Multiancestral polygenic risk score for pediatric asthma

Check for updates

Bahram Namjou, MD,[a,b] **Michael Lape,**[a,c] **Edyta Malolepsza, PhD,**[e] **Stanley B. DeVore, MD,**[b,d] **Matthew T. Weirauch, PhD,**[a,b,c,f] **Ozan Dikilitas, MD,**[g,h] **Gail P. Jarvik, MD, PhD,**[i,j] **Krzysztof Kiryluk, MD,**[k] **Iftikhar J. Kullo, MD,**[h] **Cong Liu, PhD,**[l] **Yuan Luo, PhD,**[m] **Benjamin A. Satterfield, MD, PhD,**[h] **Jordan W. Smoller, MD, ScD,**[n,o,p] **Theresa L. Walunas, PhD,**[q] **John Connolly, PhD,**[r] **Patrick Sleiman, PhD,**[r,s] **Tesfaye B. Mersha, PhD,**[b,d] **Frank D. Mentch, PhD,**[r] **Hakon Hakonarson, MD, PhD,**[r,s] **Cynthia A. Prows, MSN, APRN,**[b,t,u] **Jocelyn M. Biagini, PhD,**[b,d] **Gurjit K. Khurana Hershey, MD, PhD,**[b,d,v] **Lisa J. Martin, PhD,**[b,t] **and Leah Kottyan, PhD,**[a,b,v] **The eMERGE Network**[w] *Cincinnati, Ohio; Cambridge and Boston, Mass; Rochester, Minn; Seattle, Wash; New York, NY; Chicago, Ill; Philadelphia, Pa; and Bethesda, Md*

Background: Asthma is the most common chronic condition in children and the third leading cause of hospitalization in pediatrics. The genome-wide association study catalog reports 140 studies with genome-wide significance. A polygenic risk score (PRS) with predictive value across ancestries has not been evaluated for this important trait.

Objectives: This study aimed to train and validate a PRS relying on genetic determinants for asthma to provide predictions for disease occurrence in pediatric cohorts of diverse ancestries. **Methods:** This study applied a Bayesian regression framework method using the Trans-National Asthma Genetic Consortium genome-wide association study summary statistics to derive a

From [a]the Center for Autoimmune Genomics and Etiology, [c]the Division of Biomedical Informatics, [d]the Division of Asthma Research, [f]the Division of Developmental Biology, [t]the Division of Human Genetics, [u]the Department of Patient Services, and [v]the Division of Allergy and Immunology, Cincinnati Children's Hospital Medical Center; [b]the Department of Pediatrics, College of Medicine, University of Cincinnati; [o]the Stanley Center for Psychiatric Research, [e]the Broad Institute of Massachusetts Institute of Technology and Harvard University, Cambridge; [g]the Department of Internal Medicine, and [h]the Department of Cardiovascular Medicine, Mayo Clinic, Rochester; [i]the Division of Medical Genetics, Department of Medicine, and [j]the Department of Genome Sciences, University of Washington Medical Center, Seattle, [k]the Division of Nephrology, Department of Medicine, College of Physicians and Surgeons, and [l]the Department of Biomedical Informatics, Columbia University, New York; [m]the Department of Preventive Medicine, and [q]the Division of General Internal Medicine and Geriatrics, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago; [n]the Psychiatric and Neurodevelopmental Genetics Unit, Center for Human Genomic Medicine, Massachusetts General Hospital, and [p]the Department of Psychiatry, Harvard Medical School, Boston; [r]the Center for Applied Genomics, Department of Pediatrics, Children's Hospital of Philadelphia; [s]the Perelman School of Medicine at the University of Pennsylvania, Philadelphia; and [w]the National Human Genome Research Institute, National Institutes of Health, Bethesda.

multiancestral PRS score, used one Electronic Medical Records and Genomics (eMERGE) cohort as a training set, used a second independent eMERGE cohort to validate the score, and used the UK Biobank data to replicate the findings. A phenome-wide association study was performed using the PRS to identify shared genetic etiology with other phenotypes.
Results: The multiancestral asthma PRS was associated with asthma in the 2 pediatric validation datasets. Overall, the multiancestral asthma PRS has an area under the curve (AUC) of 0.70 (95% CI, 0.69-0.72) in the pediatric validation 1 and AUC of 0.66 (0.65-0.66) in the pediatric validation 2 datasets. We found significant discrimination across pediatric subcohorts of European (AUC, 95% CI, 0.60 and 0.66), African (AUC, 95% CI, 0.61 and 0.66), admixed American (AUC, 0.64 and 0.70), Southeast Asian (AUC, 0.65), and East Asian (AUC, 0.73) ancestry. Pediatric participants with the top 5% PRS had 2.80 to 5.82 increased odds of asthma compared to the bottom 5% across the training, validation 1, and validation 2 cohorts when adjusted for ancestry. Phenome-wide association study analysis confirmed the strong association of the identified PRS with asthma (odds ratio, 2.71, $P_{FDR} = 3.71 \times 10^{-65}$) and related phenotypes.
Conclusions: A multiancestral PRS for asthma based on Bayesian posterior genomic effect sizes identifies increased odds of pediatric asthma. (J Allergy Clin Immunol 2022;150:1086-96.)

Key words: Genetics, asthma, GWAS, polygenic risk score, PRS, PheWAS

Asthma is an inflammatory disease of the airway with symptoms including coughing, wheezing, chest tightness, and shortness of breath caused by airflow obstruction and hyperresponsiveness.[1,2] Asthma affects 7 million children across the United States, yielding a prevalence equaling 10% of children, but this prevalence can vary by sex, ancestry, and age.[1] Patients with asthma experience increased morbidity with respiratory infection, days absent from school and work, emergency department visits, hospitalizations, and even death. While asthma is a significant economic and health burden, proactive treatment and specific interventions can prevent severe disease requiring emergency or inpatient management.[3,4] Given this, a major focus is primary prevention.[5] There have been substantial efforts to develop asthma prediction models based on clinical factors,[6,7] but these tools rely on early life phenotypes that may not have been assessed. Thus, development of additional predictive tools are warranted.

The etiology of asthma is multifactorial with contributions from environmental and genetic factors.[4,8,9] Twin studies estimate a heritability of 60% to 80% of asthma susceptibility attributable to genetic factors while also highlighting the etiological contributions of shared environment.[10] At the time of development of this multiancestral asthma polygenic risk score (PRS), there were >140 studies in the genome-wide association study (GWAS) catalog with 2167 variants reported with genome-wide significant results; 50 of the reported studies resulted from studies of participants not of full European ancestry.[11] These GWASs identify genetic loci that increase risk for asthma with a common genetic variant-based narrow sense heritability of 14.9%;[12] however, any single risk locus does not provide sufficient risk discrimination to be of practical clinical use. By accounting for genotypes at risk variants in proportion to the effect size of each genetic locus, PRSs have the potential to be tools for incorporating genetic risk into clinical decision support.[13]

A PRS for asthma has been developed; however, the data used to generate the PRS were limited to individuals of European descent. The limited diversity of participants used to generate the PRS is a problem because different demographic groups may have different underlying genetic etiologies. For example, although pediatric- and adult-onset asthma have numerous shared risk loci, genetic studies have also identified distinct risk loci in children.[14] Furthermore, the risk of asthma is different across ancestral groups, with children of African descent having higher frequencies than children of European descent do,[15-20] so it is important that for a PRS to be maximally clinically informative, the PRS should be developed and validated using training datasets that are similar to the populations to whom the PRS will be applied. Thus, a PRS developed using primarily adult studies or based on limited ancestral diversity may not be optimal to predict pediatric asthma. To address the current limitations, our goal was to develop a high-quality PRS for pediatric asthma using data from a diverse cohort representing multiple ancestries at all stages of the process. To address this goal, we used an existing large-scale genome-wide association derived from diverse populations and a Bayesian framework to derive a multiancestral PRS and tested and validated this score in multiple cohorts and phenotype definitions. With the development of this pediatric asthma PRS, we are now poised to test the clinical utility.

Our intent is to evaluate the clinical utility of this PRS in a prospective study that will include individuals across ancestral groups. It is routine in clinical settings to ask about an individual's race and ethnicity. Yet, how an individual self-identifies is not equivocal to their genetic ancestral group. Furthermore, many individuals have admixed ancestry. Thus, we developed a multiancestral PRS that could be used for all individuals.

METHODS
Study overview

The study design is presented in Fig 1. We divided the combined Electronic Medical Records and Genomics (eMERGE) I, II, and III imputed dataset from over 105,000 samples (https://emerge-network.org/genomics) into 2 nonoverlapping independent datasets (eMERGE training dataset and eMERGE validation 1 dataset 2) based on date of participation in eMERGE. The UK Biobank cohort was used as a validation 2 dataset. For the training eMERGE dataset, patients with asthma and controls were identified based on a validated algorithm using structured data (*International Classification of Diseases* [ICD] *version 9* [ICD-9 493.xx] or *version 10* [ICD-10 J45, J46] codes) for asthma as well as unstructured data as previously described.[21] Participants in the eMERGE validation dataset were classified as cases by having ≥2 *ICD-9/10* codes for asthma and as controls by having no history of asthma or atopy. The minor differences in case definitions were based on improvements in clinical data extraction over the past 15 years of eMERGE. For the validation 2 dataset in the UK Biobank, we used a previously established algorithm that was developed by UK Biobank team (Resource ID 4124) to identify participants with asthma.

Study population of eMERGE and UK Biobank and phenotype ascertainment

The eMERGE Network is a National Institutes of Health–organized and –funded consortium of US medical research institutions (https://www.genome.gov/Funded-Programs-Projects/Electronic-Medical-Records-and-Genomics-Network-eMERGE). Postimputation whole genome genotyping data for participants from the eMERGE Network were made available for this study (dbGAP [phs000888.v1.p1]). The associated BAM, xml, and vcf files are available on the eMERGE Commons web portal, accessible to sites as well as outside
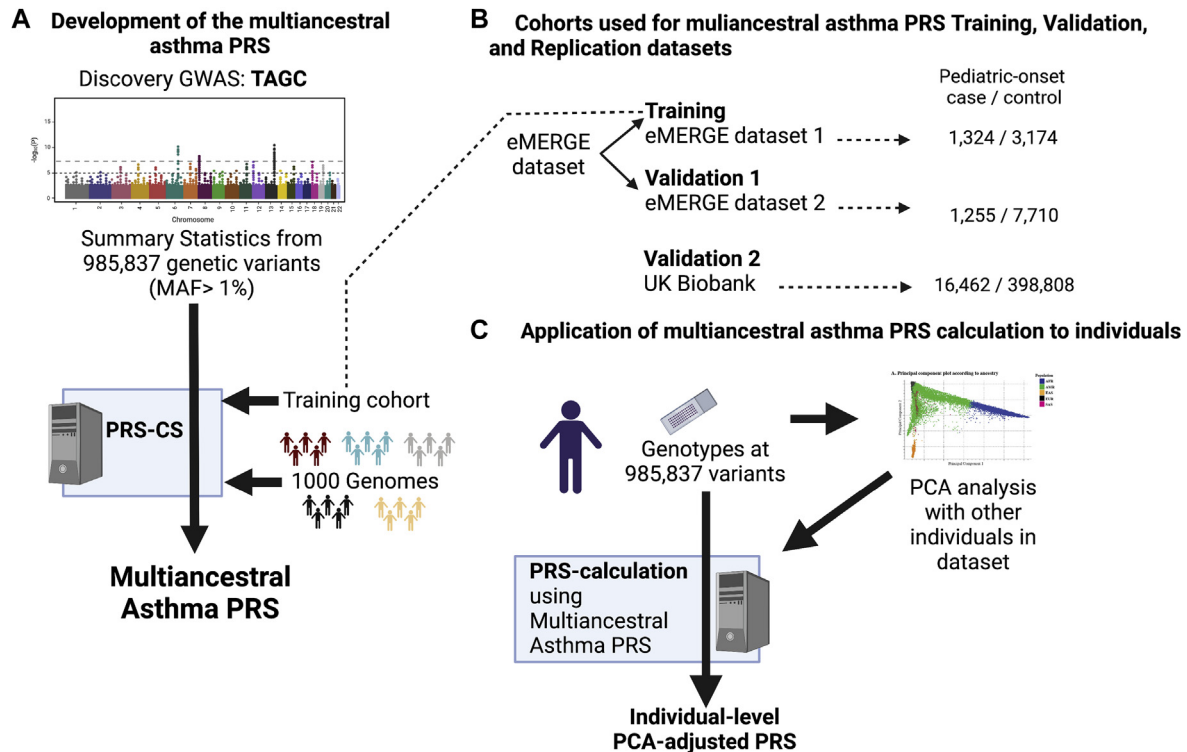
FIG 1. Study design. **A,** Summary statistics from 985,837 genetic variants with minor allele frequencies (*MAF*) >1% in the overall (cases and controls combined) TAGC discovery GWAS were used to develop the multiancestral asthma PRS in the context of the LD of the 1000 Genomes reference panel and the training dataset. **B,** The eMERGE dataset was split into 2 independent datasets (training and validation 1) and the UK Biobank was used as a validation 2 dataset. Subjects with confirmed pediatric-onset asthma and controls were used for PRS training, validation 1, and validation 2 cohorts. **C,** Individual genotypes from each subject in each dataset was used to perform a PCA. For each individual, genotypes at each of the 985,837 variants included in the multiancestral asthma PRS were used to calculate a PCA-adjusted PRS. The adjusted multiancestral PRS was applied to each individual in the training, validation 1, and validation 2 cohorts in preparation for it to be similarly applied to individuals recruited into an institutional review board–approved prospective intervention study. Figure created in BioRender.

investigators who apply for access (see eMERGE Network in Web Resources). The imputation process and genotype quality control in eMERGE followed guidelines that have been published previously.[22] Briefly, all subjects and variants had missingness <2%. Individual level genotype data were derived from 78 array batches across 12 academic medical centers. Each batch was imputed using the Michigan Imputation Server, which provides a missing single-nucleotide variant genotype imputation service using the minimac3 imputation algorithm with the Haplotype Reference Consortium genotype reference set. Only 1 genetic dataset was retained for each participant.[22]

The UK Biobank is a large long-term biobank study in the United Kingdom developed to support the investigation of the respective contributions of genetic predisposition and environmental exposure to the development of diseases[23] including asthma.[24,25] The UK Biobank postimputed data were obtained through application ID: 47377.

All participants in the eMERGE and UK Biobank cohorts provided written informed consent prior to study inclusion. The institutional review board of each contributing institution approved the eMERGE study. The North West Multicenter Research Ethics Committee and Patient Information Advisory Group approved the UK Biobank study. All analyses were conducted using deidentified data.

Prior to analyses, participant-level quality control was employed. Self-reported race and ethnicity was not used to identify ancestry. We used genetically defined ancestry for ancestry-specific analyses. Principal component analysis (PCA) of genome-wide genetic variants was used to establish ancestry (Table I). The FRAPOSA software package was used to perform PCA and assign all individuals into 5 major superpopulations (European, African, Admixed American, East Asian, and South Asian).[26] We used the phase 3 release of the 1000 Genomes data as a reference that consists of 2504 individuals from 5 superpopulations as shown in Table E1 in this article's Online Repository available at www.jacionline.org.[27] The major steps in the FRAPOSA algorithm include computing principal components of the reference dataset using the matched variants only and projecting computed principal components to the target data using an optimized implementation of the Online Augmentation, Decomposition, and Procrustes (OADP) transformation. Next, the algorithm predicts the ancestry membership by using the K[20]-nearest-neighbor method.[19] The pairwise correlation between self-reported race and genetic ancestry was 99% in UK Biobank and 97% in the eMERGE dataset, if we exclude those who self-identified as having a mixed or Hispanic race/ethnicity. All participants with sex inconsistencies were removed; additionally the dataset was pruned to remove participants to prevent duplicated individuals, twins, and first-degree relatives using PLINK's implementation of KING robust kinship coefficients.[28] In the KING pipeline, after the kinship (relationship) matrix is calculated using high-quality markers for all individuals, kinship-based pruning of samples is performed in which the program by default, randomly excludes 1 member of each pair of samples and prints all independent individuals for downstream analysis.

Results were evaluated in all individuals as well as participants who were enrolled in eMERGE as a child (age ≤ 18 years). It was not possible to identify asthma age of onset for all subjects in eMERGE who were over 18. Because the UK Biobank enrolled only adult participants, pediatric-onset asthma was identified through an assessment of the date of diagnosis in the context of the subject's current age (UK Biobank field identifiers 21003, 22147, 3786). For both eMERGE and the UK Biobank, we are confident in the identification of subjects with pediatric-onset asthma, while the true adult-onset asthma with

TABLE I. Pediatric study population

Cohort	Ancestry	Case/control	Female/male	Age, mean (SD) (y)
Training	All	1,324/3,174	2,058/2,440	11.83 (5.32)
	European	322/1,736	941/1,117	11.82 (5.20)
	African	768/727	697/798	12.55 (5.17)
	Admixed	203/642	367/478	12.14 (5.15)
	Eastern Asian*	21/40	34/27	11.68 (5.11)
	Southern Asian*	10/29	19/20	10.95 (6.01)
Validation 1	All	1,255/7,710	3,988/4,977	10.37 (5.97)
	European	386/3,982	1,883/2,485	10.90 (5.75)
	African	588/1,352	905/1,035	10.04 (6.25)
	Admixed	270/2,180	1,105/1,345	11.13 (5.50)
	Eastern Asian*	9/114	62/61	9.69 (6.17)
	Southern Asian*	2/82	33/51	10.07 (6.18)
Validation 2	All	16,462/398,808	219,728/195,542	8.40 (4.69)
	European	15,586/376,234	207,785/184,035	8.10 (4.82)†
	African	310/7,508	4,385/3,433	8.21 (4.70)†
	Admixed	205/5,020	2,717/2,508	8.16 (5.06)†
	Eastern Asian	49/1,622	877/794	7.96 (4.12)†
	Southern Asian	312/8,424	3,964/4,772	9.54 (4.74)†

A total of 70,290 cases and 467,247 controls across eMERGE and the UK Biobank in 5 superpopulations (European, Admixed American, African, East Asians, South Asian) were evaluated. Pediatric onset was defined as diagnosis before age 18 years. Please see Table E4 for a list of all participants, including those without pediatric-onset asthma.
*Due to low sample size, subgroups identified with asterisks only included in combined adjusted PRS analysis; ancestry-specific results from these subgroups are not presented.
†Mean age of onset for validation 2 cases at the time of diagnosis (see Methods).

no past medical history of asthma in childhood was unable to be accurately determined using electronic medical records.

Discovery GWAS for identifying variants to include in PRS analysis

The PRS was built using the GWAS results from a 2018 study published by the Trans-National Asthma Genetic Consortium (TAGC), which assessed 23,948 patients with adult- and pediatric-onset asthma and 118,538 controls.[29] Table E2 in this article's Online Repository at www.jacionline.org describes the studies included in TAGC. The summary statistics from 2,001,280 autosomal genetic variants that passed quality control filters were accessed through the GWAS catalog (https://www.ebi.ac.uk/gwas/home, assessed on January 8, 2021) and were used to calculate the PRS. A total of 985,837 autosomal genetic variants with minor allele frequencies >1% in the combined training and validation datasets (with cases and controls combined) were identified. These 985,837 common markers were found in the 1000 Genome reference panel, the TAGC, and in the training and validation datasets with genotyping rates of 99.9% in the training and validation 1 datasets. A total of 983,520 of these markers (99.7%) were present in the validation 2 dataset with a total genotyping rate of 98%. The complete list of the selected variants, the effect alleles, allele frequencies across ancestral groups, and posterior effect sizes (see below) are included in Table E3 in this article's Online Repository at www.jacionline.org.

Polygenic risk scores

PRSs were calculated using PRS-CS, a Bayesian polygenic prediction method that infers posterior effect sizes of genetic variants using GWAS summary statistics in the context of linkage disequilibrium between variants as assessed on an external reference panel (ie, the phase 3 release of the 1000 Genomes data).[30] The GWAS on which the PRS is derived is a multiancestral metanalysis, and the effect size for all genetic variants in the study are inverse-variance weighted with fixed effects accounting for all ancestral populations. Continuous shrinkage priors that were implemented in this pipeline allowed for marker-specific adaptive shrinkage: the amount of shrinkage applied to each genetic marker is adaptive to the strength of its association signal in the GWAS. The pipeline can accommodate diverse underlying genetic architectures. Linkage disequilibrium (LD) and an LD matrix were determined and built based on the highest number of ancestry representation in the discovery

set, which in our case was European. The training process (Fig 1) used the discovery GWAS summary statistics (multiancestral TAGC discovery GWAS), the reference population (individual-level 1000 Genomes genotype data), and the individual-level genotype and phenotype data of target population (training dataset in order to tune the hyperparameters of the prediction model using CS [auto mode]) so that the pipeline automatically learned the sparseness of the genetic architecture from data and adjusted for the LD structure accordingly.[24]

We adjusted for confounding effects due to population stratification with a linear regression model using the 10 principal components of ancestry in all participants.[31] After calculating a principal component–adjusted PRS, age and sex were used as covariates using a logistic regression fitting model implemented in R version 4.1.0.[19] The residuals from this model were used to create an ancestry-corrected PRS distribution. The distribution of unadjusted compared to the ancestry-adjusted PRSs across the 5 ancestral groups in the training and validation cohorts are presented in Fig E1 in this article's Online Repository available at www.jacionline.org.

The PRS prediction accuracy and performance was assessed by using receiver-operating characteristic area under the curve (AUC), odds ratio (OR) per 1 SD, and by variance explained (R^2 based on the Pseudo R^2 calculation based on the McFadden method as applied in Stata) in logistic regression after accounting for covariates (10 principal components, age, and sex). The median of the adjusted percentile distribution between cases and controls after ancestry standardization (ie, mean PRS of 0 and SD of 1 in each group) was assessed. As our long-term goal is to evaluate this PRS clinically, we selected the top 5% as a threshold for high risk. This threshold was selected to identify those at highest genetic risk while minimizing the number of people receiving a "high risk" result who would not develop asthma (Fig E2 in this article's Online Repository at www.jacionline.org). To measure the discrimination of the multiancestral asthma PRS, we report the top 5% of this distribution as a high polygenic score and report the increased odds of asthma by comparing the top 5% compared to both the bottom 5% and bottom 95%. After fitting the regression model, the marginal effect of sex and ancestry were also evaluated using the delta method implemented in Stata. These marginal effects measure the impact of unit change in a single variable on the prediction of asthma while all other variables of the adjusted multiancestral asthma PRS are constant.

Our primary outcome is predictive value in the pediatric cohorts based on the plan to use this PRS in a prospective study focused on children. In the tables in the Online Repository, we also report outcomes in the overall cohort

because it has additional power and allows us to compare performance in the subset of pediatric individuals.

We also benchmarked the performance of the multiancestral PRS using 2 previously published PRS algorithms.[32,33] Notably, the number of genetic variants included and the populations used in the previously published PRS are different.

PheWAS analyses

To evaluate pleiotropic effects of the multiancestral PRS for asthma against other traits, a phenome-wide association study (PheWAS) was performed using the R PheWAS software package in the training and validation cohorts.[34] Briefly, *ICD-9* codes were translated into PheWAS codes according to PheWAS map.[34] Cases and controls were identified based on ≥2 occurrences of the PheWAS code on different days in the cases and no instances in the controls.[34] For each PheWAS code, the asthma PRS was included in a logistic regression model adjusted for age, sex, and the 10 principal components. The OR is based on regression analyses using each phenotype as the dependent variable and adjusted PRS (a quantitative value calculated for each individual) as an independent variable. A false discovery rate (FDR) of 0.05 using the Benjamini–Hochberg procedure was implemented to account for multiple testing.

RESULTS

The number of participants in this study in each ancestral group that passed quality control steps (see Methods) are broken down by age and sex and presented in Table I. In total, we analyzed 70,290 participants with asthma and 467,247 controls across the multiancestral training, validation 1, and validation 2 datasets. Each dataset includes participants from each of the 5 superpopulations as defined by PCA of independent genetic variants (Table I and Table E4 in this article's Online Repository at www. jacionline.org).

An ancestry-specific asthma PRS for asthma is not optimal for clinical implementation, because individuals may align with multiple ancestries. Thus, we evaluated a single multiancestral asthma PRS that accounted for the underlying ancestral differences in the PRSs (Fig 1). The ancestry harmonization of the PRS distribution was performed to account for the density and range of each ancestry-specific distribution (Fig E1). After adjustment for ancestry, the multiancestral AUC for the validation pediatric cohorts was 0.70 (95% CI, 0.69-0.72) and 0.66 (95% CI, 0.65-0.66) in the pediatric validation 2 cohort (Fig 2, Table II, and Table E4). The discrimination of the PRS between the top 5th percentile to the bottom 5th percentile was measured in the training (OR, 2.80; 95% CI, 1.87-4.12), validation 1 (OR, 3.31; 95% CI, 2.29-4.78) and validation 2 (OR, 5.82; 95% CI, 5.19-6.53), datasets as shown in Table II for pediatric cohorts and Table E5 in this article's Online Repository at www.jacionline.org for the full datasets. A comparison of the PRS percentile distribution between cases and controls is presented in Fig 3. The risk prediction per each decile in the training, validation 1, and validation 2 cohorts are included in Fig E2.

To confirm that using multiancestral priors did not reduce the performance of the PRS, we calculated PRS for European cohorts using posterior effect sizes after training using only the European studies in TAGC (Table E6 in this article's Online Repository at www.jacionline.org) with trends toward better performance with multiancestral priors. As shown in Table III and Table E7 in this article's Online Repository at www.jacionline.org, the multiancestral PRS performance was consistent across all ancestries, with better performance in the validation pediatric cohorts.

The pediatric-only training dataset demonstrated significant discrimination using the covariate-adjusted multiancestral PRS (European: AUC, 0.67 [95% CI, 0.64-0.70], OR per SD, 1.27; African: AUC, 0.57 [95% CI, 0.54-0.60], OR per SD, 1.13; and Admixed American: AUC, 0.68 [95% CI,0.64-0.72], OR per SD, 1.61). These results were replicated in the pediatric validation 1 cohorts (European: AUC, 0.60 [95% CI, 0.57-0.63], OR per SD, 1.20; African: AUC, 0.61 [95% CI, 0.58-0.63], OR per SD, 1.27; and Admixed American: AUC, 0.64 [95% CI, 0.61-0.68], OR per SD, 1.25) (Table III [pediatric], Table E7 [overall]). The pediatric validation 2 cohort was used to further replicate the multiancestral PRS and provided the opportunity to measure the performance of the multiancestral asthma PRS in Eastern and Southern Asian cohorts (European: AUC, 0.66 [95% CI, 0.65-0.67], OR per SD, 1.57; African: AUC, 0.66 [95% CI, 0.63-0.69], OR per SD, 1.43; Admixed American: AUC, 0.70 [95% CI, 0.67-0.74], OR per SD, 1.63; Eastern Asian: AUC, 0.73 [95% CI, 0.66-0.80], OR per SD, 1.32; and Southern Asian: AUC, 0.65 [95% CI, 0.62-0.68], OR per SD, 1.32).

We compared the multiancestral asthma PRS from this study to the 2 previously published PRSs[32,33] in our 2 European training and validation datasets. The number of genetic variants in these PRSs were limited[15,22] and were developed based on genetic studies of European ancestry while the current PRS was developed based on a multiancestral GWAS. The AUCs in the full, covariate-adjusted models were lower for the Belsky et al[32] study (AUC, 0.59; 95% CI, 0.57-0.61) and Dijk et al[33] study (AUC, 0.60; 95% CI, 0.59-0.62) compared to the multiancestral PRSs in our training and validation European cohorts. Furthermore, the multiancestral PRS outperformed the previous European-derived PRS in non-European ancestries (Fig E3 in this article's Online Repository at www.jacionline.org).

In the full model logistic regression analyses of the PRSs, female sex was associated with reduced odds of asthma in pediatric cohorts across all ancestries (training-pediatric: OR, 0.74; 95% CI, 0.65-0.86; P < .0001; validation 1–pediatric: OR, 0.69; 95% CI, 0.61-0.79; P < .0001, and validation 2–pediatric: OR, 0.67; 95% CI, 0.65-0.69; P < .0001). This finding is consistent with pediatric-onset asthma being more common in males than in females.[35,36] To assess its confounding effect, we calculated the marginal effects of sex for prediction probability of asthma after fitting the logistic regression in the validation 2 cohort as a combined cohort and in each ancestry separately (Fig E4 in this article's Online Repository at www.jacionline. org). Indeed, the better predictive probability of asthma from the overall PRS model in males compared with in females is consistent with the regression analysis. Similarly, we evaluated the marginal effects of ancestry on the multiancestral PRS and found that there was substantial overlap consistent with ancestral normalization.

A PheWAS was performed in the combined full training and validation cohorts to evaluate potential pleiotropic effects of the multiancestral asthma PRS in this study with other traits. As expected, this approach confirmed the strong association of the multiancestral asthma PRS with asthma (OR, 2.71; 95% CI, 2.04-3.03; FDR-corrected P [P_{FDR}], 3.71 × 10^{-65}) (Table IV). This exploratory analysis also identified >300 pleiotropic association effects (P_{FDR} < .05) including positive associations with asthma severity and exacerbation, emphysema, and pulmonary insufficiency, as well as diabetes, eosinophilic esophagitis, food allergy, and white blood cell disorders (Fig 4, Table IV, Table E8 in this
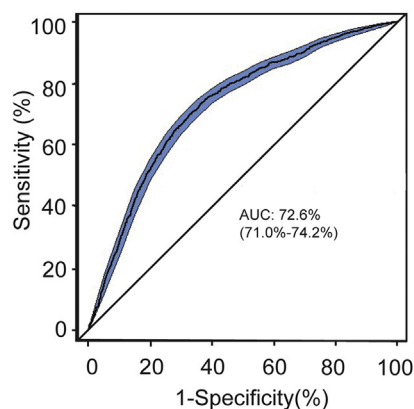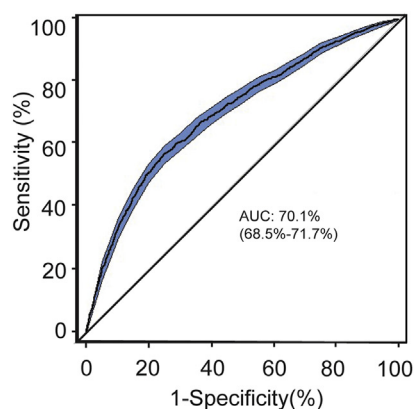
FIG 2. Multiancestral PRS performance. Overall adjusted PRS performance are shown for the training pediatric cohort **(A)**, validation 1 pediatric cohort **(B)**, and the validation 2 pediatric cohort **(C)**. The AUC and 95% CI are shown adjusted for age, sex, and 10 ancestral principal components.

TABLE II. Adjusted multiancestral PRS performance in 3 independent multiancestral pediatric cohorts

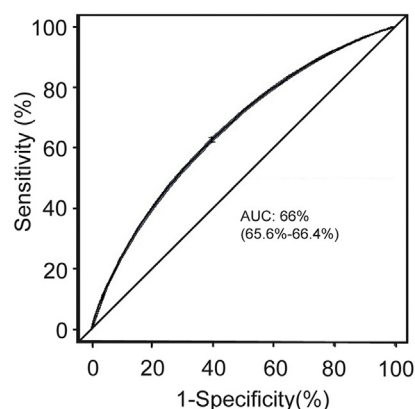| Transancestral PRS performance in pediatric cohorts	AUC	OR per SD*	Pseudo R²	OR† top 5% vs 95%	OR‡ top 5% vs bottom 5%
Training	0.73 (0.71-0.74)	1.21 (1.12-1.30)	0.11	1.84 (1.42-2.39)	2.80 (1.87-4.12)
Validation 1	0.70 (0.69-0.72)	1.22 (1.15-1.30)	0.08	2.16 (1.74-2.67)	3.31 (2.29-4.78)
Validation 2	0.66 (0.65-0.66)	1.59 (1.57-1.62)	0.04	2.37 (2.25-2.49)	5.82 (5.19-6.53)

The overall PRS multiancestry performance with 95% CIs after covariate and ancestry adjustment is presented. The OR of having asthma in patients within the top 5th percentile of the PRS distribution compared to the remaining 95% are shown with 95% CIs.
*The OR (95% CI) per SD ($P < .0001$).
†The OR (95% CI) when comparing the top 5% of standardized adjusted PRS distribution against remaining 95% ($P < .0001$).
‡The OR (95% CI) when comparing the top 5% of standardized adjusted PRS distribution against bottom 5% ($P < .0001$).

article's Online Repository at www.jacionline.org). Notably, the asthma PRS was more strongly associated with asthma than with the related phenotype allergic rhinitis (OR, 1.24; 95% CI, 1.10-1.40; P_{FDR}, 3.21×10^{-4}) (Tables E8 and E9 in this article's Online Repository at www.jacionline.org), supporting the phenotype-specificity of the asthma PRS. This approach also detected novel negative associations with traits such as hyperlipidemia and hypercholesterolemia (OR, 0.62; 95% CI, 0.55-0.69; P_{FDR}, 6.55×10^{-17}) (Fig 4, Table IV, Table E8).

DISCUSSION

Pediatric asthma affects ∼10% of the children in the United States. There is no cure, underscoring the importance of prevention and early identification. In this study, we developed a PRS for asthma using an ancestrally diverse group of individuals and trained and validated the PRS's performance using multiple independent cohorts, which included pediatric onset as well as any age of onset. We demonstrate that our asthma PRS has good discriminatory performance in people of diverse ancestries and especially
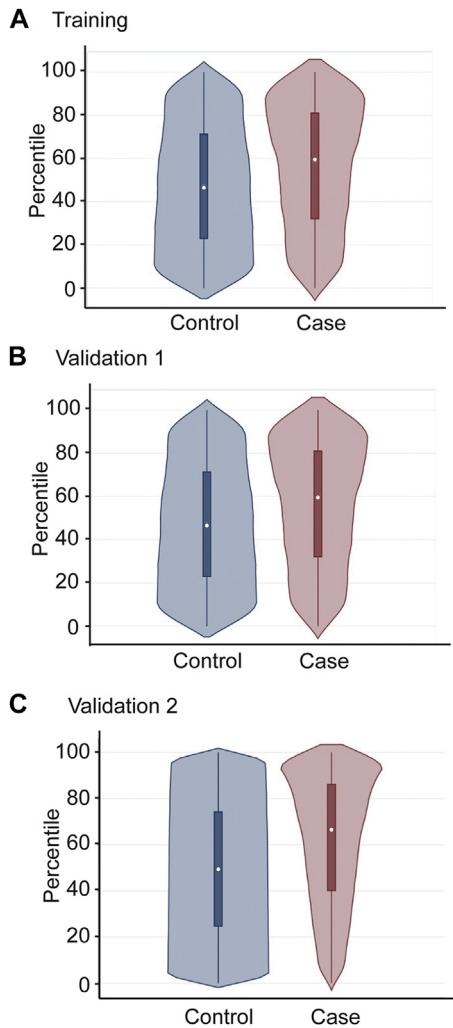
A Training

B Validation 1

C Validation 2

FIG 3. PRS percentile comparison between pediatric patients with asthma and controls. The violin plots of the median ancestry standardized PRS distributions between cases and controls in pediatric training **(A)**, validation 1 **(B)**, and validation 2 cohorts. In the boxplot inset, the dot within each box indicates the median score. In the training cohort **(A)**, the median standardized score of cases was 60% versus 47% in controls (*P* < .0001). In the validation 1 cohort **(B)**, the median standardized score of cases was 58% versus 48% in controls (*P* < .0001). In the validation 2 cohort **(C)**, the median standardized score of cases was 67% in cases versus 49% in controls (*P* < .0001). The *top* and *bottom* of the boxes indicate the 75th and 25th percentiles (interquartile range), respectively.

children, is more discriminating than prior scores, and reveals potential pleiotropic effects. Taken together, these results support the value of our multiancestral PRS.

The PRS performed well across 3 independent datasets and in 5 different ancestral groups. Because of the large number of participants assembled, we were able to evaluate the performance of the PRS overall as well as in documented pediatric cases. This evaluation revealed that our asthma PRS performed better in children than in the overall cohort of subjects with combined pediatric- and adult-onset asthma. These findings are consistent with previous genetic variant-based heritability estimates supporting a larger genetic contribution in children compared with in adults.[14] While we confidently identified individuals with pediatric-onset asthma, a limitation of this study was our inability

to identify individuals with adult-onset asthma (ie, some adults with asthma could have developed disease as a child).

Notably, our multiancestral asthma PRS performed better based on AUCs than prior asthma PRSs, especially when considering non-European populations. Based on assessment of the PGS catalog (https://www.pgscatalog.org/) in August 2021, there are currently 2 published studies focusing on PRS development for asthma alone.[32,33] These studies were limited primarily to European ancestral groups and PRS development was based on *P*-value thresholding. In complex diseases with many modest genetic effects such as asthma, the *P*-value thresholding methodology can underperform due to the omission of many variants with weaker phenotype association.[30] In contrast, the Bayesian approach used in this study incorporates genome-wide variants after considering the underlying LD population substructure and generates a posterior effect size for all variants included in the study. The benefits of the Bayesian strategy relative to other PRS approaches that use effect-size weighted additive model include[30] (1) all association data from a GWAS are used—including information that is usually not included in approaches that start with robust genetic associations; (2) the approach incorporates the differences in LD and genetic architecture between ancestral groups; and (3) the use of continuous shrinkage priors allows the model to consider robust genetic signals with large effect sizes as well as small effects with less significant association signals. This suggests that the Bayesian methodology is superior for the development of a PRS. However, because we compared different methodology and different populations (multiancestral), a broader comparison of these methods for the development of multiancestral pediatric asthma PRSs is justified.

The improved performance of our PRS may also be due to both the approaches used—Bayesian and multiancestral. Because asthma risk varies by ancestry, including individuals from diverse ancestral groups will be essential to ensure the clinical use of PRSs does not exacerbate health disparities.[37] Other approaches to develop transancestral and multiancestral PRSs differ by how they select and weigh genetic risk variants and how they integrate genetic data with other clinical and environmental data.[38-42] While we used PRS-CS, other models use best linear unbiased prediction (BLUP) and least absolute shrinkage and section operator approaches (LASSO) to estimate genetic effect sizes in joint models of multiple variants, and predictions are performed simultaneously.[43-46] There are also numerous methods that use different approaches to account for differences in LD in individuals of different ancestry.[47,48] As statistical methods are developed to improve multiancestral PRS, multiancestral asthma PRS should be continuously refined with future publications of genetic association studies of asthma in larger, admixed populations. It is possible that similarly powered ancestry-specific analyses would identify additional loci with more impactful effect sizes; however, the results of the multiancestral study might prove to be more broadly useful when applied to a heterogenous population, such as the type of people who go to a primary care setting. Furthermore, a major strength of our multiancestral PRS is that a single PRS is applicable to all ancestral groups. Thus, clinicians and researchers will not be required to *a priori* assign an individual to an ancestral group.

While TAGC might be the largest and most diverse metaanalysis from the perspective of genetic ancestry, there are several limitations to consider in the context of the multiancestral asthma

TABLE III. Multiancestral asthma PRS performance in 3 independent cohorts

Pediatric cohorts	AUC (full model)*	Pseudo R^2	OR per SD†
European ancestry			
Training	0.67 (0.64-0.70)	0.05	1.27 (1.12-1.44)
Validation 1	0.60 (0.57-0.63)	0.02	1.20 (1.08-1.34)
Validation 2	0.66 (0.65-0.67)	0.04	1.57 (1.55-1.60)
African ancestry			
Training	0.57 (0.54-0.60)	0.01	1.13 (1.01-1.26)‡
Validation 1	0.61 (0.58-0.63)	0.02	1.27 (1.14-1.42)
Validation 2	0.66 (0.63-0.69)	0.03	1.43 (1.24-1.65)
Admixed American ancestry			
Training	0.68 (0.64-0.72)	0.07	1.61 (1.28-2.02)
Validation 1	0.64 (0.61-0.68)	0.03	1.25 (1.06-1.47)
Validation 2	0.70 (0.67-0.74)	0.07	1.63 (1.36-1.95)
Eastern Asian ancestry			
Validation 2	0.73 (0.66-0.80)	0.08	1.32 (0.99-1.77)‡
Southern Asian ancestry			
Validation 2	0.65 (0.62-0.68)	0.03	1.32 (1.18-1.48)

The standardized PRS distribution was evaluated in a logistic regression model adjusted for age and sex as well as the 10 principal components that informed the 5 superpopulations. The AUC and 95% CI, OR per SD, and the total variation explained (pseudo R^2) are shown.
*AUC full model includes age, sex, and 10 principal components.
†The OR per SD of PRS distribution (logistic regression $P < .0001$).
‡For the African Ancestry eMERGE datasets 1-pediatric and Eastern Asian UK Biobank-pediatric cohorts, $P = .03$ and .07, respectively.

TABLE IV. Selected results of PheWAS applying multiancestral asthma PRS to training and validation 1 cohorts

Description*	PheWAS-code	Case	Control	OR†	95% CI	P_{FDR}
Positively associated with asthma PRS						
Asthma	495	12,963	70,020	2.71	2.41-3.02	3.17×10^{-65}
Asthma with exacerbation	495.2	3,043	70,020	4.72	3.74-5.95	3.39×10^{-39}
Emphysema	508	8,276	73,217	1.78	1.56-2.05	1.22×10^{-16}
Chronic obstructive asthma	495.1	1,433	70,020	3.23	2.35-4.45	4.90×10^{-13}
Type 1 diabetes	250.1	4,283	69,292	1.91	1.58-2.30	1.16×10^{-11}
Chronic airway obstruction	496	8,056	70,020	1.56	1.36-1.80	5.57×10^{-10}
Respiratory failure	509	6,029	73,217	1.61	1.37-1.88	3.05×10^{-9}
Wheezing	512.1	2,234	47,495	2.17	1.68-2.82	4.37×10^{-9}
Diabetes mellitus	250	19,764	69,292	1.34	1.21-1.48	9.36×10^{-9}
Eosinophilic esophagitis	530.15	607	62,044	3.85	2.38-6.20	3.40×10^{-8}
Type 2 diabetes	250.2	19,137	69,292	1.32	1.19-1.46	9.78×10^{-8}
Diseases of white blood cells	288	4,694	77,026	1.61	1.35-1.91	1.14×10^{-7}
Allergic reaction to food	930	1,688	65,842	2.25	1.66-3.04	1.49×10^{-7}
Negatively associated with asthma PRS						
Postmenopausal disorders	627	11,313	75,725	0.55	0.48-0.63	8.00×10^{-18}
Peripheral enthesopathies	726	16,894	66,608	0.64	0.57-1.71	6.31×10^{-17}
Hypercholesterolemia	272.11	19,899	51,696	0.62	0.55-0.69	6.55×10^{-17}
Hematuria	593	7,632	68,213	0.55	0.48-0.64	2.50×10^{-16}
Benign neoplasm of skin	216	11,233	79,947	0.63	0.56-0.71	1.86×10^{-14}
Disorders of lipid metabolism	272	41,493	51,696	0.73	0.67-0.80	1.04×10^{-11}
Hyperlipidemia	272.1	41,299	51,696	0.73	0.67-0.80	1.11×10^{-11}
Disorders of synovium	727	9,291	66,608	0.64	0.56-1.73	3.65×10^{-11}
Cataract	366	14,306	81,833	0.67	0.60-0.76	5.14×10^{-11}
Carbohydrate transport disorder	271	1,410	97,301	0.36	0.27-0.49	1.34×10^{-10}
Disaccharide malabsorption	271.3	1,302	97,301	0.35	0.25-0.48	1.40×10^{-10}
Elevated prostate specific antigen	796	2,746	84,778	0.53	0.41-0.67	2.92×10^{-7}

*Selected results at $P_{FDR} < .05$. The complete lists of traits are included in Table E5.
†OR < 1 indicates negative association of trait with asthma PRS.

PRS. Specifically, if there are differences in the genetic etiology by age of onset of asthma, the TAGC is not composed of a majority of pediatric cases. Furthermore, the degree to which pediatric-onset asthma cases are represented in the meta-analysis differs greatly based on the race and ethnicity of the component studies. Thus, while our PRS performed well for pediatric asthma, continued refinement of the PRS is warranted with special focus on pediatric cases and capturing more ancestral diversity in the discovery data.

To understand how underlying asthma risk may relate to a variety of conditions, we performed a PheWAS analysis to examine positive or negative association of other conditions
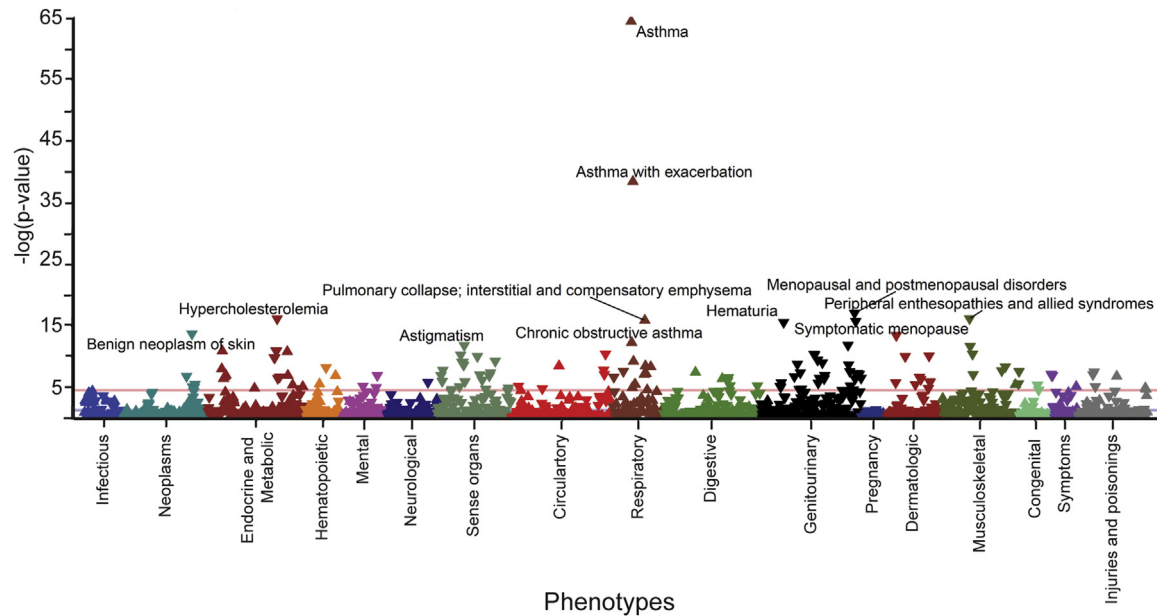
FIG 4. A plot of PheWAS analysis of the asthma PRS within the combined training and validation cohorts. Manhattan plots of phenome-wide association analyses with phecodes (x-axis) and FDR-corrected PheMap phenotype probability (y-axis). The *red line* indicates the Bonferroni level of significance (5.0×10^{-5}).

with the asthma PRS. Notably, we measured a stronger effect size for asthma with exacerbation (PheWAS code 495.2; OR, 4.72) and chronic asthma (PheWAS code 495.1; OR, 3.23) than asthma (PheWAS code 495; OR, 2.71) (Table IV). In the case of this PheWAS, the OR is based on regression analyses using each phenotype as the dependent variable and adjusted PRS as an independent variable. These findings suggest that the asthma PRS may be useful not only for prevention but also to help clinicians select treatment strategies for children already diagnosed with asthma. These findings require replication, as we do not have sufficiently uniform measurements in the subjects in the training, validation 1, and validation 2 cohorts to identify whether PRS in patients is associated with disease severity. However, these findings provide rationale for a controlled prospective study to test the association of asthma severity with PRS. Not surprisingly, other atopic conditions such as eosinophilic esophagitis and food allergy were positively associated with the asthma PRS, as these conditions have been noted to have increased rates in patients with asthma.[49,50] However, we also found positive associations with type I and type II diabetes. Intriguingly, several studies have reported a higher-than-expected co-occurrence of asthma and type I diabetes, supporting a partially shared genetic etiology.[51-53] We also found a surprising negative association between our asthma PRS and hypercholesterolemia/hyperlipidemia. In contrast to our findings, previous meta-analyses have reported that asthma is associated with worse lipid profiles at a phenotypic level.[54] One possible explanation for this discrepancy is that the use of inhaled corticosteroids (a primary treatment of asthma) is associated with a worse lipid profile in adults.[55] It is also possible that environmental factors, such as pollution, are driving increased risks for asthma as well as poor lipid profiles.[56-58] While the PheWAS analysis suggests potential pleiotropic effects (both increased and decreased risk of other diseases), additional work is required to clarify these relationships.

This study is an initial step toward developing a multiancestral PRS to be used in the eMERGE IV network (https://www.genome.gov/Funded-Programs-Projects/Electronic-Medical-Records-and-Genomics-Network-eMERGE) prospective intervention cohort study beginning in 2022. The eMERGE IV study will enroll 5000 children (underrepresented, non-European preferred) across the 10 eMERGE clinical sites. Multiancestral PRS for 4 phenotypes (asthma, obesity, type I diabetes, type II diabetes) will be calculated and returned to participants' parents and primary care providers. Parents and primary care providers of children with a high-risk asthma PRS (top 5th percentile) will also receive guideline-informed health recommendations.[59,60] We seek to understand how primary care providers, patients, and patient families change their behavior in reaction to a top 5th percentile asthma PRS. The prospective study will collect family history information and clinical factors to display along with an asthma high-risk PRS that providers can use to calculate the Pediatric Asthma Risk Score. Recent studies have validated the Pediatric Asthma Risk Score as a tool to predict asthma development in young children based on family history, eczema before age 3 years, wheezing apart from colds before age 3, African American ancestry, and sensitization to ≥2 food or aero allergens.[7]

A previous group suggested that their pediatric asthma PRS did not provide any discriminatory value above clinical risk factors.[33] Yet, assessing clinically predictive factors can be challenging due to the lack of consistent capturing of such data in cohorts with sufficient statistical power. Even if the PRS and the clinical risk prediction substantially overlap in who is identified at risk, the development of such a PRS would still be of value. This is because not all children undergo allergic sensitization testing before age 3 years, and clinical presentations such as eczema and wheezing without a cold may not be recognized by parents. Thus, alternative strategies for risk stratification are needed.

Notably, there are already preventive measures that can be prioritized if a child is identified as high risk.[61] For example, once a child is identified as high risk for asthma, families can be counseled to limit smoke exposure, identify and avoid known allergens, prevent viral infection, and limit dust and mold exposure.[4,62,63] However, we recognize that the use of genetics alone to predict asthma has inherent limitations, because both genes and environment contribute to asthma risk.

A long-term goal beyond eMERGE IV will be to create and validate a combined/integrated predictive model that includes genetic, family history, clinical, and environmental risk factors. Data collected during the eMERGE IV prospective study will provide essential elements toward our long-term goal. In addition to genotype data, family history information, and relevant clinical factors, we will also have geocodes to develop a combined/integrated predictive model.

In conclusion, we present the development and validation of a pediatric asthma PRS that performs effectively across ancestries in 3 independent cohorts and identifies novel pleiotropic relationships. In the future, this PRS will be used in the context of additional demographic and clinical risk factors as part of a genome-informed risk assessment to help families of children at high risk for asthma take preventive steps to avoid disease.

Clinical implications: This PRS will be used to identify children at increased risk for asthma across a multisite multiancestral prospective intervention cohort study.

REFERENCES

1. Holgate ST, Wenzel S, Postma DS, Weiss ST, Renz H, Sly PD. Asthma. Nat Rev Dis Primers 2015;1:15025.
2. Kuruvilla ME, Lee FE, Lee GB. Understanding asthma phenotypes, endotypes, and mechanisms of disease. Clin Rev Allergy Immunol 2019;56:219-33.
3. Ramsahai JM, Hansbro PM, Wark PAB. Mechanisms and management of asthma exacerbations. Am J Respir Crit Care Med 2019;199:423-32.
4. Castillo JR, Peters SP, Busse WW. Asthma exacerbations: pathogenesis, prevention, and treatment. J Allergy Clin Immunol Pract 2017;5:918-27.
5. Wiksten J, Toppila-Salmi S, Makela M. Primary prevention of airway allergy. Curr Treat Options Allergy 2018;5:347-55.
6. Castro-Rodriguez JA, Holberg CJ, Wright AL, Martinez FD. A clinical index to define risk of asthma in young children with recurrent wheezing. Am J Respir Crit Care Med 2000;162(4 Pt 1):1403-6.
7. Biagini Myers JM, Schauberger E, He H, Martin LJ, Kroner J, Hill GM, et al. A Pediatric Asthma Risk Score to better predict asthma development in young children. J Allergy Clin Immunol 2019;143:1803-10.e2.
8. Schoettler N, Rodriguez E, Weidinger S, Ober C. Advances in asthma and allergic disease genetics: Is bigger always better? J Allergy Clin Immunol 2019;144:1495-506.
9. Sheth KK, Lemanske RF Jr. Pathogenesis of asthma. Pediatrician 1991;18:257-68.
10. Thomsen SF. Exploring the origins of asthma: lessons from twin studies. Eur Clin Respir J 2014;1(suppl 1); https://doi.org/10.3402/ecrj.v1.25535.
11. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI GWAS catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res 2019; 47(D1):D1005-12.
12. Ferreira MA, Vonk JM, Baurecht H, Marenholz I, Tian C, Hoffman JD, et al. Shared genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. Nat Genet 2017;49:1752-7.
13. Torkamani A, Wineinger NE, Topol EJ. The personal and clinical utility of polygenic risk scores. Nat Rev Genet 2018;19:581-90.
14. Pividori M, Schoettler N, Nicolae DL, Ober C, Im HK. Shared and distinct genetic risk factors for childhood-onset and adult-onset asthma: genome-wide and transcriptome-wide studies. Lancet Respir Med 2019;7:509-22.
15. Vergara C, Murray T, Rafaels N, Lewis R, Campbell M, Foster C, et al. African ancestry is a risk factor for asthma and high total IgE levels in African admixed populations. Genet Epidemiol 2013;37:393-401.
16. Flores C, Ma SF, Pino-Yanes M, Wade MS, Perez-Mendez L, Kittles RA, et al. African ancestry is associated with asthma risk in African Americans. PLoS One 2012;7:e26807.
17. Brim SN, Rudd RA, Funk RH, Callahan DB. Asthma prevalence among US children in underrepresented minority populations: American Indian/Alaska Native, Chinese, Filipino, and Asian Indian. Pediatrics 2008;122:e217-22.
18. Pino-Yanes M, Thakur N, Gignoux CR, Galanter JM, Roth LA, Eng C, et al. Genetic ancestry influences asthma susceptibility and lung function among Latinos. J Allergy Clin Immunol 2015;135:228-35.
19. Fishe J, Zheng Y, Lyu T, Bian J, Hu H. Environmental effects on acute exacerbations of respiratory diseases: a real-world big data study. Sci Total Environ 2022; 806(Pt 1):150352.
20. Kaur S, Rosenstreich D, Cleven KL, Spivack S, Grizzanti J, Reznik M, et al. Severe asthma in adult, inner-city predominantly African-American and Latinx population: demographic, clinical and phenotypic characteristics. J Asthma 2021 Dec 2; 1-11 [Online ahead of print.]
21. Almoguera B, Vazquez L, Mentch F, Connolly J, Pacheco JA, Sundaresan AS, et al. Identification of four novel loci in asthma in European American and African American populations. Am J Respir Crit Care Med 2017;195:456-63.
22. Stanaway IB, Hall TO, Rosenthal EA, Palmer M, Naranbhai V, Knevel R, et al. The eMERGE genotype set of 83,717 subjects imputed to ~40 million variants genome wide and association with the herpes zoster medical record phenotype. Genet Epidemiol 2019;43:63-81.
23. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015;12:e1001779.
24. Zhu Z, Lee PH, Chaffin MD, Chung W, Loh PR, Lu Q, et al. A genome-wide cross-trait analysis from UK Biobank highlights the shared genetic architecture of asthma and allergic diseases. Nat Genet 2018;50:857-64.
25. Zhu Z, Zhu X, Liu CL, Shi H, Shen S, Yang Y, et al. Shared genetics of asthma and mental health disorders: a large-scale genome-wide cross-trait analysis. Eur Respir J 2019;54:1901507.
26. Zhang D, Dey R, Lee S. Fast and robust ancestry prediction using principal component analysis. Bioinformatics 2020;36:3439-46.
27. 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. Nature 2015;526:68-74.
28. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 2015;4:7.
29. Demenais F, Margaritte-Jeannin P, Barnes KC, Cookson WOC, Altmuller J, Ang W, et al. Multiancestry association study identifies new asthma risk loci that colocalize with immune-cell enhancer marks. Nat Genet 2018;50:42-53.
30. Ge T, Chen CY, Ni Y, Feng YA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. Nat Commun 2019;10:1776.
31. Khera AV, Chaffin M, Zekavat SM, Collins RL, Roselli C, Natarajan P, et al. Whole-genome sequencing to characterize monogenic and polygenic contributions in patients hospitalized with early-onset myocardial infarction. Circulation 2019; 139:1593-602.
32. Belsky DW, Sears MR, Hancox RJ, Harrington H, Houts R, Moffitt TE, et al. Polygenic risk and the development and course of asthma: an analysis of data from a four-decade longitudinal study. Lancet Respir Med 2013;1:453-61.
33. Dijk FN, Folkersma C, Gruzieva O, Kumar A, Wijga AH, Gehring U, et al. Genetic risk scores do not improve asthma prediction in childhood. J Allergy Clin Immunol 2019;144:857-60.e7.
34. Carroll RJ, Bastarache L, Denny JC. R PheWAS: data analysis and plotting tools for phenome-wide association studies in the R environment. Bioinformatics 2014;30:2375-6.
35. Frohlich M, Pinart M, Keller T, Reich A, Cabieses B, Hohmann C, et al. Is there a sex-shift in prevalence of allergic rhinitis and comorbid asthma from childhood to adulthood? A meta-analysis. Clin Transl Allergy 2017;7:44.
36. Pinart M, Keller T, Reich A, Frohlich M, Cabieses B, Hohmann C, et al. Sex-related allergic rhinitis prevalence switch from childhood to adulthood: a systematic review and meta-analysis. Int Arch Allergy Immunol 2017;172:224-35.
37. Oh SS, Galanter J, Thakur N, Pino-Yanes M, Barcelo NE, White MJ, et al. Diversity in clinical and biomedical research: a promise yet to be fulfilled. PLoS Med 2015;12:e1001918.
38. Rudolph A, Song M, Brook MN, Milne RL, Mavaddat N, Michailidou K, et al. Joint associations of a polygenic risk score and environmental risk factors for breast cancer in the Breast Cancer Association Consortium. Int J Epidemiol 2018;47:526-36.
39. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat Genet 2018;50:1219-24.

40. Mars N, Koskela JT, Ripatti P, Kiiskinen TTJ, Havulinna AS, Lindbohm JV, et al. Polygenic and clinical risk scores and their impact on age at onset and prediction of cardiometabolic diseases and common cancers. Nat Med 2020;26:549-57.

41. Mak TSH, Porsch RM, Choi SW, Zhou X, Sham PC. Polygenic scores via penalized regression on summary statistics. Genet Epidemiol 2017;41:469-80.

42. Vilhjalmsson BJ, Yang J, Finucane HK, Gusev A, Lindstrom S, Ripke S, et al. Modeling linkage disequilibrium increases accuracy of polygenic risk scores. Am J Hum Genet 2015;97:576-92.

43. Speed D, Balding DJ. MultiBLUP: improved SNP-based prediction for complex traits. Genome Res 2014;24:1550-7.

44. Zhou X, Carbonetto P, Stephens M. Polygenic modeling with Bayesian sparse linear mixed models. PLoS Genet 2013;9:e1003264.

45. Shi J, Park JH, Duan J, Berndt ST, Moy W, Yu K, et al. Winner's curse correction and variable thresholding improve performance of polygenic risk modeling based on genome-wide association study summary-level data. PLoS Genet 2016;12: e1006493.

46. Lello L, Avery SG, Tellier L, Vazquez AI, de Los Campos G, Hsu SDH. Accurate genomic prediction of human height. Genetics 2018;210:477-97.

47. Amariuta T, Ishigaki K, Sugishita H, Ohta T, Koido M, Dey KK, et al. Improving the trans-ancestry portability of polygenic risk scores by prioritizing variants in predicted cell-type-specific regulatory elements. Nat Genet 2020;52: 1346-54.

48. Marquez-Luna C, Loh PR, South Asian Type 2 Diabetes Consortium, SIGMA Type 2 Diabetes Consortium, Price AL. Multiethnic polygenic risk scores improve risk prediction in diverse populations. Genet Epidemiol 2017;41:811-23.

49. Gonzalez-Cervera J, Arias A, Redondo-Gonzalez O, Cano-Mollinedo MM, Terrehorst I, Lucendo AJ. Association between atopic manifestations and eosinophilic esophagitis: a systematic review and meta-analysis. Ann Allergy Asthma Immunol 2017;118:582-90.e2.

50. Foong RX, du Toit G, Fox AT. Asthma, food allergy, and how they relate to each other. Front Pediatr 2017;5:89.

51. Metsala J, Lundqvist A, Virta LJ, Kaila M, Gissler M, Virtanen SM, et al. The association between asthma and type 1 diabetes: a paediatric case-cohort study in Finland, years 1981-2009. Int J Epidemiol 2018;47:409-16.

52. Hsiao YT, Cheng WC, Liao WC, Lin CL, Shen TC, Chen WC, et al. Type 1 diabetes and increased risk of subsequent asthma: a nationwide population-based cohort study. Medicine (Baltimore) 2015;94:e1466.

53. Smew AI, Lundholm C, Savendahl L, Lichtenstein P, Almqvist C. Familial coaggregation of asthma and type 1 diabetes in children. JAMA Netw Open 2020;3: e200834.

54. Peng J, Huang Y. Meta-analysis of the association between asthma and serum levels of high-density lipoprotein cholesterol and low-density lipoprotein cholesterol. Ann Allergy Asthma Immunol 2017;118:61-5.

55. Fessler MB, Massing MW, Spruell B, Jaramillo R, Draper DW, Madenspacher JH, et al. Novel relationship of serum cholesterol with asthma and wheeze in the United States. J Allergy Clin Immunol 2009;124:967-74.e1-15.

56. McGuinn LA, Schneider A, McGarrah RW, Ward-Caviness C, Neas LM, Di Q, et al. Association of long-term PM2.5 exposure with traditional and novel lipid measures related to cardiovascular disease risk. Environ Int 2019;122: 193-200.

57. Keet CA, Keller JP, Peng RD. Long-term coarse particulate matter exposure is associated with asthma among children in Medicaid. Am J Respir Crit Care Med 2018;197:737-46.

58. Brunst KJ, Ryan PH, Brokamp C, Bernstein D, Reponen T, Lockey J, et al. Timing and duration of traffic-related air pollution exposure and the risk for childhood wheeze and asthma. Am J Respir Crit Care Med 2015;192:421-7.

59. Expert Panel Working Group of the National Heart, Lung, and Blood Institute administered and coordinated National Asthma Education and Prevention Program Coordinating Committee, Cloutier MM, Baptist AP, Blake KV, Brooks EG, Bryant-Stephens T, et al. 2020 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: 1217-70.

60. Cloutier MM, Teach SJ, Lemanske RF Jr, Blake KV. The 2020 focused updates to the NIH asthma management guidelines: key points for pediatricians. Pediatrics 2021;147:e2021050286.

61. Tackett AP, Farrow M, Kopel SJ, Coutinho MT, Koinis-Mitchell D, McQuaid EL. Racial/ethnic differences in pediatric asthma management: the importance of asthma knowledge, symptom assessment, and family-provider collaboration. J Asthma 2021;58:1395-406.

62. Elenius V, Jartti T. Vaccines: could asthma in young children be a preventable disease? Pediatr Allergy Immunol 2016;27:682-6.

63. Szefler SJ. Advances in pediatric asthma in 2012: moving toward asthma prevention. J Allergy Clin Immunol 2013;131:36-46.