# Genome-wide meta-analysis identifies novel maternal risk variants and enables polygenic prediction of preeclampsia and gestational hypertension

Michael C. Honigberg[1-3]*, Buu Truong[2,3]*, Raiyan R. Khan[4], Brenda Xiao[5], Laxmi Bhatta[6,7], Thi Ha Vy[8], Rafael F. Guerrero[9], Art Schuermans[2,3,10], Margaret Sunitha Selvaraj[2,3], Aniruddh P. Patel[1-3], Satoshi Koyama[2,3], So Mi Jemma Cho[2,3,11], Shamsudheen Karuthedath Vellarikkal[3,12], Mark Trinder[3,13], Sarah M. Urbut[1-3], Kathryn J. Gray[14,15], Ben M. Brumpton[6,7], Snehal Patil[16], Sebastian Zöllner[16], Genes & Health Research Team, Estonian Biobank Research Team, nuMoM2b Research Team, Richa Saxena[3,15], Girish N. Nadkarni[8], Ron Do[8], Qi Yan[17], Itsik Pe'er[4], Shefali Setia Verma[5], Rajat M. Gupta[3,12], David M. Haas[18], Hilary C. Martin[19], David A. van Heel[20], Triin Laisk[21], Pradeep Natarajan[1-3,15]

[1]Cardiology Division, Department of Medicine, Massachusetts General Hospital, Boston, MA, USA
[2]Cardiovascular Research Center, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA
[3]Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA
[4]Department of Computer Science, Columbia University, New York, NY, USA
[5]University of Pennsylvania, Philadelphia, PA, USA
[6]K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Trondheim, Norway
[7]HUNT Research Centre, Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Levanger, Norway
[8]The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA
[9]Department of Biological Sciences, North Carolina State University, Raleigh, NC, USA
[10]Faculty of Medicine, KU Leuven, Leuven, Belgium
[11]Integrative Research Center for Cerebrovascular and Cardiovascular Diseases, Seoul, Republic of Korea
[12]Cardiovascular Division, Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA
[13]Centre for Heart Lung Innovation, University of British Columbia, Vancouver, BC, Canada
[14]Department of Obstetrics and Gynecology, Brigham and Women's Hospital, Boston, MA, USA
[15]Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA
[16]Department of Biostatistics and Center for Statistical Genetics, University of Michigan School of Public Health, Ann Arbor, MI, USA
[17]Department of Obstetrics and Gynecology, Columbia University, New York, NY, USA
[18]Indiana University School of Medicine, Indianapolis, IN, USA
[19]Department of Human Genetics, Wellcome Sanger Institute, Cambridge, UK
[20]Blizard Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK
[21]Estonian Genome Centre, Institute of Genomics, University of Tartu, Tartu, Estonia

*These authors contributed equally.

**Corresponding Author:**

Pradeep Natarajan, MD MMSc
Massachusetts General Hospital
185 Cambridge Street, CPZN 3.184
Boston, MA 02114
Phone: 617-726-1843
Fax: 617-726-2203
Email: pnatarajan@mgh.harvard.edu
Twitter: @pnatarajanmd

**Abstract**

Preeclampsia and gestational hypertension are common pregnancy complications associated with adverse maternal and offspring outcomes. Current tools for prediction, prevention, and treatment are limited. In discovery analysis, we tested the association of maternal DNA sequence variants with preeclampsia in 17,150 cases and 451,241 controls and with gestational hypertension in 8,961 cases and 184,925 controls using multi-ancestry meta-analysis. We identified 12 independent loci associated with preeclampsia/eclampsia, 7 of which are novel (including *MTHFR-CLCN6*, *WNT3A*, *PGR*, *FLT1*, and *RGL3*), and 7 loci associated with gestational hypertension. Identified loci highlight the role of natriuretic peptide signaling, angiogenesis, renal glomerular function, trophoblast development, and immune dysregulation. We derived genome-wide polygenic risk scores that predicted preeclampsia/eclampsia and gestational hypertension in external datasets, independent of first-trimester risk markers. Collectively, these findings provide mechanistic insights into the hypertensive disorders of pregnancy and advance pregnancy risk stratification.

The hypertensive disorders of pregnancy (HDPs) represent a leading global cause of maternal and neonatal morbidity and mortality and cause approximately 14% of maternal deaths worldwide[1,2]. Up to 15% of child-bearing women experience an HDP in at least one pregnancy[3,4]. The HDPs include preeclampsia, defined as new-onset hypertension or worsening hypertension after 20 weeks' gestation plus proteinuria or other evidence of end-organ dysfunction; gestational hypertension, defined as new-onset hypertension without accompanying features of preeclampsia; and eclampsia, defined as progression of preeclampsia to maternal seizures[5,6]. In addition to short-term risks of end-organ failure and death in the absence of prompt recognition and treatment, those who develop HDPs have roughly two-fold long-term risk of cardiovascular disease events and premature cardiovascular mortality compared with those who experience only normotensive pregnancies for reasons that remain incompletely understood[7,8].

The pathophysiology of the HDPs is increasingly recognized to be heterogeneous with both maternal and fetal contributions. In the contemporary model of preeclampsia pathophysiology, defective trophoblast invasion in early placental development and incomplete remodeling of the maternal spiral arteries lead to placental ischemia later in gestation[1,9,10]. The distressed placenta secretes an excess of circulating anti-angiogenic proteins (e.g., soluble fms-like tyrosine kinase receptor 1 [sFlt-1] and soluble endoglin) which induce the systemic maternal endothelial dysfunction and vasoconstriction that drive the clinical manifestations of preeclampsia (hypertension and proteinuria)[11]. In addition, maternal cardiometabolic risk factors (e.g., pre-pregnancy chronic hypertension, diabetes, obesity) and pre-pregnancy kidney and autoimmune disease strongly predict preeclampsia[1,12] and influence early placentation as well as maternal vascular adaptation to pregnancy[10,13].

Genetic analysis may yield novel mechanistic insights into the pathophysiology of the HDPs and may also help improve pregnancy risk prediction. An estimated 31-35% of preeclampsia predisposition has been attributed to maternal genetics using familial aggregation-

based approaches[14,15]. However, few genetic loci linked to preeclampsia have been identified and robustly validated to date. Several fetal variants near the *FLT1* gene, which encodes placenta-derived sFlt-1, have been reported to associate with preeclampsia[16,17]. Steinthorsdottir et al. recently performed the largest maternal GWAS of preeclampsia to date and identified associations near *FTO* (the first reported obesity-associated locus)[18], *ZNF831*, and several other blood pressure-associated genes (*MECOM*, *FGF5*, and *SH2B3*) in combined meta-analysis of 12,150 cases and 164,098 controls[17] but without robust replication to date. In addition, several studies have now reported that increased maternal polygenic risk of hypertension associates with risk of HDPs[17,19-21].

In this work, we performed multi-ancestry GWAS meta-analysis for preeclampsia/eclampsia and separately for gestational hypertension. We then used these results to train and test polygenic risk scores (PRS) for each outcome in independent datasets.

## Results

*Association of common genetic variants with preeclampsia/eclampsia and gestational hypertension.*

We tested the association of common variants (minor allele frequency [MAF] ≥1%) with preeclampsia/eclampsia among 17,150 cases and 451,241 controls in discovery analysis (78.0% European, 21.2% Asian, 0.5% Hispanic, 0.3% African, **Supplementary Table 1**) using multi-ancestry fixed-effects meta-analysis in METAL[22]. Cases were identified principally using *International Classification of Diseases* (ICD) codes and phecodes corresponding to preeclampsia and, where available, eclampsia (**Supplementary Tables 1-2**); controls were generally either women with exclusively normotensive pregnancies or all women without codes corresponding to hypertension in pregnancy[17]. We did not observe inflation in test statistics (lambda genomic inflation factor: 1.049, **Supplementary Fig. 1a**). In discovery analysis, we identified 12 independent loci at the commonly used statistical significance threshold of

$P<5x10^{-8}$, including 5 previously nominated by Steinthorsdottir et al. (*MECOM* [3q26], *FGF5* [4q21], *SH2B3* [12q22], *FTO* [16q12], and *ZNF831* [20q13])[17] and 7 additional loci (*MTHFR-CLCN6* [1p36], *WNT3A* [1q42], *MICA* [6p21], *LINC00484* [9q22], *PGR* [11q22], *FLT1* [13q12], and *RGL3* [19p13]; **Fig. 1a, Table 1, Supplementary Table 3**).

We pursued replication of these GWAS results in three follow-up cohorts that collectively included 2,433 preeclampsia/eclampsia cases and 246,132 controls. Given the high observed genetic correlation between preeclampsia/eclampsia and gestational hypertension ($r_g$ = 0.71), we meta-analyzed genetic associations across these related phenotypes (including 747 gestational hypertension cases and 222,119 controls in follow-up cohorts) to maximize power to replicate findings for both outcomes (see Methods). Using this approach, we replicated 7 of 12 associations from discovery analysis with $P<0.05$ and consistent direction of effect, including the novel associations at *MTHFR-CLCN6*, *PGR*, *FLT1*, and *RGL3* (**Table 1, Supplementary Table 4**). Eleven associations had consistent direction of effect in replication cohorts. Ten of 12 associated loci retained genome-wide significance in combined meta-analysis of preeclampsia/eclampsia discovery and follow-up cohorts, and 6 loci (*MTHFR-CLCN6*, *MECOM*, *FGF5*, *PGR*, *FTO*, and *ZNF381*) replicated in follow-up analysis restricted to preeclampsia/eclampsia only (**Supplementary Table 5**). Conditional analysis in GCTA-COJO[23] identified a second independent association on chromosome 20 near *ZBTB46* (lead variant rs4809370, OR 1.08, $P=1.4x10^{-8}$). In meta-analysis of discovery and replication cohorts, 2 additional loci attained genome-wide significance (*FGL1* [8p22] and *UPB1* [22q11]; **Supplementary Fig. 2**).

We next tested the association of common variants with gestational hypertension among 8,961 cases and 184,925 controls in discovery analysis (91.3% European, 6.7% Asian, 1.2% Hispanic, 0.7% African) and among 747 cases and 222,119 controls in follow-up cohorts (**Supplementary Tables 1-2**). Gestational hypertension cases were identified based on qualifying ICD codes for gestational hypertension and an absence of qualifying codes for

preeclampsia/eclampsia. As with preeclampsia/eclampsia, we did not observe inflation in test statistics (lambda genomic inflation factor: 0.988, **Supplementary Fig. 1b**). In discovery analysis, we identified 7 independent genome-wide significant loci associated with gestational hypertension, including 4 also associated with preeclampsia/eclampsia (*MECOM*, *FGF5*, *RGL3*, and *ZNF831*) and 3 additional associations (*NPR3* [5p13], *TNS2-CSAD* [12q13], and *PREX1* [20q13]; **Table 2; Fig. 1b**). Four of 7 significant associations replicated with $P<0.05$ in follow-up cohorts (**Table 2, Supplementary Table 4**), and all 7 loci had consistent direction of effect in replication. In addition, all 7 loci retained genome-wide significance in meta-analysis of gestational hypertension discovery and follow-up cohorts (**Supplementary Table 6**). In combined meta-analysis of discovery and follow-up cohorts, the *MTHFR-CLCN6* locus (1p36) additionally reached genome-wide significance (**Supplementary Fig. 3**).

Given the high degree of genetic correlation observed between preeclampsia/eclampsia and gestational hypertension, we used multi-trait analysis of genome-wide association summary statistics (MTAG)[24] to boost power to identify additional associated variants. MTAG identified 2 additional loci with genome-wide significance (**Supplementary Table 7**), with the same loci identified for both preeclampsia/eclampsia (**Supplementary Fig. 4**) and gestational hypertension: *PLCE1* (10q23), a blood pressure-associated gene involved in glomerular podocyte development[25], which also narrowly missed statistical significance in the gestational hypertension discovery GWAS discovery GWAS ($P=5.02\times10^{-8}$); and *FURIN* (15q26), whose product is a protein convertase involved in processing pro-natriuretic peptides[26] and whose expression is decreased in preeclamptic placentas[27].

*Genetic correlation across hypertension-related phenotypes.*

We used cross-trait LD score regression[28] to assess genetic correlations among preeclampsia/eclampsia, gestational hypertension, systolic blood pressure (SBP), and diastolic blood pressure (DBP)[29]. As noted earlier, preeclampsia/eclampsia and gestational hypertension

were strongly genetically correlated ($r_g$ = 0.71, SE = 0.08). SBP demonstrated stronger genetic correlation with gestational hypertension ($r_g$ = 0.73, SE = 0.06) vs. preeclampsia/eclampsia ($r_g$ = 0.52, SE = 0.05). Of note, correlation of SBP with gestational hypertension ($r_g$ = 0.73) was larger than that of SBP with DBP ($r_g$ = 0.62, SE = 0.03, in the Million Veteran Program[29], with similar genetic correlation between SBP and DBP observed previously in the UK Biobank[30]). Genetic correlations with the HDPs were stronger for SBP vs. DBP (**Supplementary Table 8**).

*Gene prioritization at preeclampsia/eclampsia and gestational hypertension risk loci.*

To prioritize causal genes, we first performed colocalization analysis with expression quantitative trait loci (eQTLs) within ±500 kilobases across 52 tissues in the Genotype-Tissue Expression project (GTEx) (**Supplementary Tables 9-10**)[31]. Colocalization analyses implicated *FGF5* and *NPR3* as causal genes at their respective loci. The lead variant at the *MTHFR-CLCN6* locus colocalized with *CLCN6* eQTLs as well as expression of *NPPA* (which encodes the precursor to atrial natriuretic peptide [ANP]). The lead preeclampsia/eclampsia variant at the *ZNF831* locus colocalized with multiple genes but most strongly with *ZBTB46* expression, including in arterial tissue. We also observed multiple colocalizations with *WNT3A* (*WNT3A*, *GJC2,* and mitochondrial proteins *IBA57* and *MPRL55*), *MICA* (*CLIC1* and psoriasis-associated genes *TCF19*, *CCHCR1*, and *PSORS1C1*), and *RGL3* (*ZNF627* and *EPOR*). We observed no strong colocalizations with lead variants at *LINC00484*, *PGR*, *SH2B3*, *FLT1*, *FTO*, or *CSAD*.

Next, we queried variant-to-gene evidence in Open Targets Genetics v7 (**Supplementary Tables 11-12**)[32] and generated polygenic priority scores (PoPS, **Supplementary Table 13**)[33] for lead variants. Both approaches nominated *MECOM*, *FGF5*, *SH2B3*, *FTO*, and *NPR3* as the most likely causal gene and their respective loci. PoPS prioritized *NPPA* as the most likely causal gene at the *MTHFR-CLCN6* locus and *TRPC6* as the most likely causal gene at the *PGR* locus.

To further understand how identified genes might influence HDP risk, we queried nearest genes and those prioritized by colocalization, variant-to-gene scores, and/or PoPS in a publicly available database of the human placental transcriptome including preeclampsia cases and controls (**Supplementary Table 14**)[34]. Consistent with the correlation of increased circulating placental sFlt-1 with preeclampsia incidence[11,35], *FLT1* gene expression was increased in preeclamptic placentas ($\log_2$(fold change) 0.39, false discovery rate-adjusted *P*=0.003). Expression of *WNT3A,* which occurs almost exclusively in the placenta[36], was increased in preeclamptic placentas vs. healthy controls ($\log_2$(fold change) 0.21, adjusted *P*=0.029), in line with another recent study[37]. *OBSCN,* which was most strongly prioritized by PoPS at the *WNT3A* locus, was also overexpressed in preeclamptic vs. control placentas ($\log_2$(fold change) 0.18, adjusted *P*=0.037). Furthermore, preeclamptic placentas demonstrated lower expression of *ARHGAP42*—which sits adjacent to *PGR* and encodes Rho GTP-ase activating protein 42, a known regulator of vascular tone and blood pressure expressed selectively in smooth muscle cells[38]—compared with controls ($\log_2$(fold change) -0.18, adjusted *P*=0.004).

We analyzed expression of prioritized genes in a dataset of single-nuclei RNA sequencing (snRNA-seq) from non-atherosclerotic human aortic tissue. SnRNA-seq identified subpopulations of vascular smooth muscle cells, fibromyocytes, fibroblasts, endothelial cells, macrophages, natural killer T cells, and neuronal cells. The greatest enrichment was seen in the two endothelial cell (EC) populations and in macrophages (**Extended Data Fig. 1**). The EC1 subpopulation is enriched for genes in angiogenesis and lipid transport, while the EC2 subpopulation is enriched for genes in extracellular matrix production and integrin expression[39]. Relative expression in ECs vs. other cell types was strongest for *FLT1* (EC1/EC2), *ZBTB46* (EC1), and *MECOM* (EC2).

*Training and testing polygenic risk scores for preeclampsia/eclampsia and gestational hypertension.*

We used PRS-CS[40] to construct genome-wide PRS for preeclampsia/eclampsia ($PRS_{preeclampsia}$) and gestational hypertension ($PRS_{GH}$) from our corresponding discovery GWAS summary statistics. In addition, since blood pressure polygenic risk has previously been associated with HDPs[17,19,20], we used PRS-CS to derive a PRS for SBP ($PRS_{SBP}$) using the SBP GWAS from the Million Veteran Program[29] to compare prediction across scores and determine whether a linear combination of each HDP PRS and $PRS_{SBP}$ improves performance.

We tuned polygenic scores among 423 women with a history of preeclampsia/eclampsia, 482 women with a history of gestational hypertension only, and 211,129 parous women without a reported history of HDPs in the UK Biobank. The global shrinkage parameter of $1x10^{-4}$ was chosen for $PRS_{preeclampsia}$ and $PRS_{SBP}$ and of $1x10^{-6}$ for $PRS_{GH}$ as these values generated the highest $R^2$ (**Supplementary Table 15**). A linear combination of $PRS_{preeclampsia}$ and $PRS_{SBP}$ ($PRS_{preeclampsia+SBP}$) improved performance vs. each score individually for the outcome of preeclampsia (**Supplementary Table 16**).

The individual and combination PRS tuned in the UK Biobank were carried forward for external validation in two complementary datasets: a Norwegian population-based cohort linked to the Medical Birth Register of Norway (HUNT, preeclampsia/eclampsia only) and a prospective U.S. pregnancy cohort (nuMoM2b). Among 25,582 Norwegian women in HUNT (1,569 [6.1%] with preeclampsia/eclampsia), the prevalence of preeclampsia/eclampsia ranged from ~4% among women in the bottom decile of $PRS_{preeclampsia+SBP}$ to ~10% among the top decile of $PRS_{preeclampsia+SBP}$ (**Fig. 2a**). After adjustment for age at birth, $age^2$, and the first 10 principal components (PC) of ancestry, the OR corresponding to the top 10%vs. bottom 90% of $PRS_{preeclampsia+SBP}$ was 1.85 (95% CI 1.61-2.13, $P$=$6.3x10^{-18}$). $PRS_{preeclampsia+SBP}$ increased Nagelkerke's $R^2$ by 28.5% compared with the $PRS_{preeclampsia}$ alone and by 79.3% compared with $PRS_{SBP}$ alone (**Supplementary Table 17**).

We next tested PRS performance in the prospective, multi-ancestry nuMoM2b cohort of U.S. women recruited in the first trimester of their first pregnancy, including 507 (6.5%) who developed preeclampsia, 1,358 (17.3%) who developed gestational hypertension, and 5,986 with normotensive pregnancies (overall 70.8% European, 16.0% African, and 1.0% admixed American ancestry). Rates of preeclampsia/eclampsia ranged from ~4% among women in the bottom decile of $PRS_{preeclampsia+SBP}$ to ~10% among the top decile of $PRS_{preeclampsia+SBP}$ (**Fig. 2b**). Rates of gestational hypertension ranged from ~9% among women in the bottom decile of $PRS_{GH+SBP}$ to ~24% among the top decile of $PRS_{GH+SBP}$ (**Fig. 2c**). As in HUNT, incorporating SBP PRS in linear combination boosted PRS performance for HDPs, especially for gestational hypertension, although PRS performance was better for European-ancestry vs. other women (**Supplementary Table 18**). After adjustment for age, PC 1-10, and self-reported race/ethnicity, $PRS_{preeclampsia+SBP}$ and $PRS_{GH+SBP}$ each predicted their respective outcomes (preeclampsia/eclampsia: OR 1.81, 95% CI 1.39-2.33, for top 10% vs. bottom 90% $PRS_{preeclampsia+SBP}$, $P=7.0\times10^{-6}$; gestational hypertension: OR 1.52, 95% CI 1.26-1.82, for top 10% vs. bottom 90% $PRS_{GH+SBP}$, $P=9.1\times10^{-6}$). After further adjustment for first-trimester SBP, antihypertensive medication use, and body mass index (BMI)—well-established clinical predictors of HDPs—the scores both remained predictive (preeclampsia/eclampsia: OR 1.65, 95% CI 1.26-2.15, for top 10% vs. bottom 90% $PRS_{preeclampsia+SBP}$, $P=2.4\times10^{-4}$; gestational hypertension: OR 1.54, 95% CI 1.27-1.85, or top 10% vs. bottom 90% $PRS_{GH+SBP}$, $P=9.8\times10^{-6}$). Compared with a model including age, PC 1-10, self-reported race/ethnicity, first-trimester SBP, antihypertensive medication use, and first-trimester SBP, addition of $PRS_{preeclampsia+SBP}$ improved the C-statistic for preeclampsia/eclampsia from 0.688 to 0.700 (+0.012, 95% CI 0.002-0.023, Delong's $P=0.02$). Similarly, addition of $PRS_{GH+SBP}$ improved the C-statistic for gestational hypertension from 0.648 to 0.659 (+0.011, 95% CI 0.004-0.019, Delong's $P=0.003$).

Low-dose aspirin starting after 12 weeks' gestation represents an evidence-based but underutilized[41] strategy to reduce risk of preeclampsia. To further probe the potential clinical

impact of incorporating PRS to guide aspirin allocation, we examined aspirin eligibility according to current U.S. Preventive Service Task Force major criteria[42] with and without addition of $PRS_{preeclampsia+SBP}$ as an additional eligibility criterion in the nuMoM2b cohort. Among singleton, nulliparous women (i.e., the population enrolled in nuMoM2b), major criteria for aspirin eligibility are chronic pre-pregnancy hypertension, pre-gestational diabetes, kidney disease, and autoimmune disease.[42] Major risk factors identified only a small proportion of the cohort that developed HDPs (**Table 3**), and aspirin would have been recommended in only 17.2% of women who developed preeclampsia/eclampsia. Incorporating the top 10% of $PRS_{preeclampsia+SBP}$ increased identification of the aspirin-eligible proportion to 30.2% of women with preeclampsia/eclampsia. Expanding aspirin eligibility further to include the top 25% of $PRS_{preeclampsia+SBP}$ captured nearly half (46.9%) of women who developed preeclampsia/eclampsia. Addition of high $PRS_{preeclampsia+SBP}$ to major risk factors to up-classify risk of preeclampsia/eclampsia vs. other pregnancies yielded net reclassification of +4.3% (95% CI 1.7-6.8%) for top 5% $PRS_{preeclampsia+SBP}$, +4.4% (95% CI 1.4-7.5%) for top 10% $PRS_{preeclampsia+SBP}$, and +8.4% (95% CI 4.2-12.5%) for top 25% $PRS_{preeclampsia+SBP}$ (**Table 3**).

*Phenome-wide associations with preeclampsia/eclampsia and gestational hypertension polygenic risk.*

We performed sex-stratified phenome-wide association analysis (pheWAS) for $PRS_{preeclampsia}$ and $PRS_{GH}$ across 1,445 phecode-based phenotypes in the UK Biobank. $PRS_{preeclampsia}$ was significantly associated with 36 phenotypes in women and 37 phenotypes in men with Bonferroni-corrected statistical significance ($P<0.05/1,445 = 3.5\times10^{-5}$); $PRS_{GH}$ was significantly associated with 25 and 32 phenotypes in women and men, respectively (**Fig. 3**). $PRS_{preeclampsia}$ and $PRS_{GH}$ were most strongly associated with hypertension in both sexes ($PRS_{preeclampsia}$: $OR_{women}$ 1.15 per SD, 95% CI 1.14-1.16, $P=1.4\times10^{-175}$; $OR_{men}$ 1.12 per SD, 95% CI 1.11-1.13, $P=7.5\times10^{-112}$; $PRS_{GH}$: $OR_{women}$ 1.15 per SD, 95% CI 1.14-1.16, $P=5.1\times10^{-184}$; $OR_{men}$

1.13 per SD, 95% CI 1.11-1.14, P=1.3x$10^{-133}$). Other strong phenotypic associations included hypercholesterolemia, type 2 diabetes, obesity, and atherosclerotic cardiovascular disease (**Supplementary Tables 19-20**). $PRS_{preeclampsia}$ predicted ischemic heart disease in women (OR 1.09 per SD, 95% CI 1.07-1.11, *P*=4.4x$10^{-27}$) and men (OR 1.09 per SD, 95% CI 1.07-1.10, *P*=2.6x$10^{-40}$), as did $PRS_{GH}$ ($OR_{women}$ 1.07 per SD, 95% CI 1.05-1.08, *P*=4.4x$10^{-16}$; $OR_{men}$ 1.08 per SD, 95% CI 1.07-1.09, *P*=1.4x$10^{-35}$). $PRS_{preeclampsia}$ was also associated with several autoimmune phenotypes, including celiac disease, type 1 diabetes, hypothyroidism (in women), and a suggestive association with rheumatoid arthritis in women (*P*=5.7x$10^{-6}$), whereas type 1 diabetes and celiac disease were not significantly associated with $PRS_{GH}$ in either sex.

## Discussion

We present an expanded multi-ancestry maternal genome-wide association analysis of preeclampsia/eclampsia and, to our knowledge, the first distinct GWAS of gestational hypertension, two pregnancy complications on the spectrum of HDPs with substantial associated public health impact. We identified 12 independent genomic loci associated with preeclampsia/eclampsia, 7 of which are novel (including *MTFHR-CLCN6*, *WNT3A*, *PGR*, *FLT1*, and *RGL3*), and 7 loci associated with gestational hypertension, of which 4 overlap with preeclampsia/eclampsia, with the majority of loci replicating in independent cohorts. Identified loci highlight the role of angiogenesis and endothelial cell function (*FLT1*, *ZBTB46*), natriuretic peptide signaling (*NPPA*, *NPR3*, *FURIN*), renal glomerular function (*TRPC6*, *TNS2*, *PLCE1*), trophoblast development (*WNT3A*), and immune dysregulation (*MICA*, *SH2B3*) in the pathogenesis of these conditions, with some loci (*FLT1*[16,17,35], *WNT3A*[37,43]) described to influence risk via the fetal genome. Furthermore, we demonstrated that an HDP PRS augmented by a systolic blood pressure PRS stratified risk of HDP. Polygenic risk scores predicted HDP risk among nulliparous women independent of first-trimester blood pressure and BMI, indicating potential clinical utility of these scores risk for pregnancy risk stratification.

Collectively, these findings may have implications for advancing HDP prediction, prevention, and treatment.

First, our findings provide insights into mechanisms of HDP pathogenesis and underscore the causal role of blood pressure. As in previous GWAS of preeclampsia[17], most genetic associations we identified were also associated with blood pressure[25], and genetic correlation between blood pressure and the HDPs, especially gestational hypertension, was high. These findings align with prior work demonstrating heightened polygenic BP risk in those with HDPs[17,19,20]. Indeed, we observed that integrating SBP PRS with HDP PRS in linear combination further improved predictive performance for both preeclampsia/eclampsia and gestational hypertension. The recently published randomized CHAP trial of treatment for mild chronic hypertension in pregnancy demonstrated that lowering blood pressure pharmacologically reduced risk of progression to preeclampsia[44], supporting the notion that elevated blood pressure is not merely a clinical manifestation of the HDPs but also plays a causal role in disease pathogenesis.

Second, our GWAS findings suggest a key role of natriuretic peptide signaling in the pathogenesis of the HDPs. The natriuretic peptides (e.g., ANP, B-type natriuretic peptide [BNP]) promote renal sodium excretion and counteract renin-angiotensin and sympathetic nervous system activation[45]. The lead variant at the *MTHFR-CLCN6* locus colocalized with *NPPA* expression in atrial tissue (its primary site of expression). ANP (the protein product of *NPPA*) also plays a role in decidualization and spiral artery remodeling[46], a process known to be impaired in the early pathogenesis of preeclampsia[1,10]. Furthermore, ANP is cleared from the circulation by the protein product of *NPR3*[47], which was identified in the GWAS of gestational hypertension. Human data also support accelerated ANP clearance in the setting of preeclampsia[48]. Of note, our lead risk variant at the *MTHFR-CLCN6* locus has been previously associated with reduced levels of circulating N-terminal pro-BNP[49]. A recent analysis from the nuMoM2b cohort found that first-trimester levels of N-terminal pro-BNP were unexpectedly

lower among women who subsequently developed HDPs later in pregnancy after adjustment for race and BMI[50]. Collectively, these findings suggest that a relative deficiency in endogenous natriuretic peptide signaling may predispose to HDPs. Synthetic natriuretic peptides have been developed previously (e.g., nesiritide), and the natriuretic peptides may represent a future therapeutic target for direct or indirect modulation toward HDP prevention and/or treatment.

Third, our findings suggest several other potential novel mechanisms underlying HDP risk. Although *ZNF831* has been previously associated with hypertension, the mechanisms by which risk variants increase blood pressure have been unclear. In our conditional analysis at the *ZNF831* locus, we detected a second independent association with *ZBTB46*, which was also strongly colocalized with the lead *ZNF831* variant in multiple tissues (including arterial tissue). ZBTB46 is a transcription factor expressed in dendritic cells[51,52] and in vascular endothelial cells, consistent with our observed expression in aortic endothelial cells, and *ZBTB46* overexpression suppresses endothelial cell proliferation and angiogenesis *in vitro*[53]. Furthermore, ZBTB46 is sensitive to shear stress[53], which may have relevance to the hyperdynamic hemodynamic state of pregnancy.

In addition, we identified and replicated a novel association in the intergenic region between *PGR* and *TRPC6* (rs2508372), which has several plausible mechanistic links to the HDPs. Along with other newly identified HDP-associated loci (*TNS2*[54] and *PLCE1*[25]), *TRPC6* has also been linked to glomerular function. It has been implicated in focal segmental glomerulosclerosis and diabetic nephropathy[55] and mediates the proteinuria and renal dysfunction induced by exposure to hypertension and diabetes[56]. In addition, the progesterone receptor mediates the physiologic response to progesterone, including pregnancy-associated vasodilation[57]. As noted, *ARHGAP42* (adjacent to *PGR*), which was found to have reduced expression in preeclamptic placentas, regulates vascular tone and blood pressure[38]. Furthermore, this same intergenic lead variant has been associated with reduced circulating levels of matrix metalloproteinase-10[58], which mediates vascular endothelial growth factor

(VEGF)-induced angiogenesis[59]. Further research is necessary to clarify which of these potential mechanisms primarily mediate the preeclampsia/eclampsia risk associated with the *PGR/TRPC6* locus.

Fourth, associations at *MICA* and *SH2B3* highlight the role of immune function in preeclampsia pathogenesis. Differences in T-cell phenotypes and circulating pro- and anti-inflammatory cytokines in preeclampsia are well described[9,10]. We detected an association near *MICA* (encoding major histocompatibility complex class I polypeptide-related sequence A) within the *HLA* region, which has also been linked to multiple autoimmune diseases and may reflect the importance of maternal immune tolerance of fetal cells at the maternal-fetal interface in early placentation[10]. Indeed, we found that preeclampsia/eclampsia PRS predicted type 1 diabetes, celiac disease, and hypothyroidism. Subclinical thyroid dysfunction has been associated with increased risk for preeclampsia[60], and previous Finnish data suggested term preeclampsia predicts incident hypothyroidism later in life[61]. SH2B3 (also known as LNK) is expressed primarily in endothelial and hematopoietic cells and negatively regulates cytokine signaling; reduced SH2B3 function has been linked to atherosclerosis as well as several autoimmune diseases[62]. The lead variant in our preeclampsia/eclampsia GWAS at *SH2B3* is in linkage disequilibrium ($D'$=0.96, $R^2$=0.91)[63] with the well-described coronary artery disease risk allele at this locus (rs3184504)[64]. Our lead *SH2B3* variant was also associated with increased leukocyte, red cell blood, and platelet counts in the UK Biobank (nealelab.is/uk-biobank) and with heightened levels of vascular cell adhesion protein 1, platelet glycoprotein IB alpha chain, interleukin-2 receptor, and lymphocyte activation gene 3 protein, among other immune-related proteins[49]. Recent data indicate that reduced SH2B3 function promotes neutrophil extracellular trap formation, a process implicated in preeclampsia pathogenesis[65,66], and arterial thrombosis[67]. This process was dependent on the presence of oxidized phospholipids[67], which are increased in the setting of preeclampsia vs. normotensive pregnancy[68,69].

Fifth, we observed that polygenic risk scores for preeclampsia/eclampsia and gestational hypertension were predictive in multiple external test cohorts and independent of first-trimester risk indicators (blood pressure and BMI), with women in the top PRS decile demonstrating 65-85% greater risk of preeclampsia/eclampsia vs. other women. Predictive accuracy of clinical risk factors for HDPs is modest[70,71]. Meta-analysis indicates that personal history of preeclampsia, pre-pregnancy chronic hypertension, and pre-pregnancy chronic diabetes are associated with relative risks of 8.4, 5.1, and 3.7, respectively, for development of preeclampsia. However, nulliparity carries the largest population attributable fraction of associated risk for preeclampsia (approximately one-third)[12], and most affected women lack any overt pre-pregnancy risk factors other than nulliparity[6]. Improving pregnancy risk prediction, especially in the first pregnancy, therefore remains a pressing clinical need to optimize HDP prevention[42,72]. Low-dose aspirin after 12 weeks' gestation represents one evidence-based strategy to mitigate risk of preeclampsia and preterm birth in women at increased risk for preeclampsia[42]. Although our GWAS included substantially more non-European individuals than prior GWAS, >75% of individuals in discovery cohorts were European, and as such, the preeclampsia/eclampsia and gestational hypertension PRS generally performed better in European ancestry women vs. others, consistent with many prior published PRS and a well-recognized challenge in contemporary genetics[73]. However, incorporating PRS of a highly correlated phenotype (here, SBP) improved HDP prediction. Ongoing efforts to incorporate more non-European individuals in HDP genetic discovery is likely to improve both genetic discovery and cross-ancestry polygenic prediction[74].

This study should be considered in the context of other limitations. Due to HDP phenotyping limitations in large datasets using ICD code-based ascertainment, some participants may have had preeclampsia superimposed on chronic hypertension rather than *de novo* preeclampsia, which may enrich genetic associations for hypertension predilection. We were unable to examine more granular HDP subtypes such as preeclampsia with severe

features, HELLP (hemolysis, elevated liver enzymes, and low platelets) syndrome, preterm vs. term vs. postpartum onset, or HDP with vs. without accompanying fetal growth restriction. The underlying pathophysiology of the HDPs is heterogeneous and may vary across these subtypes; future adequately powered studies should examine these subtypes separately as implications for pregnancy care and long-term maternal health risk may differ[75].

Overall, multi-ancestry genome-wide meta-analysis of preeclampsia/eclampsia and gestational hypertension revealed novel distinct and overlapping risk loci and enabled polygenic prediction of the HDPs, with implications for HDP prediction, prevention, and treatment.

## Methods

*Study cohorts, phenotypes, genotyping, and association analysis.*

Preeclampsia/eclampsia and gestational hypertension case and control counts and definitions for each cohort are summarized in **Supplementary Tables 1-2**. If an individual had qualifying codes for both preeclampsia/eclampsia and gestational hypertension, she was designated as having preeclampsia/eclampsia. In multi-ancestry cohorts, association analyses were performed within each ancestry group separately and subsequently meta-analyzed.

*FinnGen:* Cases were identified using ICD-8, -9, and -10 codes; controls were women in FinnGen without codes corresponding to hypertension and/or proteinuria during pregnancy (ICD-10 O10-O16). Sample genotyping in FinnGen was performed using Illumina (Illumina Inc., San Diego, CA) and Affymetrix (Thermo Fisher Scientific, Santa Clara, CA, USA) arrays. Genotype calls were made using GenCall or zCall for Illumina and AxiomGT1 algorithm for Affymetrix data as described previously[76,77]. Individuals were removed for ambiguous sex, genotype missingness >5%, heterozygosity >±4 SD, and non-Finnish ancestry. Variants were removed for missingness >2%, Hardy-Weinberg equilibrium (HWE) $P<1x10^{-6}$, and minor allele count (MAC) <3. Pre-phasing was performed with Eagle v2.3.5 using 20,000 conditioning haplotypes. Genotypes were imputed with Beagle 4.1 using the population-specific Sequencing

Initiative Suomi v3 imputation reference panel. Association analyses were performed using SAIGE[78] with adjustment for age, genotyping batch, and PC 1-10.

*Estonian Biobank:* The Estonian Biobank is a population-based biobank with over 200,000 participants[79]. Preeclampsia/eclampsia cases were identified using ICD-10 O14-O15. Gestational hypertension cases were defined using ICD-10 O13. Controls were defined using ICD-10 O80-O84 and no codes for HDPs. All Estonian Biobank participants have been genotyped at the Core Genotyping Lab of the Institute of Genomics, University of Tartu, using the Illumina Global Screening Array v1.0, v2.0, and v2.0_EST arrays. Samples were genotyped and PLINK format files were created using Illumina GenomeStudio v2.0.4. Individuals were excluded from the analysis if their call-rate was <95% or if sex defined based on heterozygosity of X chromosome did not match sex in phenotype data. Before imputation, variants were filtered by call-rate <95%, HWE $P<1x10^{-4}$, and MAF <1%. We also used the MAC filter --minMAC=5. Variant positions were updated to genome build 37 and all variants were changed to be from TOP strand using GSAMD-24v1-0_20011747_A1-b37.strand.RefAlt.zip files from https://www.well.ox.ac.uk/~wrayner/strand/. Pre-phasing was performed with Eagle v2.3 software using 20,000 conditioning haplotypes, and imputation was done using Beagle v.28Sep18.793 with effective population size $n_e$=20,000. Population specific imputation reference of 2297 WGS samples was used. Analyses were carried out with SAIGE[78], adjusting for year of birth and PC 1-10.

*Genes & Health:* Genes & Health is a cohort of British Pakistani and Bangladeshi individuals recruited primarily in East London, England[80]. Cases and parous controls were identified using qualifying ICD-10 and SNOMED codes (**Supplementary Tables 1-2**). Genotyping was performed using the Illumina Infinium Global Screening Array v3.0, and quality control was performed using Illumina GenomeStudio and PLINK v1.9. Individuals were removed who did not have Pakistani or Bangladeshi ancestry defined as >±3 SD from the mean of PC1 and those who self-reported another ethnicity. Variants with call rate <0.99, MAF <1%, and

HWE $P<1x10^{-6}$ were removed. Imputation was performed using the Michigan Imputation Server with the GenomeAsia reference panel. Association analyses were performed using SAIGE[78] with adjustment for age, age$^2$, and PC 1-10.

*Michigan Genomics Initiative:* The Michigan Genomics Initiative enrolls participants receiving care at Michigan Medicine and links biospecimen data to electronic health record (EHR) data. Preeclampsia cases were identified in the freeze 3 dataset using phecode 642.1 and its corresponding ICD codes[81], and women lacking this phecode served as controls. Genotyping was performed using one of two versions of the Illumina Infinium CoreExome-24 bead array platform. Relatedness within the cohort was estimated using KING v2.1.3. Individuals were removed for discordant, missing, or ambiguous sex, kinship coefficient >0.45 with another participant, call rate <99%, estimated contamination >2.5%, or missingness on any chromosome >5%. Variants were excluded with poor intensity separation based on metrics from GenomeStudio (GenTrain score <0.15 or Cluster Separation score <0.3), overall call rate <99%, or HWE *P*-value $<1x10^{-4}$. Genotypes were phased using EAGLE v2.4.1 and imputed using the TOPMed reference panel. Association analysis was conducted using SAIGE[78] with adjustment for age, genotype array, and PC 1-10.

*Mass General Brigham Biobank:* The Mass General Brigham Biobank is a health system-based biobank linking genomic data to EHR data. Preeclampsia/eclampsia cases were identified using ICD-10 O14-O15, gestational hypertension cases were identified using ICD-10 O13, and controls were identified using ICD-10 O80-O82. Variants with MAF <1%, missingness per variant >1%, and HWE *P*-value $<10^{-6}$ were removed. Imputation was performed using the TOPMed reference panel. Association analysis was performed with variants filtered by MAC ≥50 and INFO score ≥0.6 using REGENIE[82] with adjustment for age, genotype batch, and PC 1-10.

*Biobank Japan:* Biobank Japan is a biobank of approximately 200,000 Japanese adults. Preeclampsia cases were identified using phecode 642 and its corresponding ICD codes, and

other women constituted the control group[83]. Genotyping was performed using the Illumina HumanOmniExpressExome BeadChip or a combination of the Illumina HumanOmniExpress and HumanExome BeadChip. Individuals with call rates <98% or closely related individuals (PI_HAT >0.175 in PLINK) were excluded. Variants with call rate <99%, HWE $P<1.0\times10^{-6}$, and number of heterozygotes <5 were excluded. Genotype data were imputed with 1000 Genomes Project Phase 3 v5 genotype data and Japanese whole-genome sequencing data. Association analysis was performed using SAIGE[78] with adjustment for age, age$^2$, and PC 1-20.

*BioMe:* Bio*Me* is a health system-based biobank at the Icahn School of Medicine at Mount Sinai in New York, NY, USA. Preeclampsia/eclampsia and gestational hypertension cases were identified using ICD-10 codes O14-O15 and O13, respectively. Controls were other women enrolled in Bio*Me* who indicated a history of childbirth on their intake questionnaire. Genotyping was performed using the Illumina Global Screening Array. Individuals with ethnicity-specific heterozygosity rate that surpassed ±3 SD of the population-specific mean, those with a call rate of ≤95%, and those with discordance between EMR-recorded and genetic sex were removed. For variant-level quality control, sites with a call rate below 95% and sites with HWE $P<1\times10^{-8}$ were excluded. Imputation of variants was then performed with the Michigan Imputation Server pipeline using the TOPMed reference panel. Association analysis was performed separately in African-, Hispanic-, and European-ancestry women using SAIGE[78] with adjustment for age, age$^2$, and PC 1-10.

*InterPregGen consortium:* We incorporated summary statistics for preeclampsia from the discovery GWAS meta-analyses of European cohorts and Central Asian cohorts from the InterPregGen consortium[17]. European-ancestry discovery cohorts included GOPEC (United Kingdom), deCODE (Iceland), the Avon Longitudinal Study of Parents and Children (United Kingdom), MoBa (Norway), SSI (Denmark), and FINRISK (Finland) (7,219 cases and 155,620 controls). Central Asian cohorts included two Kazakh cohorts and one Uzbek cohort (2,296 cases and 2,059 controls). Study-specific preeclampsia definitions are summarized in

Steinthorsdottir et al[17]. Fixed-effects inverse variance-weighted meta-analysis was performed in METAL[22].

*HUNT:* The HUNT study is a population-based cohort study in Nord-Trøndelag County, Norway. Genotyped, parous, European-ancestry women were included in the present analysis. Preeclampsia/eclampsia was ascertained by linkage to the Medical Birth Registry of Norway, which defines preeclampsia as systolic blood pressure ≥140 mmHg and/or diastolic blood pressure ≥90 mmHg accompanied by proteinuria >0.3 g/24 h or >1+ on urine dipstick[84]. The current analysis includes genetic data from approximately 90% of participants from HUNT2 (1995-1997) and HUNT3 (2006-2008) who were genotyped by genome-wide SNP arrays in 2015[85,86]. Genotyping, quality control metrics, and imputation have been described previously[85]. Briefly, one of three different Illumina HumanCoreExome arrays (HumanCoreExome12 v1.0, HumanCoreExome12 v1.1, and UM HUNT Biobank v1.0) were used for genotyping the HUNT2 and HUNT3 samples[85]. Sample sand variants with call rate <99% were excluded. Imputation was performed using 2,201 HUNT samples with whole genome sequencing, the Haplotype Reference Consortium, and TOPMed imputation panel (MAC >10). Association analysis for GWAS replication was performed using SAIGE[78] with adjustment for age, age$^2$, and PC 1-10.

*UK Biobank:* The UK Biobank is a population-based cohort study of adult residents of the UK aged 40-69 years at the time of recruitment between 2006-2010. Preeclampsia/eclampsia and gestational hypertension cases were identified using qualifying ICD codes (Supplementary Table 1); controls were women self-reporting a history of at least one live birth at the time of study enrollment without qualifying ICD codes for preeclampsia/eclampsia or gestational hypertension or any self-reported history of HDPs. Genotyping was performed using the UK BiLEVE Axiom Array or the UK Biobank Axiom Array (both Affymetrix, Santa Clara, CA). Individuals with single nucleotide variant missingness ≥10% were excluded. Imputation was performed centrally using the Haplotype Reference Consortium, UK10K, and 1000 Genome reference panels[87]. Variants were required to pass the following

quality control filters: MAF ≥1%, single nucleotide variant missingness <10%, and HWE $P \geq 10^{-15}$, MAC ≥50, and INFO score ≥0.6. Association analysis for GWAS replication was performed in European-ancestry participants using REGENIE[82] with adjustment for age, genotyping array, and PC 1-10.

*Penn Medicine Biobank (PMBB):* PMBB is a health system-based biobank at the University of Pennsylvania, Philadelphia, PA, USA. Preeclampsia/eclampsia cases were identified using ICD-10 codes O14-O15 and O13 or phecode 642.1. Gestational hypertension cases were identified using ICD-10 code O13. Controls were other women enrolled in PMBB. Genotyping was performed using the Illumina Global Sequencing Array v2.0; genotype data were imputed to the TOPMed reference panel using the Michigan Imputation Server. Variants with MAF<1%, missing rate >10%, and HWE $P<10^{-8}$ were filtered from the GWAS. Association analysis for GWAS replication was performed in REGENIE separately for African-ancestry and European-ancestry women with adjustment for age, age$^2$, and PC 1-5.

*nuMoM2b:* The nuMoM2b study is a prospective U.S. pregnancy cohort of nulliparous women enrolled in the first trimester of pregnancy between 2012-2015. HDPs were determined by chart abstraction and adjudication according to previously published definitions[88]. Gestational hypertension was defined as SBP ≥140 mmHg or DBP ≥90 mmHg on two occasions ≥6 hours apart or one occasion with subsequent antihypertensive therapy after 20 weeks' gestation, excluding blood pressures recorded during the second stage of labor, without other qualifying features for preeclampsia or eclampsia. Preeclampsia was defined according to the same blood pressure criteria plus proteinuria or other findings meeting criteria for severe features, including HELLP syndrome[88]. Women documented as meeting these same blood pressure criteria before 20 weeks' gestation were designated as having chronic hypertension. All women were at risk for development of preeclampsia/eclampsia; only those without chronic hypertension were at risk for the outcome of gestational hypertension. Genotyping was performed using the Illumina Infinium Multi-Ethnic Global D2 BeadChip. Individuals related

within two degrees by the KING algorithm were removed. Variants with MAF <1%, genotyping rate <95%, and HWE $P<5x10^{-6}$ were removed[89]. After phasing with EAGLE, imputation was performed for European-, African-, and admixed American-ancestry subjects using the TOPMed reference panel via the TOPMed Imputation Server. Women in nuMoM2b with imputed genotypes were used for external testing of optimized PRS (described below).

*Genome-wide meta-analysis and replication.*

Variants from GWAS summary statistics were matched by genome build 38 position and alleles. GWAS summary statistics that were in genome build 37 were lifted to genome build 38 using UCSC liftOver (https://genome.ucsc.edu/cgi-bin/hgLiftOver). We used METAL[22] to perform fixed-effect inverse variance-weighted meta-analysis. Correction for genomic inflation factor was carried out prior to meta-analysis. Meta-analyses were conducted among discovery cohorts, among follow-up cohorts, and across all cohorts. Given the potential overlap of 400 preeclampsia/eclampsia cases from FINRISK between the InterPregGen meta-analysis and FinnGen (8.4% of FinnGen cases and 2.3% of overall discovery cases) and 7,805 controls (5.7% of FinnGen controls and 1.7% of overall discovery controls), we conducted a sensitivity analysis excluding FinnGen (Supplementary Table 3).

Lead variants for preeclampsia/eclampsia and gestational hypertension with $P<5x10^{-8}$ in multi-ancestry meta-analysis of discovery cohorts were interrogated in multi-ancestry meta-analysis of follow-up cohorts using METAL: HUNT (preeclampsia/eclampsia only), UK Biobank, and PMBB (European- and African-ancestry women). Given strong genetic correlation between preeclampsia/eclampsia and gestational hypertension ($r_g$=0.71) in discovery cohorts, associations for preeclampsia/eclampsia and gestational hypertension were meta-analyzed in follow-up cohorts for primary replication with correction applied for overlapping control samples across phenotypes in UK Biobank and PMBB, with $P<0.05$ and consistent direction of effect in follow-up cohorts indicating replication. Secondarily, we meta-analyzed preeclampsia/eclampsia

in HUNT with gestational hypertension in UK Biobank and PMBB (Supplementary Table 4), examined preeclampsia/eclampsia and gestational hypertension in follow-up cohorts separately, and performed combined meta-analysis of discovery and follow-up cohorts for preeclampsia/eclampsia and for gestational hypertension.

*Conditional and joint analysis.*

We conducted conditional analysis using GCTA-COJO[23] on the multi-ancestry discovery meta-analyses of preeclampsia/eclampsia and gestational hypertension to identify additional association signals at the genome-wide significant loci. We used the European LD reference from a randomly selected set of 10,000 unrelated individuals. The LD panel included variants with MAF > 1% and INFO ≥0.3. The analysis was restricted to variants within ±1 Mb from lead variants ($P<5x10^{-8}$). In COJO, the lead variants were conditioned from each chromosome and independent variants were iteratively included. All variants were then simultaneously fitted in the joint analysis. Variants with $P<2x10^{-7}$ were considered genome-wide significant. One additional variant with conditional $P<2x10^{-7}$ for association with preeclampsia/eclampsia was identified on chromosome 20 ($P=1.4x10^{-8}$).

*Genetic correlation.*

We used LD score regression (LDSC, https://github.com/bulik/ldsc)[28] with pre-computed LD scores for 1.2 million HapMap3 variants after excluding MHC region in the European population (https://data.broadinstitute.org/alkesgroup/LDSCORE/eur_w_ld_chr.tar.bz2) to calculate genetic correlation between preeclampsia/eclampsia and gestational hypertension and correlation of each HDP with systolic and diastolic blood pressure. In addition, the LDSC intercept indicates potential confounding due to potential population stratification and cryptic relatedness. We observed intercepts of 1.03 for preeclampsia/eclampsia and 0.97 for gestational hypertension.

*Colocalization analysis.*

We obtained the tissue-specific gene expression from GTEx data portal for 52 tissues[31]. We used marginal effect sizes, standard errors, and MAF for all SNPs within ±500 of lead variants from discovery analysis ($P<5x10^{-8}$) as the input. We performed colocalization using the coloc.abf() function in R package coloc v4. The $H_4$ test statistic estimates the posterior probability of a shared causal variant between preeclampsia/eclampsia or gestational hypertension and expression of a particular gene. $H_4$ >0.7 indicated strong evidence of colocalization, $H_4$ 0.5-<0.7 indicated weak evidence of colocalization, and $H_4$ <0.5 indicated no colocalization.

*Polygenic prioritization of causal genes.*

We performed additional causal gene prioritization using the PoPS method[33]. Briefly, PoPS integrates GWAS summary statistics with gene expression, biology pathways, and predicted protein-protein interaction data to identify likely causal genes at genome-wide significant loci. A linear model was trained to predict gene-level association scores and estimate Z-scores indicating the confidence of the causal role at a given locus. In total, PoPS scores were calculated for 18,000 genes. The top 5 available prioritized genes within 500 kilobases of lead preeclampsia/eclampsia and gestational hypertension variants were extracted and compared with the results of other *in silico* analyses.

*Placental transcriptome data.*

Genes nearest to lead variants, colocalization hits, and genes prioritized within the top 5 by either Open Targets variant-to-gene score or PoPS score were queried in a publicly available database of placental gene expression (https://www.obgyn.cam.ac.uk/placentome/)[34]. Samples were obtained from a prospective cohort of nulliparous women in Cambridge, United Kingdom.

Differential expression analysis for preeclampsia included 82 preeclampsia cases and 82 controls samples matched on presence of labor, Cesarean section, gestational age, fetal sex, smoking status, maternal BMI, and maternal age. RNA sequencing was performed on placental biopsy specimens with a median sequencing depth of 101 million reads per sample. Differential expression (in $\log_2$-fold change) and corresponding *P*-values were generated using DESeq2; *P*-values were then adjusted for multiple comparisons using the Benjamini-Hochberg method. We report differentially expressed prioritized genes with adjusted $P < 0.05$ (Supplementary Table 11).

*Gene expression in human aortic tissue.*

We queried expression of prioritized genes in subpopulations of vascular smooth muscle cells, fibromyocytes, fibroblasts, endothelial cells, macrophages, natural killer T cells, and neuronal cells in a dataset of single-nuclei RNA sequencing from human aortic tissue. In total, 1,114 unique molecular identifiers were obtained per cell. Prioritized genes were nearest genes, genes with strong colocalization, genes with weak colocalization plus prioritization by another method (top 5 Open Targets variant-to-gene score or PoPS), and genes prioritized by another method plus statistically significant differential placental transcription in the human placental transcriptome browser[34]'; of these genes, 24 were available in the single-nuclei RNA sequencing dataset (Extended Data Fig. 1). Single-nuclei RNA sequencing was performed on non-atherosclerotic aortic root tissue obtained during coronary artery bypass graft surgery. A total of 4537 nuclei were obtained for downstream analysis. Cell types and subtypes were defined using top marker genes and pathway enrichment scores. Raw *Cellranger* output data were filtered for removal of ambient RNA using *cellbender* in 'full' running mode. The resultant filtered cell-gene matrix was used for quality control and downstream analysis. All preliminary quality-control and clustering were performed using *scanpy*. Any cells with fewer than 300 genes captured or greater than 0.1% mitochondrial reads were excluded from the analysis.

Each sample was processed with *Scrublet* to exclude doublets. The top 10,000 variable genes were used for analysis. Relative expression of queried genes in each cell type against other cell types in normal aortic tissue was quantified as Z-scores.

*Derivation and testing of genome-wide polygenic risk scores.*

We used PRS-CS to derive genome-wide PRS for preeclampsia/eclampsia and gestational hypertension from the corresponding discovery GWAS summary statistics and for systolic blood pressure from the Million Veteran Program GWAS summary statistics[29]. PRS-CS is a Bayesian approach that estimates posterior effect sizes by placing a continuous shrinkage prior on variant effect sizes[40]. The preeclampsia/eclampsia PRS included 1,087,033 HapMap3 variants, the gestational hypertension PRS included 1,087,916 HapMap3 variants, and the SBP PRS included 1,064,898 HapMap3 variants. PRS were trained on the UK Biobank European LD panel. Individual-level polygenic scores were generated in the tuning and test datasets as the sum of genotypes*weights using PLINK. We used logistic regression to test the association of each PRS with preeclampsia/eclampsia and gestational hypertension with adjustment for age, age$^2$, and PC 1-10. PRS were tuned in the UK Biobank. Specifically, we used a small-scale grid search of global shrinkage parameter □ values (1, $1\times10^{-2}$, $1\times10^{-4}$, and $1\times10^{-6}$) for each PRS as recommended to identify the □ that produced the best predictive performance as measured by $R^2$ in the tuning dataset. We then fitted a linear combination of optimized preeclampsia/eclampsia PRS and SBP PRS for the outcome of preeclampsia/eclampsia and a linear combination of gestational hypertension PRS and SBP PRS for the outcome of gestational hypertension. The optimal linear combination derived for preeclampsia/eclampsia was $0.1889*Z_{\text{preeclampsia/eclampsia}} + 0.1864*Z_{\text{SBP}}$, and the linear combination derived for gestational hypertension was $0.1662*Z_{\text{gestational_hypertension}} + 0.3050*Z_{\text{SBP}}$. We then carried these weighed linear combination scores forward for final testing in nuMoM2b (European, African, and admixed American ancestry groups) and HUNT (European ancestry). PRS performance was evaluated

using the OR for top decile vs. bottom 90% of PRS, the OR per SD of PRS, and Nagelkerke's $R^2$.

We tested whether PRS correctly re-classified nuMoM2b participants with HDPs as aspirin-eligible in comparison with the major criteria endorsed by the U.S. Preventive Services Task Force[42]. Major criteria include: history of preeclampsia, which does not apply in nuMoM2b as all women were nulliparous; multifetal gestation, which does not apply in nuMoM2b as all women had singleton pregnancies; chronic hypertension, defined as a diagnosis of hypertension before pregnancy or blood pressure ≥140/90 mmHg on two occasions at least 6 hours apart prior to 20 weeks' gestation; pre-gestational diabetes type 1 or type 2; any pre-pregnancy kidney disease; and autoimmune disease, defined here as antiphospholipid antibody syndrome, systemic lupus erythematosus, rheumatoid arthritis, inflammatory bowel disease (Crohn's disease or ulcerative colitis), or "other collagen vascular or autoimmune disease." We calculated net reclassification for composite HDPs vs. normotensive pregnancy and for preeclampsia/eclampsia vs. all other pregnancies when $PRS_{preeclampsia+SBP}$ was added to major risk factors to identify women as eligible for preventive aspirin therapy. In this context, net reclassification was calculated as $P$(up|case) - $P$(up|non-case). Bootstrap resampling performed 1,000 times was used to estimate 95% confidence intervals for net reclassification.

*Phenome-wide association analysis.*

We tested the association of preeclampsia/eclampsia PRS and gestational hypertension PRS with 1,445 phecode-based combined prevalent and incident phenotypes[90] in sex-stratified fashion among genotyped UK Biobank participants with adjustment for age and PC 1-5 using the PheWAS package[91] in R 3.6.0 (https://github.com/PheWAS/PheWAS). Bonferroni-corrected $P<0.05/1{,}445 = 3.5\times10^{-5}$ indicated statistical significance.

## References


1. Burton GJ, Redman CW, Roberts JM, Moffett A. Pre-eclampsia: pathophysiology and clinical implications. *BMJ*. 07 2019;366:l2381. doi:10.1136/bmj.l2381
2. Jiang L, Tang K, Magee LA, et al. A global view of hypertensive disorders and diabetes mellitus during pregnancy. *Nat Rev Endocrinol*. Sep 15 2022;doi:10.1038/s41574-022-00734-y
3. Garovic VD, White WM, Vaughan L, et al. Incidence and Long-Term Outcomes of Hypertensive Disorders of Pregnancy. *J Am Coll Cardiol*. 05 2020;75(18):2323-2334. doi:10.1016/j.jacc.2020.03.028
4. Fraser A, Nelson SM, Macdonald-Wallis C, et al. Associations of pregnancy complications with calculated cardiovascular disease risk and cardiovascular risk factors in middle age: the Avon Longitudinal Study of Parents and Children. *Circulation*. Mar 2012;125(11):1367-80. doi:10.1161/CIRCULATIONAHA.111.044784
5. Magee LA, Brown MA, Hall DR, et al. The 2021 International Society for the Study of Hypertension in Pregnancy classification, diagnosis & management recommendations for international practice. *Pregnancy Hypertens*. Mar 2022;27:148-169. doi:10.1016/j.preghy.2021.09.008
6. ACOG Practice Bulletin No. 202: Gestational Hypertension and Preeclampsia. *Obstet Gynecol*. Jan 2019;133(1):e1-e25. doi:10.1097/AOG.0000000000003018
7. Honigberg MC, Zekavat SM, Aragam K, et al. Long-Term Cardiovascular Risk in Women With Hypertension During Pregnancy. *J Am Coll Cardiol*. Dec 2019;74(22):2743-2754. doi:10.1016/j.jacc.2019.09.052
8. Wang YX, Arvizu M, Rich-Edwards JW, et al. Hypertensive Disorders of Pregnancy and Subsequent Risk of Premature Mortality. *J Am Coll Cardiol*. Mar 16 2021;77(10):1302-1312. doi:10.1016/j.jacc.2021.01.018
9. Ives CW, Sinkey R, Rajapreyar I, Tita ATN, Oparil S. Preeclampsia-Pathophysiology and Clinical Presentations: JACC State-of-the-Art Review. *J Am Coll Cardiol*. 10 06 2020;76(14):1690-1702. doi:10.1016/j.jacc.2020.08.014
10. Rana S, Lemoine E, Granger J, Karumanchi SA. Preeclampsia. *Circ Res*. Mar 2019;124(7):1094-1112. doi:10.1161/CIRCRESAHA.118.313276
11. Levine RJ, Maynard SE, Qian C, et al. Circulating angiogenic factors and the risk of preeclampsia. *N Engl J Med*. Feb 2004;350(7):672-83. doi:10.1056/NEJMoa031884
12. Bartsch E, Medcalf KE, Park AL, Ray JG, Group HRoP-eI. Clinical risk factors for pre-eclampsia determined in early pregnancy: systematic review and meta-analysis of large cohort studies. *BMJ*. Apr 2016;353:i1753. doi:10.1136/bmj.i1753
13. Jung E, Romero R, Yeo L, et al. The etiology of preeclampsia. *Am J Obstet Gynecol*. 02 2022;226(2S):S844-S866. doi:10.1016/j.ajog.2021.11.1356
14. Cnattingius S, Reilly M, Pawitan Y, Lichtenstein P. Maternal and fetal genetic factors account for most of familial aggregation of preeclampsia: a population-based Swedish cohort study. *Am J Med Genet A*. Nov 2004;130A(4):365-71. doi:10.1002/ajmg.a.30257
15. Nilsson E, Salonen Ros H, Cnattingius S, Lichtenstein P. The importance of genetic and environmental effects for pre-eclampsia and gestational hypertension: a family study. *BJOG*. Mar 2004;111(3):200-6. doi:10.1111/j.1471-0528.2004.00042x.x
16. McGinnis R, Steinthorsdottir V, Williams NO, et al. Variants in the fetal genome near FLT1 are associated with risk of preeclampsia. *Nat Genet*. Aug 2017;49(8):1255-1260. doi:10.1038/ng.3895
17. Steinthorsdottir V, McGinnis R, Williams NO, et al. Genetic predisposition to hypertension is associated with preeclampsia in European and Central Asian women. *Nat Commun*. 11 25 2020;11(1):5976. doi:10.1038/s41467-020-19733-6
18. Dina C, Meyre D, Gallina S, et al. Variation in FTO contributes to childhood obesity and severe adult obesity. *Nat Genet*. Jun 2007;39(6):724-6. doi:10.1038/ng2048

19. Honigberg MC, Chaffin M, Aragam K, et al. Genetic Variation in Cardiometabolic Traits and Medication Targets and the Risk of Hypertensive Disorders of Pregnancy. *Circulation*. Aug 2020;142(7):711-713. doi:10.1161/CIRCULATIONAHA.120.047936
20. Gray KJ, Kovacheva VP, Mirzakhani H, et al. Risk of pre-eclampsia in patients with a maternal genetic predisposition to common medical conditions: a case-control study. *BJOG*. 01 2021;128(1):55-65. doi:10.1111/1471-0528.16441
21. O'Kelly AC, Michos ED, Shufelt CL, et al. Pregnancy and Reproductive Risk Factors for Cardiovascular Disease in Women. *Circ Res*. 02 18 2022;130(4):652-672. doi:10.1161/CIRCRESAHA.121.319895
22. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. Sep 01 2010;26(17):2190-1. doi:10.1093/bioinformatics/btq340
23. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet*. Jan 07 2011;88(1):76-82. doi:10.1016/j.ajhg.2010.11.011
24. Turley P, Walters RK, Maghzian O, et al. Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nat Genet*. 02 2018;50(2):229-237. doi:10.1038/s41588-017-0009-4
25. Padmanabhan S, Caulfield M, Dominiczak AF. Genetic and molecular aspects of hypertension. *Circ Res*. Mar 13 2015;116(6):937-59. doi:10.1161/CIRCRESAHA.116.303647
26. Rubattu S, Forte M, Marchitti S, Volpe M. Molecular Implications of Natriuretic Peptides in the Protection from Hypertension and Target Organ Damage Development. *Int J Mol Sci*. Feb 13 2019;20(4)doi:10.3390/ijms20040798
27. Ohwaki A, Nishizawa H, Kato A, et al. Altered serum soluble furin and prorenin receptor levels in pregnancies with pre-eclampsia and fetal growth restriction. *J Gynecol Obstet Hum Reprod*. Dec 2021;50(10):102198. doi:10.1016/j.jogoh.2021.102198
28. Bulik-Sullivan BK, Loh PR, Finucane HK, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet*. Mar 2015;47(3):291-5. doi:10.1038/ng.3211
29. Giri A, Hellwege JN, Keaton JM, et al. Trans-ethnic association study of blood pressure determinants in over 750,000 individuals. *Nat Genet*. 01 2019;51(1):51-62. doi:10.1038/s41588-018-0303-9
30. Bulik-Sullivan B, Finucane HK, Anttila V, et al. An atlas of genetic correlations across human diseases and traits. *Nat Genet*. Nov 2015;47(11):1236-41. doi:10.1038/ng.3406
31. Battle A, Brown CD, Engelhardt BE, et al. Genetic effects on gene expression across human tissues. *Nature*. 10 11 2017;550(7675):204-213. doi:10.1038/nature24277
32. Ghoussaini M, Mountjoy E, Carmona M, et al. Open Targets Genetics: systematic identification of trait-associated genes using large-scale genetics and functional genomics. *Nucleic Acids Res*. 01 08 2021;49(D1):D1311-D1320. doi:10.1093/nar/gkaa840
33. Weeks EM, Ulirsch JC, Cheng NY, et al. Leveraging polygenic enrichments of gene features to predict genes underlying complex traits and diseases. *medRxiv*. 2020:2020.09.08.20190561. doi:10.1101/2020.09.08.20190561
34. Gong S, Gaccioli F, Dopierala J, et al. The RNA landscape of the human placenta in health and disease. *Nat Commun*. 05 11 2021;12(1):2639. doi:10.1038/s41467-021-22695-y
35. Maynard SE, Min JY, Merchan J, et al. Excess placental soluble fms-like tyrosine kinase 1 (sFlt1) may contribute to endothelial dysfunction, hypertension, and proteinuria in preeclampsia. *J Clin Invest*. Mar 2003;111(5):649-58. doi:10.1172/JCI17189
36. Tekola-Ayele F, Zeng X, Chatterjee S, et al. Placental multi-omics integration identifies candidate functional genes for birthweight. *Nat Commun*. 05 02 2022;13(1):2384. doi:10.1038/s41467-022-30007-1

37. Zhang Z, Wang X, Zhang L, Shi Y, Wang J, Yan H. Wnt/β-catenin signaling pathway in trophoblasts and abnormal activation in preeclampsia (Review). *Mol Med Rep*. Aug 2017;16(2):1007-1013. doi:10.3892/mmr.2017.6718
38. Bai X, Lenhart KC, Bird KE, et al. The smooth muscle-selective RhoGAP GRAF3 is a critical regulator of vascular tone and hypertension. *Nat Commun*. 2013;4:2910. doi:10.1038/ncomms3910
39. Kalluri AS, Vellarikkal SK, Edelman ER, et al. Single-Cell Analysis of the Normal Mouse Aorta Reveals Functionally Distinct Endothelial Cell Populations. *Circulation*. 07 09 2019;140(2):147-163. doi:10.1161/CIRCULATIONAHA.118.038362
40. Ge T, Chen CY, Ni Y, Feng YA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat Commun*. 04 2019;10(1):1776. doi:10.1038/s41467-019-09718-5
41. Ray JG, Abdulaziz KE, Berger H, DOH-NET (Diabetes Oe, and Hypertension in Pregnancy Research Network). Aspirin Use for Preeclampsia Prevention Among Women With Prepregnancy Diabetes, Obesity, and Hypertension. *JAMA*. 01 25 2022;327(4):388-390. doi:10.1001/jama.2021.22749
42. Davidson KW, Barry MJ, Mangione CM, et al. Aspirin Use to Prevent Preeclampsia and Related Morbidity and Mortality: US Preventive Services Task Force Recommendation Statement. *JAMA*. 09 28 2021;326(12):1186-1191. doi:10.1001/jama.2021.14781
43. Pollheimer J, Loregger T, Sonderegger S, et al. Activation of the canonical wingless/T-cell factor signaling pathway promotes invasive differentiation of human trophoblast. *Am J Pathol*. Apr 2006;168(4):1134-47. doi:10.2353/ajpath.2006.050686
44. Tita AT, Szychowski JM, Boggess K, et al. Treatment for Mild Chronic Hypertension during Pregnancy. *N Engl J Med*. Apr 02 2022;doi:10.1056/NEJMoa2201295
45. Daniels LB, Maisel AS. Natriuretic peptides. *J Am Coll Cardiol*. Dec 18 2007;50(25):2357-68. doi:10.1016/j.jacc.2007.09.021
46. Zhang W, Li S, Lou J, et al. Atrial natriuretic peptide promotes uterine decidualization and a TRAIL-dependent mechanism in spiral artery remodeling. *J Clin Invest*. 10 15 2021;131(20)doi:10.1172/JCI151053
47. Maack T, Suzuki M, Almeida FA, et al. Physiological role of silent receptors of atrial natriuretic factor. *Science*. Oct 30 1987;238(4827):675-8. doi:10.1126/science.2823385
48. Gu Y, Thompson D, Xu J, et al. Aberrant pro-atrial natriuretic peptide/corin/natriuretic peptide receptor signaling is present in maternal vascular endothelium in preeclampsia. *Pregnancy Hypertens*. Jan 2018;11:1-6. doi:10.1016/j.preghy.2017.12.001
49. Sun BB, Maranville JC, Peters JE, et al. Genomic atlas of the human plasma proteome. *Nature*. 06 2018;558(7708):73-79. doi:10.1038/s41586-018-0175-2
50. Hauspurg A, Marsh DJ, McNeil RB, et al. Association of N-Terminal Pro-Brain Natriuretic Peptide Concentration in Early Pregnancy With Development of Hypertensive Disorders of Pregnancy and Future Hypertension. *JAMA Cardiol*. Mar 01 2022;7(3):268-276. doi:10.1001/jamacardio.2021.5617
51. Satpathy AT, KC W, Albring JC, et al. Zbtb46 expression distinguishes classical dendritic cells and their committed progenitors from other immune lineages. *J Exp Med*. Jun 04 2012;209(6):1135-52. doi:10.1084/jem.20120030
52. Meredith MM, Liu K, Kamphorst AO, et al. Zinc finger transcription factor zDC is a negative regulator required to prevent activation of classical dendritic cells in the steady state. *J Exp Med*. Aug 27 2012;209(9):1583-93. doi:10.1084/jem.20121003
53. Wang Y, Sun HY, Kumar S, Puerta MDM, Jo H, Rezvan A. ZBTB46 is a shear-sensitive transcription factor inhibiting endothelial cell proliferation via gene expression regulation of cell cycle proteins. *Lab Invest*. 03 2019;99(3):305-318. doi:10.1038/s41374-018-0060-5

54. Ashraf S, Kudo H, Rao J, et al. Mutations in six nephrosis genes delineate a pathogenic pathway amenable to treatment. *Nat Commun*. 05 17 2018;9(1):1960. doi:10.1038/s41467-018-04193-w
55. Hall G, Wang L, Spurney RF. TRPC Channels in Proteinuric Kidney Diseases. *Cells*. 12 23 2019;9(1)doi:10.3390/cells9010044
56. Wang Z, Fu Y, do Carmo JM, et al. Transient receptor potential cation channel 6 contributes to kidney injury induced by diabetes and hypertension. *Am J Physiol Renal Physiol*. 01 01 2022;322(1):F76-F88. doi:10.1152/ajprenal.00296.2021
57. Pang Y, Dong J, Thomas P. Progesterone increases nitric oxide synthesis in human vascular endothelial cells through activation of membrane progesterone receptor-α. *Am J Physiol Endocrinol Metab*. May 15 2015;308(10):E899-911. doi:10.1152/ajpendo.00527.2014
58. Folkersen L, Gustafsson S, Wang Q, et al. Genomic and drug target evaluation of 90 cardiovascular proteins in 30,931 individuals. *Nat Metab*. 10 2020;2(10):1135-1148. doi:10.1038/s42255-020-00287-2
59. Heo SH, Choi YJ, Ryoo HM, Cho JY. Expression profiling of ETS and MMP factors in VEGF-activated endothelial cells: role of MMP-10 in VEGF-induced angiogenesis. *J Cell Physiol*. Sep 2010;224(3):734-42. doi:10.1002/jcp.22175
60. Toloza FJK, Derakhshan A, Männistö T, et al. Association between maternal thyroid function and risk of gestational hypertension and pre-eclampsia: a systematic review and individual-participant data meta-analysis. *Lancet Diabetes Endocrinol*. 04 2022;10(4):243-252. doi:10.1016/S2213-8587(22)00007-9
61. Männistö T, Karumanchi SA, Pouta A, et al. Preeclampsia, gestational hypertension and subsequent hypothyroidism. *Pregnancy Hypertens*. Jan 01 2013;3(1):21-27. doi:10.1016/j.preghy.2012.09.001
62. Wang W, Tang Y, Wang Y, et al. LNK/SH2B3 Loss of Function Promotes Atherosclerosis and Thrombosis. *Circ Res*. Sep 02 2016;119(6):e91-e103. doi:10.1161/CIRCRESAHA.116.308955
63. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics*. Nov 01 2015;31(21):3555-7. doi:10.1093/bioinformatics/btv402
64. Deloukas P, Kanoni S, Willenborg C, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet*. Jan 2013;45(1):25-33. doi:10.1038/ng.2480
65. Gupta AK, Hasler P, Holzgreve W, Hahn S. Neutrophil NETs: a novel contributor to preeclampsia-associated placental hypoxia? *Semin Immunopathol*. Jun 2007;29(2):163-7. doi:10.1007/s00281-007-0073-4
66. Palomo M, Youssef L, Ramos A, et al. Differences and similarities in endothelial and angiogenic profiles of preeclampsia and COVID-19 in pregnancy. *Am J Obstet Gynecol*. 08 2022;227(2):277.e1-277.e16. doi:10.1016/j.ajog.2022.03.048
67. Dou H, Kotini A, Liu W, et al. Oxidized Phospholipids Promote NETosis and Arterial Thrombosis in LNK(SH2B3) Deficiency. *Circulation*. 12 14 2021;144(24):1940-1954. doi:10.1161/CIRCULATIONAHA.121.056414
68. Anand S, Young S, Esplin MS, et al. Detection and confirmation of serum lipid biomarkers for preeclampsia using direct infusion mass spectrometry. *J Lipid Res*. Apr 2016;57(4):687-96. doi:10.1194/jlr.P064451
69. He B, Liu Y, Maurya MR, et al. The maternal blood lipidome is indicative of the pathogenesis of severe preeclampsia. *J Lipid Res*. 2021;62:100118. doi:10.1016/j.jlr.2021.100118
70. Wright D, Syngelaki A, Akolekar R, Poon LC, Nicolaides KH. Competing risks model in screening for preeclampsia by maternal characteristics and medical history. *Am J Obstet Gynecol*. Jul 2015;213(1):62.e1-62.e10. doi:10.1016/j.ajog.2015.02.018

71. Al-Rubaie ZTA, Hudson HM, Jenkins G, et al. Prediction of pre-eclampsia in nulliparous women using routinely collected maternal characteristics: a model development and validation study. *BMC Pregnancy Childbirth*. Jan 06 2020;20(1):23. doi:10.1186/s12884-019-2712-x
72. Myatt L. The prediction of preeclampsia: the way forward. *Am J Obstet Gynecol*. 02 2022;226(2S):S1102-S1107.e8. doi:10.1016/j.ajog.2020.10.047
73. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet*. 04 2019;51(4):584-591. doi:10.1038/s41588-019-0379-x
74. Graham SE, Clarke SL, Wu KH, et al. The power of genetic diversity in genome-wide association studies of lipids. *Nature*. Dec 2021;600(7890):675-679. doi:10.1038/s41586-021-04064-3
75. Roberts JM, Rich-Edwards JW, McElrath TF, Garmire L, Myatt L, Collaboration GP. Subtypes of Preeclampsia: Recognition and Determining Clinical Usefulness. *Hypertension*. 05 05 2021;77(5):1430-1441. doi:10.1161/HYPERTENSIONAHA.120.14781
76. Kurki MI, Karjalainen J, Palta P, et al. FinnGen: Unique genetic insights from combining isolated population and national health register data. *medRxiv*. 2022:2022.03.03.22271360. doi:10.1101/2022.03.03.22271360
77. Sun BB, Kurki MI, Foley CN, et al. Genetic associations of protein-coding variants in human disease. *Nature*. 03 2022;603(7899):95-102. doi:10.1038/s41586-022-04394-w
78. Zhou W, Nielsen JB, Fritsche LG, et al. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat Genet*. 09 2018;50(9):1335-1341. doi:10.1038/s41588-018-0184-y
79. Leitsalu L, Haller T, Esko T, et al. Cohort Profile: Estonian Biobank of the Estonian Genome Center, University of Tartu. *Int J Epidemiol*. Aug 2015;44(4):1137-47. doi:10.1093/ije/dyt268
80. Finer S, Martin HC, Khan A, et al. Cohort Profile: East London Genes & Health (ELGH), a community-based population genomics and health study in British Bangladeshi and British Pakistani people. *Int J Epidemiol*. 02 01 2020;49(1):20-21i. doi:10.1093/ije/dyz174
81. Wei WQ, Bastarache LA, Carroll RJ, et al. Evaluating phecodes, clinical classification software, and ICD-9-CM codes for phenome-wide association studies in the electronic health record. *PLoS One*. 2017;12(7):e0175508. doi:10.1371/journal.pone.0175508
82. Mbatchou J, Barnard L, Backman J, et al. Computationally efficient whole-genome regression for quantitative and binary traits. *Nat Genet*. 07 2021;53(7):1097-1103. doi:10.1038/s41588-021-00870-7
83. Sakaue S, Kanai M, Tanigawa Y, et al. A cross-population atlas of genetic associations for 220 human phenotypes. *Nat Genet*. 10 2021;53(10):1415-1424. doi:10.1038/s41588-021-00931-x
84. Honigberg MC, Riise HKR, Daltveit AK, et al. Heart Failure in Women With Hypertensive Disorders of Pregnancy: Insights From the Cardiovascular Disease in Norway Project. *Hypertension*. 11 2020;76(5):1506-1513. doi:10.1161/HYPERTENSIONAHA.120.15654
85. Brumpton BM, Graham S, Surakka I, et al. The HUNT study: A population-based cohort for genetic research. *Cell Genomics*. 2022;2(10)doi:10.1016/j.xgen.2022.100193
86. Åsvold BO, Langhammer A, Rehn TA, et al. Cohort Profile Update: The HUNT Study, Norway. *International journal of epidemiology*. 2022:dyac095. doi:10.1093/ije/dyac095
87. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 10 2018;562(7726):203-209. doi:10.1038/s41586-018-0579-z
88. Facco FL, Parker CB, Reddy UM, et al. Association Between Sleep-Disordered Breathing and Hypertensive Disorders of Pregnancy and Gestational Diabetes Mellitus. *Obstet Gynecol*. 01 2017;129(1):31-41. doi:10.1097/AOG.0000000000001805

89. Guerrero RF, Khan RR, Wapner RJ, et al. Genetic Polymorphisms Associated with Adverse Pregnancy Outcomes in Nulliparas. *medRxiv*. 2022:2022.02.28.22271641. doi:10.1101/2022.02.28.22271641
90. Denny JC, Ritchie MD, Basford MA, et al. PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. *Bioinformatics*. May 2010;26(9):1205-10. doi:10.1093/bioinformatics/btq126
91. Carroll RJ, Bastarache L, Denny JC. R PheWAS: data analysis and plotting tools for phenome-wide association studies in the R environment. *Bioinformatics*. Aug 15 2014;30(16):2375-6. doi:10.1093/bioinformatics/btu197

**Data availability**

GWAS summary statistics for preeclampsia/eclampsia and gestational hypertension and corresponding genome-wide polygenic scores will be made publicly available upon publication. Summary statistics used in this meta-analysis are publicly available for FinnGen r6 (https://www.finngen.fi/en/access_results) and for BioBank Japan (https://pheweb.jp/pheno/PreEclampsia). Summary statistics from the InterPregGen consortium are available at https://ega-archive.org. Placental transcriptome data are publicly available at https://www.obgyn.cam.ac.uk/placentome/.

**Acknowledgements**

This work was supported by grants from the American Heart Association 940166 (M.C.H.), 979465 (M.C.H.); the Korea Health Industry Development Institute HI19C1330 (S.M.J.C.); Harvard Catalyst Medical Research Investigator Training Program (A.P.P.); National Human Genome Research Institute U01HG011719 (A.P.P. and P.N.); the Belgian American Educational Foundation (A.S.); the U.S. National Institute of General Medical Sciences R35GM147197 (R.F.G.), R35GM124836 (R.D.); the U.S. National Heart Lung and Blood Institute K08HL146963 (K.J.G.), R01HL139865 (R.D.), R01HL155915 (R.D.), DP2HL152423 (R.M.G.), U01HL166060 (R.M.G.), R03HL148483 (R.M.G.), R01HL142711 (P.N.), R01HL127564 (P.N.), R01HL148050 (P.N.), R01HL151283 (P.N.), R01HL148565 (P.N.), R01HL135242 (P.N.), R01HL151152 (P.N.); National Institute of Diabetes and Digestive and Kidney Diseases R01DK125782 (P.N.); National Institute of Child Health and Human Development R01HD101246 (D.M.H.); Preeclampsia Foundation (K.J.G., R.S.); Fondation Leducq TNE-18CVD04 (P.N.); and the Massachusetts General Hospital Paul and Phyllis Fireman Endowed Chair in Vascular Medicine (P.N.). The work was conducted under UK Biobank application 7089. The Trøndelag Health Study (HUNT) is a collaboration between HUNT Research Centre (Faculty of Medicine and Health Sciences, NTNU, Norwegian

University of Science and Technology), Trøndelag County Council, Central Norway Regional Health Authority, and the Norwegian Institute of Public Health. The genotyping in HUNT was financed by the National Institutes of Health; University of Michigan; the Research Council of Norway; the Liaison Committee for Education, Research and Innovation in Central Norway; and the Joint Research Committee between St Olavs hospital and the Faculty of Medicine and Health Sciences, NTNU. The genetic investigations of the HUNT Study is a collaboration between researchers from the K.G. Jebsen Center for Genetic Epidemiology, NTNU and the University of Michigan Medical School and the University of Michigan School of Public Health. The K.G. Jebsen Center for Genetic Epidemiology is financed by Stiftelsen Kristian Gerhard Jebsen; Faculty of Medicine and Health Sciences, NTNU, Norway. DNA extraction and genotyping of nuMoM2b was supported by the Indiana University Precision Health Initiative. The authors thanks other members of the Estonian Biobank research team: Lili Milani, Andres Metspalu, Tõnu Esko, Mari Nelis, Reedik Mägi, and Georgi Hudjashov. Genes & Health is core-funded by Wellcome (WT102627, WT210561), the Medical Research Council (UK) (M009017, MR/X009777/1), Higher Education Funding Council for England Catalyst, Barts Charity (845/1796), Health Data Research UK (for London substantive site), and research delivery support from the NHS National Institute for Health Research Clinical Research Network (North Thames). Genes & Health is/has recently been funded by Alnylam Pharmaceuticals, Genomics PLC; and a Life Sciences Industry Consortium of Astra Zeneca PLC, Bristol-Myers Squibb Company, GlaxoSmithKline Research and Development Limited, Maze Therapeutics Inc, Merck Sharp & Dohme LLC, Novo Nordisk A/S, Pfizer Inc, Takeda Development Centre Americas Inc. We thank Social Action for Health, Centre of The Cell, members of our Community Advisory Group, and staff who have recruited and collected data from volunteers. We thank the NIHR National Biosample Centre (UK Biocentre), the Social Genetic & Developmental Psychiatry Centre (King's College London), Wellcome Sanger Institute, and Broad Institute for sample processing, genotyping, sequencing and variant annotation. We thank: Barts Health NHS Trust,

NHS Clinical Commissioning Groups (City and Hackney, Waltham Forest, Tower Hamlets, Newham, Redbridge, Havering, Barking and Dagenham), East London NHS Foundation Trust, Bradford Teaching Hospitals NHS Foundation Trust, Public Health England (especially David Wyllie), Discovery Data Service/Endeavour Health Charitable Trust (especially David Stables), NHS Digital for GDPR-compliant data sharing backed by individual written informed consent, and the volunteers participating in Genes & Health. The Genes & Health research team includes Shaheen Akhtar, Mohammad Anwar, Elena Arciero, Omar Asgar, Samina Ashraf, Gerome Breen, Raymond Chung, Charles J Curtis, Shabana Chaudhary, Maharun Chowdhury, Grainne Colligan, Panos Deloukas, Ceri Durham, Faiza Durrani, Fabiola Eto, Sarah Finer, Ana Angel Garcia, Chris Griffiths, Joanne Harvey, Teng Heng, Qin Qin Huang, Matt Hurles, Karen A Hunt, Shapna Hussain, Kamrul Islam, Ben Jacobs, Ahsan Khan, Amara Khan, Cath Lavery, Sang Hyuck Lee, Robin Lerner, Daniel MacArthur, Daniel Malawsky, Hilary Martin, Dan Mason, Mohammed Bodrul Mazid, John McDermott, Sanam McSweeney, Shefa Miah, Sabrina Munir, Bill Newman, Elizabeth Owor, Asma Qureshi, Samiha Rahman, Nishat Safa, John Solly, Farah Tahmasebi, Richard C Trembath, Karen Tricker, Nasir Uddin, David A van Heel, Caroline Winckley, and John Wright.

**Ethics declarations**

M.C.H. reports consulting fees from CRISPR Therapeutics, advisory board service for Miga Health, and grant support from Genentech, all unrelated to this work. K.J.G. has served as a consultant for BillionToOne, Aetion, and Roche for projects unrelated to this work. R.S. is a co-founder of Magnet Biomedicine, unrelated to this work. R.D. reports receiving grants from AstraZeneca, grants and nonfinancial support from Goldfinch Bio, being a scientific co-founder, consultant, and equity holder for Pensieve Health (pending), and being a consultant for Variant Bio, all unrelated to this work. P.N. reports grant support from Amgen, Apple, AstraZeneca, Boston Scientific, and Novartis, spousal employment and equity at Vertex, consulting income

from Apple, AstraZeneca, Novartis, Genentech / Roche, Blackstone Life Sciences, Foresite Labs, and TenSixteen Bio, and is a scientific advisor board member and shareholder of TenSixteen Bio and geneXwell, all unrelated to this work. The work conducted in HUNT was approved by the Regional Committee for Ethics in Medical Research, Norway (2018/2488). All participants signed informed consent for participation and the use of data in research.

**Table 1. Lead maternal variants associated with preeclampsia/eclampsia.**

| | | | | | Discovery analysis | | | | Replication analysis* | |
|---|---|---|---|---|---|---|---|---|---|---|
| **rsID** | **CHR** | **POS** | **Nearest gene** | **Effect allele / other allele** | **Weighted average EAF** | **OR** | **P-value** | **P(het)^** | **OR** | **P-value** |
| rs149764880 | 1 | 11820674 | *MTHFR / CLCN6* | G / T | 0.85 | 1.11 | 1.8e-8 | 0.30 | 1.15 | 7.4e-4 |
| rs708119 | 1 | 228015567 | *WNT3A* | C / G | 0.67 | 1.08 | 2.3e-8 | 0.92 | 1.07 | 0.09 |
| rs9855086 | 3 | 169413542 | *MECOM* | A / T | 0.53 | 1.08 | 5.7e-10 | 0.21 | 1.10 | 0.001 |
| rs16998073 | 4 | 80263187 | *FGF5* | T / A | 0.32 | 1.11 | 3.1e-14 | 0.15 | 1.09 | 0.04 |
| rs2442752 | 6 | 31383987 | *MICA* | C / T | 0.36 | 1.08 | 2.3e-9 | 0.57 | 1.02 | 0.78 |
| rs5899121 | 9 | 91145498 | *LINC00484* | TG / T | 0.85 | 1.12 | 4.6e-8 | 0.31 | 0.99 | 1.00 |
| rs2508372 | 11 | 101397592 | *PGR* | A / G | 0.79 | 1.10 | 1.3e-9 | 0.14 | 1.09 | 0.03 |
| rs10774624 | 12 | 111395984 | *SH2B3* | G / A | 0.47 | 1.10 | 2.9e-12 | 0.40 | 1.03 | 0.54 |
| rs7318880 | 13 | 28564148 | *FLT1* | T / C | 0.48 | 1.09 | 6.7e-12 | 0.79 | 1.07 | 0.02 |
| rs1421085 | 16 | 53767042 | *FTO* | C / T | 0.40 | 1.10 | 4.7e-12 | 0.61 | 1.06 | 0.66 |
| rs167479 | 19 | 11416089 | *RGL3* | G / T | 0.55 | 1.10 | 7.6e-10 | 0.20 | 1.07 | 0.04 |
| rs259983 | 20 | 59160402 | *ZNF831* | C / A | 0.16 | 1.13 | 1.0e-12 | 0.40 | 1.15 | 9.7e-4 |

*Given strong genetic correlation between preeclampsia/eclampsia and gestational hypertension ($r_g$=0.71) in discovery cohorts, associations for preeclampsia/eclampsia and gestational hypertension were meta-analyzed in follow-up cohorts for primary replication. P-values were calculated with correction for overlapping control samples.
^P(het) indicates P-value for heterogeneity across discovery cohorts.

**Table 2. Lead maternal variants associated with gestational hypertension.**

| | | | | | Discovery analysis | | | | Replication analysis* | |
|---|---|---|---|---|---|---|---|---|---|---|
| **rsID** | **CHR** | **POS** | **Nearest gene** | **Effect allele / other allele** | **Weighted average EAF** | **OR** | **P-value** | **P(het)^** | **OR** | **P-value** |
| rs9855086 | 3 | 169413542 | *MECOM* | A / T | 0.54 | 1.10 | 1.4e-9 | 0.51 | 1.10 | 0.001 |
| rs16998073 | 4 | 80263187 | *FGF5* | T / A | 0.32 | 1.12 | 4.6e-11 | 0.54 | 1.09 | 0.04 |
| rs13154066 | 5 | 32831564 | *NPR3* | C / T | 0.58 | 1.11 | 6.1e-10 | 0.60 | 1.05 | 0.07 |
| rs7139122 | 12 | 53173006 | *TNS2 / CSAD* | A / G | 0.02 | 1.46 | 1.1e-9 | 0.13 | 1.22 | 0.19 |
| rs167479 | 19 | 11416089 | *RGL3* | G / T | 0.56 | 1.11 | 6.9e-9 | 0.76 | 1.07 | 0.04 |
| rs2208589 | 20 | 48791877 | *PREX1* | G / A | 0.84 | 1.15 | 2.3e-8 | 0.17 | 1.05 | 0.56 |
| rs260017 | 20 | 59160402 | *ZNF831* | C / A | 0.18 | 1.16 | 7.0e-13 | 0.32 | 1.15 | 9.7e-4 |

*Given strong genetic correlation between preeclampsia/eclampsia and gestational hypertension ($r_g$=0.71) in discovery cohorts, associations for preeclampsia/eclampsia and gestational hypertension were meta-analyzed in follow-up cohorts for primary replication. P-values were calculated with correction for overlapping control samples.
^P(het) indicates P-value for heterogeneity across discovery cohorts.

**Table 3. Aspirin eligibility to prevent preeclampsia using major clinical risk factors and polygenic risk.** This Table reports the proportion of participants in nuMoM2b (nulliparous women with singleton gestations recruited in the first trimester of pregnancy) who experienced each pregnancy outcome (normotensive pregnancy, gestational hypertension, or preeclampsia/eclampsia) who would have been identified as candidates for low-dose aspirin to prevent preeclampsia according to U.S. Preventive Services Task Force major criteria, with or without addition of preeclampsia/eclampsia and systolic blood pressure combination polygenic risk score ($PRS_{preeclampsia+SBP}$) as an additional criterion for aspirin eligibility. The two right-most columns display net reclassification using $PRS_{preeclampsia+SBP}$ to up-classify risk for composite hypertensive disorders of pregnancy (HDPs) vs. normotensive pregnancy and for preeclampsia/eclampsia vs. other pregnancies when added to major risk factors.

| | **Normotensive pregnancy (n=5,986)** | **Gestational hypertension (n=1,358)** | **Preeclampsia/ eclampsia (n=507)** | **Net reclassification for HDPs vs. normotensive pregnancy (95% CI)** | **Net reclassification for preeclampsia vs. other pregnancy (95% CI)** |
|---|---|---|---|---|---|
| **Major risk factors only** | 496 (8.3%) | 97 (7.1%) | 87 (17.2%) | Ref. | Ref. |
| **Major risk factors or top 5% PRS** | 728 (12.2%) | 187 (13.8%) | 131 (23.9%) | +3.3% (2.0 to 4.6%) | +4.3% (1.7 to 6.8%) |
| **Major risk factors or top 10% PRS** | 966 (16.1%) | 258 (19.0%) | 153 (30.2%) | +4.3% (2.7 to 6.0%) | +4.4% (1.4 to 7.5%) |
| **Major risk factors or top 25% PRS** | 1,686 (28.2%) | 476 (35.1%) | 238 (46.9%) | +8.5% (6.3 to 10.9%) | +8.4% (4.2 to 12.5%) |

**Figure 1. Manhattan plots of preeclampsia/eclampsia and gestational hypertension in discovery cohorts.** Manhattan plots (chromosomal position on the X-axis and -log(10) of the P-value on the Y-axis) are displayed for **(A)** preeclampsia/eclampsia in 17,150 cases and 451,241 controls and **(B)** gestational hypertension in 8,961 cases and 184,925 controls. Analyses included multi-ancestry meta-analysis of common variants (minor allele frequency ≥1%). Loci are labeled by the gene nearest to the lead variant.

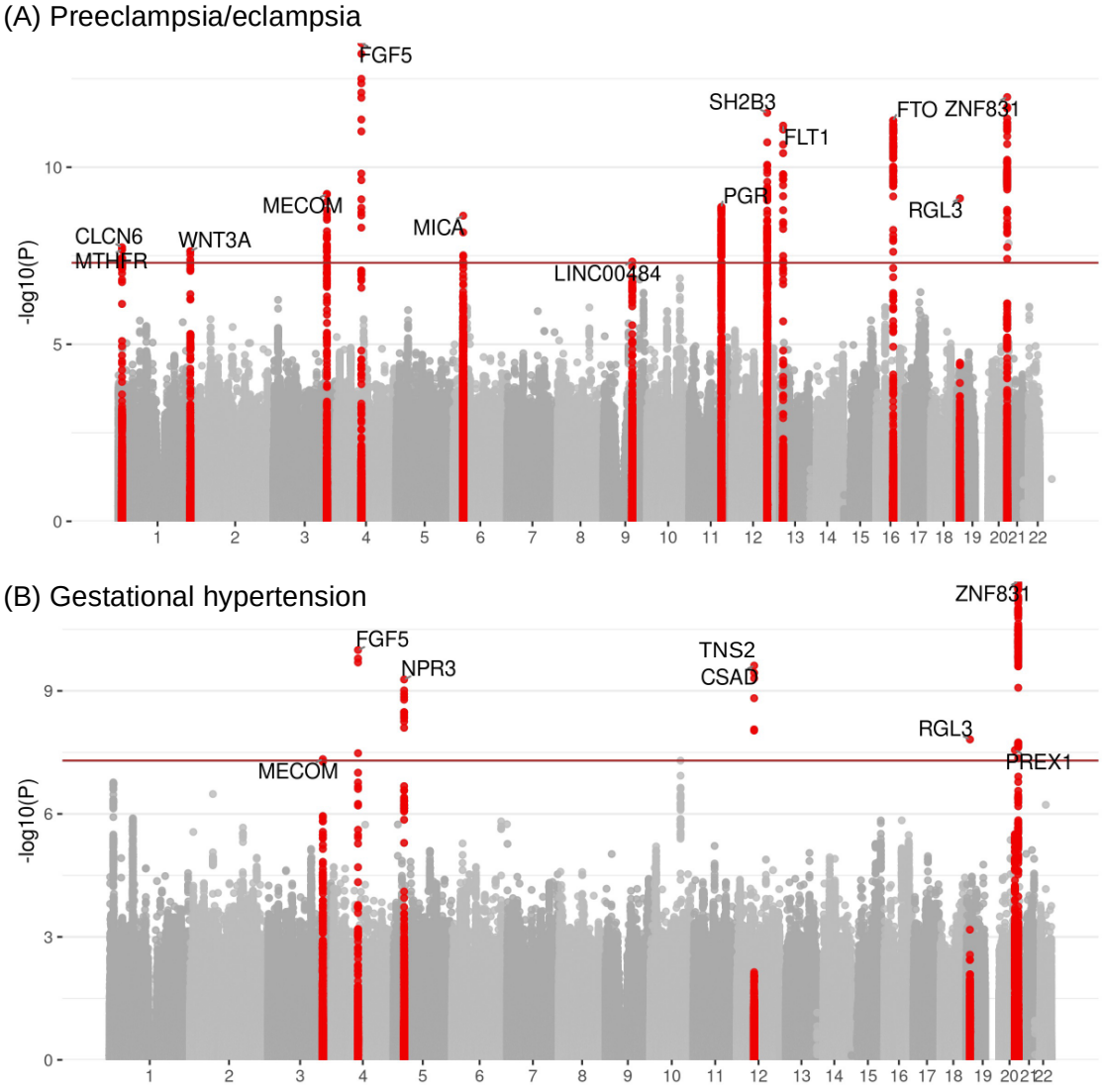


**Figure 2. Polygenic prediction of preeclampsia/eclampsia and gestational hypertension in test cohorts.** Polygenic risk scores (PRS) for preeclampsia/eclampsia and gestational hypertension were derived from our discovery genome-wide meta-analyses, tuned in the UK Biobank, and carried forward for testing in independent cohorts. Prevalence of preeclampsia/eclampsia vs. percentile of preeclampsia/eclampsia and SBP combination PRS in **(A)** HUNT and **(B)** nuMoM2b. **(C)** Prevalence of gestational hypertension vs. percentile of gestational hypertension and SBP combination PRS in nuMoM2b. Distribution of preeclampsia/eclampsia and SBP combination PRS by preeclampsia/eclampsia status in **(D)** HUNT and **(E)** nuMoM2b. **(F)** Distribution of gestational hypertension and SBP combination PRS by gestational hypertension status in nuMoM2b.

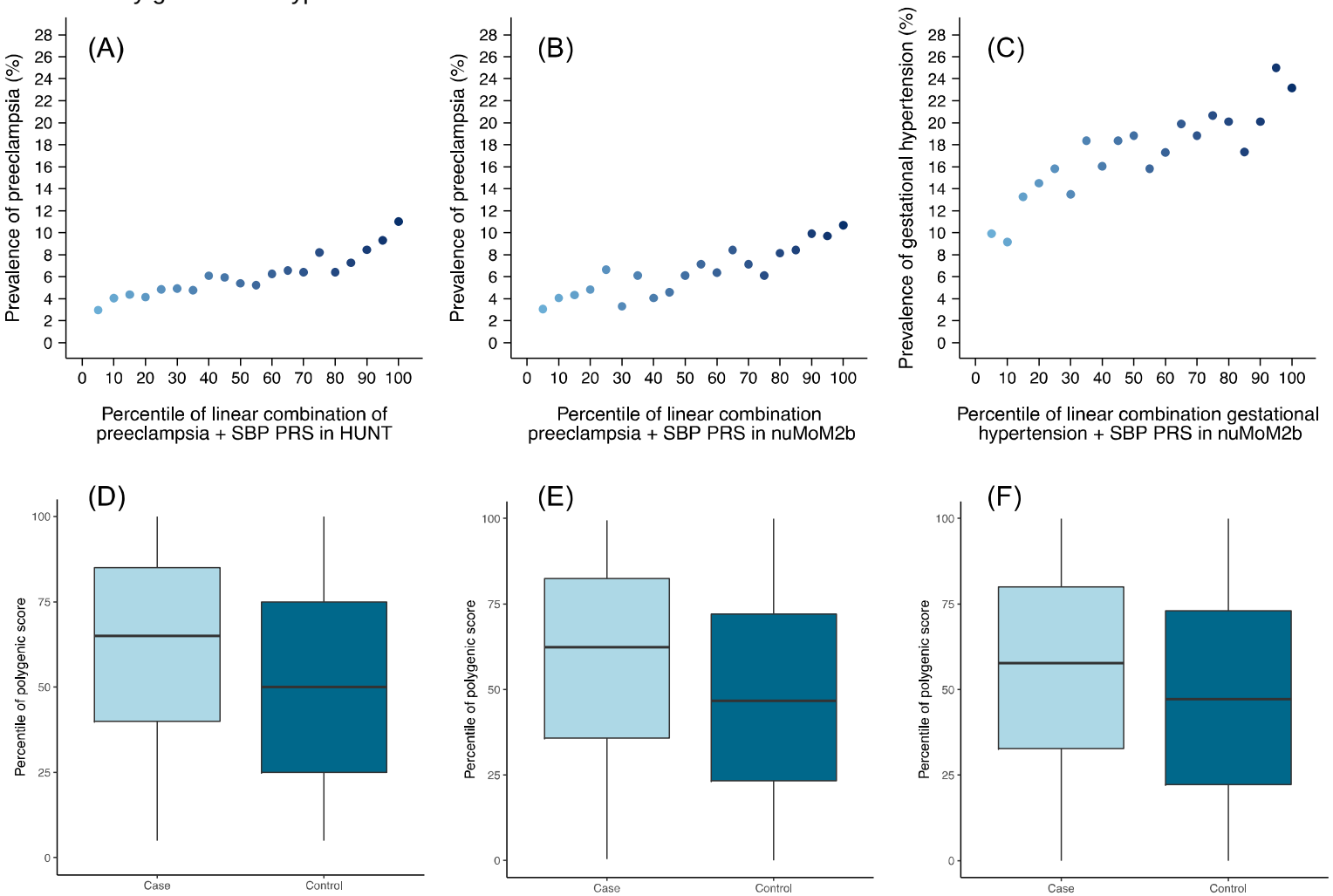
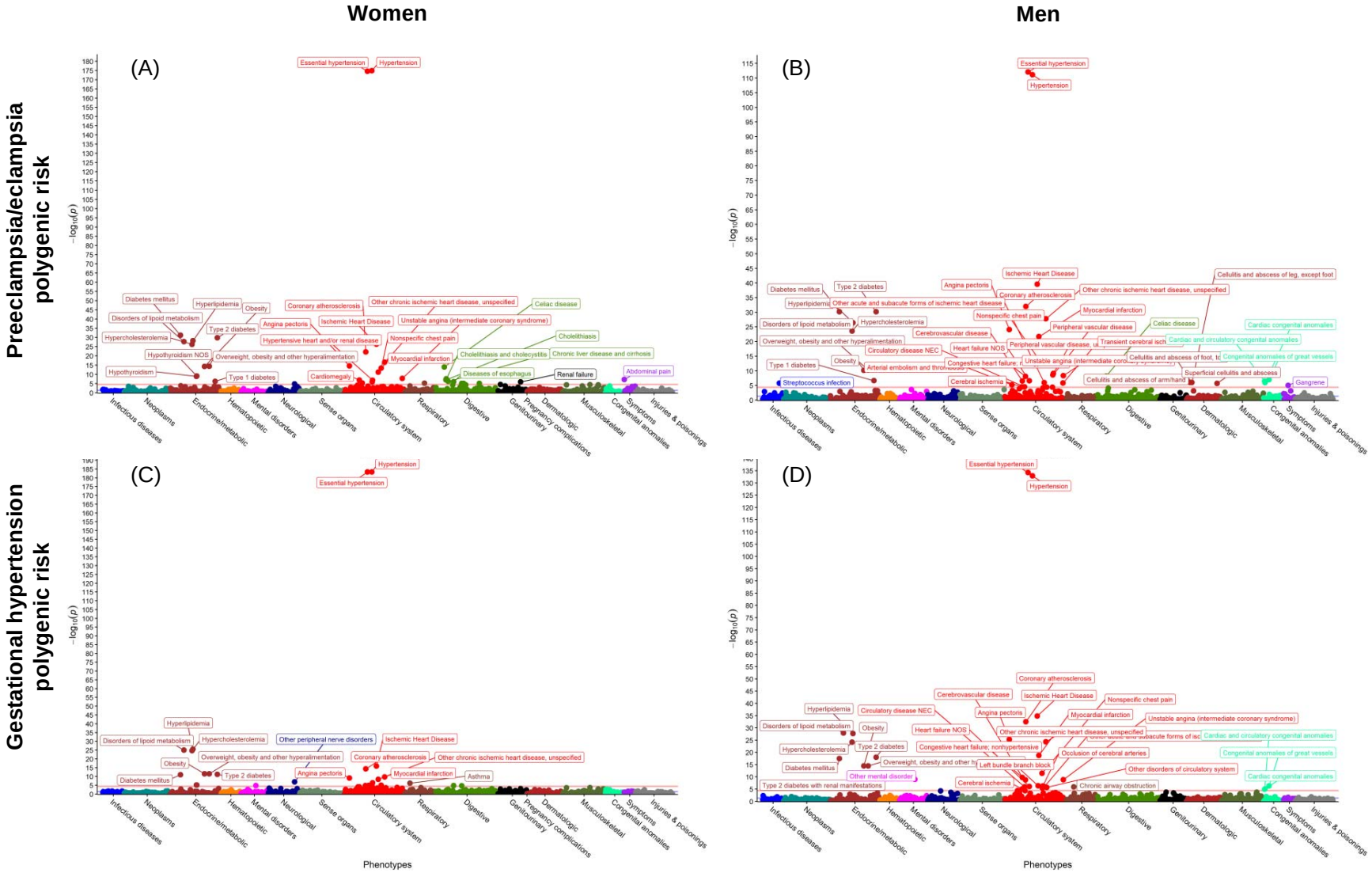
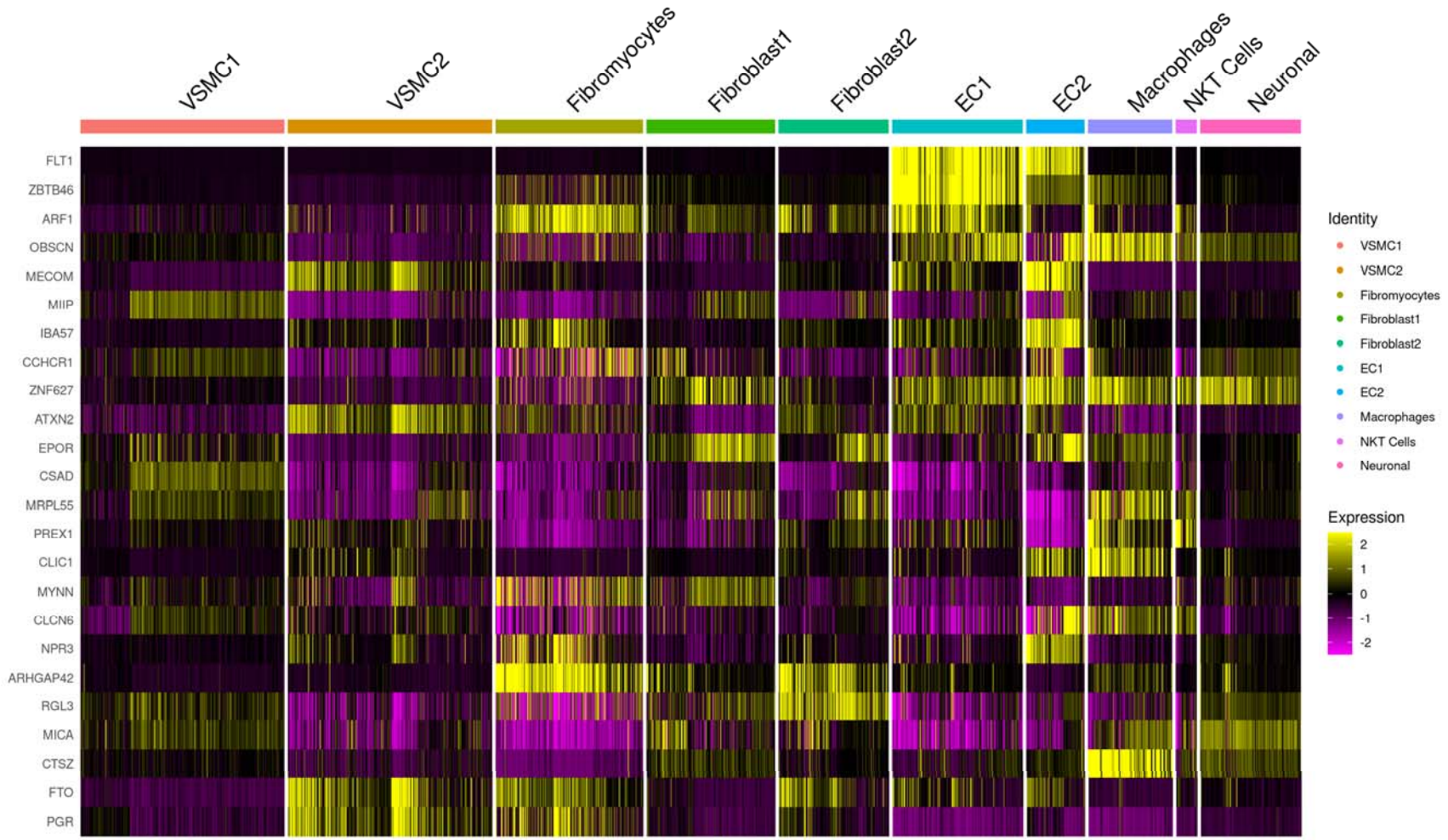


**Figure 3. Sex-stratified phenome-wide association study of preeclampsia/eclampsia and gestational hypertension polygenic risk in the UK Biobank.** Preeclampsia/eclampsia polygenic risk **(A-B)** and gestational hypertension polygenic risk **(C-D)** were associated with 1,445 phenotypes among women (A, C) and men (B, D) in the UK Biobank. Associations with phenotypes were tested using logistic regression with adjustment for age and PC 1-5.



## Extended Data

**Extended Data Figure 1. Relative expression of prioritized genes in human aortic cells with single-nuclei RNA sequencing.** We analyzed expression of genes prioritized by genome-wide meta-analysis of preeclampsia/eclampsia and gestational hypertension and secondary *in silico* analyses in a dataset of single-nuclei RNA sequencing from two normal human flash-frozen aortic specimens. Most prioritized genes were enriched in endothelial cell populations and/or macrophages.