Article [Working Title]: A multi-ancestry genome-wide association study identifies 1 novel candidate loci in the RARB gene associated with hypertensive disorders of pregnancy

Short Title: Multi-ancestry GWAS of hypertension in pregnancy

Jasmine A. Mack1,2, Adam Burkholder1, Farida S. Akhtari1, John S. House1, Ulla Sovio2, Gordon C.S. Smith2, Charles P. Schmitt1, David C. Fargo1, Janet E. Hall1, Alison A. Motsinger-Reif1

1 National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709, USA
2 University of Cambridge, Department of Obstetrics and Gynaecology, Cambridge, UK, CB2 0SW

Corresponding author: Alison Motsinger-Reif
alison.motsinger-reif@nih.gov

Word count: 7917
Abstract

Background

Genetic factors related to pregnancy-related traits are understudied, especially among ancestrally diverse cohorts. This study assessed maternal contributions to hypertensive disorders of pregnancy (HDP) in multi-ancestry cohorts.

Methods

We performed a genome-wide association study of HDP using data from the Personalized Environment and Genes Study (PEGS) cohort (USA) with validation in the UK Biobank (UKBB). We performed gene-level and gene-set analyses and tested the association of polygenic scores (PGS) for systolic blood pressure (SBP), preeclampsia (PE), and gestational hypertension (GH).

Results

We identified two novel maternal genome-wide significant associations with HDP. The lead independent variants were rs114954125 on chromosome 2 (near LRP1B; OR (95% CI): 3.03 (2.05, 4.49); P=3.19 x 10^{-8}) and rs61176331 on chromosome 3 (near RARB; OR (95% CI): 3.09 (2.11, 4.53); P=7.97 x 10^{-9}). We validated rs61176331 in the UKBB (P=3.73 x 10^{-2}). When aggregating SNPs by genes, RARB (P=1.36 x 10^{-5}) and RN7SL283P (P=2.56 x 10^{-2}) were associated with HDP. Inflammatory and immunological biological pathways were most strongly related to HDP-associated genes. While all blood pressure and HDP-related PGS were significantly associated with HDP in PEGS, the SBP PGS was a stronger predictor of HDP (area under the curve (AUC): 0.57; R^2=0.7%) compared to the PE PGS (AUC: 0.53; R^2=0.2%).

Conclusion

Our study is the first to identify and validate maternal genetic variants near RARB associated with HDP. The findings demonstrate the power of multi-ancestry studies for genetic discovery and highlight the relationship between immune response and HDP and the utility of PGS for risk prediction.

ClinicalTrials.gov Identifier for PEGS: NCT00341237
### Non-standard Abbreviations and Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFR</td>
<td>African ancestry</td>
</tr>
<tr>
<td>AMR</td>
<td>Admixed American ancestry</td>
</tr>
<tr>
<td>EAF</td>
<td>Effect allele frequency</td>
</tr>
<tr>
<td>EAS</td>
<td>East Asian ancestry</td>
</tr>
<tr>
<td>EUR</td>
<td>European ancestry</td>
</tr>
<tr>
<td>GH</td>
<td>Gestational hypertension</td>
</tr>
<tr>
<td>GWAS</td>
<td>Genome-wide association study</td>
</tr>
<tr>
<td>HDP</td>
<td>Hypertensive disorders of pregnancy</td>
</tr>
<tr>
<td>KEGG</td>
<td>Kyoto Encyclopedia of Genes and Genomes</td>
</tr>
<tr>
<td>MAC</td>
<td>Minor allele count</td>
</tr>
<tr>
<td>MIM</td>
<td>Mendelian Inheritance in Man</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal components analysis</td>
</tr>
<tr>
<td>PE</td>
<td>Preeclampsia</td>
</tr>
<tr>
<td>PEGS</td>
<td>Personalized Environment and Genes Study</td>
</tr>
<tr>
<td>PGS</td>
<td>Polygenic score</td>
</tr>
<tr>
<td>SAS</td>
<td>South Asian ancestry</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>SNV</td>
<td>Single nucleotide variant</td>
</tr>
<tr>
<td>UKBB</td>
<td>UK Biobank</td>
</tr>
</tbody>
</table>
Introduction

Hypertensive disorders of pregnancy (HDP) are a heterogeneous set of adverse pregnancy outcomes that occur in the perinatal period comprising gestational hypertension (GH), preeclampsia/eclampsia (PE), chronic, pre-existing hypertension with or without superimposed PE, and postpartum PE. HDP occur in approximately 5-10% of pregnancies worldwide although prevalence varies by disease subtype and region. HDP incidence increased by 11% from 1990 to 2019. Globally, an estimated 50,000-75,000 deaths occur annually due to PE. HDP are the primary cause of maternal morbidity and mortality in high-resource countries and a secondary cause globally.

There is racial and ethnic disparity regarding HDP prevalence and incidence as well as the contributing risk factors. Black, Hispanic, and Indigenous individuals experience disproportionately high rates of HDP compared to White individuals. Risk factors for HDP include pre-pregnancy body mass index (BMI) > 25 kg/m², nulliparity, prior GH, maternal age of more than 35 years, family history of HDP, and pre-existing hypertension.

The pathophysiology of HDP is not fully understood, but dysfunction in placental growth factors and regulators of angiogenesis are suggested mechanisms, with maternal and fetal contributions. This mechanism is hypothesized to be associated with diminished placental perfusion causing systemic vascular endothelial dysfunction through the
release of antiangiogenic factors from the placenta\textsuperscript{11,12}. Placental hypoperfusion is thought to be caused by failure of transformation of uterine spiral arteries into low-resistance conduits, which is a normal feature of invasion by extravillous cytotrophoblast\textsuperscript{13,14}. Failure of spiral artery transformation has been linked to fetal growth restriction, miscarriage, and PE\textsuperscript{14}. Moreover, placental factors can interact with maternal genetics, epigenetics, and environmental factors to influence the response to placental dysfunction\textsuperscript{15}. Maternal immune system changes are involved in the development and pathogenesis of PE\textsuperscript{16–22}, but the mechanisms are incompletely understood.

Early genetic studies of HDP have been predominantly focused on individuals of European ancestry\textsuperscript{23,24}. In research in large cohorts with genetic data, participants from underrepresented racial/ethnic groups are consistently excluded from published analyses due to small sample sizes and the use of sparse methods to account for population structure\textsuperscript{25–27}.

Prior genome-wide association studies (GWAS) of PE have revealed strong signals near the FLT1 gene (MIM:165070) based on a meta-analysis of fetal genotypes from five cohorts, of which all participants were of European ancestry: Genetics of Pre-Eclampsia (GOPEC), Avon Longitudinal Study of Parents and Children (ALSPAC), deCODE, Norwegian Mother and Child (MoBa), and Finnish Genetics of Pre-eclampsia Consortium (FINNPEC)\textsuperscript{24}. The largest maternal genome-wide multi-ancestry meta-analysis of PE and GH was conducted in cohorts comprising FinnGen, Estonian
Biobank, Genes and Health, Michigan Genomics Initiative, Mass General Brigham Biobank, Biobank Japan, BioMe Biobank, InterPregGen Consortium, Trøndelag Health Study (HUNT), UKBB, Penn Medicine Biobank, and Nulliparous Pregnancy Outcomes Study: Monitoring Mothers-to-Be (nuMoM2b)\(^28\). Among 31,091 HDP cases assessed in the meta-analysis, 83% were of European ancestry. Multi-ancestry meta-analyses have been conducted in diverse cohorts such as the Pan UK Biobank (UKBB), but the control groups of the GWAS also included female UKBB participants who have never been pregnant\(^29\).

To increase diversity and support the representation of non-European populations in genomics research, GWAS that include minority racial and ethnic groups can be conducted in currently available data from diverse cohorts. The results can be stratified by ancestry, and meta-analyses can jointly model effects in a multi-ancestry analysis. Building upon previous multi-ancestry genetic studies, we conducted multi-ancestry GWAS in whole-genome sequencing (WGS) data from a discovery cohort from the North Carolina-based Personalized Environment and Genes Study (PEGS) using a case-control study design. The joint modeling exhibited improved power for genetic discovery, enabling the first discovery of a maternal genetic association near the retinoic acid receptor (\(RARB\)) gene (MIM: 180220) with HDP. We validated this association in the UKBB and a meta-analysis in PEGS and the UKBB. In downstream analyses, we performed gene-level analysis and gene-set analyses to elucidate the biological mechanisms involved in HDP. Further, in the PEGS dataset, we assessed the predictive
ability of existing polygenic scores (PGS) for PE and GH, with HDP in predicting risk of these conditions.

Methods
The data used in this study are available to researchers through application to dbGAP the UK Biobank (https://www.ukbiobank.ac.uk/register-apply/). Reference panel data from the 1000 Genomes Project and polygenic scores from the PGS Catalog are publicly available. The Major Resources Table provides details. The summary statistics that support the study findings are available from the corresponding author upon reasonable request.

Study Populations
PEGS Cohort
For the discovery cohort, we selected participants from PEGS, a long-term, prospective cohort based in North Carolina30. PEGS collects extensive clinical, genetic, environmental, and self-reported health information to investigate disease etiology and risk factors across phenotypes in a diverse group of 19,685 individuals.

Participants are administered the Health and Exposure Survey, External Exposome Survey, and Internal Exposome Survey that request information on sociodemographic factors, individual and family medical history, lifestyle factors, medication use, diet, and physical environmental exposures from childhood to adulthood. While PEGS exclusion criteria include women who self-report being pregnant, a small number of participants
were pregnant when they completed the Internal Exposome Study, as this was administered later in the data collection process.

For a subset of 4,737 PEGS participants who provided blood and urine samples, WGS data were processed, with the optional return of results for select health conditions. The Broad Institute sequenced all samples on the NovaSeq 6000 platform with a target genome-wide read depth of 30x (24 samples per flow cell).

For downstream analyses, we removed related individuals in the PEGS genomic cohort by analyzing the genotypes of all individuals using KING 2.2.5\textsuperscript{31} to generate pairwise kinship estimation. We excluded second-degree or closer relatives using a kinship threshold of 0.0884. We estimated genome-wide local ancestry on the autosomes and X chromosome using RFMix version 2, a discriminative modeling approach\textsuperscript{32}. We classified genomic positions by five genetic similarity groups: African (AFR), Admixed American (AMR), East Asian (EAS), European (EUR), and South Asian (SAS). We classified individuals into one of the genetic similarity groups if any genome-wide local ancestry proportion was greater than 0.50. Otherwise, we classified participants as “None”.

The PEGS cohort is a registered clinical trial, and all participants provided written informed consent (ClinicalTrials.gov Identifier: NCT00341237)\textsuperscript{30}. We obtained the data used for the analyses from PEGS Data Freeze version 2.0 created in August 2021. The
Supplemental Methods section provides details on genome sequencing, quality control, and local ancestry inference.

UK Biobank

For the validation cohort, we selected participants from the UK Biobank (UKBB). UKBB is a large, deeply phenotyped genomic cohort comprising 488,366 participants aged 40-69 years at recruitment from 2006 to 2010. About 6% of participants (>30,000) identified as having a non-White ethnic background. Genetic data are linked to electronic health records and survey data, and UKBB has information on diagnosis history, brain imaging, medication use, lifestyle factors, deprivation indices, environmental exposure data, and sociodemographic characteristics.

Ethical approval for UKBB was obtained from the North West Centre for Research Committee (11/NW/0382). Informed consent was obtained for all participants. The following work described was approved by the UKBB team under Application No. 57849. We obtained predominant ancestry assignments from UKBB Returned Dataset 2442. The Supplemental Methods provides details on genotyping and genetic similarity groups in UKBB.

Case-Control Definition

Figure S1 outlines the steps taken to define cases and controls in the PEGS cohort. Because the data available for PEGS and UKBB are not standardized, we defined cases and controls separately for each cohort and harmonized the case-control sets as
much as possible. The **Supplemental Methods** section provides details on case-control definition.

**PEGS**

We defined cases and controls for 4,933 PEGS participants who self-reported at least one pregnancy. After filtering, there were 202 GH cases, which is the primary outcome, 92 PE cases, 78 GH only cases, 223 GH/PE cases, and 1,569 controls. **Table S1** lists the phenotypes and survey questions used to define cases and controls.

**UKBB**

We defined cases and controls in the UKBB population using the International Classification of Diseases (ICD) codes (ICD10) for pregnancy-induced hypertension with and without significant proteinuria (PE and hemolysis, elevated liver enzymes and low platelets (HELLP) syndrome), eclampsia, and unspecified maternal hypertension (**Table S1**). After filtering, there were 1,319 HDP cases and 2,050 controls.

**Statistical Analysis**

For each cohort, we calculated frequencies and proportions for categorical variables and mean/standard deviation for normally distributed continuous variables, including covariates and sociodemographic characteristics. For continuous variables without normal distribution, we report median and interquartile ranges. We estimated $P$ values by Fisher’s exact test to assess associations among covariates and performed t-tests.
for continuous variables. For variables without normal distribution, such as BMI, we
used the Mann-Whitney U test.

**Genome-wide Association Study for Common Variants**

We conducted GWAS (MAF $\geq 0.03$) using multivariable logistic regression to model the
odds of HDP, assuming an additive allelic genetic model. We adjusted for the top 10
genetic principal components, self-reported race/ethnicity, age and BMI at enrollment,
and gravidity. We used PLINK2.0$^{35}$ for all GWAS with Firth-corrected logistic regression
modeling to account for case-control imbalance$^{36}$. Because information on gravidity was
available for few UKBB participants, we excluded this covariate in the UKBB analyses.

We included genotyping batch as a covariate in the UKBB analyses given that two types
of arrays were used in genotyping with 106 batches$^{25}$. We used a threshold of $P <$
$5 \times 10^{-8}$ for genome-wide significance and $P < 1 \times 10^{-6}$ for suggestive significance.

In addition to the main outcome of HDP, we performed GWAS for three other
phenotypes in the PEGS discovery cohort (PE, GH without PE, and GH and/or PE) for
all individuals and separately for the AFR and EUR similarity groups. We adjusted for
within-group variation using the top 10 principal components, age and BMI at
enrollment, and gravidity. The AMR, EAS, and SAS similarity groups had insufficient
numbers of participants to perform stratified analyses (**Table 1**).

To validate our findings, we filtered GWAS summary statistics from the PEGS multi-
ancestry analysis to include only variants with suggestive association ($P < 1 \times 10^{-6}$).
We assessed these variants in the UKBB using harmonized cases and controls. We used GCTA-COJO for conditional and joint analyses to identify independently associated loci\textsuperscript{37}. We used METAL to perform inverse variance-weighted meta-analyses for variants with suggestive association in PEGS and UKBB\textsuperscript{38}.

**Gene-based Association Analysis and Gene-set Analysis**

To increase statistical power for discovery, we used MAGMA to perform gene-set association analysis, adjusting for the top 10 genetic principal components, self-reported race/ethnicity, age and BMI at enrollment, and gravidity. We used Ensembl annotation on GRCh38 to map the SNVs from the GWAS to 31,427 genes\textsuperscript{39}. We utilized the results from the gene-level analysis to perform gene-set analysis in MAGMA\textsuperscript{40,41}. We used the Kyoto Encyclopedia of Genes and Genomes (KEGG) to define gene sets by biological pathways and conditions\textsuperscript{42,43}. To control the false discovery rate (FDR), we used the Benjamini-Hochberg method to adjust $P$ values from the statistical analyses for multiple comparisons. We define statistical significance by an adjusted $P$ value of 0.05 and suggestive FDR-corrected significance by an adjusted $P$ value of 0.10.

**Application of Polygenic Scores**

To assess the predictive value of SBP, PE, and GH for HDP in PEGS, we applied an SBP PGS derived from the UKBB female sub-cohort in Kauko et al.\textsuperscript{44} and separate scores for SBP, PE, and GH from meta-analyses in Honigberg et al.\textsuperscript{28}. We separately computed the four PGS in the PEGS pregnancy ($n = 1,771$) and female sub-cohorts ($n$...
to model HDP and essential hypertension, respectively. We standardized each PGS. The Supplemental Methods section provides details.

We performed logistic regression with Firth correction to assess the association between HDP risk and the standardized PGS, adjusting for the top ten genetic principal components, self-reported race/ethnicity, age and BMI at enrollment, and gravidity. For further validation, we tested the association of each PGS with essential hypertension in the PEGS female sub-cohort, adjusting for the same covariates except gravidity. Additionally, we performed stratified analyses for the AFR and EUR genetic similarity groups. For each PGS, we report odds ratios per standard deviation with a 95% CI. We calculated the AUC for each model using the pROC package in R to assess discriminative ability. All PGS analyses were performed in R 4.3.1 with a significance threshold of 0.05.

The Major Resources Table in the Supplemental Materials provides details on data and code availability.

**Results**

In the PEGS cohort, participants with HDP (11.4% of the study sample) were younger, had a higher BMI at enrollment, and had more pregnancies compared to participants in the control group ($P<0.001$; Table 1). Among HDP cases, 4.5% experienced recurrent diagnoses of PE, and 18.8% experienced GH during two or more pregnancies.

Race/ethnicity distribution, education level, and income were similar for the case and
control groups (Table 1). In the UKBB cohort, race/ethnicity, predominant genetic similarity group, and education level were similar for cases and controls (Table S2). Similar to PEGS, UKBB participants with HDP were younger ($P=0.024$) and had a higher BMI at recruitment ($P<0.001$) compared to controls.

**Novel Maternal Loci Identified on RARB**

In the PEGS GWAS, 11 variants in two loci met genome-wide significance (Figure 1A; Table S3) for HDP. An association was found for a significant novel candidate region in the maternal genome (202 cases and 1,569 controls) consisting of intergenic variants near the low-density lipoprotein (LDL) receptor-related protein 1B (LRP1B; MIM:608766) gene on chromosome 2 (Figure 1B). A stronger signal was observed for a region identified as intronic to the RARB gene (MIM:180220) on chromosome 3 (Figure 1C). The lead variant was rs61176331, with an OR of 3.09 for HDP (EAF: 0.048; 95% confidence interval (95% CI): (2.11, 4.53); $P=7.97 \times 10^{-9}$). The second most significant variant was rs114954125, with an OR of 3.03 (EAF: 0.051; 95% CI: (2.05, 4.49); $P=3.19 \times 10^{-8}$) (Table 2). A genomic inflation factor of 1.02 indicated sufficient adjustment for population stratification (Figure S2A). Conditional and joint analyses confirmed that rs61176331 ($P_{\text{joint}}=9.56 \times 10^{-9}$) and rs114954125 ($P_{\text{joint}}=3.79 \times 10^{-8}$) are independent loci (Table 2). In results stratified by genetic similarity group, these loci had larger effect sizes although the $P$ values were attenuated. rs61176331 had an OR of 6.98 and 2.37 for the AFR and EUR genetic similarity groups, respectively (Table 2; Figure S2B). rs114954125 had a similar effect size as the multi-ancestry GWAS for the EUR genetic similarity group, but this variant did not reach the quality control threshold
in the AFR genetic similarity group. We validated rs61176331 in the UKBB analysis (OR: 1.29; 95% CI: (1.02, 1.65); \( P=3.73 \times 10^{-2} \)) (Figure S2B). In the meta-analysis including PEGS and UKBB, strong associations remained for the candidate region near \( RARB \) (Table 2; Table S3).

In a sensitivity analysis with three HDP subtypes as the outcome (PE, GH only, and GH and/or PE; Figure S3), there was distinct attenuation of signals in the multi-ancestry HDP GWAS with PE as the outcome (21 cases) (Figure S3A). In this GWAS, the candidate signal on \( RARB \) was attenuated below the genome-wide significance threshold while rs114954125 on chromosome 2 remained significant. (Figure S3E).

**Significant Gene-wide Associations for \( RARB \)**

**Figure 2** is a gene-based Manhattan plot. Two gene signals met the FDR-corrected significance threshold—\( RARB \) \( (P=1.36 \times 10^{-3}) \) and \( RN7SL283P \) \( (P=2.56 \times 10^{-3}) \). Among the top 15 most significant genes, there were strong peaks on chromosome 6 (\( HLA-DRB5, BTNL2, RNU1-61P \)) and chromosome 17 (\( LRRC59, EME1, ACSF2; Table S4 \)) although these associations did not reach gene-wide significance.

**Immune-related Pathways Linked to HDP**

Among the ten most significant KEGG gene sets associated with HDP were the pentose phosphate pathway and the hedgehog signaling pathway (Figure 3). Gene-set analyses revealed additional immune-related diseases and processes related to HDP-associated genes (Table S5), but these sets, which included autoimmune thyroid disease, graft-
versus-host disease, adipocytokine signaling pathway, and allograft rejection
(uncorrected P<0.05), did not meet the FDR-corrected significance threshold.

Polygenic Score for Systolic Blood Pressure is Associated with HDP

We tested the association of HDP with PGS for SBP, PE, GH with HDP, and essential hypertension in the PEGS female sub-cohort. Table S6 shows the number of variants missing in the PEGS cohort for each PGS. In the overall model, all four PGS were significantly associated with HDP (P<0.05). The AUC (0.57; 95% CI: (0.52-0.61); Figure 4; Figure S5) and pseudo-$R^2$ (0.7%; Table S7) were the highest for the SBP PGS [1] based on the UKBB. One standard deviation increase in the SBP PGS [1] was associated with a 2.09-fold increase in HDP odds (95% CI: 1.40, 3.13). GH-PGS had the highest OR for HDP (Figure 4). Figure S6 shows HDP prevalence by percentile for each PGS. When the results were stratified by genetic similarity group, the four PGS were not significantly associated with HDP in the AFR genetic similarity group, but there were similar effect sizes for the overall models and for the EUR genetic similarity group (Figure 4).

For essential hypertension, the two SBP PGS were strongly associated with HDP in both the overall and stratified analyses (Figure S4). The pseudo-$R^2$ for SBP PGS [2] was the highest at 5.6% (Table S7). The receiver operator characteristic curves indicate that SBP PGS [2] is the most predictive of essential hypertension in PEGS (Figure S7), evidenced by the clear linear trend in the percentile plot (Figure S8).
Discussion

Our GWAS is the first to identify and validate genetic variants that are significantly associated with HDP in the maternal genome in RARB on chromosome 3. By leveraging a multi-ancestry study design, we increased statistical power for genetic discovery, which enabled us to confirm our findings in the clinically phenotyped UKBB cohort.

The two genome-wide significant loci in PEGS were also significantly associated with HDP in the gene-level analysis, where RARB and RN7SL283P were the top two genes. Retinoic acid receptors are essential for healthy placental and fetal development. The receptor binds retinoic acid, which is the biologically active form of vitamin A that facilitates cellular signaling in embryonic morphogenesis, cell proliferation, differentiation, and apoptosis\textsuperscript{45–47}. Maternal retinol is transferred across the placenta to embryos\textsuperscript{47}. All-trans-retinoic acid plays a significant role in maintaining immune balance in some tissues and is linked to the etiology of PE, where greater dysregulation in retinoic acid levels in the decidua (maternal endometrium) is observed in PE cases compared to those with a normal pregnancy\textsuperscript{48}. A link has been suggested between decidualization and the anti-angiogenic protein soluble fms-like tyrosine kinase-1 (sFLT1), with retinoic acid directly reducing sFLT1 production in the decidua, which affects early vascular development\textsuperscript{48}. Furthermore, in GWAS in a Peruvian cohort focused on maternal and offspring genetic effects on PE, rs4241542, a variant that is 80 kb from the lead variant rs61176331, was associated with proteinuria in offspring (\(P=8.48 \times 10^{-6}\))\textsuperscript{49}, further supporting the biological validity of this region in RARB for...
HDP. Future research is needed to investigate the interaction at the maternal-fetal interface, including the influence of gene-gene interaction in placental pathophysiology.

While the variant near *LRP1B* was not validated in UKBB, rs114954125 is near *RN7SL283P*, which is a pseudogene. *LRP1B* belongs to the LDL receptor family and is a tumor suppressor gene\(^{50}\). In a previous candidate SNP study, the T-allele variant in *LRP1B*, rs35821928, was associated with increased maternal height and decreased maternal weight\(^{51}\).

While the MAGMA gene-set analysis did not yield pathways that met the stringent FDR-corrected significance threshold, the top ten most significant gene sets provide intriguing insights into the probable molecular mechanisms of HDP. The pentose phosphate pathway is among these notable gene sets. This metabolic pathway is involved in cellular redox homeostasis, which is related to oxidative stress as a factor in HDP\(^{11,12}\). The identification of immune-response pathways and conditions such as autoimmune thyroid disease, graft-versus-host disease, and allograft rejection points to an immunological dysregulation component of HDP. This is consistent with a growing body of research focused on the immunological basis of HDP\(^{52,19,22,48}\).

The gene-level analysis revealed a strong peak of genes at chromosome 6 that did not meet the gene-wide significance threshold. *HLA-DRB5* and *BTNL2* belong to the human major histocompatibility (HLA) complex, where they are involved in the generation of an immune response. Additionally, the adipocytokine signaling pathway connects to
signaling molecules originating from adipose tissue, which has been linked to various aspects of cardiovascular health. This could indicate adipose tissue malfunction plays a role in HDP and other cardiometabolic disorders of pregnancy such as gestational diabetes. These suggestive findings related to immune response, inflammation, and metabolic dysfunction call for further analysis and interpretation in large pregnancy cohorts.

The differing patterns of suggestive associations in the GWAS of the composite HDP outcome, GH, and PE in PEGS may indicate differences in etiology (Figure S3). The heterogeneous nature of HDP has led to the understanding that there are several distinct PE subtypes. A two-stage placental model related to the development of PE demonstrated that an early-onset, more severe subtype may be associated with underlying genetic or environmental factors that result in abnormal placentation. In contrast, a later-onset, less severe subtype may be a consequence of factors such as obesity, diabetes, cardiovascular disorders, or multi-fetal pregnancy. Further, risk factors for GH and PE vary, supporting evidence for differing etiology across the collection of hypertensive disorders. However, a holistic assessment of HDP that combined PE with other maternal hypertensive disorders as a phenotype for GWAS yielded novel, overlapping risk loci identified in models of PE only and PE or other HDP.

Considering polygenic prediction, using a blood pressure PGS to predict HDP has been found to be more effective than using a PE PGS for women who are non-hypertensive.
before pregnancy\textsuperscript{57}. In this study, we assessed the predictive utility of four PGS in the context of HDP and essential hypertension. While these scores were developed in the White British sub-cohort of UKBB\textsuperscript{28,44}, our results demonstrate that an SBP PGS is a strong predictor of HDP. In the stratified results, none of the scores exhibited predictive value for HDP for the AFR genetic similarity group, and only the SBP scores were significantly associated with essential hypertension. These findings highlight the need for customizing PGS development to be inclusive of varying genetic similarity backgrounds to improve the accuracy of HDP risk assessment, as the study's effectiveness in predicting HDP using a PE PGS or GH PGS does not apply to all populations. Furthermore, while the SBP PGS showed promise for HDP prediction and prevention, it is worth investigating how these genetic scores could be combined with other clinical and environmental factors to develop a more holistic approach to HDP prediction and prevention.

This study has several limitations. First, a composite outcome like HDP can obscure disparate etiologies, for example, between PE and GH. Because of the limited sample size, we grouped the heterogeneous conditions that comprise HDP to increase power to detect genetic variant associations. Second, the results are based on data from a particular time point for each participant. The lack of longitudinal data does not allow us to account for the continuum of risk due to factors such as maternal age and gestational age at the onset of HDP. Further, the study design does not consider the recurrence of complications in subsequent pregnancies. Third, the outcomes are dichotomized due to the nature of the questionnaires completed, which may result in the oversimplification of
adverse pregnancy outcomes. Fourth, misclassification bias may be present, since participants who enrolled in PEGS before a pregnancy complication occurred were included in the control group. Fourth, there is a lack of harmonization of phenotypes between study cohorts. While PEGS data are limited, the UKBB uses a wide variety of diagnosis codes for phenotyping, and thus not all pregnancy-related adverse outcomes were excluded from the control group.

Perspectives (max 250 words - Conclusion)

Variants near *RAR* were significantly associated with HDP in a multi-ancestry cohort. We validated this association in a clinically phenotyped cohort. In gene-set analyses, probable mechanisms associated with HDP included inflammatory and immunological pathways. These findings are supported by recent research focused on the relationship between retinoic acid dysregulation and PE. In-depth functional genomics investigations can elucidate the molecular pathways by which genetic variations near *RAR* contribute to the etiology of HDP, such as the influence of these variants on *RAR* expression, epigenetic alterations, and downstream signaling pathways within the placenta and maternal vascular system. The study demonstrates the strong predictive value of SBP PGS for HDP, but further work should investigate approaches to improve the accuracy and inclusivity of risk assessment for HDP.

HDP is a complex and challenging area of obstetrics, with a significant, disproportionate impact on maternal and fetal health across populations. While we focus on the genetic contribution to HDP, examining the role of social determinants of health is also
important, in part because environmental factors can influence genetic risk and the development of HDP. Addressing social, lifestyle, and genetic factors is imperative to reduce the burden of HDP and improve maternal and fetal outcomes. Further research in large, diverse pregnancy cohorts is needed to fully understand the underlying mechanisms and develop effective preventative and therapeutic strategies.

Novelty and Relevance (max 100 words)

What is new?

- Genetic studies of hypertensive disorders of pregnancy (HDP) have predominantly focused on individuals of European ancestry.
- Our multi-ancestry genome-wide association study is the first to identify and validate genetic variants that are significantly associated with HDP in the maternal genome in the $RARB$ gene.

What is relevant?

- We identified two novel genome-wide significant associations with HDP in the maternal genome, namely rs114954125 on chromosome 2 and rs61176331 on chromosome 3. The variant rs61176331 was validated in the UK Biobank. Gene-set analyses revealed a suggestive association between HDP and inflammatory/immunological pathways such as the pentose phosphate pathway and autoimmune thyroid disease.

Clinical/Pathophysiological Implications?
• Retinoic acid, which binds to \textit{RARB}, inhibits the expression of sFLT1 in decidua.

The effects of genetic variants near RARB on the pathophysiology of HDP could provide insights into the underlying molecular mechanisms.

• Existing genetic scores for SBP, PE, and GH were predictive of HDP and essential hypertension in the overall PEGS cohort but were not predictive of HDP for those of predominantly African ancestry. Future directions include the development of genetic risk prediction modeling inclusive of large, ancestrally diverse populations to improve utility.

\textbf{Acknowledgements}

We would like to acknowledge the participants in the PEGS and UK Biobank cohorts for making this work possible. We would also like to thank Hannah Collins Cakar for assistance with manuscript preparation.

\textbf{Sources of Funding}

This study was supported by the National Institutes of Health Intramural Research Program and the NIH-Oxford-Cambridge Scholars program.

\textbf{Disclosures}

None
Supplemental Material

Supplemental Methods

Tables S1–S7

Figures S1–S8

References #58–#90

References


29. Pan UKBB | Pan UKBB [Internet]. [cited 2021 Mar 30];Available from: https://pan.ukbb.broadinstitute.org/


34. Return 2442 [Internet]. [cited 2023 Jul 19];Available from: https://biobank.ndph.ox.ac.uk/ukb/dset.cgi?id=2442


46. RARB retinoic acid receptor beta [Homo sapiens (human)] - Gene - NCBI
   [Internet]. [cited 2022 Sep 24];Available from:

   2012;139:843–858.


49. Nieves-Colón MA, Badillo Rivera KM, Sandoval K, Villanueva Dávalos V,
   Enríquez Lencinas LE, Mendoza-Revilla J, Adhikari K, González-Buenfil R, Chen
   JW, Zhang ET, et al. Clotting factor genes are associated with preeclampsia in

   Gibbs R, et al. WGSA: an annotation pipeline for human genome sequencing

   genetic risk factors on chromosome 2q22. Pregnancy Hypertens. Int. J. Womens

52. LaMarca B. Endothelial dysfunction; an important mediator in the Pathophysiology

53. Trivett C, Lees ZJ, Freeman DJ. Adipose tissue function in healthy pregnancy,
   1756.

54. Staff AC. The two-stage placental model of preeclampsia: An update. J. Reprod.

   and differences between the risk factors for gestational hypertension and
   preeclampsia: A population based cohort study in south China. Pregnancy

56. Tyrmi JS, Kaartokallio T, Lokki AI, Jääskeläinen T, Kortelainen E, Ruotsalainen S,
   with preeclampsia and hypertensive disorders of pregnancy. JAMA Cardiol.
   2023;8:674–683.

   S, Bello NA, Aittokallio J, et al. Associations of polygenic risk scores for
preeclampsia and blood pressure with hypertensive disorders of pregnancy. J. Hypertens. 2023;41:380.

58. gatk4-genome-processing-pipeline [Internet]. 2022 [cited 2022 Sep 6];Available from: https://github.com/gatk-workflows/gatk4-genome-processing-pipeline


61. Picard Tools - By Broad Institute [Internet]. [cited 2022 Sep 6];Available from: https://broadinstitute.github.io/picard/

62. applybqsr — Clara Parabricks v3.7 documentation [Internet]. [cited 2022 Sep 6];Available from: https://docs.nvidia.com/clara/parabricks/3.7.0/Documentation/ToolDocs/man_applybqsr.html

63. DeepVariant [Internet]. 2022 [cited 2022 Sep 6];Available from: https://github.com/google/deepvariant


73. IGSR | samples [Internet]. [cited 2022 May 9];Available from: https://www.internationalgenome.org/data-portal/sample


75. : Category 100319 [Internet]. [cited 2022 Sep 9];Available from: https://biobank.ndph.ox.ac.uk/ukb/label.cgi?id=100319

76. : Resource 1671 [Internet]. [cited 2022 Sep 9];Available from: https://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=1671

77. : Return 2442 [Internet]. [cited 2023 Jul 19];Available from: https://biobank.ndph.ox.ac.uk/ukb/dset.cgi?id=2442

78. Quality Control (QC) | Pan UKBB [Internet]. [cited 2021 Apr 9];Available from: https://pan-dev.ukbb.broadinstitute.org/docs/qc


Tables
<table>
<thead>
<tr>
<th>Hypertensive Disorders of Pregnancy, n (%)</th>
<th>Study Sample n = 1,771</th>
<th>HDP Cases n = 202</th>
<th>Controls n = 1,569</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gestational hypertension</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of pregnancies with gestational</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hypertension</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>97 (5.5)</td>
<td>97 (48.0)</td>
<td>0 (0.0)</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>2</td>
<td>24 (1.4)</td>
<td>24 (11.9)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>3 or more</td>
<td>14 (0.9)</td>
<td>14 (6.9)</td>
<td>0 (0.0)</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Missing</td>
<td>1636 (92.4)</td>
<td>67 (33.2)</td>
<td>1569 (100.0)</td>
<td></td>
</tr>
<tr>
<td>Gestational hypertension without</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>preeclampsia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of pregnancies with preeclampsia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>complication</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>58 (3.3)</td>
<td>58 (28.7)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4 (0.2)</td>
<td>4 (2.0)</td>
<td>0 (0.0)</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>3 or more</td>
<td>5 (0.3)</td>
<td>5 (2.5)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>1704 (96.2)</td>
<td>135 (66.8)</td>
<td>1569 (100.0)</td>
<td></td>
</tr>
<tr>
<td>Gestational hypertension and/or</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>preeclampsia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of pregnancies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>202 (11.4)</td>
<td>202 (100.0)</td>
<td>0 (0.0)</td>
<td>&lt;0.001*</td>
</tr>
</tbody>
</table>

**Covariates**

- **Age at enrollment, mean (SD)**: 47.2 (13.6) for Study Sample, 44.8 (12.7) for Controls, 47.5 (13.7) for HDP Cases, *p* = 0.0071†
- **BMI at enrollment, median (IQR)**: 27.4 (23.4, 32.6) for Study Sample, 28.7 (25.0, 34.9) for Controls, 27.3 (23.2, 32.3) for HDP Cases, *p* < 0.001‡
- **Gravidity, median (IQR)**: 3.0 (2.0, 4.0) for Study Sample, 4.0 (3.0, 5.0) for Controls, 3.0 (2.0, 4.0) for HDP Cases, *p* < 0.001‡
- **Race/ethnicity, n(%)**
  - American Indian/Alaskan Native: 16 (0.9) for Study Sample, 1 (0.5) for Controls, 15 (1.0) for HDP Cases, *p* = 0.1878*
  - Asian: 39 (2.2) for Study Sample, 3 (1.5) for Controls, 36 (2.3) for HDP Cases
  - Black: 358 (20.2) for Study Sample, 45 (22.3) for Controls, 313 (20.0) for HDP Cases
  - Hispanic/Latino: 47 (2.7) for Study Sample, 3 (1.5) for Controls, 44 (2.8) for HDP Cases, *p* = 0.1878*
  - White: 1256 (70.9) for Study Sample, 138 (68.3) for Controls, 1118 (71.3) for HDP Cases
  - Other (Multiple or Native Hawaiian/Pacific Islander): 22 (1.2) for Study Sample, 5 (2.5) for Controls, 17 (1.1) for HDP Cases
  - Unknown: 33 (1.9) for Study Sample, 7 (3.5) for Controls, 26 (1.7) for HDP Cases
### Predominant genetic similarity group

<table>
<thead>
<tr>
<th>Group</th>
<th>African</th>
<th>Admixed American</th>
<th>East Asian</th>
<th>European</th>
<th>South Asian</th>
<th>No group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>382 (21.6)</td>
<td>8 (0.5)</td>
<td>25 (1.4)</td>
<td>1328 (75.0)</td>
<td>15 (0.8)</td>
<td>13 (0.7)</td>
</tr>
<tr>
<td></td>
<td>49 (24.3)</td>
<td>0 (0.0)</td>
<td>1 (0.5)</td>
<td>150 (74.3)</td>
<td>2 (1.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td></td>
<td>333 (21.2)</td>
<td>8 (0.5)</td>
<td>24 (1.5)</td>
<td>1178 (75.1)</td>
<td>13 (0.8)</td>
<td>13 (0.8)</td>
</tr>
</tbody>
</table>

- Predominant genetic similarity group

### Socioeconomic Characteristics, n (%)

#### Educational Level at Enrollment

<table>
<thead>
<tr>
<th>Level</th>
<th>No group</th>
<th>High school education or less</th>
<th>Some college or technical/vocational school</th>
<th>Bachelor's degree</th>
<th>Graduate or professional degree</th>
<th>Missing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>13 (0.8)</td>
<td>183 (10.3)</td>
<td>549 (31.0)</td>
<td>530 (29.9)</td>
<td>495 (28.0)</td>
<td>14 (0.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14 (6.9)</td>
<td>73 (36.1)</td>
<td>57 (28.2)</td>
<td>57 (28.2)</td>
<td>1 (0.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>169 (10.8)</td>
<td>476 (30.3)</td>
<td>473 (30.2)</td>
<td>438 (27.9)</td>
<td>13 (0.8)</td>
</tr>
</tbody>
</table>

#### Income at Enrollment (USD)

<table>
<thead>
<tr>
<th>Income Range</th>
<th>No group</th>
<th>High school education or less</th>
<th>Some college or technical/vocational school</th>
<th>Bachelor's degree</th>
<th>Graduate or professional degree</th>
<th>Missing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than $50,000</td>
<td>13 (0.8)</td>
<td>659 (37.2)</td>
<td>441 (24.9)</td>
<td>603 (34.0)</td>
<td>68 (3.8)</td>
<td>6 (0.3)</td>
</tr>
<tr>
<td>$50,000 - $79,999</td>
<td>13 (0.8)</td>
<td>69 (34.2)</td>
<td>62 (30.7)</td>
<td>66 (32.7)</td>
<td>5 (2.5)</td>
<td>6 (0.3)</td>
</tr>
<tr>
<td>$80,000 or more</td>
<td>13 (0.8)</td>
<td>590 (37.6)</td>
<td>379 (24.2)</td>
<td>537 (34.2)</td>
<td>63 (4.0)</td>
<td>6 (0.3)</td>
</tr>
<tr>
<td>Missing</td>
<td>13 (0.8)</td>
<td>6 (0.3)</td>
<td>6 (0.3)</td>
<td>6 (0.3)</td>
<td>6 (0.3)</td>
<td>6 (0.3)</td>
</tr>
</tbody>
</table>

- **Table 1.** Demographic characteristics of participants in the PEGS discovery cohort.
  - * The \( P \) values were computed by Fisher’s exact test.
  - † The \( P \) value was computed by \( t \)-test.
  - ‡ The \( P \) values were computed by performing the Mann-Whitney \( U \) test.
<table>
<thead>
<tr>
<th>Nearest Gene</th>
<th>Lead Variant</th>
<th>Chr:Position</th>
<th>Effect Allele/Other Allele</th>
<th>EAF</th>
<th>OR</th>
<th>95% CI</th>
<th>P</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>LRP1B</em></td>
<td>rs114954125</td>
<td>2:139978665</td>
<td>G/A</td>
<td>0.051</td>
<td>3.03</td>
<td>2.05-4.49</td>
<td>3.19E-08</td>
<td>3.79E-08</td>
</tr>
<tr>
<td><em>RARB</em></td>
<td>rs61176331</td>
<td>3:24930847</td>
<td>A/C</td>
<td>0.048</td>
<td>3.09</td>
<td>2.11-4.53</td>
<td>7.97E-09</td>
<td>9.56E-09</td>
</tr>
</tbody>
</table>

**Genetic Similarity Group for African Descent**

(49 cases/333 controls)

<table>
<thead>
<tr>
<th>Nearest Gene</th>
<th>Lead Variant</th>
<th>Chr:Position</th>
<th>Effect Allele/Other Allele</th>
<th>EAF</th>
<th>OR</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>RARB</em></td>
<td>rs61176331</td>
<td>3:24930847</td>
<td>A/C</td>
<td>0.042</td>
<td>6.98</td>
<td>2.97-16.42</td>
<td>8.32E-06</td>
</tr>
</tbody>
</table>

**Genetic Similarity Group for European Descent**

(150 cases/1,178 controls)

<table>
<thead>
<tr>
<th>Nearest Gene</th>
<th>Lead Variant</th>
<th>Chr:Position</th>
<th>Effect Allele/Other Allele</th>
<th>EAF</th>
<th>OR</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>LRP1B</em></td>
<td>rs114954125</td>
<td>2:139978665</td>
<td>G/A</td>
<td>0.067</td>
<td>2.95</td>
<td>1.97-4.42</td>
<td>1.47E-07</td>
</tr>
<tr>
<td><em>RARB</em></td>
<td>rs61176331</td>
<td>3:24930847</td>
<td>A/C</td>
<td>0.051</td>
<td>2.37</td>
<td>1.51-3.74</td>
<td>1.92E-04</td>
</tr>
</tbody>
</table>

**Validation**

(1,319 cases/2,050 controls)
<table>
<thead>
<tr>
<th>Nearest Gene</th>
<th>Lead Variant</th>
<th>Chr:Position</th>
<th>Other Allele</th>
<th>Effect Allele/ Other Allele</th>
<th>EAF</th>
<th>OR</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRP1B</td>
<td>rs114954125</td>
<td>2:139978665</td>
<td>G/A</td>
<td>EAF/Effect Allele/Other Allele</td>
<td>0.061</td>
<td>1.02</td>
<td>0.83-1.27</td>
<td>8.34E-01</td>
</tr>
<tr>
<td>RARB</td>
<td>rs61176331</td>
<td>3:24930847</td>
<td>A/C</td>
<td>EAF/Effect Allele/Other Allele</td>
<td>0.045</td>
<td>1.29</td>
<td>1.02-1.65</td>
<td>3.73E-02</td>
</tr>
</tbody>
</table>

GWAS + Validation
(1,521 cases/3,619 controls)

<table>
<thead>
<tr>
<th>Nearest Gene</th>
<th>Lead Variant</th>
<th>Chr:Position</th>
<th>Other Allele</th>
<th>Effect Allele/ Other Allele</th>
<th>EAF</th>
<th>OR</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRP1B</td>
<td>rs114954125</td>
<td>2:139978665</td>
<td>G/A</td>
<td>EAF/Effect Allele/Other Allele</td>
<td>0.059</td>
<td>1.31</td>
<td>1.09-1.58</td>
<td>4.64E-03</td>
</tr>
<tr>
<td>RARB</td>
<td>rs61176331</td>
<td>3:24930847</td>
<td>A/C</td>
<td>EAF/Effect Allele/Other Allele</td>
<td>0.046</td>
<td>1.66</td>
<td>1.35-2.04</td>
<td>1.27E-06</td>
</tr>
</tbody>
</table>

Table 2. Summary statistics for rs114954125 (LRP1B) and rs61176331 (RARB). Chr = chromosome, EAF = effect allele frequency.
Figure 1. Manhattan plot of hypertensive disorders of pregnancy GWAS results and LocusZoom plots of genome-wide significant loci in the PEGS cohort (202 cases and 1,569 controls). The genome coordinates for B and C reflect GRCh38 assembly.
Figure 2. Manhattan plot of gene-based association analysis of hypertensive disorders of pregnancy. Labeled genes represent the nearest gene within 50 kb upstream or downstream of the gene.
Figure 3. Gene-set analysis using KEGG with MAGMA. The ten most significant KEGG sets are shown on the y-axis. Corresponding unadjusted -log_{10}(p-values) are shown on the x-axis. P values were estimated by linear regression model-based testing using MAGMA.
Figure 4. Polygenic score summary statistics and forest plot for odds ratio (ORs) per standard deviation (SD) for hypertensive disorders of pregnancy in multi-ancestry models and African and European genetic similarity groups. AUC: area under the curve.