Phenotypes of Asthma. J Allergy Clin Immunol Pract. 2020;8(2):429–440. [PubMed: 32037107]

8. Conrad LA, Cabana MD, Rastogi D. Defining pediatric asthma: Phenotypes to Endotypes and Beyond. Pediatr Res. 2021;90(1):45–51. [PubMed: 33173175]

9. Heaney L, Robinson DS. Severe asthma treatment: need for characterizing patients. The Lancet. 2005;365(9463):974–976.

10. Haldar P, Pavord ID, Shaw DE, Berry MA, Thomas M, Brightling CE, Wardlaw AJ, Green RH. Cluster analysis and clinical asthma phenotypes. Am J Respir Crit Care Med. 2008;178(3):218–224. [PubMed: 18480428]

11. Fishe JN, Labilloy G, Higley R, Casey D, Ginn A, Baskovich B, Blake KV. Single Nucleotide Polymorphisms (SNPs) in PRKG1 & SPATA13-AS1 are associated with bronchodilator response: a pilot study during acute asthma exacerbations in African American children. Pharmacogenet Genomics. 2021;31(7):146–154. [PubMed: 33851947]

12. Expert Panel Working Group of the National Heart, Lung, and Blood Institute (NHLBI) administered and coordinated National Asthma Education and Prevention Program Coordinating Committee (NAEPPCC), Cloutier MM, Baptist AP, Blake KV, Brooks EG, Bryant-Stephens T, DiMango E, Dixon AE, Elward KS, Hartert T, Krishnan JA, Lemanske RF Jr, Ouellette DR, Pace WD, Schatz M, Skolnik NS, Stout JW, Teach SJ, Umscheid CA, Walsh CG. 2020

13. Focused Updates to the Asthma Management Guidelines: A Report from the National Asthma Education and Prevention Program Coordinating Committee Expert Panel Working Group. J Allergy Clin Immunol. 2020;146(6):1217–1270. [PubMed: 33280709] Erratum in: J Allergy Clin Immunol. 2021 Apr;147(4):1528–1530.

14. Shenkman E, Hurt M, Hogan W, Carrasquillo O, Smith S, Brickman A, et al. OneFlorida Clinical Research Consortium: linking a clinical and translational science institute with a community based distributive medical education model. Academic Medicine. 2018;93(3):451. [PubMed: 29045273]

15. Li Q, He Z, Guo Y, Zhang H, George TJ, Hogan W, et al. Assessing the Validity of an *a priori* Patient-Trial Generalizability Score using Real-world Data from a Large Clinical Data Research Network: A Colorectal Cancer Clinical Trial Case Study. AMIA Annu Symp Proc. 2019;2019:1101–10. [PubMed: 32308907]

16. Zhang Zhongyuan, et al. "Binary matrix factorization with applications." Seventh IEEE international conference on data mining (ICDM 2007). IEEE, 2007.

17. Pehkonen P, Wong G, Törönen P. Theme discovery from gene lists for identification and viewing of multiple functional groups. BMC bioinformatics. 2005 Dec;6(1):1–8. [PubMed: 15631638]

18. Kodinariya Trupti M., and Makwana Prashant R.. "Review on determining number of Cluster in K-Means Clustering." International Journal 1.6 (2013): 90–95.

19. Asthma prevalence, health care use and mortality: United States, 2003-05." 2015. Available from: http://www.cdc.gov/nchs/data/hestat/asthma03-05/asthma03-05.htm. Accessed May 24, 2022.

20. Naqvi M, Thyne S, Choudhry S, Tsai HJ, Navarro D, Castro RA, Nazario S, Rodriguez-Santana JR, Casal J, Torres A, et al. Ethnic-specific differences in bronchodilator responsiveness among African Americans, Puerto Ricans, and Mexicans with asthma. J Asthma. 2007;44(8):639–48. [PubMed: 17943575]

21. Fishe JN, Labilloy G, Higley R, Casey D, Ginn A, Baskovich B, Blake KV. "Single nucleotide polymorphisms (SNPs) in PRKG1 and SPATA13-AS1 are associated with bronchodilator response: a pilot study during acute asthma exacerbations in African American children." Pharmacogenomics and Genetics. 2021;31:146–154.

22. Locke BW, Lee JJ, Sundar KM. OSA and Chronic Respiratory Disease: Mechanisms and Epidemiology. Int J Environ Res Public Health. 2022 Apr 30;19(9):5473. DOI: 10.3390/ijerph19095473. [PubMed: 35564882]

23. Cardet JC, Bulkhi AA, Lockey RF. Non-respiratory Comorbidities in Asthma. J Allergy Clin Immunol Pract. 2021 Nov;9(11):3887–3897. DOI: 10.1016/j.jaip.2021.08.027. Epub 2021 Sep 4. [PubMed: 34492402]
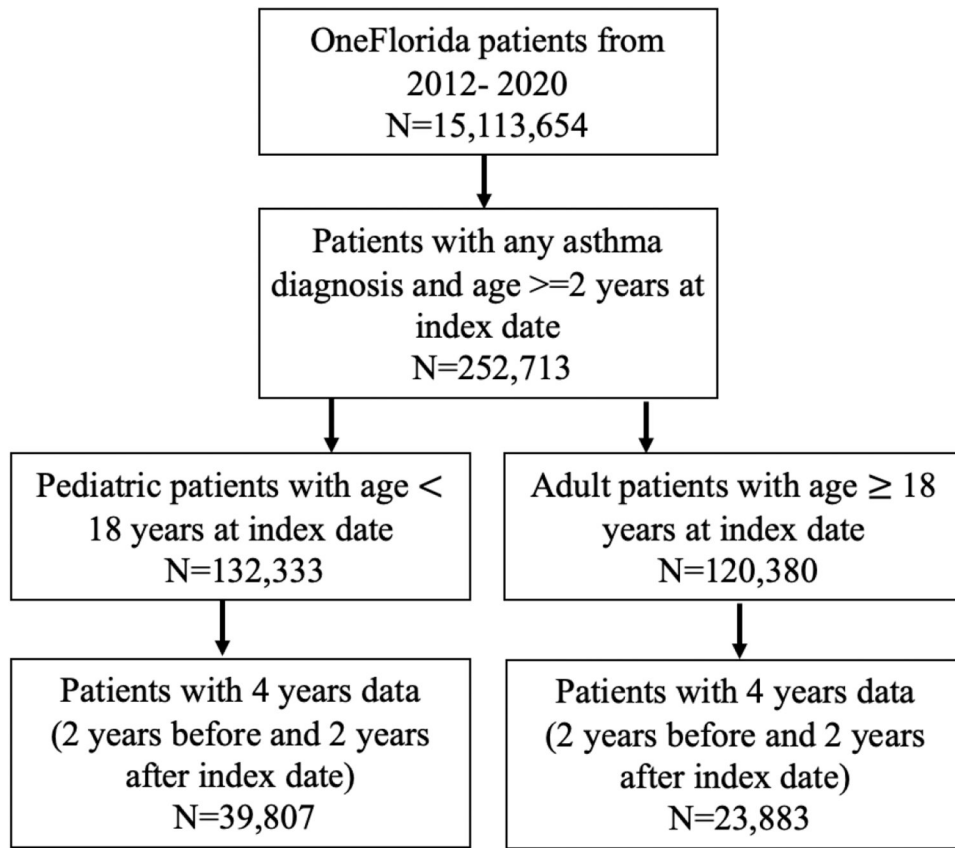
Figure 1: Inclusion and Exclusion Criteria Applied to Study Population
*Index date: first asthma diagnosis date
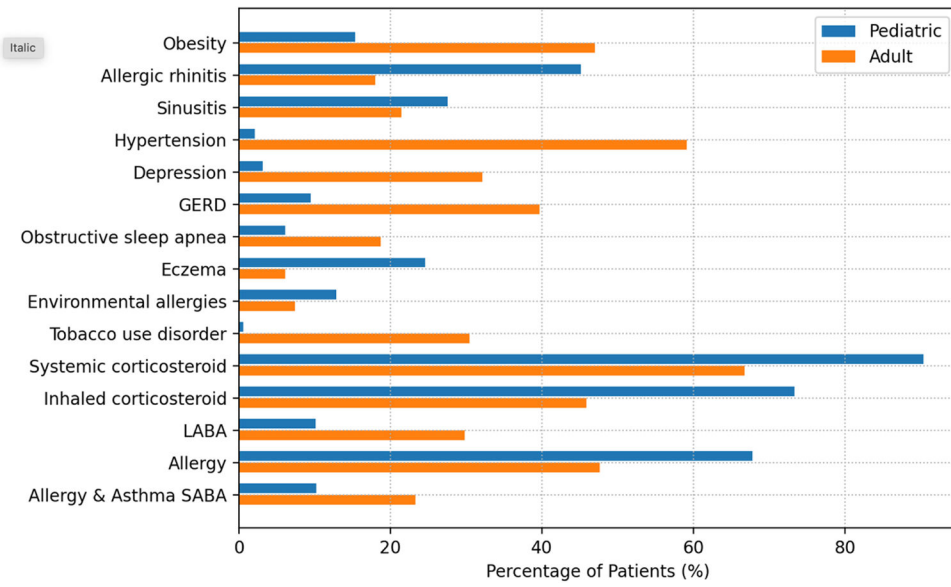
Figure 2: Relevant Comorbidities of Adult and Pediatric Study Patients
GERD = gastroesophageal reflux disease; LABA = long-acting beta agonist; SABA = short-acting beta agonist
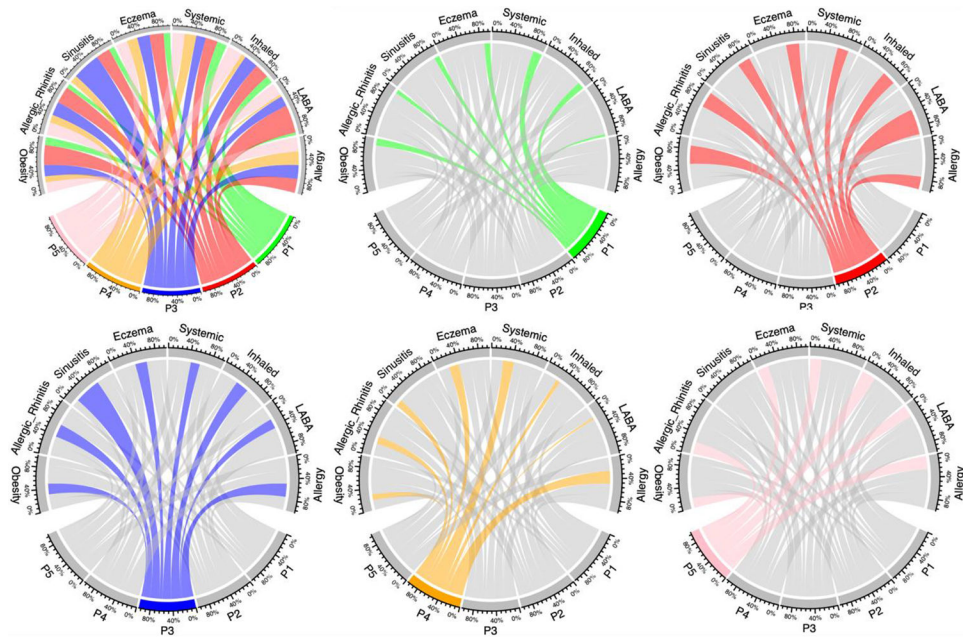
Figure 3: Graphical Depiction of Pediatric Subtypes

From top left to right bottom: All subtypes, P1, P2, P3, P4, P5.

For each individual figure, the color shading is anchored to the subtype represented. The projections of that color go to the characteristics found in the subtype, and the width of those projections represents the relative proportion of patients in that subtype who have that characteristic.
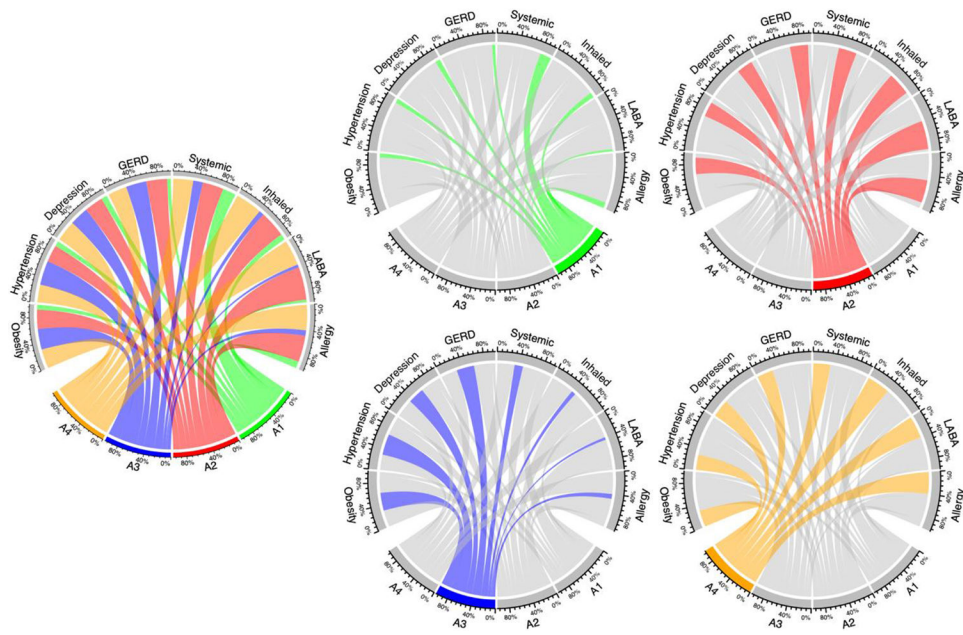
Figure 4: Graphical Depiction of Adult Subtypes

Left center: All subtypes; then from top left to bottom right: A1, A2, A3, A4.

For each individual figure, the color shading is anchored to the subtype represented. The projections of that color go to the characteristics found in the subtype, and the width of those projections represents the relative proportion of patients in that subtype who have that characteristic.

Table 1.

Characteristics of the 5 Pediatric Subtypes

| Characteristic | Total | Subtypes | | | | | χ^2 p-value |
		P1	P2	P3	P4	P5	
No. of patients (%)	39807 (100)	12621 (31.7)	3139 (7.9)	7299 (18.3)	5269 (13.2)	11479 (28.8)	
Age, Mean (SD), yr	7.1 (4.1)	7.1 (4.3)	7.3 (4.4)	7.3 (4.2)	6.3 (3.8)	7.3 (3.9)	
Sex, No. (%)							
Female	16381 (41.2)	5239 (41.5)	1328 (42.3)	3041 (41.7)	2160 (41)	4613 (40.2)	***
Race,[#] No. (%)							
Asian	239 (0.6)	65 (0.5)	20 (0.6)	49 (0.7)	29 (0.6)	76 (0.7)	***
Black	16127 (40.5)	5913 (46.9)	889 (28.3)	2495 (34.2)	2264 (43)	4566 (39.8)	***
White	11223 (28.2)	3092 (24.5)	1126 (35.9)	2140 (29.3)	1882 (35.7)	2983 (26)	***
American Indian or Alaska Native	71 (0.2)	21 (0.2)	7 (0.2)	13 (0.2)	7 (0.1)	23 (0.2)	ns
Native Hawaiian or Other Pacific Islander	3 (< 0.1)	1 (<0.1)	0 (0)	2 (<0.1)	2 (<0.1)	2 (<0.1)	ns
Multiple races	113 (2.8)	24 (0.2)	19 (0.6)	14 (0.2)	29 (0.6)	27 (0.2)	*
Other[^]	9517 (23.9)	2781 (22)	707 (22.5)	2044 (28)	859 (16.3)	3126 (27.2)	***
Comorbidities, No. (%)							
Obesity	6117 (15.4)	1533 (12.1)	967 (30.8)	1270 (17.4)	500 (9.5)	1847 (16.1)	ns
Allergic rhinitis	17972 (45.1)	2486 (19.7)	2427 (77.3)	4513 (61.8)	1881 (35.7)	6665 (58.1)	***
Sinusitis	10983 (27.6)	2138 (16.9)	1330 (42.4)	6201 (85)	1182 (22.4)	132 (1.1)	***
Hypertension	853 (2.1)	208 (1.6)	267 (8.5)	166 (2.3)	71 (1.3)	141 (1.2)	***
Second hand smoke exposure	1665 (4.2)	412 (3.3)	206 (6.6)	259 (3.5)	312 (5.9)	476 (4.1)	***
Depression	1260 (3.2)	300 (2.4)	244 (7.8)	278 (3.8)	126 (2.4)	312 (2.7)	***
GERD	3781 (9.5)	690 (5.5)	2342 (74.6)	389 (5.3)	360 (6.8)	0 (0)	***
Obstructive sleep apnea	2453 (6.2)	341 (2.7)	1191 (37.9)	424 (5.8)	182 (3.5)	315 (2.7)	***
Eczema	9783 (24.6)	1981 (15.7)	1071 (34.1)	2173 (29.8)	1330 (25.2)	3228 (28.1)	***
Pulmonary hypertension	178 (0.4)	45 (0.4)	99 (3.2)	14 (0.2)	9 (0.2)	11 (0.1)	***
Disorders relating to short gestation and low birthweight	668 (1.7)	152 (1.2)	257 (8.2)	77 (1.1)	87 (1.7)	95 (0.8)	***
Environmental allergies	5130 (12.9)	689 (5.5)	656 (20.9)	1188 (16.3)	890 (16.9)	1707 (14.9)	***
Allergic bronchopulmonary aspergillosis	17 (0)	3 (0)	6 (0.2)	4 (0.1)	1 (0)	3 (0)	*
Drug Class, No. (%)							
Systemic corticosteroid	35967 (90.4)	10382 (82.3)	3126 (99.6)	5875 (80.5)	5105 (96.9)	11479 (100)	***
Inhaled corticosteroid	29203 (73.4)	6188 (49)	3127 (99.6)	6848 (93.8)	1561 (29.6)	11479 (100)	***
Long-acting beta agonist	4023 (10.1)	355 (2.8)	1005 (32)	872 (11.9)	126 (2.4)	1665 (14.5)	***
Allergy	27004 (67.8)	67 (0.5)	3087 (98.3)	7266 (99.5)	5105 (96.9)	11479 (100)	**
Allergy & Asthma shortacting beta agonist	4063 (10.2)	776 (6.1)	830 (26.4)	489 (6.7)	630 (12)	1338 (11.7)	***
Biologic agent	116 (0.3)	67 (0.5)	49 (1.6)	0 (0)	0 (0)	0 (0)	ns

Characteristic	Total	Subtypes					χ^2 p-value
		P1	P2	P3	P4	P5	
Other controller	315 (0.8)	18 (0.1)	266 (8.5)	8 (0.1)	23 (0.4)	0 (0)	ns
Other short-acting beta agonist	4 (0)	0 (0)	1 (0)	0 (0)	1 (0)	2 (0)	ns

*
p<0.05

**
p<0.01

***
p<0.001

#
ns=not significant, GERD = gastroesophageal reflux disease, Race - missing data not shown

^
Other = Other as defined by PCORnet common data model v6.0.

Table 2.

Characteristics of the 4 Adult Subtypes

Characteristic	Total	Subtypes				χ^2 p-value
		A1	A2	A3	A4	
No. of patients (%)	23883 (100)	9261 (38.8)	129 (0.5)	6705 (28.1)	7788 (32.6)	
Age, Mean (SD), yr	44.4 (18.2)	36.3 (16.4)	44.8 (15.5)	55.1 (17.7)	44.7 (15.7)	
Sex, No. (%)						
Female	19305 (80.8)	7302 (78.8)	106 (82.2)	5387 (80.3)	6510 (83.6)	
Race,[#] No. (%)						
Asian	173 (0.7)	71 (0.8)	1 (0.8)	53 (0.8)	48 (0.6)	ns
Black	8037 (33.7)	3199 (34.5)	37 (28.7)	2077 (31)	2724 (35)	***
White	9770 (40.9)	3925 (42.4)	46 (35.7)	2841 (42.4)	2958 (38)	***
American Indian or Alaska Native	54 (0.2)	23 (0.2)	1 (0.8)	11 (0.2)	19 (0.2)	*
Native Hawaiian or Other Pacific Islander	7 (0.1)	7 (0.1)	0 (0)	0 (0)	0 (0)	ns
Multiple races	135 (5.7)	52 (0.6)	1 (0.8)	53 (0.8)	29 (0.4)	ns
Other [^]	4208 (17.6)	1662 (17.9)	29 (22.5)	1191 (17.8)	1326 (17)	***
Comorbidities, No. (%)						
Obesity	11236 (47.0)	1520 (16.4)	82 (63.6)	4612 (68.8)	5022 (64.5)	***
Allergic rhinitis	4310 (18.0)	670 (7.2)	84 (65.1)	729 (10.9)	2827 (36.3)	***
Sinusitis	5136 (21.5)	1284 (13.9)	66 (51.2)	1022 (15.2)	2764 (35.5)	***
Hypertension	14118 (59.1)	1916 (20.7)	89 (69)	6543 (97.6)	5570 (71.5)	***
Second hand smoke exposure	284 (1.2)	56 (0.6)	4 (3.1)	67 (1)	157 (2)	ns
Depression	7683 (32.2)	1212 (13.1)	54 (41.9)	2821 (42.1)	3596 (46.2)	***
GERD	9485 (39.7)	997 (10.8)	80 (62)	4065 (60.6)	4343 (55.8)	***
Obstructive sleep apnea	4476 (18.7)	134 (1.4)	60 (46.5)	2034 (30.3)	2248 (28.9)	***
Eczema	1468 (6.1)	306 (3.3)	20 (15.5)	328 (4.9)	814 (10.5)	***
Pulmonary hypertension	1825 (7.6)	59 (0.6)	13 (10.1)	955 (14.2)	798 (10.2)	***
Environmental allergies	1768 (7.4)	424 (4.6)	31 (24)	324 (4.8)	989 (12.7)	***
Tobacco use disorder	7284 (30.5)	2515 (27.2)	20 (15.5)	1769 (26.4)	2980 (38.3)	***
Allergic bronchopulmonary aspergillosis	26 (0.1)	4 (0)	2 (1.6)	5 (0.1)	15 (0.2)	ns
Drug Class, No. (%)						
Systemic corticosteroid	15947 (66.8)	5668 (61.2)	125 (96.9)	3033 (45.2)	7121 (91.4)	***
Inhaled corticosteroid	10955 (45.9)	2020 (21.8)	125 (96.9)	1212 (18.1)	7598 (97.6)	***
Long-acting beta agonist	7130 (29.9)	711 (7.7)	122 (94.6)	581 (8.7)	5716 (73.4)	***
Allergy	11372 (47.6)	2390 (25.8)	128 (99.2)	1413 (21.1)	7441 (95.5)	***
Allergy & Asthma short-acting beta agonist	5570 (23.3)	926 (10)	79 (61.2)	692 (10.3)	3873 (49.7)	***
Biologic agent	129 (0.5)	0 (0)	129 (100)	0 (0)	0 (0)	***
Other controller	1949 (8.2)	81 (0.9)	65 (50.4)	138 (2.1)	1665 (21.4)	***
Other short-acting beta agonist	20 (0.1)	0 (0)	1 (0.8)	0 (0)	19 (0.2)	ns

*
p<0.05

**
p<0.01

***
p<0.001

#
ns=not significant, GERD = gastroesophageal reflux disease, Race - missing data not shown

^
Other = Other as defined by PCORnet common data model v6.0.