Cumulative Genetic Risk for Asthma Contributes to Disease Severity in Children with Asthma living in Urban Environments

Matthew Dapas[1], William Wentworth-Sheilds[2], Emma E. Thompson[2], Rajesh Kumar[3], Elizabeth Lippner[3], Robert A. Wood[4], George T. O'Connor[5], Gurjit K. Khurana Hershey[6], Rebecca S. Gruchalla[7], Andrew H. Liu[8], Edward M. Zoratti[9], Leonard B. Bacharier[10], Stephanie Lovinsky-Desir[11], Michele A. Gill[12], William J. Sheehan[13], Shilpa J. Patel[14], Matthew C. Altman[15], James E. Gern[16], Cynthia M. Visness[17], Peter J. Gergen[18], Patrice M. Becker[18], Daniel J. Jackson[16*], Carole Ober[2*]

1. Division of Rheumatology, Department of Medicine; Center for Human Immunobiology; Center for Genetic Medicine; Northwestern University Feinberg School of Medicine, Chicago, IL
2. Department of Human Genetics, University of Chicago, Chicago, IL
3. Ann and Robert H. Lurie Children's Hospital; Division of Allergy and Immunology, Department of Pediatrics, Northwestern University Feinberg School of Medicine, Chicago, IL
4. Department of Pediatrics, Johns Hopkins University Medical Center, Baltimore, MD
5. Pulmonary Center, Boston University School of Medicine, Boston, MA
6. Division of Asthma Research, Cincinnati Children's Hospital, Cincinnati, OH
7. Department of Internal Medicine, University of Texas Southwestern Medical Center, Dallas, TX
8. Children's Hospital Colorado; Department of Allergy and Immunology, University of Colorado School of Medicine, Aurora, CO
9. Departments of Medicine and Public Health Sciences, Henry Ford Health and Michigan State University College of Human Medicine, Detroit, MI
10. Monroe Carell Jr. Children's Hospital at Vanderbilt University Medical Center, Nashville, TN
11. Department of Pediatrics and Environmental Health Sciences, Columbia University, New York, NY
12. Department of Pediatrics, Washington University School of Medicine, St. Louis, MO
13. Children's National Hospital; Division of Allergy and Immunology, George Washington University School of Medicine and Health Sciences, Washington, DC
14. Division of Emergency Medicine, Children's National Hospital, Washington, DC
15. Department of Allergy and Infectious Diseases, University of Washington, Seattle, WA
16. Departments of Pediatrics and Medicine, University of Wisconsin School of Medicine and Public Health, Madison, WI
17. Rho Inc., Durham, NC
18. National Institute of Allergy and Infectious Diseases, Bethesda, MD

* These authors jointly supervised this work.

ABSTRACT

Background: Childhood-onset asthma is highly heritable, with nearly 200 risk loci identified in genome-wide association studies. Aggregated polygenic risk scores can be used to quantify genetic predisposition to asthma, but their power to predict asthma severity in multi-ancestral groups has not been determined.

Objective: Our aim was to examine the predictive power of biobank-derived asthma polygenic risk scores in children with asthma living in urban environments.

Methods: We generated polygenic risk scores for asthma, derived from a large-scale genome-wide association meta-analysis, in four multi-ancestry asthma study cohorts of children living in urban environments. We assessed genetic predictions across different subphenotypes of asthma and tested for associations between genetic asthma risk and measures of asthma severity.

Results: Genetic asthma prediction was significantly stronger for more symptomatic asthma phenotypes ($P<0.001$). Polygenic risk scores were significantly higher in difficult-to-control vs. easy-to-control asthma ($P=0.02$). Genetic risk was also significantly associated with more frequent exacerbations ($P=0.03$), higher blood eosinophil levels ($P=0.01$), and lower lung function ($P<0.001$).

Conclusion: Cumulative genetic risk for asthma is associated with disease severity and exacerbation risk in children with asthma living in urban environments.

INTRODUCTION

Children living in urban environments face an excess burden of asthma morbidity and mortality[1]. Urban environments are enriched for environmental exposures and sociodemographic factors that are associated with increased asthma prevalence and severity[2,3]. Genetics also contributes to the overall risk of developing asthma, as evidenced by nearly 200 risk loci uncovered in a large genome-wide association study (GWAS)[4]. Understanding how genetic risk contributes to asthma severity in children living in urban environments could lead to more accurate risk stratification and ultimately help guide personalized management strategies.

Genetic risk can be modeled using aggregated polygenic risk scores (PRSs). Many studies have developed predictive PRSs for asthma[5], and phenome-wide association studies with asthma PRSs have suggested that genetic risk may correlate with asthma severity[5,6]. However, the relationship between cumulative genetic asthma risk and disease severity has not been fully explored[6]. Here, we generated PRSs for asthma in four multi-ancestry asthma study cohorts of children living in urban environments. We then assessed genetic predictions for different subphenotypes of asthma and tested for associations between genetic asthma risk and multiple measures of asthma severity.

RESULTS

We utilized genome-wide genotype data from four multi-ancestry asthma cohorts of children living in urban environments (APIC[7], URECA[8], MUPPITS-1[9], and MUPPITS-2[1]), as summarized in **Table 1**, including 1,108 children with childhood-onset asthma. The APIC (Asthma Phenotypes in Inner-City Children) study was an observational study in which different phenotypes of childhood-onset asthma (clusters A-E) were identified according to varying symptoms, biomarkers, and disease severity[7]. The URECA (Urban Environment and Childhood Asthma) study is a birth cohort enriched for family history of asthma or allergic disease[8]. The MUPPITS (Mechanisms Underlying Asthma Exacerbations Prevented and Persistent with Immune Based Therapy: A Systems Approach) studies investigated exacerbations in children with exacerbation-prone eosinophilic asthma[1,9]. Nearly all of the participants in these studies were parent-identified as Black or Hispanic, and the cohorts featured similar overall distributions of race/ethnicity (**Table 1**). Non-asthma population-based controls from the Consortium on Asthma in African-ancestry Populations in the Americas (CAAPA)[10] were used as a comparator group (n=205). The distribution of ancestry PCs by cohort are shown in **Figure S1**.

Table 1. Demographic characteristics of childhood-onset asthma cohorts

Characteristic	All	APIC	URECA	MUPPITS 1 & 2
Number	1462	507	527	428
Age, years, median (range)	10 (6-17)	11 (6-17)	10 (7-10)	11 (6-17)
Female sex	665 (45%)	216 (43%)	261 (50%)	188 (44%)
Race/Ethnicity				
Black (non-Hispanic)	966 (66%)	318 (63%)	377 (72%)	269 (63%)
White (non-Hispanic)	23 (2%)	7 (1%)	7 (1%)	9 (2%)
Hispanic	378 (26%)	154 (30%)	104 (20%)	120 (28%)
Other/mixed	93 (6%)	26 (5%)	38 (7%)	29 (7%)
Unknown	4 (<1%)	2 (<1%)	1 (<1%)	1 (<1%)
Site				
Baltimore	233 (16%)	84 (17%)	149 (28%)	-
Boston	228 (16%)	65 (13%)	124 (24%)	39 (9%)
Chicago	97 (7%)	62 (12%)	-	35 (8%)
Cincinnati	84 (6%)	45 (9%)	-	39 (9%)
Dallas	86 (6%)	38 (7%)	-	48 (11%)
Denver	103 (7%)	59 (12%)	-	44 (10%)
Detroit	106 (7%)	50 (10%)	-	56 (13%)
New York	227 (16%)	65 (13%)	99 (19%)	63 (15%)
St. Louis	207 (14%)	-	155 (29%)	52 (12%)
Washington, D.C	91 (6%)	39 (8%)	-	52 (12%)
Clinical traits				
Asthma Diagnosis	1108 (75%)	507 (100%)	173 (33%)	428 (100%)
Eosinophils/μL, median (IQR)	300 (190-500)	300 (160-485.5)	200 (100-350)	400 (300-600)
FEV_1/FVC, median (IQR)	0.80 (0.73-0.86)	0.79 (0.72-0.85)	0.84 (0.79-0.88)	0.76 (0.68-0.82)

APIC, Asthma Phenotypes in Inner-City Children; FEV1, Forced expiratory volume in one second; FVC, Forced vital capacity; IQR, interquartile range; MUPPITS, Mechanisms Underlying Asthma Exacerbations Prevented and Persistent with Immune Based Therapy: A Systems Approach; URECA, Urban Environment and Childhood Asthma

Polygenic risk for asthma was modeled using PRS-CS[11] on genome-wide association study summary statistics, as in our previous study[5], from a large multi-ancestry genetic association meta-analysis from the Global Biobank Meta-analysis Initiative (GBMI; 153,763 cases and 1,647,022 controls)[4]. PRSs were derived from the intersect of genotyped single nucleotide variants (SNVs) across all genotyped cohorts (n=791,608 SNVs).

PRS distributions differed significantly by cohort (P<0.001, Kruskal-Wallis), with higher median PRSs observed in cohorts characterized by more severe forms of asthma. (**Figures 1A, 1B**). Notably, the PRSs were significantly higher in children without asthma but enriched for family history of asthma than in population-based controls without asthma, likely owing to inherited disease risk (P=0.03, Wilcoxon). The

PRS predicted asthma diagnosis more accurately in cohorts characterized by greater symptom burden, according to relative areas under the receiver operating characteristic curve (AUCs, **Figure 1C**). Overall, the PRS yielded significantly better prediction for moderate-severe asthma (AUC=0.65) than for the other asthma phenotypes (AUC=0.53; P_{Δ}=<0.001, **Figure 1D**). These results demonstrate that genetic asthma prediction is stronger for more symptomatic phenotypes and further suggest that genetic risk correlates with increasing disease severity in children.

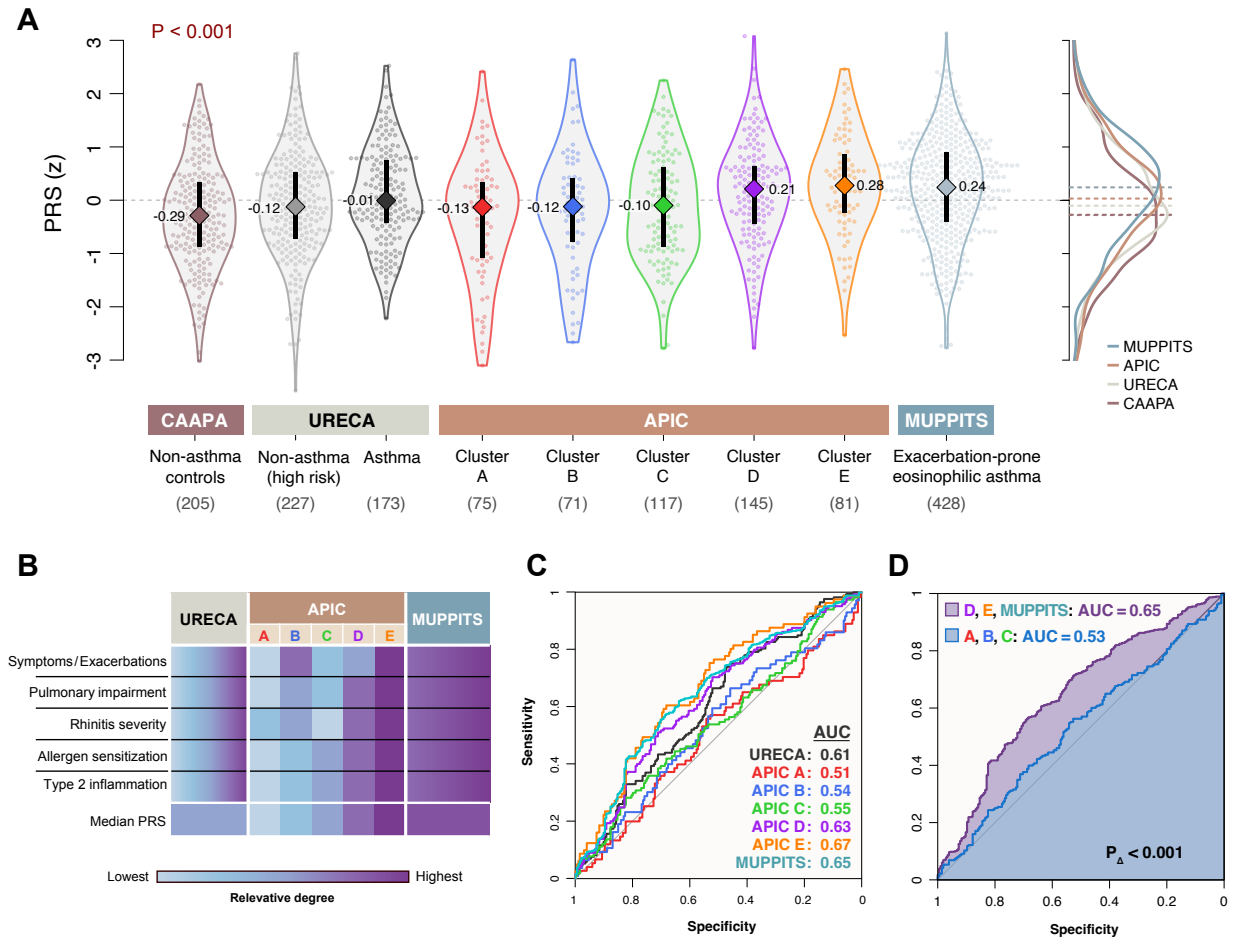

Figure 1. Genetic asthma prediction is stronger in more symptomatic phenotypes. A) Distributions and median values of asthma PRSs by genotyped cohort, including APIC's previously defined sub-phenotypes (clusters A-E)[7]. B) Relative degrees of categorical asthma severity for each cohort. C) Receiver operating characteristic (ROC) curves with areas under the curve (AUC) for predicting asthma using the PRS for each cohort. D) Prediction of asthma in the combined moderate-to-severe asthma cohorts (APIC D, E & MUPPITS) was significantly more accurate than for the milder asthma subphenotypes (APIC A, B, C), combined.

To more directly test whether polygenic asthma risk correlates with disease severity, we tested PRS associations with severity metrics within individual study cohorts. In APIC, PRSs were significantly higher

in difficult-to-control vs. easy-to-control asthma[12] (P=0.02; **Figure 2A**), indicating that children with more genetic risk alleles tended to require higher treatment levels to control their asthma. In the MUPPITS-2 trial, which investigated the effect of an anti-IL-5 inhibitor (mepolizumab) in children with exacerbation-prone eosinophilic asthma[1], polygenic risk was significantly associated with higher exacerbation rates. However, the association was only observed in the placebo-treated group (P=0.03; **Figure 2B**). The correlation was not significant in the study arm treated with mepolizumab (P=0.53), suggesting that the biologic effectively reduced the genetic effect on the exacerbation rate.

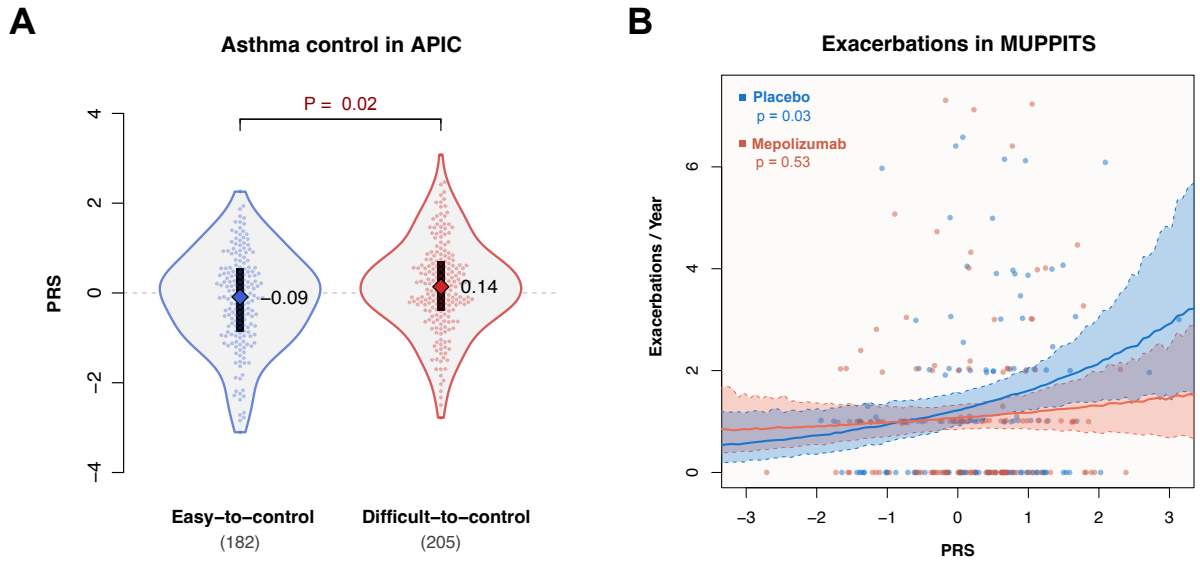

Figure 2. Polygenic risk is associated with asthma severity and exacerbation risk. A) Distributions and median values of asthma PRSs in APIC by required treatment step[12]. B) Exacerbation rate vs. PRSs in MUPPITS-2, split by treatment arm.

We next assessed associations between the PRS and quantitative biomarkers of asthma severity, adjusting for age, sex, study, and genetic ancestry (**Figure E1** in the Online Repository)[13]. In children with asthma across all cohorts, the PRS was significantly correlated with higher blood eosinophil levels (P=0.01; **Figure 3A**) and lower lung function (P<0.001; **Figure 3C**), as measured by the ratio of forced expiratory volume in one second to forced vital capacity (FEV_1/FVC). Comparing the highest to lowest PRS deciles, these effects amounted to significant median differences of 101 eosinophils per μL (P=0.02; **Figure 3B**), and -3.0% FEV_1/FVC (P=0.005; **Figure 3D**). These findings further demonstrate that genetic asthma risk is correlated with quantitative measures relevant to asthma severity.

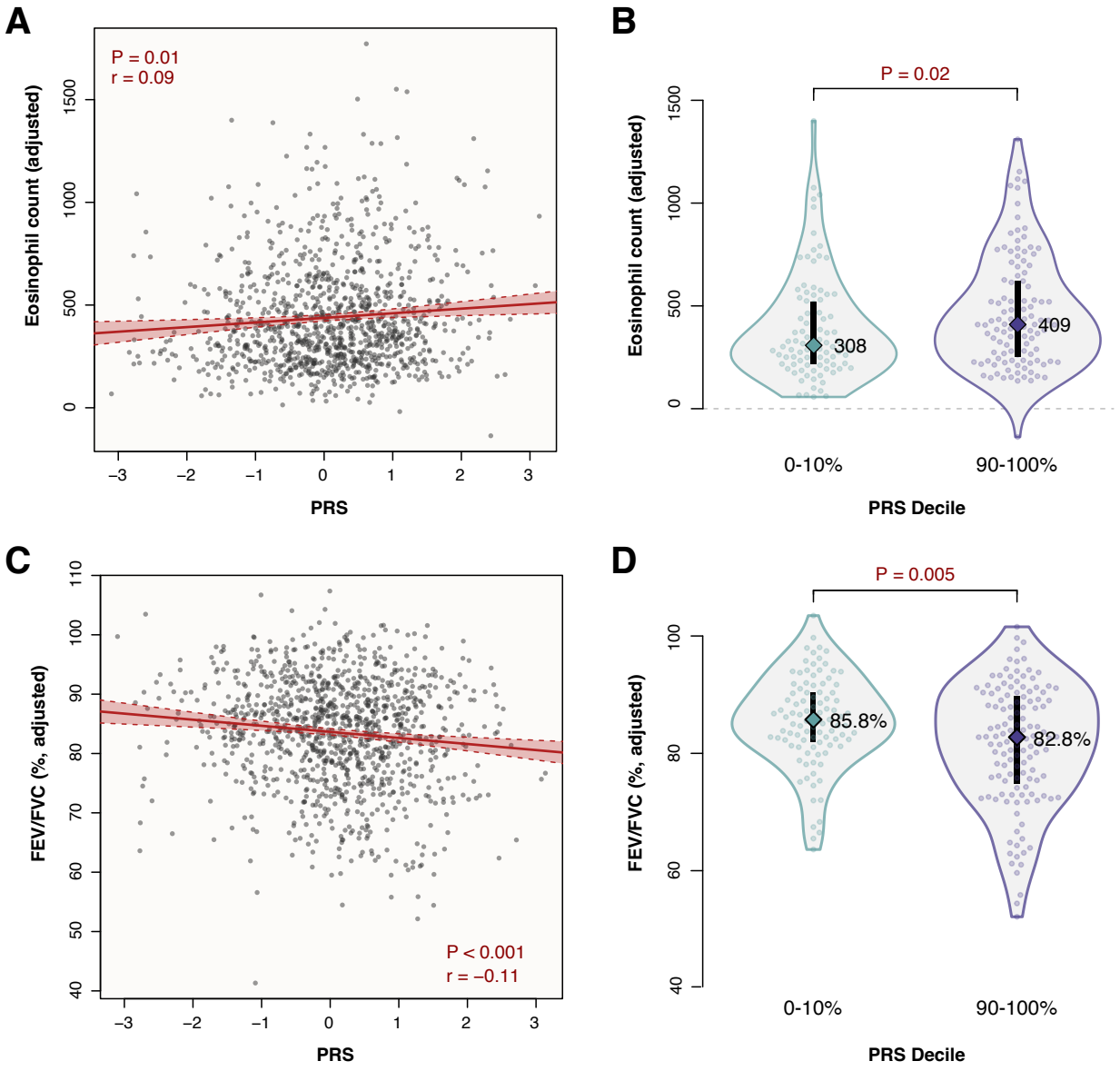

Figure 3. Polygenic risk is associated with quantitative biomarkers of asthma severity. A) Eosinophil count per μL, adjusted for age, sex, study, and ancestry, plotted against the PRS in a combined sample from APIC, URECA, and MUPPITS (n=1,062). The modeled correlation fit with 95% confidence interval is drawn in red. B) Distributions and median values of adjusted eosinophil counts in the lowest and highest PRS deciles. C) Adjusted FEV_1/FVC plotted against the PRS in a combined sample from APIC, URECA, and MUPPITS (n=1,098). D) Distributions and median values of adjusted FEV_1/FVC in the lowest and highest PRS deciles.

DISCUSSION

Genetic disease risk depends on phenotype definitions, ancestries, and environmental exposures. The GBMI asthma GWAS that we used for modeling our PRS drew predominantly from European and East Asian adult

cohorts and defined asthma according to specific International Classification of Diseases (ICD)-9 and ICD-10 codes in participant electronic health records[14]. Therefore, the asthma phenotypes, ancestries, and environments captured in the GBMI GWAS differ overall from the four cohorts studied here, which were limited to childhood-onset asthma cases in urban environments in the United States, consisting largely of Black and Hispanic populations, as well as from the CAAPA controls, which were African American. PRS predictive performance deteriorates across ancestries[15], and we previously demonstrated how modeling genetic risk from different asthma GWASs yielded PRSs with different trait associations[5]. Furthermore, asthma cases in the GBMI GWAS consisted largely of adult-onset disease[4,5], and the genetic architectures of childhood- and adult-onset asthma are partly distinct, with reported genetic correlations ranging from 0.67-0.78[16,17]. Although the GBMI meta-analysis is the largest and most diverse asthma GWAS conducted to date, and the GBMI-derived PRS predicted asthma in each of the four cohorts tested here with comparable accuracy to previous asthma PRS studies[5], the PRS is not a perfect proxy for genetic asthma risk in these cohorts. Thus, the actual correlation between genetic asthma risk and disease severity is likely stronger than the effects we observed here.

Among patients with physician-diagnosed asthma according to ICD-10 codes, the vast majority have mild asthma, and the prevalence of severe asthma is less than 5%[18]. Asthma in the GBMI GWAS was primarily defined according to ICD codes in the constituent biobanks. Therefore, it is unlikely that the PRS correlations we observed with asthma severity are due to asthma in GBMI being disproportionately severe. Instead, our results indicate that the cumulative burden of genetic risk variants for asthma in general contributes to disease severity. This conclusion is supported by GWASs of moderate-to-severe asthma, which found that 88% of association signals that reached genome-wide significance had been previously reported in studies that predominantly assessed samples from patients with mild asthma[19,20]. Our findings expand on those from a population-based longitudinal study in European ancestry individuals, in which asthma cases with a higher polygenic burden among 17 risk variants were more likely to manifest atopy, airway hyperresponsiveness, and incompletely reversible airflow obstruction, and more often experienced hospitalizations for breathing problems[21]. Our results also align with another study that found polygenic risk for asthma was significantly associated with asthma exacerbations in European ancestry individuals[22]. Similar observations, where cumulative genetic risk from low-effect variants were associated with disease severity. have been reported in other complex diseases, as well, including cardiovascular disease[23] and autoimmune disease[24]. Such genetic effects combine with those from rare variants[19] and environmental exposures[25] to contribute to disease severity through different pathways[2].

In summary, the results of our study indicate that the cumulative effect of asthma-associated genetic variants not only increase disease risk but also contribute to disease severity. The effects observed in this study

transcend ancestry, asthma phenotypes, and environments. Genetic risk models that are more closely aligned to demographic and clinical profiles of study populations could more precisely capture genetic influences on severity. Future studies should consider incorporating genetic risk for developing more comprehensive and accurate prediction models of asthma severity to inform strategies for clinical prevention, intervention, and management.

METHODS

Cohorts

We analyzed samples and phenotypes from several National Institutes of Allergy and Infectious Diseases (NIAID)-funded asthma studies conducted by the Inner-City Asthma Consortium (ICAC)[26]: the Asthma Phenotypes in the Inner City (APIC) study[7,12], the Urban Environment and Childhood Asthma (URECA) birth cohort study[8], and the Mechanisms Underlying Asthma Exacerbations Prevented and Persistent with Immune Based Therapy: A Systems Approach (MUPPITS 1 & 2) studies[1,9]. Non-asthma, population-based controls from the Consortium on Asthma among African-ancestry Populations in the Americas (CAAPA) study[10] was used as a comparator group to assess the distribution of the polygenic risk score (PRS) in an unselected healthy population with similar ancestry to the APIC, URECA, and MUPPITS participants. Ethical approval for this work was granted by a central institutional review board, WGC IRB[27].

APIC

APIC was a prospective, observational study in children with asthma age 6-17 years old from low-income areas (≥20% of residents below federal poverty level) of nine U.S. cities. Hierarchical clustering was performed on a large collection of baseline demographic, clinical, and environmental variables to identify different phenotypes of childhood-onset asthma. Five clusters were described (clusters A-E) according to varying symptoms, biomarkers, and disease severity[7]. APIC participants were also grouped based on required asthma controller treatment levels. Easy-to-control asthma was defined as requiring ≤100mcg/day of fluticasone at a majority of study visits, whereas difficult-to-control asthma was defined as requiring ≥500mcg/day[12]. APIC participants were genotyped using whole-genome sequencing (WGS) as previously described[13].

URECA

URECA was a prospective, observational study of a birth cohort enriched for family history of asthma or allergic disease in four U.S. cities[8]. All families lived in neighborhoods where at least 20% of the population had incomes below the poverty level. Clinical measurements from age 10 were used when available;

128 otherwise, the most recent measurement after age 5 were used. URECA participants were genotyped using
129 WGS as previously described[13].
130
131 *MUPPITS 1 & 2*
132 The MUPPITS studies investigated exacerbations in children with exacerbation-prone eosinophilic asthma
133 who were living in urban environments across nine U.S. cities. Participants in MUPPITS were 6-17 years
134 old, had blood eosinophils of ≥150 cells/μL, required at least twice-daily treatment with inhaled
135 corticosteroids, and had at least two exacerbations treated with systemic corticosteroids in the previous
136 year. All participants lived in census tracts where at least 10% of families had incomes below the poverty
137 level[28]. The MUPPITS-1 study was a prospective, longitudinal study that compared transcriptome profiles
138 in children who experienced cold symptoms between those that subsequently experienced an exacerbation
139 and those that did not[9]. The MUPPITS-2 study was a randomized controlled trial to assess the safety and
140 efficacy of mepolizumab for reducing exacerbations in this population[1]. Genotypes for MUPPITS 1 & 2
141 were determined in 461 subjects with available DNA using the Illumina Genome Diversity Array (GDA,
142 8v1.0, hg19) at the University of Chicago. Individual samples were removed if they met any of the
143 following quality control (QC) criteria: genotyped vs. self-reported sex mismatches, unexpected sample
144 duplication, >2% missing genotypes across all variants, or autosomal heterozygosity >5 standard deviations
145 from the mean. Variants were excluded under the the following QC criteria: >4% missing genotypes across
146 all samples or exact Hardy-Weinberg p-value $<5\times10^{-8}$. A total of 1,661,770 directly genotyped single-
147 nucleotide variants remained following QC. Variants were lifted from human reference hg19 to hg38 using
148 CrossMap[29]. Eagle[30] was used for genotype phasing and minimac4[31] was used for genotype imputation.
149 Imputation was performed using the TOPMed reference panel (r2 version 1.0.0) of >100,000 whole genome
150 sequences (human genome build 38) and the Michigan Imputation Server[32]. Following QC, 428 samples
151 from MUPPITS 1 & 2 remained for analysis.
152
153 *CAAPA*
154 CAAPA[10] performed WGS on approximately 1,000 individuals with African ancestry from 19 populations
155 across the Americas and Caribbean as previously described[33]. The WGS data is available through dbGAP
156 (accession phs001123.v2.p1). In this study, we utilized WGS data from CAAPA non-asthma control
157 samples from among the African American populations (n=205). Variant QC followed the same criteria as
158 applied to APIC and URECA WGS data[13].
159
160 **Modeling Genetic Asthma Risk**

Polygenic risk for asthma was modeled on association summary statistics from a large-scale, multi-ancestry genome-wide association study (GWAS) conducted as part of the Global Biobank Meta-analysis Initiative (GBMI). Summary statistics were available via https://www.globalbiobankmeta.org. The GWAS was a meta-analysis across 18 biobanks, consisting of 153,763 cases and 1,647,022 controls. Primary ancestry among asthma cases was 79.3% European, 12.1% East Asian, 3.3% African, 2.6% Admixed American, 2.6% Central and Southern Asian, and 0.01% Middle Eastern[4]. Asthma case and control statuses were assigned primarily according to "PheCodes" mapped from International Classification of Disease (ICD) codes, but several biobanks used self-reported data.

The polygenic risk model was generated using PRS-CS[11] on the intersect of genotyped single nucleotide variants from GBMI and across all genotyped cohorts. PRS-CS applies a Bayesian regression framework with continuous shrinkage priors on variant effect sizes. The default half-Cauchy priors, $\alpha = 1$ and $\beta = 0.5$, were used for determining local shrinkage parameters (ψj). We used the "PRS-CS-auto" algorithm to estimate the global shrinkage parameter (φ) directly from the GWAS summary statistics. For the Gibbs sampling used to derive the global shrinkage parameter and posterior variant effect sizes, we ran 10,000 Markov Chain Monte Carlo iterations with 5,000 burn-in steps. Each autosomal chromosome was modeled separately. We used the EUR linkage disequilibrium reference panel constructed by the PRS-CS developers using UK Biobank data. Posterior variant effects were aggregated into scores per individual using the Plink (v2.00a5.8LM AVX2 Intel, 21 Nov 2023) "score" function[34], with missing genotypes imputed as the posterior effect size multiplied by the effect allele frequency. Cumulative PRSs were converted to Z scores (across all genotyped cohorts) for analysis.

Ancestry estimation

Ancestry principal components (PCs) were calculated on the intersect of high-quality single-nucleotide variants genotyped in the asthma cohorts and several reference panels from the 1000 Genomes Project (1KG; n=156)[35] and the Human Genome Diversity Project (HGDP; n=52)[36], as in our previous study[13]. Ancestry PCs were calculated accounting for subject relatedness using PC-Air[37] and PC-Relate[38], with initial kinship estimates derived using KING[39]. Kinship was estimated in iterations until resultant PCs stabilized (n=3).

Association Testing

Differences in PRS distributions between groups were assessed using Wilcoxon rank-sum tests for pairwise comparisons and Kruskal-Wallis tests for more than two groups. Statistical differences in receiver operating characteristic areas under the curve were evaluated using DeLong's test. PRS effects on exacerbation rates in MUPPITS 2 were determined using negative binomial generalized linear models, adjusting for age, sex, and the first three PCs of global ancestry (**Figure E1**). PRS associations with quantitative traits across

194 multiple studies were assessed using a two-stage approach[13,40,41]. First, traits were regressed on age, sex,
195 study, and the first three ancestry PCs. Then, model residuals were rank-normalized and regressed again on
196 the same covariates with the addition of the PRS as a predictor. Adjusted trait values for plotting PRS
197 associations were derived by adding first-stage model intercepts to residuals.

198 **SUPPLEMENTAL FIGURES**

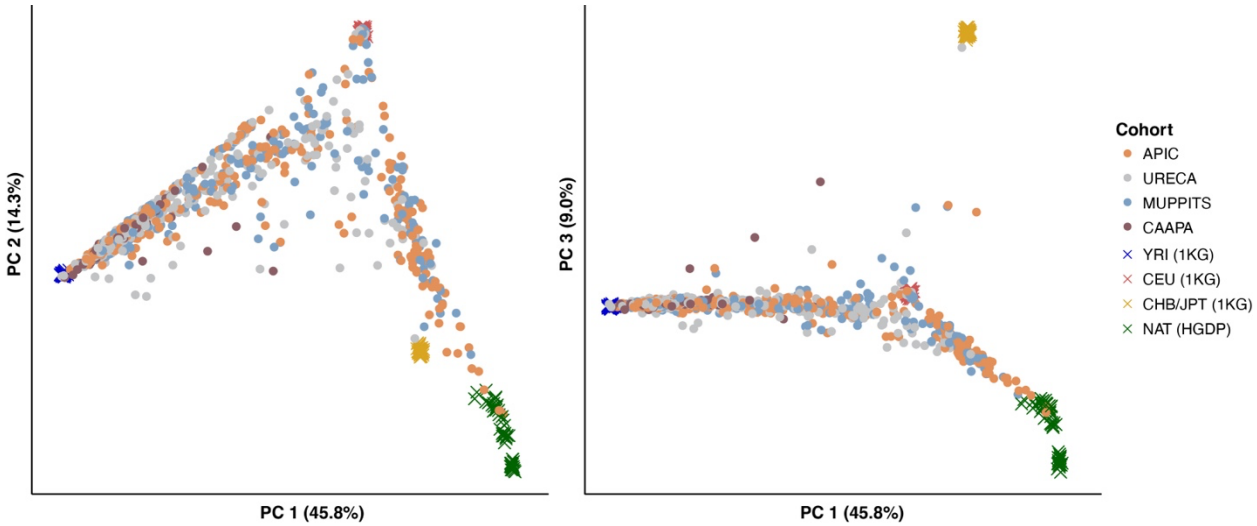

Figure S1. Principal components of ancestry by cohort. The top three principal components (PCs) of ancestry are plotted for sequenced APIC, URECA, MUPPITS, and CAAPA participants, colored by study, along with the four ancestry reference populations used for determining ancestry.

199

ACKNOWLEDGEMENTS

URECA V has been funded in whole or in part with Federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, and Department of Health and Human Services under Cooperative Agreement number 1UM1AI160040, HHSN272200900052C, HHSN272201000052I, 1UM1AI114271-01, and UM2AI117870. Additional support was provided by the National Center for Research Resources, National Center for Advancing Translational Sciences, and National Institutes of Health, under grants NCRR/NIHRR00052, NCRR/NIHM01RR00533, NCRR/NIH1UL1RR025771, NCRR/NIHM01RR00071, 1UL1RR024156, NCRR/NIH RR00052, UL1TR001079, NCRR/NIHM01RR00533, NCRR/NIH1UL1RR025771, NCRR/NIHM01RR00071, NCRR/NIH1UL1RR024156, NCATS/NIHUL1TR000040, 5UL1RR024992-02, NCATS/NIHUL1TR001079, NCATS/NIH1UL1TR001430, NCATS/NIHUL1TR001873, and NCATS/NIH UL1TR002345.

The APIC project has been funded in whole or in part with Federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, and Department of Health and Human Services under Contract numbers HHSN272200900052C, HHSN272201000052I, 1UM1AI114271-01, and UM2AI117870. Additional support was provided by the National Center for Research Resources, and National Center for Advancing Translational Sciences, National Institutes of Health, under grants NCATS/NIH UL1TR000150, NCRR/NIH UL1TR000451, UL1TR001105, NCRR/NIH 1UL1RR025780 (CCTSI is supported in part by Colorado CTSA Grant UL1TR000154 from NCATS/NIH UL1TR001082), NCATS/NIH UL1TR000040, NCATS/NIH UL1TR000075, NCRR/NCAT/NIH UL1TR000077-04. GlaxoSmithKline (GSK) provided Ventolin HFA, Flonase, Flovent 50 mcg, Flovent 100 mcg, Flovent 250 mcg , Advair 250/50 mcg, Advair 500/50 mcg under a clinical trial agreement with NIH NIAID. None of these companies had a role in the development or approval of the protocol, conduct of the trial, data analysis, manuscript preparation, or the decision to submit for publication.

The MUPPITS-1 and MUPPITS-2 projects have been funded in whole or in part with Federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, and Department of Health and Human Services under Contract numbers 1UM1AI114271-01 and UM2AI117870. Additional support was provided by the National Center for Research Resources, National Center for Advancing Translational Sciences, and National Institutes of Health, under grants UL1TRG01422 (05/19/08 – 04/29/12), CTSA Grant UL1RR025741 (04/30/12-08/11/15), CTSA Grant UL1TR000150 (08/12/15-03/31/19), CTSA Grant UL1TR001422, NIH/NCATS Colorado CTSA UL1 TR002535, NCATS/NIH UL1TR001876, and NIH/CTSA 5UL1TR001425- 03.

Dr. Becker's and Gergen's co-authorship of this report does not necessarily represent the views of the National Institute of Allergy and Infectious Diseases, the National Institutes of Health or any other agency of the United States Government.

DATA AVAILABILITY

Phenotype, genotype, GWAS summary statistics, and whole-genome sequencing files are available in dbGaP under accession phs002921 for the APIC and URECA studies and under accession phs004273 for the MUPPITS studies.

REFERENCES

1. Jackson DJ, Bacharier LB, Gergen PJ, Gagalis L, Calatroni A, Wellford S, et al. Mepolizumab for urban children with exacerbation-prone eosinophilic asthma in the USA (MUPPITS-2): a randomised, double-blind, placebo-controlled, parallel-group trial. Lancet. 2022;400(10351):502-11.
2. Liu AH, Babineau DC, Krouse RZ, Zoratti EM, Pongracic JA, O'Connor GT, et al. Pathways through which asthma risk factors contribute to asthma severity in inner-city children. J Allergy Clin Immunol. 2016;138(4):1042-50.
3. Poowuttikul P, Saini S, Seth D. Inner-City Asthma in Children. Clin Rev Allergy Immunol. 2019;56(2):248-68.
4. Tsuo K, Zhou W, Wang Y, Kanai M, Namba S, Gupta R, et al. Multi-ancestry meta-analysis of asthma identifies novel associations and highlights the value of increased power and diversity. Cell Genom. 2022;2(12):100212.
5. Dapas M, Lee YL, Wentworth-Sheilds W, Im HK, Ober C, Schoettler N. Revealing polygenic pleiotropy using genetic risk scores for asthma. HGG Adv. 2023;4(4):100233.
6. Namjou B, Lape M, Malolepsza E, DeVore SB, Weirauch MT, Dikilitas O, et al. Multiancestral polygenic risk score for pediatric asthma. J Allergy Clin Immunol. 2022;150(5):1086-96.
7. Zoratti EM, Krouse RZ, Babineau DC, Pongracic JA, O'Connor GT, Wood RA, et al. Asthma phenotypes in inner-city children. J Allergy Clin Immunol. 2016;138(4):1016-29.
8. Gern JE, Visness CM, Gergen PJ, Wood RA, Bloomberg GR, O'Connor GT, et al. The Urban Environment and Childhood Asthma (URECA) birth cohort study: design, methods, and study population. BMC Pulm Med. 2009;9:17.
9. Altman MC, Gill MA, Whalen E, Babineau DC, Shao B, Liu AH, et al. Transcriptome networks identify mechanisms of viral and nonviral asthma exacerbations in children. Nat Immunol. 2019;20(5):637-51.
10. Mathias RA, Taub MA, Gignoux CR, Fu W, Musharoff S, O'Connor TD, et al. A continuum of admixture in the Western Hemisphere revealed by the African Diaspora genome. Nat Commun. 2016;7:12522.
11. Ge T, Chen CY, Ni Y, Feng YA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. Nat Commun. 2019;10(1):1776.
12. Pongracic JA, Krouse RZ, Babineau DC, Zoratti EM, Cohen RT, Wood RA, et al. Distinguishing characteristics of difficult-to-control asthma in inner-city children and adolescents. J Allergy Clin Immunol. 2016;138(4):1030-41.
13. Dapas M, Thompson EE, Wentworth-Sheilds W, Clay S, Visness CM, Calatroni A, et al. Multi-omic association study identifies DNA methylation-mediated genotype and smoking exposure effects on lung function in children living in urban settings. PLoS Genet. 2023;19(1):e1010594.
14. Zhou W, Kanai M, Wu KH, Rasheed H, Tsuo K, Hirbo JB, et al. Global Biobank Meta-analysis Initiative: Powering genetic discovery across human disease. Cell Genom. 2022;2(10):100192.

279 15. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current
280 polygenic risk scores may exacerbate health disparities. Nat Genet. 2019;51(4):584-91.

281 16. Ferreira MAR, Mathur R, Vonk JM, Szwajda A, Brumpton B, Granell R, et al. Genetic
282 Architectures of Childhood- and Adult-Onset Asthma Are Partly Distinct. Am J Hum
283 Genet. 2019;104(4):665-84.

284 17. Pividori M, Schoettler N, Nicolae DL, Ober C, Im HK. Shared and distinct genetic risk
285 factors for childhood-onset and adult-onset asthma: genome-wide and transcriptome-wide
286 studies. Lancet Respir Med. 2019;7(6):509-22.

287 18. Sato K, Ohno T, Ishii T, Ito C, Kaise T. The Prevalence, Characteristics, and Patient
288 Burden of Severe Asthma Determined by Using a Japan Health Care Claims Database. Clin
289 Ther. 2019;41(11):2239-51.

290 19. Chang D, Hunkapiller J, Bhangale T, Reeder J, Mukhyala K, Tom J, et al. A whole genome
291 sequencing study of moderate to severe asthma identifies a lung function locus associated
292 with asthma risk. Sci Rep. 2022;12(1):5574.

293 20. Shrine N, Portelli MA, John C, Soler Artigas M, Bennett N, Hall R, et al. Moderate-to-
294 severe asthma in individuals of European ancestry: a genome-wide association study.
295 Lancet Respir Med. 2019;7(1):20-34.

296 21. Belsky DW, Sears MR, Hancox RJ, Harrington H, Houts R, Moffitt TE, et al. Polygenic
297 risk and the development and course of asthma: an analysis of data from a four-decade
298 longitudinal study. Lancet Respir Med. 2013;1(6):453-61.

299 22. Moll M, Sordillo JE, Ghosh AJ, Hayden LP, McDermott G, McGeachie MJ, et al.
300 Polygenic risk scores identify heterogeneity in asthma and chronic obstructive pulmonary
301 disease. J Allergy Clin Immunol. 2023;152(6):1423-32.

302 23. Sherafati A, Norland K, Naderian M, Schaid DJ, Kullo IJ. Polygenic Risk and Coronary
303 Artery Disease Severity. Circ Genom Precis Med. 2024;17(5):e004470.

304 24. Reid S, Alexsson A, Frodlund M, Morris D, Sandling JK, Bolin K, et al. High genetic risk
305 score is associated with early disease onset, damage accrual and decreased survival in
306 systemic lupus erythematosus. Ann Rheum Dis. 2020;79(3):363-9.

307 25. Lezmi G, Lejeune S, Pin I, Blanchon S, Bouazza N, Jolaine V, et al. Factors Associated
308 with Asthma Severity in Children: Data from the French COBRAPed Cohort. J Allergy
309 Clin Immunol Pract. 2021;9(5):1969-79.

310 26. Togias A, Fenton MJ, Gergen PJ, Rotrosen D, Fauci AS. Asthma in the inner city: the
311 perspective of the National Institute of Allergy and Infectious Diseases. J Allergy Clin
312 Immunol. 2010;125(3):540-4.

313 27. Stoffel B, Sorkness C, Pech C. Use of a Single, Independent IRB: Case Study of an NIH
314 Funded Consortium. Contemp Clin Trials Commun. 2017;8:114-21.

315 28. Altman MC, Flynn K, Rosasco MG, Dapas M, Kattan M, Lovinsky-Desir S, et al.
316 Inducible expression quantitative trait locus analysis of the MUC5AC gene in asthma in
317 urban populations of children. J Allergy Clin Immunol. 2021;148(6):1505-14.

318 29. Zhao H, Sun Z, Wang J, Huang H, Kocher JP, Wang L. CrossMap: a versatile tool for
319 coordinate conversion between genome assemblies. Bioinformatics. 2014;30(7):1006-7.

30. Loh PR, Palamara PF, Price AL. Fast and accurate long-range phasing in a UK Biobank cohort. Nat Genet. 2016;48(7):811-6.

31. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. Nat Genet. 2012;44(8):955-9.

32. Das S, Forer L, Schonherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. Nat Genet. 2016;48(10):1284-7.

33. Daya M, Rafaels N, Brunetti TM, Chavan S, Levin AM, Shetty A, et al. Association study in African-admixed populations across the Americas recapitulates asthma risk loci in non-African populations. Nat Commun. 2019;10(1):880.

34. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4.

35. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. Nature. 2015;526(7571):68-74.

36. Bergstrom A, McCarthy SA, Hui R, Almarri MA, Ayub Q, Danecek P, et al. Insights into human genetic variation and population history from 929 diverse genomes. Science. 2020;367(6484).

37. Conomos MP, Miller MB, Thornton TA. Robust inference of population structure for ancestry prediction and correction of stratification in the presence of relatedness. Genet Epidemiol. 2015;39(4):276-93.

38. Conomos MP, Reiner AP, Weir BS, Thornton TA. Model-free Estimation of Recent Genetic Relatedness. Am J Hum Genet. 2016;98(1):127-48.

39. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. Bioinformatics. 2010;26(22):2867-73.

40. McCaw ZR, Lane JM, Saxena R, Redline S, Lin X. Operating characteristics of the rank-based inverse normal transformation for quantitative trait analysis in genome-wide association studies. Biometrics. 2020;76(4):1262-72.

41. Sofer T, Zheng X, Gogarten SM, Laurie CA, Grinde K, Shaffer JR, et al. A fully adjusted two-stage procedure for rank-normalization in genetic association studies. Genet Epidemiol. 2019;43(3):263-75.