

Revealing polygenic pleiotropy using genetic risk scores for asthma

Matthew Dapas,¹ Yu Lin Lee,^{1,2} William Wentworth-Sheilds,¹ Hae Kyung Im,³ Carole Ober,^{1,*} and Nathan Schoettler^{4,5,*}

Summary

In this study we examined how genetic risk for asthma associates with different features of the disease and with other medical conditions and traits. Using summary statistics from two multi-ancestry genome-wide association studies of asthma, we modeled polygenic risk scores (PRSs) and validated their predictive performance in the UK Biobank. We then performed phenome-wide association studies of the asthma PRSs with 371 heritable traits in the UK Biobank. We identified 228 total significant associations across a variety of organ systems, including associations that varied by PRS model, sex, age of asthma onset, ancestry, and human leukocyte antigen region alleles. Our results highlight pervasive pleiotropy between asthma and numerous other traits and conditions and elucidate pathways that contribute to asthma and its comorbidities.

Introduction

Asthma is a common respiratory disorder, affecting 3%–4% of people globally, with prevalences exceeding 15% in many regions among different groups.^{1–3} The disease is characterized by episodic airflow obstruction caused by airway narrowing from bronchial constriction and inflammation, which can be induced by inhaled exposures, such as viruses, allergens, and pollutants.⁴ Genetic factors account for a substantial portion of overall disease liability,⁵ with over 200 different genetic risk loci reported for asthma in large genome-wide association studies (GWASs).^{6,7} These GWAS findings have implicated many genes and pathways in the pathogenesis of asthma,⁸ but the significant GWAS alleles contribute only modestly to asthma heritability⁹ and individually paint a limited picture of the overall genetic risk for asthma.¹⁰ By instead considering the cumulative contribution from all common genetic variants and modeling a comprehensive polygenic risk score (PRS), we can measure how genetic risk for asthma differs by sex, age of disease onset, and ancestry and the extent to which it is associated with other traits and diseases, generating new insights into asthma etiology and its shared biology with other traits.

Previous studies that have developed PRSs for asthma have focused on predictive performance^{11–15} or associations with specific traits,^{15–21} while broader studies of cross-trait genetic correlations between asthma and other conditions offer less interpretable findings than PRS associations in terms of relative disease risk.^{7,22–29} In this study, we aimed to provide a quantitative overview of the genetic landscape for asthma in terms of other heritable traits and

conditions and their associations with genetic asthma risk. We first modeled PRSs of asthma using GWAS summary statistics from the Trans-National Asthma Genetic Consortium (TAGC; 23,948 cases, 118,538 controls)³⁰ and the Global Biobank Meta-analysis Initiative (GBMI; 153,763 cases, 1,647,022 controls)⁷ and then derived scores for individuals in the UK Biobank (UKB; $n = 398,744$).³¹ We assessed the predictive performance of our models in several UKB populations stratified by age of disease onset, sex, and ancestry. We then performed phenome-wide association studies (PheWASs) of genetic asthma risk with 371 heritable traits reported in the UKB using both GBMI and TAGC PRS models and analyzed sensitivities to asthma status, age of disease onset, ancestry, and human leukocyte antigen (HLA) region alleles. Our results highlight shared genetic architecture between asthma and numerous other traits and conditions, including some with sex, age, and/or ancestry dependencies, elucidating the affected pathways that contribute to asthma and its comorbidities.

Subjects and methods

Study populations and asthma phenotyping

The polygenic risk models used in this study were derived from summary statistics of the TAGC and GBMI meta-analyses of asthma.^{7,30,32} The TAGC meta-analysis integrated data from 66 GWASs from multiple populations: European ancestry (19,954 cases, 107,715 controls), African ancestry (2,149 cases, 6,055 controls), Japanese (1,239 cases, 3,976 controls), and Latino (606 cases, 792 controls). Summary statistics from TAGC were available for the entire multi-ancestry sample (23,948 cases, 118,538 controls) and for the European-ancestry subset (19,954 cases,

¹Department of Human Genetics, University of Chicago, Chicago, IL, USA; ²Biological Sciences Collegiate Division, University of Chicago, Chicago, IL, USA; ³Section of Genetic Medicine, Department of Medicine, University of Chicago, Chicago, IL, USA; ⁴Section of Pulmonary and Critical Care Medicine, Department of Medicine, University of Chicago, Chicago, IL, USA

⁵Lead contact

*Correspondence: cober@uchicago.edu (C.O.), nschoettle@bsd.uchicago.edu (N.S.)

<https://doi.org/10.1016/j.xhgg.2023.100233>.

© 2023 The Author(s). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).



107,715 controls). Twenty-seven of the 66 studies included in TAGC only included cases of asthma with age of onset ≤ 16 years (8,976 cases, 18,399 controls). Cases of asthma were based on doctor's diagnosis and/or standardized questionnaires.

The GBMI includes combined health and genomic data from over 2.2 million participants across 23 biobanks.³³ The discovery GWAS dataset used to model the GBMI asthma PRS consisted of 197,342 cases and 1,903,937 controls from eight biobanks (BioBank Japan,³⁴ BioVU,³⁵ Lifelines,³⁶ Ontario Health Study,³⁷ Estonian Biobank,³⁸ FinnGen,³⁹ Michigan Genomics Initiative,⁴⁰ and the HUNT Study⁴¹). The cohort was 66.3% European (141,957/1,251,642 cases/controls), 27.7% East Asian (33,312/549,116 cases/controls), 2.4% African (8,661/42,279 cases/controls), 2.1% South Asian (6,162/38,103 cases/controls), and 1.4% admixed American (7,250/22,797 cases/controls).³² Asthma case and control statuses were assigned according to specific International Classification of Diseases (ICD)-9 and ICD-10 codes in participant electronic health records,³³ except for BioBank Japan, which used physician's diagnosis or past medical history (including self-report) to assign asthma status, and the Lifelines study, which used doctor's diagnosis or current symptoms/treatment.

To evaluate the TAGC and GBMI polygenic risk models, PRSs were derived for genotyped individuals in the UKB.³¹ We tested our model in three different population samples defined by ancestry: white British ($n = 376,237$), white non-British ($n = 27,026$), and African ($n = 7,196$). The white British and white non-British populations were as defined by the UKB,³¹ while the African-ancestry population combined samples with African ancestries from multiple UKB cohorts. For this African-ancestry population, we included individuals from the "Black or Black British," "African," "Caribbean," and "Any other Black background" cohorts with a first principal component (PC) of ancestry value >150 (UKB data field 22009), which indicated significant African admixture and greater genetic dissimilarity from the white UKB cohorts. UKB data are available following a process described at <https://www.ukbiobank.ac.uk/enable-your-research>.

We used quality-control metrics derived centrally by the UKB to process the genotype data, including ancestry PC values and ancestry group designation.³¹ We excluded individuals with poor-quality genotypes based on excess heterozygosity, missingness, and/or ambiguous genetic sex assignment. For pairs of individuals who were first- or second-degree relatives, only the individual with the smallest number of missing genotypes was included. We performed additional quality control of genotypes, requiring an imputation quality score (INFO) >0.8 , a minor allele frequency (MAF) $>0.1\%$, variant call rate $>95\%$, and a Hardy-Weinberg equilibrium p value $>0.1 \times 10^{-10}$.

We defined asthma in the UKB according to age of asthma diagnosis fields (UKB data fields 3786 and 22147), with individuals with a diagnosis before age 66 years considered to have had asthma. As in our previous study,⁹ we defined childhood-onset asthma (COA) and adult-onset asthma (AOA) in the UKB using strict age-of-onset criteria to minimize the likelihood of misclassification; COA was defined as an asthma diagnosis before 12 years of age ($n = 9,611$) and AOA was defined as asthma onset after age 25 years and before 66 years ($n = 21,427$). Participants older than 38 years of age without an asthma diagnosis (self-reported or ICD-10 codes) were included as controls for these fields. Individuals with chronic obstructive pulmonary disease (COPD), emphysema, or chronic bronchitis (self-reported or ICD-10 codes) were excluded from the AOA and control groups.

PRS modeling

Polygenic risk models were generated from each of the TAGC and GBMI results, separately. For TAGC, we derived polygenic risk models from the GWAS summary statistics. For GBMI, we adopted the model published by Wang and colleagues³² from the GBMI leave-UKB-out multi-ancestry GWAS deposited on the Polygenic Score Catalog (PGS001787).⁴² All polygenic risk models were generated using PRS-CS,⁴³ a Python-based program that applies a high-dimensional Bayesian regression framework with continuous shrinkage priors placed on single-nucleotide polymorphism (SNP) effect sizes. The default half-Cauchy priors, $\alpha = 1$ and $\beta = 0.5$, were used for determining the local shrinkage parameters in the PRS-CS model (ψ_j). To determine the global shrinkage parameter (ϕ), the "PRS-CS-auto" algorithm was applied to estimate the value directly from the GWAS summary statistics. Each autosomal chromosome was modeled separately, producing different parameters for each. For the Gibbs sampling used to derive the global shrinkage parameter and posterior SNP effect sizes, we ran 10,000 Markov Chain Monte Carlo iterations with 5,000 burn-in steps for the TAGC models, and the GBMI models used the default parameters of 1,000 and 500, respectively.³² For each set of TAGC summary statistics, we trained two different models, incorporating European linkage disequilibrium (LD) panels from either the 1000 Genomes (1KG) Project⁴⁴ or from the UKB, as prepared by the PRS-CS developers.⁴³ The GBMI model utilized the European 1KG LD panel. To improve computational efficiency and standardize the reference sets, the LD panels were created for common (minor allele frequency $>1\%$), non-ambiguous (no A/T and C/G) HapMap3 SNPs with INFO >0.8 (UKB = 1,117,425 SNPs; 1KG = 1,120,696 SNPs).⁴³ LD block boundaries were determined using LDetect on European samples from 1KG ($n = 1,703$ blocks).⁴⁵ Previous studies have shown that European-based LD panels can adequately approximate LD in multi-ancestry GWAS with majority European ancestry.^{32,46} Posterior SNP effect sizes were combined into aggregate scores for individuals in the UKB using the Plink (v2.00a2 AFX2) "score" function.⁴⁷ Effects for missing genotypes were imputed as the posterior SNP effect size multiplied by the effect allele frequency. We converted the PRSs to Z scores for analysis.

Asthma prediction in the UKB

We evaluated scores from five different polygenic risk models, four from TAGC utilizing either the 1KG or UKB LD reference panels with each of the European- and multi-ancestry summary statistics and one from GBMI excluding UKB participants and utilizing the 1KG LD panel.^{7,32} We assessed the predictive performance of the five PRSs in the white British, white non-British, and African-ancestry populations in the UKB using receiver-operating characteristic curves.⁴⁸ We estimated the variance explained by each PRS using R^2 on liability scale, using bootstrap sampling ($n = 1,000$) to derive 95% confidence intervals (CIs).⁴⁹ We selected the PRS with the highest area under the receiver-operating-characteristic curve (AUC) for all subsequent primary statistical analyses. PRS associations with asthma were evaluated using logistic regressions, adjusting for sex and the first 10 PCs of ancestry (first 20 PCs in African-ancestry samples). In assessing relative risks between PRS quantiles, we set individuals with scores between the 40th and 60th percentiles as the reference group. Interaction effects with sex for PRS associations were tested by adding a PRS-by-sex interaction term to the logistic regression models.

PheWAS

Traits included in the PheWAS were selected from 4,178 UKB phenotypes with additive SNP-heritability estimates from the Neale Lab UKB SNP-Heritability Browser (<http://www.nealelab.is/uk-biobank/>), as updated on October 19, 2019. SNP heritability was estimated using partitioned LD score regression⁵⁰ and assigned confidence ratings based on effective sample sizes, standard errors, sex biases, and ordinal coding. We considered traits with estimated SNP heritability >0.10 with low, medium, or high confidence estimates (n = 519). We then omitted those that were not interpretable as biomedical phenotypes or laboratory values, such as occupations or food preferences. We extracted phenotypes from the UKB using the ukbREST tool on a local data server.¹⁰ Phenotypes derived from ICD-10 codes in participant hospital inpatient records correspond to primary diagnoses (UKB data field 41202).

Some phenotypes were refined to improve interpretability. We removed the self-reported (UKB data field 20002) and doctor-diagnosed diabetes (UKB data field 2443) fields, which did not distinguish between type 1 and type 2 and instead used primary and secondary ICD-10 codes (UKB data fields 41202 and 41204) to indicate either disease. For forced expiratory volume in 1 s (FEV₁), we only considered the percentage predicted phenotype (UKB data field 20154). For COPD, we expanded the definition to account for all self-reported or doctor-diagnosed COPD, emphysema, or chronic bronchitis (UKB data fields 20002, 41202, 41204, 22128, 22129, 22130, 22148, 22149, 22150, 3992). Following filtering and refinement, 371 traits remained (208 binary phenotypes, 163 quantitative traits) for association testing with the PRS. Twenty-nine of the traits were measured in only one sex, including traits that could apply to either sex but were assessed in sex-specific survey questions (e.g., UKB data fields 6153 and 6177). Traits were manually categorized into 13 groups according to their affiliated biological system: anthropometric (n = 45), asthma and allergic disease (n = 23); cardiovascular (n = 39), gastrointestinal (n = 12); hematologic and blood chemistry (n = 46); musculoskeletal and skin (n = 34); neoplasms (n = 14); neurologic and behavioral (n = 48); ocular (n = 32); other (n = 14); pulmonary (n = 17); renal and urologic (n = 9); and reproductive, endocrine, and metabolic (n = 38).

We performed separate PheWAS in the following UKB cohorts: white British, white British excluding individuals with diagnosis or self-report of asthma, white non-British, and African ancestry. We tested associations between each of the PheWAS traits and the asthma PRS using logistic regressions for binary phenotypes and linear regressions for quantitative traits, adjusting for sex and the first 10 PCs of ancestry (first 20 PCs in African-ancestry samples). Quantitative traits were standardized using a Z score transformation but were not normalized prior to association testing, as it was assumed sample sizes were sufficiently large to invoke the central limit theorem. Phenome-wide significance was determined using Bonferroni multiple-test correction ($p < 0.05/371$), which was conservative considering the high correlations between many pairs of traits (e.g., left arm fat mass, right arm fat mass). For display purposes, we labeled highly correlated trait groupings as single traits, such as for anthropometric measures with values from both the left and right sides, resulting in 253 non-redundant traits, and displayed the strongest trait association statistics from within each grouping. Trait associations with p values less than the minimum subnormal double-precision floating-point number, 2^{-1074} , were labeled as $p < 4.9 \times 10^{-324}$.

In defining the non-asthma PheWAS population in the UKB, we excluded individuals with a report of asthma in any of the

following UKB fields: 3786, 22147, 6152, 20002, 22127, and 41202. We did not further exclude individuals based on COPD, emphysema, or chronic bronchitis, as we did in defining the AOA and control groups, so that we would compare the PRS association effects for these traits across PheWAS cohorts.

When comparing individual trait associations between different PRS models, we determined bias toward one PRS or the other based on the absolute differences between regression effect sizes. Therefore, for quantitative traits, the biases equated to the total differences in standardized effect sizes [$\Delta\beta_2$] between PRS associations, whereas, for binary traits, the biases were equivalent to the relative change in odds ratio (OR) (OR_1/OR_2). To evaluate whether there were differences in association strengths between PRSs for specific phenotypic categories, we regressed the effect sizes or ORs of all nominally significant ($p < 0.05$) PheWAS associations of one PRS against the other with a fixed intercept at zero for each category and then performed one-sample t tests of the resultant slopes against a value of 1. When applicable, we combined the binary and quantitative results into one p value for each category using Stouffer's method weighted by the number of traits in each category by variable type (binary or quantitative).

To test for significant differences in individual PRS-trait association effects, whether between different populations for the same PRS or between different PRSs for the same population, we used the Wald statistic:

$$W = \frac{\beta_1 - \beta_2}{\sqrt{se(\beta_1)^2 + se(\beta_2)^2}}$$

where β_1 and β_2 are the estimated effect sizes for each PRS association, and W asymptotically follows a z distribution for binary traits and a t distribution for quantitative traits. For the PheWAS in non-asthmatic individuals, we used inverse probability weighting to mitigate potential selection bias induced from conditioning on asthma status.⁵¹ Regression weights were set as the inverse predicted probability of never having asthma, as determined from a logistic regression of asthma (union of all asthma phenotypes) against the PRS, sex, ancestry PCs 1–10, and four additional environmental confounders: body mass index (BMI), the Townsend deprivation index, disability allowance, and pack years smoking.

HLA effects

To evaluate the extent to which the HLA region contributed to PRS-trait associations, we calculated modified PRSs for individuals in the white British cohort, in which all contributions from SNPs in the extended HLA region (chr6:28477797-33448354, GRCh37) were removed. We then performed a PheWAS using the HLA-removed scores and compared the association effect sizes with those generated from the initial PRS PheWAS in the white British population. Statistical probabilities for differences in AUCs between models were estimated using 1,000 bootstrap permutations.

Trait mediation analysis

To quantify the relative contributions of eosinophilia, lung function, and allergic disease on the overall genetic risk of asthma, we performed mediation analyses⁵² with each factor modeled as a mediator between the PRS and asthma. We performed separate tests for asthma, COA, and AOA. We used blood eosinophil count (UKB field 30150), ratio of FEV₁ to forced vital capacity (FVC) (UKB field 20258), and age of hay fever diagnosis (UKB field 3761) as the mediators in our models, denoting eosinophilia, lung function, and allergic disease, respectively. Z score transformation was first

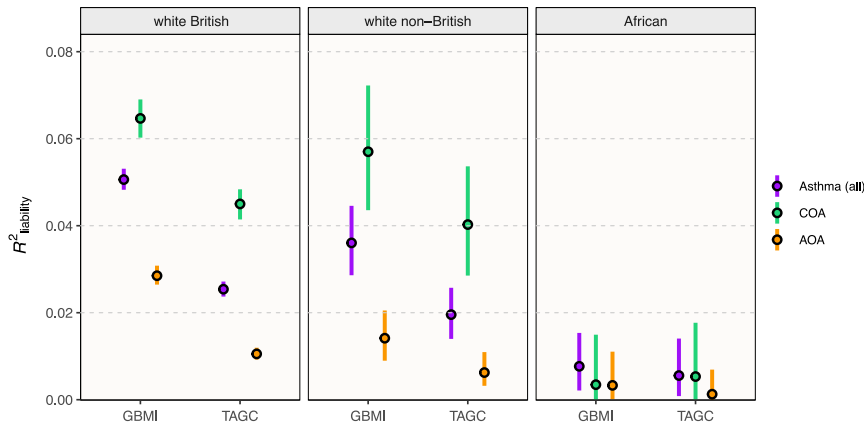


Figure 1. Asthma liability explained by GBMI and TAGC polygenic risk scores
 R^2 on the liability scale with 95% CIs are shown for the GBMI PRS and best-performing TAGC PRS for all asthma, COA (<12 years old), and AOA (≥ 25 and <66 years old) in three different UKB populations according to ancestry. AUC, area under the curve.

The GBMI PRS was the most accurate in all populations, with an AUC of 0.623 in white British, 0.608 in white non-British, and 0.545 in African-ancestry individuals. Because

the GBMI model was the best performing in each group (Figure 1), we used the individual risk scores generated from that model for the remainder of the study, except where we further compare trait associations between models below (see section “trait association differences by PRS model”). The OR for asthma per standard deviation (SD) increase in the PRS was 1.56 (95% CI = 1.54–1.58). Individuals with a PRS in the top decile were 4.69 times (95% CI = 4.43–4.95) more likely to have had asthma than those in the lowest decile (Figure 2A).

Genetic risk by sex and age of onset

Many features of asthma differ by sex and age at disease onset. For example, COA is more prevalent in males and has a high rate of remission, while AOA is more prevalent in females and is associated with more severe symptoms.^{54,55} Moreover, genetic risk factors contribute more to the development of COA relative to AOA.^{9,56} We therefore tested PRS associations with asthma by age of onset and by sex to quantify these differences and identify potential interaction effects. The mean PRS was significantly higher in individuals with COA ($\mu = 0.55$) than in individuals with AOA ($\mu = 0.31$; $p = 3.4 \times 10^{-87}$; Figure 2B). The OR for each SD increase in the PRS was 1.83 (95% CI = 1.79–1.87) for COA and 1.43 (95% CI = 1.41–1.45) for AOA (Figure 2C). Individuals with a PRS in the top decile were 8.3 times (95% CI = 7.4–9.3) more likely to have had COA and 3.5 times (95% CI = 3.3–3.8) more likely to have had AOA than those in the lowest decile. Accordingly, the asthma PRS was better at predicting COA (AUC = 0.665) than AOA (AUC = 0.600; Figure 2D). These results reinforce how genetic risk factors contribute more to COA. Indeed, AOA was more prevalent than COA in individuals with a lower asthma PRS, and COA was more common in individuals with a higher PRS (Figure 3A).

We also observed significant interaction effects between the PRS and sex, depending on age of onset. While the overall risk of COA was higher in males than in females, regardless of genetic risk, the PRS was slightly more correlated with disease risk in females (interaction $p = 1.0 \times 10^{-2}$; ratio of ORs = 1.06; Figure 3B). We observed the opposite effects in AOA, with a higher overall risk in females but a

applied to each trait within the white British UKB population. The mediation analysis consisted of three regression models:

$$Y = i_1 + cX + \beta_1 \mathbf{Z} + e_1 \quad (\text{Equation 1})$$

$$M = i_2 + aX + \beta_2 \mathbf{Z} + e_2 \quad (\text{Equation 2})$$

$$Y = i_3 + c'X + bM + \beta_3 \mathbf{Z} + e_3 \quad (\text{Equation 3})$$

where Y is the dependent asthma phenotype; X is the PRS; M is the mediating trait; c , a , c' , and b are coefficients relating the independent variables to the dependent variables; \mathbf{Z} is a matrix of baseline covariates with β coefficients; i is the intercept; and e are the residuals. We included sex and the first 10 PCs of ancestry as fixed covariates. We also included BMI (UKB field 21001), the Townsend deprivation index (UKB field 189), and estimated pack years of smoking (UKB fields 2867, 2887, 2897, 2907, 3436, 3456, 3486, 6183, 6194, 20160, 21022) as fixed covariates to mitigate against unmeasured confounding. The proportion of the PRS effect on asthma mediated by each mediating trait was determined using the standardized logistic solution proposed by MacKinnon et al.⁵³:

$$c_s = c \sqrt{1 + \frac{b^2 \sigma_{MX}^2}{\pi^2/3}} \quad (\text{Equation 4})$$

$$\text{Proportion mediated} = \frac{ab}{c_{\text{standardized}}} \quad (\text{Equation 5})$$

where c_s is the standardized form of the c coefficient from Equation 1 and σ_{MX}^2 is the residual variance from Equation 2. We performed bootstrap sampling ($n = 1,000$) to derive 95% CIs.

Results

Genetic risk prediction of asthma

We evaluated the predictive performance of five genetic risk models for asthma, derived using the PRS-CS polygenic prediction method,⁴³ on summary statistics from either the TAGC³⁰ or GBMI⁷ meta-analyses of asthma GWASs. The four TAGC scores featured between 694,173 SNPs and 699,291 SNPs, depending on the LD panel. The GBMI-based PRS consisted of 637,438 SNPs.⁷ We compared the performance of these five scores in predicting asthma in white British, white non-British, and African-ancestry population samples in the UKB (Table S1).

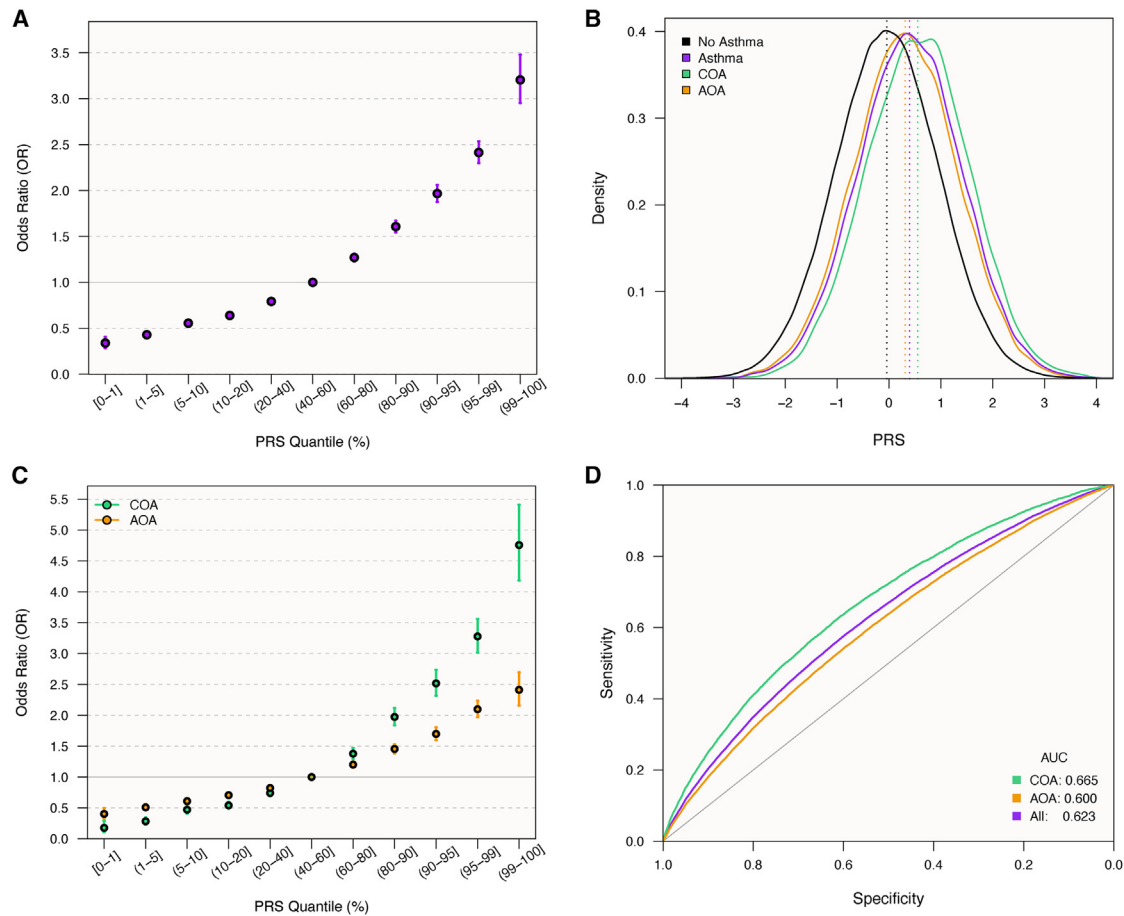


Figure 2. PRS prediction of asthma in white British UKB participants

(A) Odds ratios (ORs) with 95% CIs of asthma risk for individuals in different GBMI PRS quantiles relative to those with scores in the 40th–60th quantile.

(B) Density plot showing distributions of PRS scores by asthma status, with mean scores indicated by vertical lines.

(C) ORs with 95% CIs of COA and AOA risk for individuals in different PRS quantiles relative to those with scores in the 40th–60th quantile.

(D) Receiver-operating-characteristic curves for the PRS as a predictor of asthma, COA, and AOA.

greater PRS correlation in males (interaction $p = 4.6 \times 10^{-2}$; ratio of ORs = 1.03; Figure 3C). Overall, these results showed that genetic risk is higher in COA than AOA and that sex interacted with the age of asthma onset.

PheWAS of asthma PRS

To quantify the polygenic pleiotropy between asthma and other traits and conditions, we performed a PheWAS of the GBMI asthma PRS with 371 heritable biomedical traits in the UKB white British population ($n = 376,237$). After applying a conservative Bonferroni correction for multiple testing, 214 traits were significantly associated with the asthma PRS across a variety of biological systems (Figure 4; Table S2). As expected, asthma and its related traits and comorbidities were the most strongly associated phenotypes, including wheezing ($p < 4.9 \times 10^{-324}$, OR = 1.30), treatment with albuterol ($p < 4.9 \times 10^{-324}$, OR = 1.59), FEV₁/FVC ($p < 4.9 \times 10^{-324}$, $\beta_z = -0.08$), diagnosis of hay fever/allergic rhinitis or eczema ($p < 4.9 \times 10^{-324}$, OR = 1.24), and eosinophil count ($p < 4.9 \times 10^{-324}$, standardized effect of 1 SD increase in PRS [β_z] = 0.11). In line

with our earlier GWAS findings,⁹ the PRS was significantly associated with younger age of onset as a quantitative trait ($p = 3.7 \times 10^{-111}$, $\beta_z = -0.12$), where each SD increase in the PRS corresponded to 2.1 years (95% CI = 1.9–2.3) younger age of onset. Other significant disease associations of note include several immune-mediated diseases, such as celiac disease ($p = 1.5 \times 10^{-40}$, OR = 1.38), type 1 diabetes (T1DM; $p = 8.2 \times 10^{-41}$, OR = 1.26), and hypothyroidism ($p = 2.3 \times 10^{-36}$, OR = 1.10).

Twenty-nine of the UKB traits included in our study were measured in only one sex (Table S2), and 17 of those were significantly associated with the asthma PRS (Figure 4). In females, age at first live birth ($p = 2.5 \times 10^{-46}$, $\beta_z = -0.04$, $\beta = -0.17$ years), age at menopause ($p = 7.1 \times 10^{-22}$, $\beta_z = -0.03$, $\beta = -0.14$ years), and hormone-replacement therapy ($p = 5.6 \times 10^{-19}$, OR = 1.26) were the top three sex-specific associations with the asthma PRS. In males, insulin treatment ($p = 3.0 \times 10^{-29}$, OR = 1.26), hypertension treatment ($p = 1.1 \times 10^{-9}$, OR = 1.03), and male pattern baldness ($p = 1.7 \times 10^{-7}$, OR = 1.03) were the top three sex-specific associations.

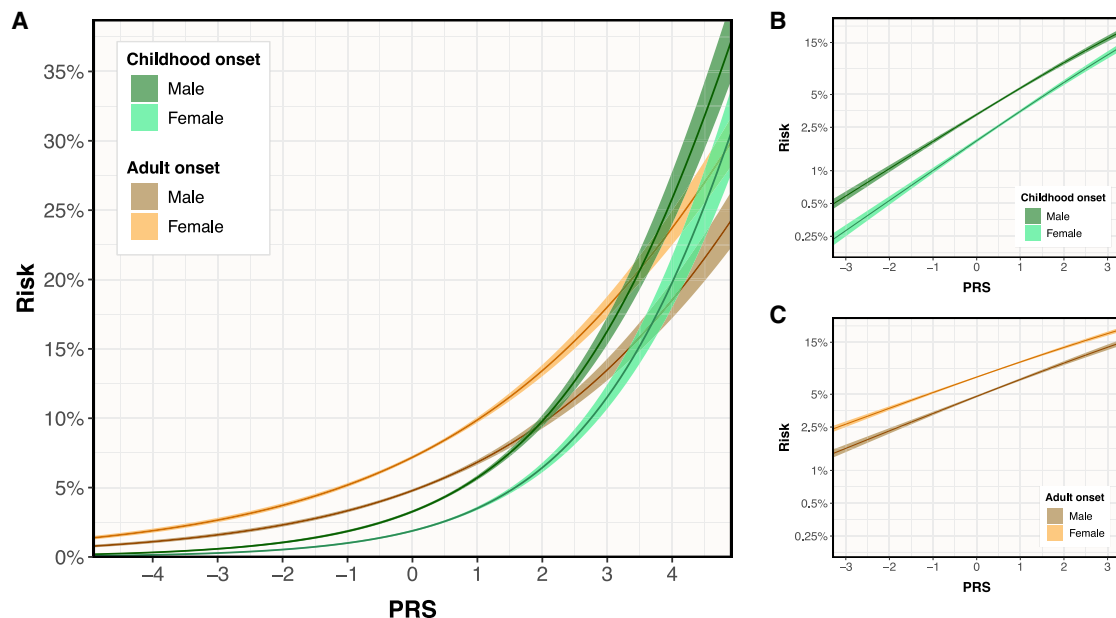


Figure 3. Asthma risk vs. GBMI PRS by sex and age of onset

(A) Predicted risks of COA and AOA are plotted against the PRS score separately for males and females in the white British UKB population with 95% CIs.

(B) Predicted risks of COA are plotted on a log scale with 95% CIs, with the different slopes demonstrating the PRS-sex interaction for COA.

(C) Predicted risks of AOA are plotted on a log scale with 95% CIs, demonstrating the PRS-sex interaction for AOA.

Trait association differences by PRS model

The GBMI and TAGC PRSs were derived from different cohorts that used different phenotype definitions for asthma. The GBMI meta-analysis defined asthma predominantly according to ICD codes from health records of adult participants,^{7,33} whereas TAGC relied more on physician diagnosis and over 40% of its studies were limited to COA.³⁰ To explore how these differences might affect pleiotropic associations with asthma, we repeated the PheWAS using the best-performing TAGC PRS (multi-ancestry cohort, UKB LD panel), and compared the association results with those from the GBMI PRS PheWAS. The association effect size and OR correlations between PheWASs were $r = 0.87$ and $r = 0.90$, respectively (Figure 5).

The TAGC PRS was significantly associated with 110 traits after Bonferroni correction (Figure S1; Table S3), including 14 that were uniquely associated with the TAGC PRS. The TAGC PRS was not significantly associated with 118 of the traits significantly associated with the GBMI PRS. Of the 45 binary traits that were significantly associated with both PRSs, celiac disease (UKB field 20002) had the largest relative effect size bias toward the TAGC PRS ($OR_{TAGC} = 1.81$, $OR_{GBMI} = 1.38$), and breathing problems at work (UKB field 22616) had the largest effect size bias toward the GBMI PRS ($OR_{TAGC} = 1.31$, $OR_{GBMI} = 1.57$; Figure 5A). Of the 51 quantitative phenotypes significantly associated with both PRSs, age of asthma diagnosis (UKB field 22147) had the largest difference in association effect sizes biased toward the TAGC PRS ($\beta_{z,TAGC} = -0.11$, $\beta_{z,GBMI} = -0.10$), whereas FEV₁ (percentage predicted; UKB field 20154) had the largest bias to-

ward the GBMI PRS ($\beta_{z,TAGC} = -0.05$, $\beta_{z,GBMI} = -0.09$; Figure 5B).

Categorically, anthropometric trait associations were the most biased toward the GBMI PRS ($p = 4.4 \times 10^{-14}$), followed by asthma and allergic disease ($p = 1.0 \times 10^{-11}$), pulmonary ($p = 1.1 \times 10^{-6}$), musculoskeletal and skin ($p = 5.4 \times 10^{-8}$), and other ($p = 1.8 \times 10^{-3}$) traits (Figures 5C and 5D). Notably, the TAGC PRS was not significantly associated with any of the 24 neurologic and behavioral traits that were significantly associated with the GBMI PRS, including five measures of smoking. Most of the remaining traits that were significantly associated with the GBMI PRS and not the TAGC PRS were related to cardiovascular and metabolic health (Tables S2 and S3).

Trait associations in a non-asthma population

Trait associations with asthma genetic risk scores do not necessarily arise from shared genetic architecture but could be secondary to disease-induced physiological and/or behavioral changes. To evaluate the degree to which the trait associations with asthma genetic risk were independent of asthma, we repeated the PheWAS of the GBMI asthma PRS in the white British population, excluding all individuals with a report or diagnosis of asthma (remaining $n = 328,353$). We used inverse probability weighting to mitigate selection bias induced from conditioning on asthma (see section “subjects and methods”). Of the 205 non-asthma-defining significant traits in the initial PheWAS, 54 (26.3%) were no longer significantly associated with the asthma PRS after correcting for multiple testing, and 17 had significantly weaker association effects



Figure 4. Phenome-wide association test of GBMI asthma PRS in the UKB

The p values for associations with the asthma PRS are shown for 257 non-redundant traits out of 371 heritable phenotypes tested in the UKB. Traits to the right of the dashed line were significantly associated after correcting for multiple testing ($p < 1.35 \times 10^{-4}$). Traits with association p values $< 1 \times 10^{-10}$ are labeled. Traits are grouped by corresponding organ system. The direction of each arrow corresponds to the direction of the association effect. ALP, alkaline phosphatase; AST, aspartate aminotransferase; BMI, body mass index; CRP, C-reactive protein; COPD, chronic obstructive pulmonary disease; DED, diabetes-related eye disease; FEV₁, forced expiratory volume in 1 s; FVC, forced vital capacity; GGT, gamma-glutamyl transferase; HbA1C, glycated hemoglobin; Hgb, hemoglobin; HDL, high-density lipoprotein; HRT, hormone-replacement therapy; MCH, mean corpuscular hemoglobin; RBC, red blood cell; T1DM, type 1 diabetes mellitus; T2DM, type 2 diabetes mellitus; Tx, treatment.

in the non-asthmatic cohort (Figure 6; Table S4). Of the 17 traits with weaker association effects, five were related to asthma and allergic disease, eight were related to pulmonary function, and four were related to leukocyte counts. PRS associations with FEV₁ ($\Delta\beta = -35\%$, $p_{\Delta\beta} = 8.4 \times 10^{-15}$), FVC ($\Delta\beta = -25\%$, $p_{\Delta\beta} = 2.0 \times 10^{-5}$), FEV₁/FVC ($\Delta\beta = -42\%$, $p_{\Delta\beta} = 1.4 \times 10^{-38}$), and COPD prevalence ($\Delta\beta = -49\%$, $p_{\Delta\beta} = 8.6 \times 10^{-10}$) were all significantly weaker in the non-asthmatic cohort. The association with total leukocyte counts was also significantly weaker ($\Delta\beta = -31\%$, $p_{\Delta\beta} = 7.9 \times 10^{-5}$), including significant reductions in association strength for eosinophil counts ($\Delta\beta = -18\%$, $p_{\Delta\beta} = 1.1 \times 10^{-15}$) and neutrophil counts ($\Delta\beta = -38\%$, $p_{\Delta\beta} = 2.4 \times 10^{-5}$), specifically. These results highlight specific trait associations with genetic asthma risk that arise largely due to disease-induced changes, but most of the traits with significantly reduced effect sizes ($n = 13/17$, 76%) nonetheless remained significantly associated with the asthma PRS. Importantly, for the vast majority of traits, their association effect sizes did not change

significantly in the non-asthmatic cohort relative to the cohort with asthmatics, indicating that most trait associations with genetic asthma risk are due to pleiotropic effects arising from shared biology.

Replication in subjects with different ancestries

To assess the robustness of the asthma PRS-trait associations, we performed PheWASs of the GBMI PRS in two other UKB cohorts, a white non-British population ($n = 27,026$; Table S5) and a population of individuals with African ancestries ($n = 7,196$; Table S6). In the white non-British population, 208 (97%) of the trait associations had consistent directions of effect with the white British cohort, of which 139 (65%) replicated at $p < 0.05$. Considering asthma prediction was similar in the white British and white non-British cohorts (Table S1), the non-replicated traits were likely due to the loss of power in the smaller cohort.

In the population with African ancestries, 109 (52%) of the significant trait associations from the white British

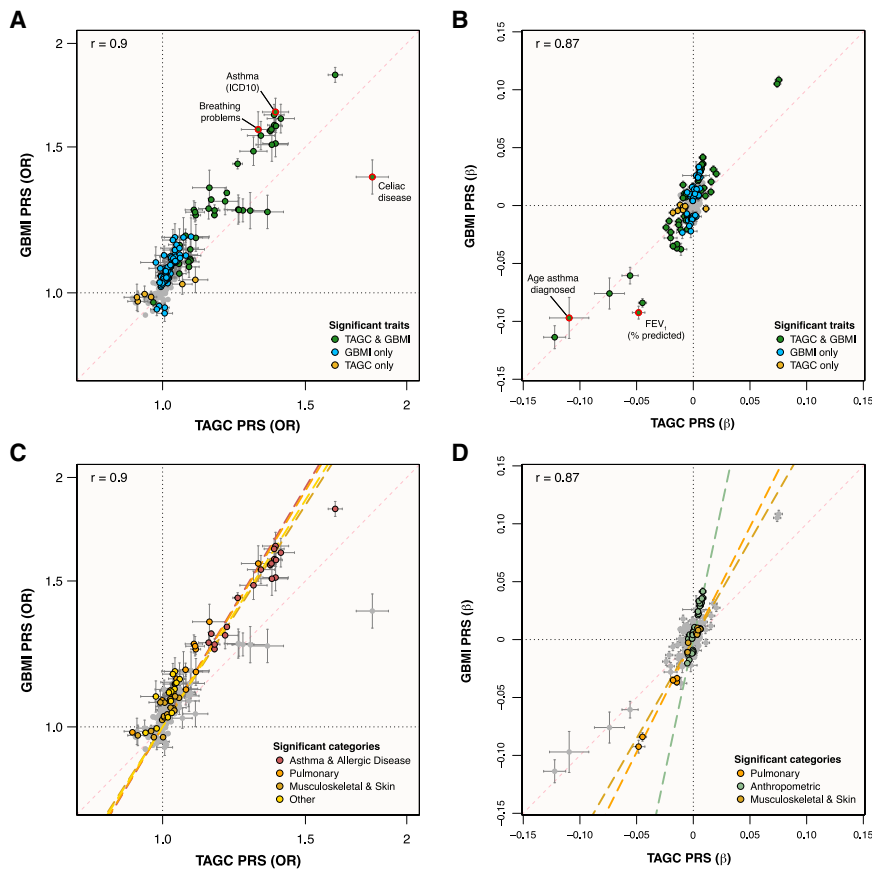


Figure 5. Scatterplots of GBMI PRS vs. TAGC PRS-trait associations in the UKB (A and B) Binary (A) and quantitative (B) traits association effects are shown with 95% CIs, colored according to which PRSs had significant associations. Traits highlighted in red had the largest effect bias toward one or the other PRS models. (C and D) Binary (C) and quantitative (D) traits association effects are shown with 95% CIs, colored according to which trait categories were significantly biased toward one of the PRS models. ICD-10, International Classification of Diseases, 10th revision.

PRS-trait associations after removing HLA region variants from the asthma PRS (Table S7). The region contained only 0.7% ($n = 4,694$) of the total number of modeled SNPs but included 39 of the top 100 by scoring weight. The correlation between the whole-genome PRS and the HLA-removed PRS was 0.88.

Removal of the HLA region had a negligible effect on the PRS association and predictive performance with asthma ($\Delta\beta=0\%$, $\Delta_{AUC} = -0.003$). We observed no difference for COA

PheWAS yielded the same direction of effect, with 20 (11%) at $p < 0.05$. Although this cohort was relatively small, the PRS was also less predictive of asthma in this cohort relative to the others (Table S1), demonstrating the limited portability of PRSs across populations,⁵⁷ even when using a PRS based on multi-ancestry GWASs. For the traits that replicated at $p < 0.05$ in the white non-British and African-ancestry cohorts ($n = 18$), we tested for differences in association effect sizes between African ancestry and the white British and white non-British cohorts. We identified four traits with significantly different effects, including wheezing in the past year ($p = 9.7 \times 10^{-5}$), and three measures of self-reported asthma (UKB data fields 20002, $p = 8.8 \times 10^{-4}$; 6152 [asthma], $p = 4.5 \times 10^{-4}$; 6152 [none], $p = 7.2 \times 10^{-4}$). In each case, the PRS association effect sizes were significantly smaller in the African-ancestries population (Table S6) and likely reflect the lower predictive accuracy of the PRS in the population with African ancestries.

Influence of the HLA region

The HLA region on chromosome 6p21.3 is the most gene-dense and polymorphic region in the human genome,⁵⁸ encoding proteins with critical roles in immune responses. Genetic variants in the HLA region have been robustly associated with risk for both autoimmune and allergic diseases.^{59,60} To investigate the influence of HLA region alleles on asthma PRS associations, we measured GBMI

($\Delta\beta=0\%$, $\Delta_{AUC} = 0.000$). The HLA-removed PRS was marginally worse at predicting AOA ($\Delta\beta = -4\%$, $\Delta_{AUC} = -0.004$) compared to the whole-genome PRS, but none of these differences were statistically significant.

In contrast, PRS associations with autoimmune conditions were dramatically reduced after removing the HLA region variants. Associations with celiac disease ($\Delta\beta = -68\%$, $p_{\Delta\beta} = 1.5 \times 10^{-10}$), T1DM ($\Delta\beta = -60\%$, $p_{\Delta\beta} = 1.5 \times 10^{-8}$), hypothyroidism ($\Delta\beta = -51\%$, $p_{\Delta\beta} = 6.5 \times 10^{-6}$), and their corresponding treatments were significantly weaker with the HLA-removed PRS (Figure 7). The only trait with a significantly stronger association with the asthma PRS after removing the HLA region was eosinophil levels ($\Delta\beta = 10\%$, $p_{\Delta\beta} = 4.2 \times 10^{-6}$), suggesting an opposite effect of the HLA region compared to the rest of the genome's shared architecture with blood eosinophil levels. No other traits demonstrated significantly different association effect sizes with the HLA-removed PRS. Overall, these results reveal that the pleiotropy between asthma and immune-mediated traits is largely concentrated within the HLA region.

Mediation of cardinal asthma-associated traits on genetic risk for asthma

Allergic disease, lung function, and eosinophils are all highly correlated with asthma, yet there is considerable heterogeneity among these traits in asthmatic patients. To quantify the relative extent to which each of these traits

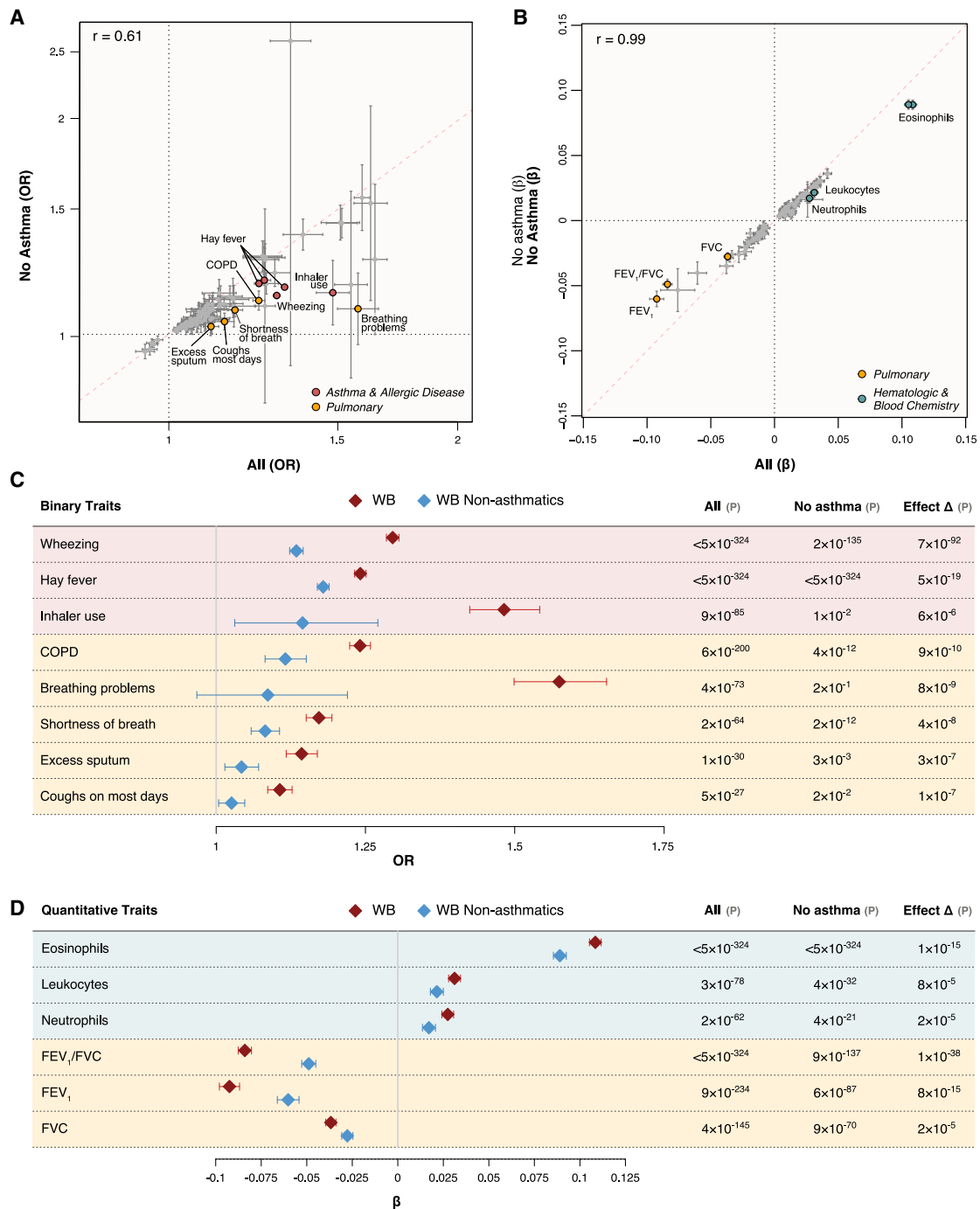


Figure 6. GBMI PRS-trait associations in non-asthmatics

(A and B) Scatterplots of binary (A) and quantitative (B) trait associations with the GBMI PRS in UKB white British for individuals with no history of asthma vs. all individuals are shown with 95% CIs. Traits with significant differences in association effect sizes are labeled and colored according to trait category.

(C and D) Forest plots of binary (C) and quantitative (D) trait associations with the GBMI PRS with significant effect size differences between the non-asthmatic population and the entire white British cohort are shown with 95% CIs. COPD, chronic obstructive pulmonary disease; WB, white British.

mediate the genetic risk of asthma, we performed mediation analyses⁵² with each trait modeled as a mediator between the GBMI PRS and asthma, stratified by COA and AOA. Blood eosinophil count, FEV₁/FVC, and age of hay fever diagnosis were used as the metrics for eosinophilia, lung function, and allergic disease, respectively (Figure 8).

For COA, age of hay fever diagnosis mediated the largest proportion of genetic risk (12%) relative to eosinophil count (6%) and FEV₁/FVC (4%). In AOA, however, age of hay fever diagnosis did not mediate any of the genetic risk (−1%), while eosinophils mediated the greatest proportion (8%), followed by FEV₁/FVC (5%). These results

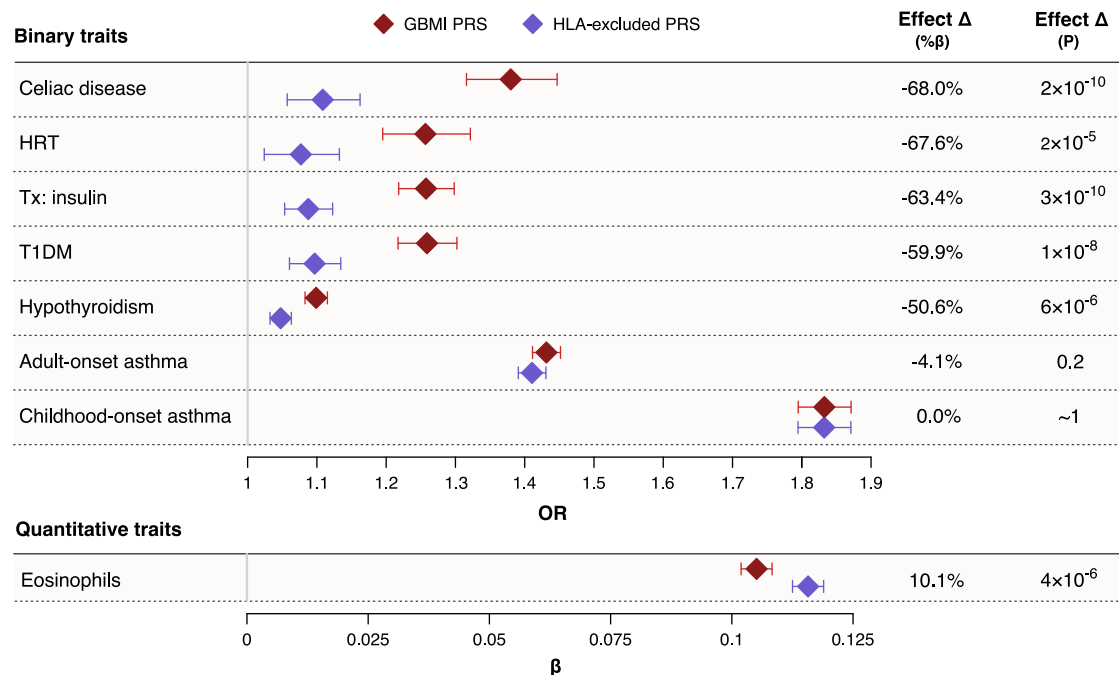


Figure 7. Changes in asthma PRS associations after removing HLA region

Changes in PRS association effect sizes after removing the HLA region from the PRS ($\Delta\beta$) are shown with 95% confidence intervals for traits with significant effect reductions and for COA and AOA. Traits with a negative $\Delta\beta$ had a weaker association with the asthma PRS, and traits with a positive $\Delta\beta$ had a stronger association with the asthma PRS after removing HLA region alleles from consideration.

are consistent with clinical observations of allergic disease being a prominent comorbidity with COA and further indicates that this may be due to shared genetics.

Discussion

This study provides a quantitative overview of the genetic landscape for asthma, revealing polygenic pleiotropy between asthma and hundreds of other traits and conditions across a variety of biological systems and highlights differences by sex and age of onset. Previous studies have developed PRSs for asthma,^{6,7,11–21} with comparable AUC values when reported,^{12,14,21} but most were more limited by just focusing on predictive performance^{11–15} or by testing associations with specific traits.^{15–21} Our study features a more comprehensive evaluation of the traits and conditions associated with genetic asthma risk, including dependencies on age of onset, sex, ancestry, disease status, and HLA alleles, thereby revealing novel features of asthma and its associated traits.

Utilization of a PRS enabled us to perform direct quantification of relative disease risks and facilitated additional analyses that can account for genetic risk strata, disease status, mediating effects, and/or confounding factors. In contrast, previous studies of cross-trait genetic correlations with asthma provide estimates of shared genetic effects among common variants but are less interpretable in terms of relative disease risk.^{7,22–29} The relative risks between different disease pairs with similar genetic correlation esti-

mates can vary dramatically.⁶¹ PRS-based associations therefore offer a different view of multimorbidity with additional utility than cross-trait genetic correlation estimates alone.

Age and sex both significantly alter genetic effects on asthma risk. Asthma that begins in childhood (COA) is more often diagnosed in males and is associated with different risk factors compared to asthma that develops in adulthood (AOA), which is more commonly diagnosed in females and is associated with increased respiratory symptoms and a lower frequency of quiescent disease.^{54,55,62} In fact, clinically, COA and AOA are often considered different subtypes of asthma. Recent studies have revealed that the genetic etiologies of COA and AOA are indeed partly distinct,^{9,56} with an estimated genetic correlation of $r_g = 0.67$.⁵⁶ We previously reported that the genetic risks for AOA were largely a subset of those for COA but with overall smaller effect sizes.⁹ Consistent with this, SNP-based heritability estimates for COA were approximately three times greater than those estimated for AOA.^{9,56} These combined findings indicate that genetics plays a more prominent role in asthma that presents earlier in life and, conversely, risk of AOA is more environmentally mediated. In the present study, the genetic risk for asthma was significantly associated with a younger age of onset and was a better predictor of COA than AOA.

Our results further reinforced that genetic risk for asthma acts through different mechanisms depending on the age of onset. The mediation analyses revealed that

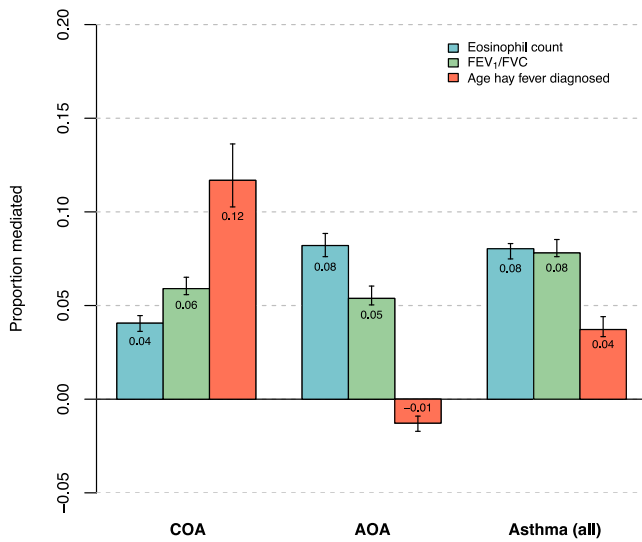


Figure 8. Mediation of cardinal asthma-associated traits on genetic risk for asthma

For the PRS associations with asthma, the estimated proportions mediated by eosinophil count (blue), FEV₁/FVC (green), and age of hay fever diagnosis (orange) are shown with 95% confidence intervals. Mediation analysis was conducted separately for COA, AOA, and all asthma. In each instance, the asthma PRS was the independent variable, asthma status was the dependent variable, and one of the three traits was the mediator.

genetic risk for COA is more related to allergic disease, whereas genetic risk for AOA is more related to eosinophilia. These findings align with observations that allergic asthma is the predominant endotype in children.^{54,62,63}

Mechanistic differences between COA and AOA may also be contributing to the trait association differences we observed between the GBMI PRS PheWAS and the TAGC PRS PheWAS. The GBMI PRS was modeled primarily on ICD codes in adults and is therefore likely identifying more of an adult-onset phenotype than the PRS from TAGC, in which the cases in 40% of the studies were COA. Indeed, the binary UKB trait with the second largest association bias toward the GBMI PRS was the primary asthma ICD-10 code. Of the 1,667 individuals in the UKB with a primary asthma ICD-10 code and information on age of asthma diagnosis, only 471 (28%) reported an asthma diagnosis in childhood, whereas 945 (57%) reported a first asthma diagnosis in adulthood (>25 and <66 years old). Anthropometric traits, consisting almost entirely of body-fat measures, were the most biased toward the GBMI PRS, and the TAGC PRS was not associated with any of the 24 neurologic and behavioral traits associated with the GBMI PRS, including five different measures of smoking behavior. These findings accord with genetic correlation results reported by Ferreira et al.,⁵⁶ in which COA had a larger allergic component than AOA, whereas only AOA had significant genetic overlap with obesity-related traits and smoking history. Of the traits more strongly associated with the TAGC PRS, celiac disease stood out as having the greatest effect size difference. Previous studies have

reported correlations between COA and celiac disease incidence,^{64–68} including a study by Hemminki et al.⁶⁵ that found the risk of celiac disease following an asthma-related hospitalization was greater only in individuals that were diagnosed with asthma before 20 years of age. Importantly, while the relative composition of COA and AOA cases in GBMI and TAGC likely contributed to trait association differences between their respective PRS PheWASs, other factors, such as relative training data sample size, overall age distribution, ancestry composition, phenotype accuracy, and/or overlap between asthma cases and COPD, could also have contributed. Therefore, we cannot be certain what underlies the differences between the GBMI and TAGC PRS PheWASs. Nonetheless, our results demonstrate how compositional differences in PRS training datasets inevitably produce models with different genetic architectures for the same disease.

In addition to allergic diseases (hay fever/allergic rhinitis, eczema), pulmonary traits were among those most strongly associated with the asthma PRSs. Multiple studies have previously reported significant genetic overlap between asthma and various measures of lung function,^{7,21,24,56,69,70} but their causal relationships are complex.⁷¹ In our study, most asthma PRS associations with pulmonary traits remained significant after excluding individuals with an asthma diagnosis, indicating that asthma and lung function share a genetic architecture independent of asthma diagnosis. However, the corresponding effect sizes upon excluding asthmatics were greatly reduced, and half of the traits that were no longer significantly associated with the asthma PRS were pulmonary traits, suggesting that asthma itself leads to lower lung function. For example, after removing the asthma cases, the PRS association with COPD was still statistically significant, but its effect was substantially weaker in the non-asthmatic cohort ($\Delta\beta = -48\%$, $p_{\Delta\beta} = 2.5 \times 10^{-9}$). This combination of shared genetics and disease-induced changes could explain the contradictory results regarding the causal order of the association between lung function and asthma.⁷¹

Many risk loci for asthma and allergy are also associated with one or more autoimmune diseases.^{57,59,72} We observed significant associations with the GBMI asthma PRS and celiac disease, hypothyroidism, and T1DM, and additionally for ulcerative colitis with the TAGC PRS. Our results showed that the genetic overlap between asthma and immune-mediated conditions was primarily due to genetic variation at the HLA locus, as their PRS associations were significantly reduced after removing HLA region variants from the PRS. Interestingly, genetic risk of asthma, according to the TAGC PRS, had a protective effect on risk for ulcerative colitis, which is consistent with reports that many alleles associated with both asthma and autoimmune diseases have opposite effects on pathogenesis.^{59,60}

The asthma PRSs were also significantly associated with immune cell types. Eosinophils, in particular, had the most significant quantitative trait associations with both

the GBMI ($p < 4.9 \times 10^{-324}$, $\beta_z = 0.11$) and TAGC ($p < 4.9 \times 10^{-324}$, $\beta_z = 0.07$) PRSs. Eosinophils are granulocytes that contribute to immunity during multicellular parasitic infection, but, in geographical regions where parasites are less abundant, eosinophils can contribute to type 2 inflammation and subtypes of severe asthma.⁷³ Notably, the PRS association with eosinophil levels was the only association that was significantly stronger after removing the HLA region from the PRS ($\Delta\beta = 10\%$, $p_{\Delta\beta} = 4.2 \times 10^{-6}$), suggesting that the shared architecture between asthma and blood eosinophil levels resides outside of the HLA region.

We also found links between asthma, obesity, and insulin resistance, an observation that is well established in both children and adults.⁷⁴ Mendelian randomization studies indicated that BMI causally increases the risk of asthma,^{28,75,76} but findings from prospective studies suggest that asthma is also a risk factor for subsequent obesity.^{77–80} Our study suggested that the PRS associations with body fat were not due to asthma onset: most of the associations with body-fat-related traits remained significant after removing the asthma cases from the PheWAS. The GBMI PRS had greater association effects with obesity-related traits than the TAGC PRS did, which could be due to significant genetic correlations between obesity-related traits and AOA but not COA.^{28,56} Insulin resistance has been reported to modify the association between obesity and asthma in adults,⁸¹ suggesting that features of metabolic syndrome may contribute to asthma risk independent of obesity. A recent study found that asthma was associated with metabolic syndrome independent of BMI, but only in adults.⁸² In our study, body-fat traits had stronger associations than T2DM with the GBMI PRS but weaker associations than T2DM for the TAGC PRS. Together these findings showed that both insulin and obesity are associated with asthma risk, but the strength of their associations depend on the population and/or how asthma is defined.

Another notable finding was the significant association between the asthma PRS and bilirubin, a metabolite of heme. A study by Turi et al.⁸³ identified dose-dependent protective associations between unconjugated bilirubin measured in plasma at age 1 year and subsequent diagnoses of recurrent wheeze and COA.⁸³ The study found that higher concentrations of bilirubin within the normal physiological range were associated with protection from asthma. Our study extends the link between asthma and bilirubin by showing that higher genetic risk for asthma is significantly correlated with lower total bilirubin levels and further demonstrates that this relationship is due at least in part to shared genetics.

There are some important caveats and limitations of our study. First, the genetic risk for asthma captured by the PRSs were limited to HapMap3 SNPs genotyped in the training data and the UKB. The genetic risk modeled in this study was therefore limited to that captured by common tag SNPs and could not account for independent contributions from lower-frequency variants or from a more

comprehensive SNP coverage. Second, the PRSs were modeled using summary statistics from multi-ancestry GWASs,^{30,32} but most samples nonetheless consisted of individuals of European ancestries. Residual population stratification in GWASs can lead to biased PRS associations in test cohorts with overlapping population structure.⁸⁴ Therefore, any stratification biases within the TAGC or GBMI GWASs could have confounded certain PRS associations, and we did not model this potential effect. Moreover, performances of PRSs suffer when applied to populations with different ancestries, due to differences in LD structure, allele frequencies, effect sizes, and heritability.⁸⁵ Unsurprisingly, the accuracy of the PRSs was lowest in the African-ancestry cohort. We had less power to detect PRS-trait associations in this cohort as well, due to its smaller sample size. Larger non-European training samples and the development of new methods to improve bias correction and cross-population PRS accuracy are needed to prevent PRS use from exacerbating health disparities.⁸⁶ Third, the PRSs applied in this study were modeled on asthma phenotypes that included both COA and AOA cases. The models therefore likely best represent genetic risk common to both COA and AOA. PRSs modeled on COA or AOA alone may associate with different sets of traits than those identified in this study and better capture genetic differences between COA and AOA. Fourth, we did not account for potential assortative mating effects, which can inflate genetic correlation estimates for etiologically independent traits,⁸⁷ although the impact of assortative mating on PRS-trait associations has not yet been thoroughly investigated.⁸⁸ Fifth, the phenotypes based on ICD codes included in our PheWAS correspond to principal diagnoses responsible for hospital admission and do not include secondary diagnoses that coexist at admission or subsequently develop. Use of ICD codes to denote disease phenotypes is also subject to misclassification.^{89,90} Consequently, the associations reported in our study for such traits may not be representative of overall risk. Sixth, the UKB captured all phenotype data from adults. Therefore, for traits measured at the time of study, such as BMI or lung function or leukocyte proportions, we are unable to make conclusions about potential associations earlier in life. For example, eosinophil count measured in childhood would likely mediate more of the genetic risk of COA than would the eosinophil counts recorded in the UKB. Furthermore, although our previous sensitivity analysis on COA in the UKB suggested that recall bias did not contribute to our GWAS results,⁹ we cannot exclude recall bias among participants for early-life diagnoses, such as COA, which could have inflated associations in the non-asthmatic cohort if a proportion of COA cases were mis-specified as controls. Seventh, estimation of direct and indirect effects within our mediation analyses inherently assumed no unmeasured mediator-outcome confounders affected by the PRS. We modeled for confounding with BMI, the Townsend deprivation index, disability, and smoking history, but, considering the polygenic pleiotropy

demonstrated by the PRS, it is unlikely that this assumption was strictly met for all models. Therefore, our mediation results serve to illustrate relative correlations rather than provide strict causal inferences. Finally, associations with the asthma PRS do not necessarily represent causal biological relationships. We removed traits that were not interpretable as biomedical phenotypes or laboratory values, but any number of intermediate or correlated factors could contribute to associations observed with the asthma PRS.

In summary, we quantified the probabilities of developing asthma based on relative genetic risk and identified numerous traits across a wide spectrum of biological systems that were significantly associated with genetic risk for asthma. We identified trait associations that were limited to specific PRS models, demonstrating how different PRS training data produce models with different architectures. Most PRS-trait associations remained significant after controlling for asthma status, indicating shared genetics, but lung function traits and inflammatory markers were significantly less associated with genetic asthma risk in non-asthmatics, suggesting disease-state-induced changes of these traits. We found the genetic overlap with autoimmune-related phenotypes was largely limited to the HLA region and showed etiological differences between COA and AOA, with genetic risk of COA mediated more by allergic disease and AOA more closely tied to eosinophilia. Overall, this study sheds light on how asthma is related to other traits and conditions and opens new avenues for investigation, where further research may guide improved risk management and more effective therapies.

Data and code availability

The TAGC GWAS summary statistics are publicly available via the GWAS catalog (<https://www.ebi.ac.uk/gwas/downloads/summary-statistics>). GBMI PRSs are publicly available via the Polygenic Score Catalog (<https://www.pgscatalog.org/publication/PGP000262/>). UKB data are available following a process described at <https://www.ukbiobank.ac.uk/enable-your-research>. The TAGC PRS allele weights and all original code are available on GitHub (https://github.com/mdapas/Asthma_PRS_PheWAS_UKB) and were published on Zenodo (<https://doi.org/10.5281/zenodo.8237286>).

Supplemental information

Supplemental information can be found online at <https://doi.org/10.1016/j.xhgg.2023.100233>.

Acknowledgments

This research was conducted using the UKB resource under application number 44300. This research was supported by NIH grants U19 AI62310 and P30 DK20595. M.D. was supported by NIH TL1

TR002388 and T32 HL007605; N.S. was supported by NIH K08 HL153955.

Author contributions

M.D., C.O., and N.S. conceived and designed the study. M.D. and Y.L.L. performed all analyses. W.W.S. managed all biobank data and performed phenotype derivations. H.K.I. facilitated biobank data access and contributed to the interpretation of results with M.D., C.O., and N.S. M.D. wrote the manuscript. All authors read and approved the manuscript.

Declaration of interests

The authors declare no competing interests.

Received: June 2, 2023

Accepted: August 11, 2023

References

1. GBD Chronic Respiratory Disease Collaborators (2020). Prevalence and attributable health burden of chronic respiratory diseases, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Respir. Med.* 8, 585–596. [https://doi.org/10.1016/S2213-2600\(20\)30105-3](https://doi.org/10.1016/S2213-2600(20)30105-3).
2. Gupta, R.S., Zhang, X., Sharp, L.K., Shannon, J.J., and Weiss, K.B. (2008). Geographic variability in childhood asthma prevalence in Chicago. *J. Allergy Clin. Immunol.* 121, 639–645.e1. <https://doi.org/10.1016/j.jaci.2007.11.036>.
3. Pate, C.A., Zahran, H.S., Qin, X., Johnson, C., Hummelman, E., and Maililay, J. (2021). Asthma Surveillance - United States, 2006-2018. *MMWR. Surveill. Summ.* 70, 1–32. <https://doi.org/10.15585/mmwr.ss7005a1>.
4. Castillo, J.R., Peters, S.P., and Busse, W.W. (2017). Asthma Exacerbations: Pathogenesis, Prevention, and Treatment. *J. Allergy Clin. Immunol. Pract.* 5, 918–927. <https://doi.org/10.1016/j.jaip.2017.05.001>.
5. Polderman, T.J.C., Benyamin, B., de Leeuw, C.A., Sullivan, P.F., van Bochoven, A., Visscher, P.M., and Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat. Genet.* 47, 702–709. <https://doi.org/10.1038/ng.3285>.
6. Han, Y., Jia, Q., Jahani, P.S., Hurrell, B.P., Pan, C., Huang, P., Gukasyan, J., Woodward, N.C., Eskin, E., Gilliland, F.D., et al. (2020). Genome-wide analysis highlights contribution of immune system pathways to the genetic architecture of asthma. *Nat. Commun.* 11, 1776. <https://doi.org/10.1038/s41467-020-15649-3>.
7. Tsuo, K., Zhou, W., Wang, Y., Kanai, M., Namba, S., Gupta, R., Majara, L., Nkambule, L.L., Morisaki, T., Okada, Y., et al. (2022). Multi-ancestry meta-analysis of asthma identifies novel associations and highlights the value of increased power and diversity. *Cell Genom.* 2, 100212. <https://doi.org/10.1016/j.xgen.2022.100212>.
8. Kim, K.W., and Ober, C. (2019). Lessons Learned From GWAS of Asthma. *Allergy Asthma Immunol. Res.* 11, 170–187. <https://doi.org/10.4168/aaair.2019.11.2.170>.
9. Pividori, M., Schoettler, N., Nicolae, D.L., Ober, C., and Im, H.K. (2019). Shared and distinct genetic risk factors for childhood-onset and adult-onset asthma: genome-wide and

- transcriptome-wide studies. *Lancet Respir. Med.* 7, 509–522. [https://doi.org/10.1016/S2213-2600\(19\)30055-4](https://doi.org/10.1016/S2213-2600(19)30055-4).
10. Pividori, M., and Im, H.K. (2019). ukbREST: efficient and streamlined data access for reproducible research in large bio-banks. *Bioinformatics* 35, 1971–1973. <https://doi.org/10.1093/bioinformatics/bty925>.
 11. Cavazos, T.B., and Witte, J.S. (2021). Inclusion of variants discovered from diverse populations improves polygenic risk score transferability. *HGG Adv.* 2, 100017. <https://doi.org/10.1016/j.xhgg.2020.100017>.
 12. Sordillo, J.E., Lutz, S.M., Jorgenson, E., Iribarren, C., McGeachie, M., Dahlin, A., Tantisira, K., Kelly, R., Lasky-Su, J., Sakornsakolpat, P., et al. (2021). A polygenic risk score for asthma in a large racially diverse population. *Clin. Exp. Allergy* 51, 1410–1420. <https://doi.org/10.1111/cea.14007>.
 13. Dijk, F.N., Folkersma, C., Gruziova, O., Kumar, A., Wijga, A.H., Gehring, U., Kull, I., Postma, D.S., Vonk, J.M., Melén, E., and Koppelman, G.H. (2019). Genetic risk scores do not improve asthma prediction in childhood. *J. Allergy Clin. Immunol.* 144, 857–860.e7. <https://doi.org/10.1016/j.jaci.2019.05.017>.
 14. Kothalawala, D.M., Kadalayil, L., Curtin, J.A., Murray, C.S., Simpson, A., Custovic, A., Tapper, W.J., Arshad, S.H., Rezwan, F.I., Holloway, J.W.; and On Behalf Of Stelar/Unicorn Investigators (2022). Integration of Genomic Risk Scores to Improve the Prediction of Childhood Asthma Diagnosis. *J. Personalized Med.* 12, 75. <https://doi.org/10.3390/jpm12010075>.
 15. Belsky, D.W., Sears, M.R., Hancox, R.J., Harrington, H., Houts, R., Moffitt, T.E., Sugden, K., Williams, B., Poulton, R., and Caspi, A. (2013). Polygenic risk and the development and course of asthma: an analysis of data from a four-decade longitudinal study. *Lancet Respir. Med.* 1, 453–461. [https://doi.org/10.1016/S2213-2600\(13\)70101-2](https://doi.org/10.1016/S2213-2600(13)70101-2).
 16. Zhu, Z., Hasegawa, K., Ma, B., Fujiogi, M., Camargo, C.A., Jr., and Liang, L. (2020). Association of asthma and its genetic predisposition with the risk of severe COVID-19. *J. Allergy Clin. Immunol.* 146, 327–329.e4. <https://doi.org/10.1016/j.jaci.2020.06.001>.
 17. Hoang, T.T., Sikdar, S., Xu, C.J., Lee, M.K., Cardwell, J., Forno, E., Imboden, M., Jeong, A., Madore, A.M., Qi, C., et al. (2020). Epigenome-wide association study of DNA methylation and adult asthma in the Agricultural Lung Health Study. *Eur. Respir. J.* 56, 2000217. <https://doi.org/10.1183/13993003.00217-2020>.
 18. Leffa, D.T., Horta, B., Barros, F.C., Menezes, A.M.B., Martins-Silva, T., Hutz, M.H., Bau, C.H.D., Grevet, E.H., Rohde, L.A., and Tovo-Rodrigues, L. (2022). Association between Polygenic Risk Scores for ADHD and Asthma: A Birth Cohort Investigation. *J. Atten. Disord.* 26, 685–695. <https://doi.org/10.1177/10870547211020111>.
 19. Lareau, C.A., DeWeese, C.F., Adrianto, I., Lessard, C.J., Gaffney, P.M., Iannuzzi, M.C., Rybicki, B.A., Levin, A.M., and Montgomery, C.G. (2017). Polygenic risk assessment reveals pleiotropy between sarcoidosis and inflammatory disorders in the context of genetic ancestry. *Gene Immun.* 18, 88–94. <https://doi.org/10.1038/gene.2017.3>.
 20. Hüls, A., Vanker, A., Gray, D., Koen, N., MacIsaac, J.L., Lin, D.T.S., Ramadori, K.E., Sly, P.D., Stein, D.J., Kobor, M.S., and Zar, H.J. (2020). Genetic susceptibility to asthma increases the vulnerability to indoor air pollution. *Eur. Respir. J.* 55, 1901831. <https://doi.org/10.1183/13993003.01831-2019>.
 21. Namjou, B., Lape, M., Malolepsza, E., DeVore, S.B., Weirauch, M.T., Dikilitas, O., Jarvik, G.P., Kiryluk, K., Kullo, I.J., Liu, C., et al. (2022). Multiancestral polygenic risk score for pediatric asthma. *J. Allergy Clin. Immunol.* 150, 1086–1096. <https://doi.org/10.1016/j.jaci.2022.03.035>.
 22. Zhu, Z., Lee, P.H., Chaffin, M.D., Chung, W., Loh, P.R., Lu, Q., Christiani, D.C., and Liang, L. (2018). A genome-wide cross-trait analysis from UK Biobank highlights the shared genetic architecture of asthma and allergic diseases. *Nat. Genet.* 50, 857–864. <https://doi.org/10.1038/s41588-018-0121-0>.
 23. Guo, H., An, J., and Yu, Z. (2020). Identifying Shared Risk Genes for Asthma, Hay Fever, and Eczema by Multi-Trait and Multiomic Association Analyses. *Front. Genet.* 11, 270. <https://doi.org/10.3389/fgene.2020.00270>.
 24. John, C., Guyatt, A.L., Shrine, N., Packer, R., Olafsdottir, T.A., Liu, J., Hayden, L.P., Chu, S.H., Koskela, J.T., Luan, J., et al. (2022). Genetic Associations and Architecture of Asthma-COPD Overlap. *Chest* 161, 1155–1166. <https://doi.org/10.1016/j.chest.2021.12.674>.
 25. Ballard, J.L., and O'Connor, L.J. (2022). Shared components of heritability across genetically correlated traits. *Am. J. Hum. Genet.* 109, 989–1006. <https://doi.org/10.1016/j.ajhg.2022.04.003>.
 26. Lehto, K., Pedersen, N.L., Almqvist, C., Lu, Y., and Brew, B.K. (2019). Asthma and affective traits in adults: a genetically informative study. *Eur. Respir. J.* 53, 1802142. <https://doi.org/10.1183/13993003.02142-2018>.
 27. Zhu, Z., Zhu, X., Liu, C.L., Shi, H., Shen, S., Yang, Y., Hasegawa, K., Camargo, C.A., Jr., and Liang, L. (2019). Shared genetics of asthma and mental health disorders: a large-scale genome-wide cross-trait analysis. *Eur. Respir. J.* 54, 1901507. <https://doi.org/10.1183/13993003.01507-2019>.
 28. Zhu, Z., Guo, Y., Shi, H., Liu, C.L., Panganiban, R.A., Chung, W., O'Connor, L.J., Himes, B.E., Gazal, S., Hasegawa, K., et al. (2020). Shared genetic and experimental links between obesity-related traits and asthma subtypes in UK Biobank. *J. Allergy Clin. Immunol.* 145, 537–549. <https://doi.org/10.1016/j.jaci.2019.09.035>.
 29. Zhou, Y., Liang, Z.S., Jin, Y., Ding, J., Huang, T., Moore, J.H., Zheng, Z.J., and Huang, J. (2021). Shared Genetic Architecture and Causal Relationship Between Asthma and Cardiovascular Diseases: A Large-Scale Cross-Trait Analysis. *Front. Genet.* 12, 775591. <https://doi.org/10.3389/fgene.2021.775591>.
 30. Demenais, F., Margaritte-Jeannin, P., Barnes, K.C., Cookson, W.O.C., Altmüller, J., Ang, W., Barr, R.G., Beaty, T.H., Becker, A.B., Beilby, J., et al. (2018). Multiancestry association study identifies new asthma risk loci that colocalize with immune-cell enhancer marks. *Nat. Genet.* 50, 42–53. <https://doi.org/10.1038/s41588-017-0014-7>.
 31. Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L.T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J., et al. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature* 562, 203–209. <https://doi.org/10.1038/s41586-018-0579-z>.
 32. Wang, Y., Namba, S., Lopera, E., Kerminen, S., Tsuo, K., Läll, K., Kanai, M., Zhou, W., Wu, K.H., Favé, M.J., et al. (2023). Global Biobank analyses provide lessons for developing polygenic risk scores across diverse cohorts. *Cell Genom.* 3, 100241. <https://doi.org/10.1016/j.xgen.2022.100241>.
 33. Zhou, W., Kanai, M., Wu, K.H.H., Rasheed, H., Tsuo, K., Hirbo, J.B., Wang, Y., Bhattacharya, A., Zhao, H., Namba, S., et al. (2022). Global Biobank Meta-analysis Initiative: Powering genetic discovery across human disease. *Cell Genom.* 2, 100192. <https://doi.org/10.1016/j.xgen.2022.100192>.

34. Nagai, A., Hirata, M., Kamatani, Y., Muto, K., Matsuda, K., Kiyohara, Y., Ninomiya, T., Tamakoshi, A., Yamagata, Z., Mushiroda, T., et al. (2017). Overview of the BioBank Japan Project: Study design and profile. *J. Epidemiol.* *27*, S2–S8. <https://doi.org/10.1016/j.je.2016.12.005>.
35. Bowton, E.A., Collier, S.P., Wang, X., Sutcliffe, C.B., Van Driest, S.L., Couch, L.J., Herrera, M., Jerome, R.N., Slebos, R.J.C., Alborn, W.E., et al. (2015). Phenotype-Driven Plasma Biobanking Strategies and Methods. *J. Personalized Med.* *5*, 140–152. <https://doi.org/10.3390/jpm5020140>.
36. Sijtsma, A., Rienks, J., van der Harst, P., Navis, G., Rosmalen, J.G.M., and Dotinga, A. (2022). Cohort Profile Update: Lifelines, a three-generation cohort study and biobank. *Int. J. Epidemiol.* *51*, e295–e302. <https://doi.org/10.1093/ije/dyab257>.
37. Dummer, T.J.B., Awadalla, P., Boileau, C., Craig, C., Fortier, I., Goel, V., Hicks, J.M.T., Jacquemont, S., Knoppers, B.M., Le, N., et al. (2018). The Canadian Partnership for Tomorrow Project: a pan-Canadian platform for research on chronic disease prevention. *CMAJ (Can. Med. Assoc. J.)* *190*, E710–E717. <https://doi.org/10.1503/cmaj.170292>.
38. Leitsalu, L., Haller, T., Esko, T., Tammesoo, M.L., Alavere, H., Snieder, H., Perola, M., Ng, P.C., Mägi, R., Milani, L., et al. (2015). Cohort Profile: Estonian Biobank of the Estonian Genome Center, University of Tartu. *Int. J. Epidemiol.* *44*, 1137–1147. <https://doi.org/10.1093/ije/dyt268>.
39. Kurki, M.I., Karjalainen, J., Palta, P., Sipilä, T.P., Kristiansson, K., Donner, K.M., Reeve, M.P., Laivuori, H., Aavikko, M., Kainisto, M.A., et al. (2023). FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature* *613*, 508–518. <https://doi.org/10.1038/s41586-022-05473-8>.
40. Zawistowski, M., Fritsche, L.G., Pandit, A., Vanderwerff, B., Patil, S., Schmidt, E.M., VandeHaar, P., Willer, C.J., Brummett, C.M., Kheterpal, S., et al. (2023). The Michigan Genomics Initiative: A biobank linking genotypes and electronic clinical records in Michigan Medicine patients. *Cell Genom.* *3*, 100257. <https://doi.org/10.1016/j.xgen.2023.100257>.
41. Åsvold, B.O., Langhammer, A., Rehn, T.A., Kjelvik, G., Grøntvedt, T.V., Sørgerd, E.P., Fenstad, J.S., Heggland, J., Holmen, O., Stuijbergen, M.C., et al. (2023). Cohort Profile Update: The HUNT Study, Norway. *Int. J. Epidemiol.* *52*, e80–e91. <https://doi.org/10.1093/ije/dyac095>.
42. Lambert, S.A., Gil, L., Jupp, S., Ritchie, S.C., Xu, Y., Buniello, A., McMahon, A., Abraham, G., Chapman, M., Parkinson, H., et al. (2021). The Polygenic Score Catalog as an open database for reproducibility and systematic evaluation. *Nat. Genet.* *53*, 420–425. <https://doi.org/10.1038/s41588-021-00783-5>.
43. Ge, T., Chen, C.Y., Ni, Y., Feng, Y.C.A., and Smoller, J.W. (2019). Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat. Commun.* *10*, 1776. <https://doi.org/10.1038/s41467-019-09718-5>.
44. 1000 Genomes Project Consortium, Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., and Abecasis, G.R. (2015). A global reference for human genetic variation. *Nature* *526*, 68–74. <https://doi.org/10.1038/nature15393>.
45. Berisa, T., and Pickrell, J.K. (2016). Approximately independent linkage disequilibrium blocks in human populations. *Bioinformatics* *32*, 283–285. <https://doi.org/10.1093/bioinformatics/btv546>.
46. Yengo, L., Vedantam, S., Marouli, E., Sidorenko, J., Bartell, E., Sakaue, S., Graff, M., Eliassen, A.U., Jiang, Y., Raghavan, S., et al. (2022). A saturated map of common genetic variants associated with human height. *Nature* *610*, 704–712. <https://doi.org/10.1038/s41586-022-05275-y>.
47. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* *4*, 7. <https://doi.org/10.1186/s13742-015-0047-8>.
48. Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.C., and Müller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinf.* *12*, 77. <https://doi.org/10.1186/1471-2105-12-77>.
49. Lee, S.H., Goddard, M.E., Wray, N.R., and Visscher, P.M. (2012). A better coefficient of determination for genetic profile analysis. *Genet. Epidemiol.* *36*, 214–224. <https://doi.org/10.1002/gepi.21614>.
50. Finucane, H.K., Bulik-Sullivan, B., Gusev, A., Trynka, G., Reshef, Y., Loh, P.R., Anttila, V., Xu, H., Zang, C., Farh, K., et al. (2015). Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* *47*, 1228–1235. <https://doi.org/10.1038/ng.3404>.
51. Monsees, G.M., Tamimi, R.M., and Kraft, P. (2009). Genome-Wide Association Scans for Secondary Traits Using Case-Control Samples. *Genet. Epidemiol.* *33*, 717–728. <https://doi.org/10.1002/gepi.20424>.
52. MacKinnon, D.P. (2008). *Introduction to Statistical Mediation Analysis* (Lawrence Erlbaum Associates).
53. MacKinnon, D.P., Lockwood, C.M., Brown, C.H., Wang, W., and Hoffman, J.M. (2007). The intermediate endpoint effect in logistic and probit regression. *Clin. Trials* *4*, 499–513. <https://doi.org/10.1177/1740774507083434>.
54. Trivedi, M., and Denton, E. (2019). Asthma in Children and Adults-What Are the Differences and What Can They Tell us About Asthma? *Front. Pediatr.* *7*, 256. <https://doi.org/10.3389/fped.2019.00256>.
55. Naeem, A., and Silveyra, P. (2019). Sex Differences in Paediatric and Adult Asthma. *Eur. Med. (Edicion Espanola) J.* *4*, 27–35.
56. Ferreira, M.A.R., Mathur, R., Vonk, J.M., Szwajda, A., Brumpton, B., Granell, R., Brew, B.K., Ullemer, V., Lu, Y., Jiang, Y., et al. (2019). Genetic Architectures of Childhood- and Adult-Onset Asthma Are Partly Distinct. *Am. J. Hum. Genet.* *104*, 665–684. <https://doi.org/10.1016/j.ajhg.2019.02.022>.
57. Martin, A.R., Gignoux, C.R., Walters, R.K., Wojcik, G.L., Neale, B.M., Gravel, S., Daly, M.J., Bustamante, C.D., and Kenny, E.E. (2017). Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *Am. J. Hum. Genet.* *100*, 635–649. <https://doi.org/10.1016/j.ajhg.2017.03.004>.
58. Carapito, R., Radosavljevic, M., and Bahram, S. (2016). Next-Generation Sequencing of the HLA locus: Methods and impacts on HLA typing, population genetics and disease association studies. *Hum. Immunol.* *77*, 1016–1023. <https://doi.org/10.1016/j.humimm.2016.04.002>.
59. Kreiner, E., Waage, J., Standl, M., Brix, S., Pers, T.H., Couto Alves, A., Warrington, N.M., Tiesler, C.M.T., Fuertes, E., Franke, L., et al. (2017). Shared genetic variants suggest common pathways in allergy and autoimmune diseases. *J. Allergy Clin. Immunol.* *140*, 771–781. <https://doi.org/10.1016/j.jaci.2016.10.055>.
60. Li, X., Ampleford, E.J., Howard, T.D., Moore, W.C., Torgerson, D.G., Li, H., Busse, W.W., Castro, M., Erzurum, S.C., Israel, E., et al. (2012). Genome-wide association studies of asthma indicate opposite immunopathogenesis direction

- from autoimmune diseases. *J. Allergy Clin. Immunol.* *130*, 861–868.e7. <https://doi.org/10.1016/j.jaci.2012.04.041>.
61. Dong, G., Feng, J., Sun, F., Chen, J., and Zhao, X.M. (2021). A global overview of genetically interpretable multimorbidities among common diseases in the UK Biobank. *Genome Med.* *13*, 110. <https://doi.org/10.1186/s13073-021-00927-6>.
 62. Sood, A., Qualls, C., Schuyler, M., Arynchyn, A., Alvarado, J.H., Smith, L.J., and Jacobs, D.R., Jr. (2013). Adult-onset asthma becomes the dominant phenotype among women by age 40 years. the longitudinal CARDIA study. *Ann. Am. Thorac. Soc.* *10*, 188–197. <https://doi.org/10.1513/Annals-ATS.201212-115OC>.
 63. Kaur, R., and Chupp, G. (2019). Phenotypes and endotypes of adult asthma: Moving toward precision medicine. *J. Allergy Clin. Immunol.* *144*, 1–12. <https://doi.org/10.1016/j.jaci.2019.05.031>.
 64. Kero, J., Gissler, M., Hemminki, E., and Isolauri, E. (2001). Could TH1 and TH2 diseases coexist? Evaluation of asthma incidence in children with coeliac disease, type 1 diabetes, or rheumatoid arthritis: a register study. *J. Allergy Clin. Immunol.* *108*, 781–783. <https://doi.org/10.1067/mai.2001.119557>.
 65. Hemminki, K., Li, X., Sundquist, J., and Sundquist, K. (2010). Subsequent autoimmune or related disease in asthma patients: clustering of diseases or medical care? *Ann. Epidemiol.* *20*, 217–222. <https://doi.org/10.1016/j.annepidem.2009.11.007>.
 66. Canova, C., Pitter, G., Ludvigsson, J.F., Romor, P., Zanier, L., Zanotti, R., and Simonato, L. (2015). Coeliac disease and asthma association in children: the role of antibiotic consumption. *Eur. Respir. J.* *46*, 115–122. <https://doi.org/10.1183/09031936.00185714>.
 67. Assa, A., Frenkel-Nir, Y., Tzur, D., Katz, L.H., and Shamir, R. (2017). Large population study shows that adolescents with celiac disease have an increased risk of multiple autoimmune and nonautoimmune comorbidities. *Acta Paediatr.* *106*, 967–972. <https://doi.org/10.1111/apa.13808>.
 68. Patel, B., Wi, C.I., Hasassri, M.E., Divekar, R., Absah, I., Almalouhi, E., Ryu, E., King, K., and Juhn, Y.J. (2018). Heterogeneity of asthma and the risk of celiac disease in children. *Allergy Asthma Proc.* *39*, 51–58. <https://doi.org/10.2500/aap.2018.39.4100>.
 69. Hobbs, B.D., de Jong, K., Lamontagne, M., Bossé, Y., Shrine, N., Artigas, M.S., Wain, L.V., Hall, I.P., Jackson, V.E., Wyss, A.B., et al. (2017). Genetic loci associated with chronic obstructive pulmonary disease overlap with loci for lung function and pulmonary fibrosis. *Nat. Genet.* *49*, 426–432. <https://doi.org/10.1038/ng.3752>.
 70. Chang, D., Hunkapiller, J., Bhangale, T., Reeder, J., Mukhyala, K., Tom, J., Cowgill, A., Vogel, J., Forrest, W.F., Khan, Z., et al. (2022). A whole genome sequencing study of moderate to severe asthma identifies a lung function locus associated with asthma risk. *Sci. Rep.* *12*, 5574. <https://doi.org/10.1038/s41598-022-09447-8>.
 71. Sánchez-Solís, M. (2019). Early Lung Function and Future Asthma. *Front. Pediatr.* *7*, 253. <https://doi.org/10.3389/fped.2019.00253>.
 72. Ferreira, M.A., Vonk, J.M., Baurecht, H., Marenholz, I., Tian, C., Hoffman, J.D., Helmer, Q., Tillander, A., Ullemer, V., van Dongen, J., et al. (2017). Shared genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. *Nat. Genet.* *49*, 1752–1757. <https://doi.org/10.1038/ng.3985>.
 73. Busse, W.W., and Sedgwick, J.B. (1992). Eosinophils in asthma. *Ann. Allergy* *68*, 286–290.
 74. Peters, U., Dixon, A.E., and Forno, E. (2018). Obesity and asthma. *J. Allergy Clin. Immunol.* *141*, 1169–1179. <https://doi.org/10.1016/j.jaci.2018.02.004>.
 75. Granel, R., Henderson, A.J., Evans, D.M., Smith, G.D., Ness, A.R., Lewis, S., Palmer, T.M., and Sterne, J.A.C. (2014). Effects of BMI, fat mass, and lean mass on asthma in childhood: a Mendelian randomization study. *PLoS Med.* *11*, e1001669. <https://doi.org/10.1371/journal.pmed.1001669>.
 76. Skaaby, T., Taylor, A.E., Thuesen, B.H., Jacobsen, R.K., Friedrich, N., Møllehave, L.T., Hansen, S., Larsen, S.C., Völker, U., Nauck, M., et al. (2018). Estimating the causal effect of body mass index on hay fever, asthma and lung function using Mendelian randomization. *Allergy* *73*, 153–164. <https://doi.org/10.1111/all.13242>.
 77. Chen, Z., Salam, M.T., Alderete, T.L., Habre, R., Bastain, T.M., Berhane, K., and Gilliland, F.D. (2017). Effects of Childhood Asthma on the Development of Obesity among School-aged Children. *Am. J. Respir. Crit. Care Med.* *195*, 1181–1188. <https://doi.org/10.1164/rccm.201608-1691OC>.
 78. Contreras, Z.A., Chen, Z., Roumeliotaki, T., Annesi-Maesano, I., Baiz, N., von Berg, A., Bergström, A., Crozier, S., Duijts, L., Ekström, S., et al. (2018). Does early onset asthma increase childhood obesity risk? A pooled analysis of 16 European cohorts. *Eur. Respir. J.* *52*, 1800504. <https://doi.org/10.1183/13993003.00504-2018>.
 79. Hasler, G., Gergen, P.J., Ajdacic, V., Gamma, A., Eich, D., Rössler, W., and Angst, J. (2006). Asthma and body weight change: a 20-year prospective community study of young adults. *Int. J. Obes.* *30*, 1111–1118. <https://doi.org/10.1038/sj.ijo.0803215>.
 80. Moitra, S., Carsin, A.E., Abramson, M.J., Accordini, S., Amaral, A.F.S., Anto, J., Bono, R., Casas Ruiz, L., Cerveri, I., Chatzi, L., et al. (2023). Long-term effect of asthma on the development of obesity among adults: an international cohort study. *Thorax* *78*, 128–135. <https://doi.org/10.1136/thoraxjnl-2021-217867>.
 81. Cardet, J.C., Ash, S., Kusa, T., Camargo, C.A., Jr., and Israel, E. (2016). Insulin resistance modifies the association between obesity and current asthma in adults. *Eur. Respir. J.* *48*, 403–410. <https://doi.org/10.1183/13993003.00246-2016>.
 82. de Boer, G.M., Tramper-Stranders, G.A., Houweling, L., van Zelst, C.M., Pouw, N., Verhoeven, G.T., Boxma-de Klerk, B.M., Braunstahl, G.J., In 't Veen, J.C.C.M., van Rossum, E.F.C., and Hendriks, R.W. (2021). Adult but not childhood onset asthma is associated with the metabolic syndrome, independent from body mass index. *Respir. Med.* *188*, 106603. <https://doi.org/10.1016/j.rmed.2021.106603>.
 83. Turi, K.N., McKennan, C., Gebretsadik, T., Snyder, B., Seroogy, C.M., Lemanske, R.F., Jr., Zoratti, E., Havstad, S., Ober, C., Lynch, S., et al. (2021). Unconjugated bilirubin is associated with protection from early-life wheeze and childhood asthma. *J. Allergy Clin. Immunol.* *148*, 128–138. <https://doi.org/10.1016/j.jaci.2020.12.639>.
 84. Blanc, J., and Berg, J.J. (2023). Testing for differences in polygenic scores in the presence of confounding. Preprint at bioRxiv. <https://doi.org/10.1101/2023.03.12.532301>.
 85. Weissbrod, O., Kanai, M., Shi, H., Gazal, S., Peyrot, W.J., Khara, A.V., Okada, Y., Biobank Japan Project, Martin, A.R., Finucane, H.K., and Price, A.L. (2022). Leveraging fine-mapping and multipopulation training data to improve cross-population polygenic risk scores. *Nat. Genet.* *54*, 450–458. <https://doi.org/10.1038/s41588-022-01036-9>.

86. Martin, A.R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B.M., and Daly, M.J. (2019). Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat. Genet.* *51*, 584–591. <https://doi.org/10.1038/s41588-019-0379-x>.
87. Border, R., Athanasiadis, G., Buil, A., Schork, A.J., Cai, N., Young, A.I., Werge, T., Flint, J., Kendler, K.S., Sankararaman, S., et al. (2022). Cross-trait assortative mating is widespread and inflates genetic correlation estimates. *Science* *378*, 754–761. <https://doi.org/10.1126/science.abo2059>.
88. Border, R., O'Rourke, S., de Candia, T., Goddard, M.E., Visscher, P.M., Yengo, L., Jones, M., and Keller, M.C. (2022). Assortative mating biases marker-based heritability estimators. *Nat. Commun.* *13*, 660. <https://doi.org/10.1038/s41467-022-28294-9>.
89. Joo, J., Hobbs, B.D., Cho, M.H., and Himes, B.E. (2020). Trait Insights Gained by Comparing Genome-Wide Association Study Results using Different Chronic Obstructive Pulmonary Disease Definitions. *AMIA Jt. Summits Transl. Sci. Proc.* *2020*, 278–287.
90. Prieto-Centurion, V., Rolle, A.J., Au, D.H., Carson, S.S., Henderson, A.G., Lee, T.A., Lindenauer, P.K., McBurnie, M.A., Mularski, R.A., Naureckas, E.T., et al. (2014). Multicenter study comparing case definitions used to identify patients with chronic obstructive pulmonary disease. *Am. J. Respir. Crit. Care Med.* *190*, 989–995. <https://doi.org/10.1164/rccm.201406-1166OC>.