PREECLAMPSIA RISK, MATERNAL 25-HYDROXYVITAMIN D CONCENTRATION, AND VARIATION IN VITAMIN D METABOLISM PATHWAY GENES

by

Katharyn Maria Baca

BA, Clark University, 2009

MA, Clark University, 2010

MPH, University of Pittsburgh, 2015

Submitted to the Graduate Faculty of

the Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2015

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This dissertation was presented

by

Katharyn M Baca

It was defended on

November 30, 2015

and approved by

Mary L. Marazita, PhD, Professor and Vice Chair of the Department of Oral Biology, Professor, Department of Human Genetics, Graduate School of Public Health and School of Medicine, University of Pittsburgh

Hyagriv N Simhan, Department of Obstetrics, Gynecology, and Reproductive Sciences, School of Medicine, University of Pittsburgh, Magee-Womens Research Institute, Pittsburgh, Pennsylvania

Joseph M. Zmuda, PhD, Associate Professor, Departments of Epidemiology and Human Genetics, Graduate School of Public Health, University of Pittsburgh

Dissertation Advisor: Lisa M. Bodnar, PhD, MPH, RD, Associate Professor, Departments of Epidemiology, Obstetrics, Gynecology, and Reproductive Sciences, Graduate School of Public Health and School of Medicine, University of Pittsburgh

Lisa M. Bodnar, PhD, MPH, RD

PREECLAMPSIA RISK, MATERNAL 25-HYDROXYVITAMIN D CONCENTRATION, AND VARIATION IN VITAMIN D METABOLISM PATHWAY GENES

Katharyn M. Baca, PhD

University of Pittsburgh, 2015

ABSTRACT

OBJECTIVE: Our objectives were to study the relationships between 25-hydroxyvitamin D (25(OH)D) and preeclampsia risk, maternal genetic variation in 3 vitamin D metabolism genes (*GC, CYP27B1, VDR*) and preeclampsia risk, and variation in the same genes and 25(OH)D. **METHODS:** We used two racially diverse pregnancy cohorts (EVITA and Collaborative Perinatal Project (CPP)) to achieve these objectives. We estimated the association between log-transformed 25(OH)D and preeclampsia risk in EVITA by using log-binomial regression with restricted cubic splines. In EVITA and CPP, we used multivariable logistic and linear regression models to estimate the associations between allelic variation and preeclampsia risk, and genotype and log-transformed 25(OH)D, respectively. Meta-analyses were conducted to calculate estimates of association between and within cohorts. **RESULTS:** Dose-response associations of 25(OH)D were observed for both severe and mild preeclampsia. Trends of associations were observed in genetic variation and preeclampsia risk. Compared with major allele carriers, Black mothers in EVITA who carried the minor allele for rs11732451 *GC* single nucleotide polymorphism (SNP) and 2 *VDR* SNPs (rs4340112, rs10459217) had increased odds of preeclampsia, while the odds were lowered for those who carried the minor allele for 1 *GC* SNP

(rs1099028) and 2 *VDR* SNPs(rs757344, rs12721364). In the meta-analysis, two *VDR* SNPs (rs886441 and rs2853561) had trends of decreased odds of preeclampsia for all Black mothers. For the 25(OH)D analysis, statistically significant associations were observed. Compared with those with major allele genotypes, mothers with minor allele genotypes of rs1844885 (GC) and rs11168275 (*VDR*) had increased 25(OH)D and of rs11732451 (GC) had lowered 25(OH)D. In the meta-analysis on all Black mothers, rs1844885 (*GC*) was associated with increased 25(OH)D while there was a trend of decreased 25(OH)D for rs10877016 (*CYP27B1*). **CONCLUSIONS:** Low 25(OH)D may be enough to reduce risk of preeclampsia. If our findings are confirmed in a replication study, genetic variation may be an independent risk factor for maternal 25(OH)D, making the findings of this research relevant to public health.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

This dissertation is a result of collaboration with a number of outstanding researchers who offered their time and expertise. First, I am grateful to my dissertation and academic mentor, Lisa Bodnar, for her patience, for being inspirational, and for her commitment to this project. Lisa has driven me to strive for the best, while also encouraging me to fulfill my academic and career goals. I am also grateful to the rest of my committee for their perceptive suggestions and commitment to the project. They have challenged me to strengthen and build skills in epidemiology which will serve me well in my future endeavors. I would also like to acknowledge Manika Govil for her outstanding mentorship and support.

Finally, I would also like to thank my family and friends for their support throughout the process. My parents, sister, brother-in-law, and friends have supported and encouraged me throughout this program.

1. INTRODUCTION

1.0 BACKGROUND

Vitamin D deficiency is a major public health concern due to its widespread prevalence and its association with several diseases, including osteoporosis, osteopenia, cancer, heart disease, and multiple sclerosis[1-11]. The role of vitamin D in optimal pregnancy outcomes is contentious. Important traditional and non-traditional roles of vitamin D in pregnancy may include skeletal mineralization, cell differentiation and proliferation, hormone regulation, and regulation of inflammation[12]. Although several observational studies have reported vitamin D deficient women are at increased risk of preeclampsia compared with sufficient women, the associations have been inconsistent [13-25].

The discrepancy may be explained by the presence of common genetic variations in vitamin D metabolizing genes like *GC*, *CYP27B1* and vitamin D receptor (*VDR*) genes. Encoding the vitamin D binding protein (DBP), the *GC* gene is important for modulating the bioavailability of vitamin D metabolites, as well as transporting vitamin D and its metabolites in serum[26]. These metabolites include 25(OH)D which needs to be hydroxylated by an enzyme encoded by *CYP27B1* to become the active form of vitamin D ($1,25(OH)_2D$ or 1,25-dihydroxyvitamin D) [27]. $1,25(OH)_2D$ can influence biological mechanisms by binding to *VDR*, triggering a genomic response [28]. The activated *VDR* is involved in several biological roles in

pregnancy, such as placental hormone regulation and prevention of inflammation[29]. These roles may be disrupted in mothers with allelic variations of the *GC*, *CYP27B1*, and *VDR* genes, increasing their risk of vitamin D deficiency or adverse pregnancy outcomes like preeclampsia[29]. Past studies have reported common polymorphisms in the *GC*, *CYP27B1*, and *VDR* genes are associated with vitamin D deficiency and adverse health outcomes in non-pregnant adults [30-35]. However, there is limited research on the effects of these variants on poor pregnancy outcomes.

1.1 RESEARCH AIMS

We have the unique opportunity to use existing data and banked samples to determine if polymorphisms in *GC*, *CYP27B1* and *VDR* genes are associated with preeclampsia or vitamin D deficiency in pregnancy. We will use two multi-center cohorts to achieve our goal. The Collaborative Perinatal Project (CPP) (1959-1966) is the largest pregnancy cohort study in the U.S. (n=717 cases of preeclampsia). Maternal blood samples at ≤26 weeks of gestation have been assayed for 25-hydroxyvitamin D (25(OH)D), the indicator of vitamin D nutritional status. EVITA (1999-2010) is a retrospective cohort of mothers who received prenatal genetic screening at Magee-Womens Hospital in Pittsburgh, PA (n= 650 cases of preeclampsia). Maternal samples were assayed for 25(OH)D at ≤20 weeks. We will use 2514 and 2357 DNA samples from EVITA and CPP cohorts, respectively. The common haplotypes of these genes will be tagged to avoid redundancy in the genotyping of samples. Our specific aims are the following:

Aim 1: To determine the effect of vitamin D deficiency on risk of preeclampsia.

Hypothesis 1: Women with vitamin D deficiency will have increased risk of preeclampsia compared with women with vitamin D sufficiency.

Aim 2: To determine the relationship between maternal genetic variation in *GC*, *CYP27B1*, and *VDR* on risk of preeclampsia.

Hypothesis 2: Risk variant carriers will have an increased susceptibility to preeclampsia.

Aim 3: To determine the effect of *GC*, *CYP27B1*, and *VDR* gene on vitamin D deficiency.

Hypothesis 3: Carriers of risk variants will have increased risk of vitamin D deficiency.

Overall impact: Estimating associations between variation in the *GC*, *CYP27B1*, and *VDR* genes, vitamin D status, and preeclampsia may help to improve our understanding of the inconsistencies in past vitamin D research. Therefore, this study can make a significant impact on a major public health problem.

2. LITERATURE REVIEW

2.0 INTRODUCTION

The high prevalence of maternal vitamin D deficiency may be one factor contributing to adverse pregnancy outcomes [36, 37]. Experimental and observational studies show vitamin D metabolites have important roles in pregnancy, including regulating the transcription and function of genes associated with placental invasion and angiogenesis, modulating immune function and inflammatory response, as well as regulating blood pressure [38-40]. Any or all of these mechanisms can potentially increase risk of preeclampsia [41]. Several observational studies have noted vitamin D deficient women have an increased risk of preeclampsia compared with sufficient women; however the associations remain equivocal [13-25]. Associations between genetic variations and preeclampsia risk may help explain the inconsistencies in observational studies. To date, only one published study has researched the effect of vitamin D metabolizing genetic variants on 25(OH)D levels in pregnancy [42] and preeclampsia risk [43]. We will investigate 371 SNPs in three metabolizing genes that either synthesize (*CYP27B1*), carry (*GC*), or mediate downstream signaling (*VDR*) by serum 25(OH)D, and the association of these SNPs with preeclampsia and vitamin D deficiency risks. Our goal is to determine whether low 25(OH)D is associated with risk of preeclampsia, and if there are risk alleles in three vitamin

D candidate genes. We will use samples and data from two large pregnancy cohort studies to address this gap in knowledge.

2.1 VITAMIN D

2.1.1 Metabolism

Vitamin D, a fat-soluble pro-hormone, originates in two forms: vitamin D_2 (ergocalciferol) and vitamin D_3 (cholecalciferol). Vitamin D_2 is found naturally in food like oily fish, plants, and vegetables, while vitamin D_3 is found in fortified foods, such as dairy products, and dietary supplements [44-48]. Most vitamin D_3 synthesis occurs in the skin when UVB light is absorbed and interacts with 7-dehydrocholesterol. After skin synthesis and food absorption, vitamin D_2 and D_3 in serum are transported to the liver for hydroxylating reactions by chylomicrons and the vitamin D binding protein (DBP, encoded by *GC* gene) [29, 48, 49].

Once vitamin D_2 and D_3 enter the liver, enzyme-driven hydroxylating reactions convert the prohormones to 25(OH)D (25-hydroxyvitamin D or calcidiol). An enzyme called 25-hydroxylase mediates these reactions, which is an enzyme encoded by either *CYP2R1* or *CYP27A1* (cytochrome P450, family 2, subfamily R, polypeptide 1; cytochrome P450, family 27, subfamily A, polypeptide 1) [29]. Circulating 25(OH)D is the best marker for vitamin D nutritional status. It reflects dietary intake and skin synthesis, and it has a half-life of 3 weeks in serum. Vitamin D_3 has a turnover rate of 2-3 months in adipose tissue [48, 50]. About 85-90%, 10-15%, and 0.1% of 25(OH)D are attached to DBP, albumin or is in free form in maternal circulation, respectively [51]. Circulating 25(OH)D enters the kidney, as well as other types of

cells, for further metabolism. Recent studies observed DBP may have a role in maintaining stable stores of vitamin D metabolites by modulating bioavailability and the rate of vitamin D activation [52, 53].

Other hydroxylating reactions on 25(OH)D produces the active form of vitamin D, $1,25(OH)_2D$ (1,25-dihydroxyvitamin D or calcitriol). These reactions occur predominately in the kidney, as well as in cells in the skin, breast, colon, prostate, and brain and in monocytes and macrophages, catalyzed by 1 alpha-hydroxylase (encoded by *CYP27B1*; cytochrome P450 enzyme, family 27, subfamily A, polypeptide 1) [27]. Both $1,25(OH)_2D$ and 25(OH)D can be catabolized by proteins encoded by *CYP24A1*, serving as a feedback loop to avoid vitamin D toxicity. Specifically, toxicity is avoided when the protein hydrolyzes 25(OH)D and $1,25(OH)_2D$ into $24,25(OH)_2D$, and $1,24,25(OH)_3D$, respectively [29].

$1,25(OH)_2D$ has nongenomic and genomic roles in the body by attaching to the vitamin D receptor (VDR) expressed in target organs and cells [54]. In a genomic response, a heterodimer complex is formed by *VDR*, $1,25(OH)_2D$, and retinoid-x receptors (*RXRA*). The heterodimer complex binds to hormone response elements on DNA, up-regulating gene expression at the transcriptional level and influencing important biological mechanisms in the human body [29]. For nongenomic responses, $1,25(OH)_2D$ attaches to *VDR* exerting rapid effects in a variety of cell types. These nongenomic roles include regulating calcium and phosphate metabolism, renal reabsorption, and bone remodeling [54].

2.1.2 Metabolism in pregnancy

Compared with non-pregnant adults, vitamin D metabolism is more complex in pregnancy [55]. Maternal 25(OH)D concentration levels are not affected in pregnancy [56, 57]. From animal

studies we know 25(OH)D enters placental cells by endocytosis of DBP-25OHD, by diffusion of free hormone, or by both mechanisms [58]. Indeed, human cord blood concentrations of 25(OH)D equal 75 to 90% of maternal 25(OH)D concentration [56, 59]. In contrast, the physiologically active metabolite $1,25(OH)_2D$ does not readily cross the placenta. Instead 25(OH)D that passes the placenta can be hydrolyzed into $1,25(OH)_2D$ by enzymes encoded by *CYP27B1* expressed in maternal decidua cells [55]. Unlike maternal 25(OH)D, maternal $1,25(OH)_2D$ levels increase in the first and second trimester. This change may be due to the increased circulating vitamin D binding protein (DBP increases from 7 to 152% in pregnancy) and increased hydroxylation reactions in maternal kidneys [59].

2.1.3 Determinants of Vitamin D status

Serum 25(OH)D concentration is influenced by exposure to solar radiation and oral vitamin D intake. Approximately 80% of serum 25(OH)D is derived from exposure to UVB radiation [60]. When UVB is absorbed, thermal isomerization converts 7-dehydrocholesterol to precholecalciferol and to cholecalciferol. Decreased skin absorption of UVB occurs in the winter months at higher latitudes, and with use of sunscreen or UVB-protective clothing [61, 62]. Additionally, UVB is absorbed by melanin, the primary determinant of skin pigmentation [45]. This absorption decreases the amount of UVB available for photolysis in vitamin D synthesis; therefore there is a higher risk of vitamin D deficiency among individuals with deeper skin pigmentation. When sun exposure is limited, oral intake of vitamin D sources such as vitamin supplementation, fortified milk or orange juice, as well as fatty fish, can be major sources of vitamin D [63]. The recommended daily allowance (RDA) of vitamin D is 600 IU/day for men and women up to 70 years of age [64]. Prenatal vitamins contain less than the RDA; instead they

typically contain 400 IU of vitamin D_3. Additionally, serum 25(OH)D concentrations are lower in individuals who are older or obese [65].

2.1.4 Vitamin D deficiency definitions

The concentration levels of 25(OH)D to achieve optimal health is controversial. According to the Institute of Medicine (IOM), evidence has not established the amount 25(OH)D concentration needed to achieve non-skeletal health status [66]. Therefore, the IOM base their recommendation of 25(OH)D serum concentration levels on bone health which is 25(OH)D 40 nmol/L-50 nmol/L [66]. In comparison, the Endocrine Society and several experts in the field advise that 25(OH)D ≥75 nmol/L is required for overall health [60, 67-69]. Due to this controversy, studies on vitamin deficiency use various cut-offs of 25(OH)D.

2.1.5 Prevalence of Vitamin D Deficiency

Prevalence of vitamin D deficiency depends on the definition used. In the United States, the prevalence of vitamin D deficiency, defined as 25(OH)D <50 nmol/L, among adults is between 35-40% [70, 71]. Furthermore, 82.1%, 69.2%, and 30.9% of blacks, Hispanics, and whites are deficient. Overall, there has been an increase in the prevalence of vitamin D deficiency in the United States. According to data from two National Health and Nutrition Examination Surveys (NHANES III 1988-1994 and NHANES 2001-2004), the prevalence of 25(OH)D concentration levels below 50 nmol/L increased from 22% in 1988-1994 to 36% in 2001 -2004 [72]. From the 1988-1994 data collection, prevalence of deficiency was 5 times more likely in blacks compared to whites. However, the difference was 10-fold in the 2001-2004 data collection.

2.1.6 Prevalence of Vitamin D Deficiency in Pregnancy

In the 2001-2004 NHANES data collection, 28% of pregnant women had 25(OH)D <50 nmol/l compared to 37% in non-pregnant women ages 12-44 years of age [73]. Several studies have noted a racial disparity in 25(OH)D deficiency in pregnancy. In a cohort from South Carolina, 75%, 30%, and 12% of African American, Hispanic, and Caucasian pregnant women, respectively, were vitamin D deficient in South Carolina (defined as 25(OH)D <50nmol/L) [74]. In Pittsburgh, 29.2% and 54.1% of black mothers were vitamin D deficient and insufficient, while 5% and 42.1% of white mothers were deficient and insufficient (defined as 25(OH)D <37.5nmol/L and 25(OH)D 37.5nmol/L – 80 nmol/L) [75]. Overall, nationally vitamin D deficiency is common in pregnancy and there is strong evidence of a racial disparity.

2.1.7 Traditional and Nontraditional roles of Vitamin D

Vitamin D has various roles in the human body which can be divided into classical and non-classical actions. Classical actions include regulating calcium and phosphate homeostasis, and developing and maintaining bone health [76]. Non-classical actions include proliferating and differentiating cells, apoptosis, and immunomodulatory and anti-proliferative effects through the autocrine and paracrine pathways [77]. Vitamin D deficiency has been associated with traditional outcomes such as skeletal mineralization, but also non-traditional outcomes involving the renin-angiotensin system and inflammation (e.g., asthma, depression, bronchiectasis, bone health, cancers, cardiovascular disease, and diabetes) [1-8]. One of these outcomes includes preeclampsia.

2.2 PREECLAMPSIA

2.2.1 Definition of Preeclampsia

Preeclampsia is defined by new-onset hypertension and proteinuria after 20 weeks gestation. Specifically, preeclamptic women have systolic pressure greater than 140 mm Hg and/or diastolic pressure above 90 mm Hg. Proteinuria in preeclamptic women is characterized as having protein greater than 300 mg in a 24-hour urine sample. Preeclampsia can be further classified as early-and late-onset, as well by severity (mild and severe). Early-onset preeclampsia occurs at or before 33 weeks of gestation while late-onset occurs at 34 weeks of gestation or later. Severe forms of preeclampsia include cases with additional symptoms including severe hypertension, increased liver enzymes in blood, kidney dysfunction, high levels of proteinuria, or stroke.

2.2.2 Prevalence of preeclampsia

Currently, preeclampsia occurs in 2-8% of pregnancies in the United States [78]. Despite the low prevalence, preeclampsia is a public health concern due to the 25% increase in prevalence from 1987 to 2004 [79]. More specifically, the prevalence of severe preeclampsia in 2010 has increased by 322% (0.3% to 1.4%) since 1987 while prevalence has decreased for mild preeclampsia (3.1% to 2.5%) [80]. In addition, late-onset is more common in pregnancy (90%) compared to early forms of the disease [81].

2.2.3 Consequences

Preeclampsia is a multi-systemic condition which may cause organ failure, seizure, stroke, or even death to the mother [78]. However, fetal and maternal outcomes depend on severity and onset of disease. Women with early-onset and severe forms of preeclampsia have an increased risk of preterm birth, fetal death, recurrence of preeclampsia and increased risk of stillbirth in a subsequent pregnancy [82, 83]. In early-onset cases, there is a higher risk of infants being born small-for-gestational age, while there is a higher risk of large-for-gestational age in late-onset cases [84, 85]. In severe cases, preeclampsia can also be a life-threatening disease which is responsible for 16% of maternal deaths in high income countries. The risk of mortality is 3.1 times greater in blacks with preeclampsia as whites with preeclampsia [86].

2.2.4 Risk Factors

Risk factors for preeclampsia include chronic hypertension, kidney disease, black race, obesity, diabetes, younger age, and nulliparity [87]. In addition to these factors, there is a 2- to 5-fold increase in preeclampsia occurrence in women with maternal history of preeclampsia, which is evidence of a genetic component for preeclampsia [41].

2.2.5 Prevention and Treatment

Preeclampsia is a life-threatening disease for the mother and fetus due to a lack of management leading to preeclampsia [88]. Preventive agents have been tested in preeclamptic women, such as low-dose aspirin, vitamin E and C, and calcium supplements [89, 90]. In high risk women, there

is evidence that low-dose aspirin and calcium decreases risk of preeclampsia. Unfortunately, delivery of the placenta is the only cure for preeclampsia [91].

2.2.6 Pathophysiology

Preeclampsia is a two-stage process. The first stage occurs in early pregnancy during the development of the placenta. In preeclampsia, poor placental perfusion results from abnormal trophoblast invasion, poorly remodeled spiral arteries and limited utero-placental circulation [92]. In addition, there may be a maternal-fetal immunologic maladaptation that may cause poor placental perfusion [93]. The reduced utero-placental blood flow can lead to ischemia, oxidative stress, inflammation and endothelial dysfunction. Hypoxia increases production on several angiogenic and growth factors, favoring an antiangiogenic state [92]. This includes an increase in anti-angiogenesis factors sVEGFR-1 and sEndogin, and a decrease in factors that promote angiogenesis like PIGF and VEGF. Clinical symptoms of preeclampsia occur in stage 2 when anti-angiogenic factors cross the placenta and enter maternal circulation causing a systemic inflammatory response [94]. These factors can impair endothelial cell function resulting in a reduced stimulation of the renin-angiotensin system, vasoconstriction, and hypertension.

2.2.7 Vitamin D and Preeclampsia

Vitamin D deficiency may impact preeclampsia risk through several biological mechanisms. These mechanisms include placentation, endothelial function, as well as the renin-angiotensin-aldosterone system [38, 95, 96]. There are several mechanisms important to the development of a healthy placenta that involves vitamin D metabolites. Vitamin D metabolites regulate genes

important for implantation (HOXA10), trophoblast invasion, as well as anti-inflammatory responses in maternal decidua and fetal trophoblasts [38]. Vitamin D sufficiency helps maintain an appropriate inflammatory response and maternal-fetal interface by promoting an innate response to infection and by preventing an over-reactive inflammatory adaptive immunity [97]. Also significant in the development of a normal placenta, $1,25(OH)_2D$ has a role in decidualization by transforming endometrial cells into decidual cells[38]. This process is vital for proper implantation in early pregnancy. Vitamin D sufficiency also maintains endothelial function by limiting production of proinflammatory cytokines [98]. Limiting cytokine production in the presence of vitamin D sufficiency may reduce endothelial damage and reduce the amount of factors released into maternal circulation. In addition, vitamin D regulates angiogenesis through a direct effect on vascular endothelial growth factor (VEGF) gene transcription impacting endothelial cell function [99]. More specifically, $1,25(OH)_2D$ increases VEGF production in vascular smooth muscle cells by binding to response elements in the promoter region of the gene[100]. In vitamin D deficient women, this regulation may be disrupted causing endothelial dysfunction. Lastly, $1,25(OH)_2D$ has an important role in regulating the renin-angiotensin-aldosterone system (RAAS) by suppressing renin and angiotensin levels through VDR. In vitamin D deficient people, studies have found evidence of deregulated RAAS and increased oxidative stress and high blood pressure [96, 101].

2.2.8 Studies on Vitamin D and preeclampsia

2.2.8.1 Vitamin D supplementation

There is limited research on vitamin D supplementation in pregnancy. Hollis et al did not find a significant difference in preeclampsia by vitamin D supplementation group (400 IU/day, 2000 IU/day or 4000 IU/day)[102]. The lack of significance may be due to starting supplementation too late in pregnancy to reduce the risk of preeclampsia. In the follow-up trial, there were only 4 cases of hypertension related pregnancy outcomes which included preeclampsia cases [103]. Due to this small number of cases, no significant difference between supplementation groups could be detected. Future randomized control trials (RCT) are needed using vitamin D supplementation in earlier pregnancy and using larger sample sizes.

2.2.8.2 Observational studies on vitamin D and preeclampsia

Two meta-analyses of observational studies determined an association between low 25(OH)D and risk of preeclampsia. However, both noted heterogeneity among studies [17, 18]. In one of these meta-analysis, there was an indication for a strong between-study heterogeneity on 8 observational studies (I^2 = 52.7%; P = 0.039)[13, 14, 16, 20, 23, 24, 104, 105]. Indeed, a sub-meta-analysis reported a significant association between 25(OH)D <50 nmol/L cut-point for deficiency and preeclampsia risk[20, 24, 104, 105], and no significant association between women with 25(OH)D <37.5 nmol/L and risk of preeclampsia [13, 16, 23, 75]. In addition, studies conducted in the U.S had a significant association while those outside the U.S were not significant. There were no differences by study type (cohort, nested-case-control, cross-sectional, and case-control studies). In another meta-analysis, investigators reported a pooled odds ratio of 0.52 (95% CI 0.30-0.89) with evidence of heterogeneity between studies (I^2=60%, P=0.02)[18].

Only studies that reported adjusted estimates were included in the meta-analysis [13, 14, 20, 23, 24]. All 5 of these studies used the same definition of preeclampsia. When subdividing the studies by latitude, the investigators only determined an association for studies located < 45 ◦N; however there was no difference by study size, assay or trimester of blood collection. In contradiction to the previous meta-analysis, there was no difference in effects for studies using 25(OH)D <37.5 or 25(OH)D <50 nmol/L as 25(OH)D cut-points. Overall, both studies suggest an association between higher serum 25(OH)D and a reduced risk of preeclampsia despite the heterogeneity between studies.

A recent nested case-control study was not included in either meta-analyses [25]. These investigators measured 25(OH)D at <20 weeks gestation among 169 preeclampsia cases and 1,975 controls. They reported a dose-response relationship between increasing maternal 25(OH)D levels and decreasing risk of preeclampsia. The relationship had a plateau at 25(OH)D =50 nmol/L after adjusting for confounders. In addition, there was a 2-fold increased risk of preeclampsia in women with 25(OH)D <30 nmol/L compared with those who had 25(OH)D ≥ 50 nmol/L. These results indicate very low 25(OH)D (<30 nmol/l) may be an important modifiable risk factor to preeclampsia.

2.2.8.3 Severe Preeclampsia

Vitamin D deficiency may be important in severe preeclamptic cases. Baker et al conducted a nested case-control study with 51 severe preeclampsia cases and 201 controls[24]. The odds increased 5-fold for women with 25(OH)D<50nmol/L at 15-20 weeks of gestation compared with women with 25(OH)D >50 nmol/L (OR 5.41 95% CI (2.02, 14.52)). In a separate case-cohort study, Bodnar et al reported vitamin D deficiency to be a risk factor for severe preeclampsia in a study population with 57% 25(OH)D < 50 nmol/L[15]. There was a 40%

reduced risk for severe preeclampsia in women who had 25(OH)D ≥ 50nmol/L compared with 25(OH)D < 50 nmol/L (OR 0.65 95% CI (0.43, 0.98)). A study further subdivided severe preeclampsia into early-onset preeclampsia[105]. Investigators used 50 participants with early-onset preeclampsia with severe symptoms and 100 controls matched on race and gestational age of sample collection. There was a 63% decreased odds of preeclampsia per every 25 nmol/L increase in 25(OH)D concentration, after adjusting for BMI, maternal age, race, and gestational age (OR 0.37 95% CI 0.22, 0.62). These findings may indicate the importance of investigating 25(OH)D and preeclampsia risk by preeclampsia subtype.

2.2.9 Conclusion

Overall, observational studies suggest low 25(OH)D concentration in serum is associated with increased risks of preeclampsia and severe preeclampsia. However, there is evidence of heterogeneity of the association between the studies. Future studies should explore the association of vitamin D deficiency and preeclampsia by subtype due to the heterogeneity of the disease.

2.3 VITAMIN D AND GENETICS

There is evidence of a genetic contribution of circulating 25(OH)D concentration. Data on twin studies and family based studies estimated heritability of vitamin D status to be 23-80% [106, 107]. Due to the importance of vitamin D in pregnancy health, there needs to be effort to determine the genetic determinants of vitamin D deficiency in pregnancy as well as outcomes

like preeclampsia. The following reviews three important vitamin D metabolism candidate genes: *GC, CYP27B1*, and *VDR*.

2.3.1 *GC* gene

The vitamin D metabolites 25(OH)D and 1,25$(OH)_2$D circulate in blood by attaching to vitamin D binding protein (DBP). DBP is a protein encoded by *GC* gene or group-specific component gene which is located on 4q12-q13 in the human genome on the reverse strand. The protein can be found in plasma, ascitic fluid, and cerebrospinal fluid and on the surface of many cells, which bind and transport vitamin D and metabolites of vitamin D to target tissues. In serum, half-life of DBP is 2.5-3 days. In the 'free-hormone-hypothesis', only the unbound 25(OH)D is bioavailable while DBP modulates and circulating free and bioavailable 25(OH)D concentration [108]. However, this might be a simplification for vitamin D [109]. There are reservations for this hypothesis in relation to vitamin D metabolites, including evidence that serum concentrations of free 1,25$(OH)_2$D is much less than the concentrations normally bound to VDR [110, 111]. Indeed, evidence support DBP is important in maintaining levels of 25(OH)D and 1,25$(OH)_2$D in serum, as well as modulating the bioavailability of these metabolites to peripheral tissues [52]. In addition, DBP bound 25(OH)D may be needed for 25(OH)D acquisition in some cells [112-114], and a randomized control trial reported serum 25(OH)D concentration is still the best metric of vitamin D status despite the free hormone hypothesis [115].

Transportation of vitamin D metabolites may be affected by variations in the *GC* gene. DBP is highly polymorphic in humans. Variations in *GC* gene is associated with varying 25(OH)D levels, and modulates DBP levels and affinity for 25(OH)D [116]. Poor binding may

reduce the 25(OH)D concentration and other vitamin D metabolites in serum, while low DBP concentration may increase risk of vitamin D deficiency when vitamin D sources are scarce [117]. In DBP null mice, vitamin D deficiency is rapidly developed when fed a low vitamin D diet [118]. This phenomenon may be important for disease risk by reducing the amount of $1,25(OH)_2D$ in serum. Studies have reported increased risk of disease for some variants in *GC* gene, including breast cancer, prostate cancer, colorectal neoplasm, COPD, and Grave's disease [119-123]. Below is a summary on the effects of *GC* variants on 25(OH)D levels in non-pregnancy. To date, there are no studies on 25(OH)D concentration and *GC* variants in pregnancy, as well as studies on the effects of *GC* variants on pregnancy outcomes.

2.3.1.1 Association with 25(OH)D in non-pregnant adults

Nine studies examined *GC* variants and vitamin D concentration in non-pregnant adults (Appendix Section: Table 1) including two genome-wide studies [31, 119, 122-136]. The most consistent results occurred for two missense mutations: rs4588 and rs7041. The mutation variant rs4588 results in a threonine to lysine amino acid change at position 420, while rs7041 results in an aspartic acid to glutamic acid change at amino acid position 416 [137]. Two other variants, rs2282679 and rs1155563, are commonly used as a proxy to rs4588 due to high linkage disequilibrium [124, 126].

Consistently, carriers of the rs4588 C allele had lower 25(OH)D concentration. The SNP rs4588 was associated with serum 25(OH)D levels in Polish [138], Canadian, and American (white, Hispanic and African Americans) [22, 133, 139], Dutch [122], British [140], and Han Chinese [141] adults. Mean 25(OH)D concentrations differed by genotype and per risk allele. In addition, both rs4588 and rs2282679 genotypes had genome-wide significance in two genome-wide association studies [124, 126]. These studies provide stronger evidence that these variants,

or variants in linkage disequilibrium with these variants, are important in predicting 25(OH)D levels. More specifically, one GWAS reported a 49% increased risk for vitamin D deficiency for the minor allele of rs2282679 in Caucasian participants [124].

There is consistent evidence indicating 25(OH)D concentration is lower in those who carry the rs7041 T allele [126, 129-131, 133, 142]. In two studies, participants carrying two rs7041 variants had the lowest mean concentration of 25(OH)D while wild-type carriers had the highest mean[126, 131]. Furthermore, there was a decreased 25(OH)D concentration per risk allele in three separate studies [129, 133, 141]. The variant also reached genome-wide significance in two genome-wide association studies, supporting a role of rs7041 in predicting vitamin D status [124, 126].

Three studies have noted 25(OH)D concentration can differ significantly by three common haplotypes of rs4588 and rs7041 variants [22, 119, 130]. In these studies, participants had lowest 25(OH)D concentration when they carried both variant alleles of rs7041 and rs4588. Using cross-sectional data on community-dwelling adults, investigators reported rs7041 and rs4588 variants accounted for 9.9% of the variation in total 25(OH)D levels, while season and race explained 10.5% and 7.3%[22]. Furthermore, there was large heterogeneity of rs7041 and rs4588 between blacks and whites in the study population. The T/C genotype (Gc1F) was reported in 92.7% in blacks versus 6.0% of whites, and the T/A genotype (Gc2) was reported in 2.1% of blacks compared to 76.0% of whites. Gc1F haplotype is known for higher affinity for 25(OH)D and low DBP, while Gc2 has a lower affinity for 25(OH)D and is associated with higher DBP levels[109]. These relationships may be important when low levels of DBP may lead to an increased risk of 25(OH)D deficiency when vitamin D sources are rare, including increased skin pigmentation, lack of supplementation, or food sources.

The remaining variants were associated with 25(OH)D for one study each: rs12512631 (3' downstream region of *GC*)[125], rs222020[128], rs2298849[31], rs3755967[131], and rs17467825[131]. Some of these variants are in LD with rs4588 or rs7041, while others are not. Unfortunately, the functional consequences of these variants are still unclear. Finding causal variants among the coding and promoter variants has been unsuccessful.

2.3.1.2 Conclusions

Overall, there is strong evidence that genetic variants in *GC* may have an important role in determining 25(OH)D levels and health outcomes. There are two common missense variants that impact 25(OH)D concentration in non-pregnant adults. However, studies in pregnancy are sparse. Due to the physiological changes in pregnancy, including increased DBP levels, studies in pregnancy may be informative. Researching these variants in pregnant adults is important since there is consistent evidence that common *GC* variant genotypes impact 25(OH)D concentration in non-pregnant adults.

2.3.2 *CYP27B1* gene

Absorbed and synthesized vitamin D needs to hydroxylated multiple times before becoming the active metabolite of vitamin D. Hydroxylation of 25(OH)D to form $1,25(OH)_2D$ is done by an enzyme encoded by *CYP27B1*. *CYP27B1* encodes a protein from a family of cytochrome 450 enzymes. Variations in the gene may influence 25(OH)D and $1,25(OH)_2D$ serum concentration by changing the rate of hydroxylation[143]. Variants in *CYP27B1* have been noted to cause a deficiency in $1,25(OH)_2D$ leading to a disorder known as vitamin D-dependent rickets[144]. Other diseases related to CYP27B1 include multiple sclerosis, type 1

diabetes, and Addison's disease [140, 145-147]. Summarized below are studies on 25(OH)D concentration in non-pregnant adults, as well as one pregnancy related study.

2.3.2.1 Association with 25(OH)D in non-pregnant adults

The relationship of *CYP27B1* variants with 25(OH)D were assessed in several RCTs and observational studies (Appendix Section: Table 2)[31, 128, 129, 134, 135, 148-150]. However, only a few studies reported differences in 25(OH)D concentration by *CYP27B1* genotype. Other studies had low coverage of variants included in the genotyping process [125, 129, 139]. Low coverage limits the ability of the study to find an association. Also, small sample sizes limited the ability of these studies to detect a difference by genotype.

Rs10877012 is one variant in *CYP27B1* that was associated with varying 25(OH)D concentration levels [31, 151]. This variant is a promoter polymorphism (position 1260) which may affect 25(OH)D concentration by reducing synthesis of $1,25(OH)_2D$[152]. In one study, wild-type carriers had higher concentration of 25(OH)D compared to the variant carriers in summer and autumn seasons[151]. There was no association in winter or spring. This seasonal difference may be due to an increased 25(OH)D load in summer months when skin synthesis of vitamin D is highest[151]. A separate study described African-American variant carriers as having decreased 25(OH)D concentration compared to wild-type carriers[31]. The allelic frequency for the variant is highest only in African-Americans which may increase the power of detecting an association in African-Americans compared to other races. One study on pregnant women with gestational diabetes mellitus also suggest a difference in 25(OH)D concentration by *CYP27B1* genotype (rs10877012) (Appendix Section: Table 4)[148]. Overall, there is evidence that promoter region variant of *CYP27B1* gene may be an important risk factor for vitamin D deficiency.

Carriers of two other variants in *CYP27B1* were reported to have different 25(OH)D concentration by genotype[153]. 25(OH)D concentration was increased for carriers of rs703842(C/T) or carriers of rs4646536(C/T) in a Canadian multiple sclerosis study. Unfortunately, the effects of these variants were not repeated in other studies which may have had poor coverage of the two variants as well different ethnicities/races (Hispanic and African Americans).

2.3.2.2 Conclusions

Only one variant in *CYP27B1* has been repeatedly associated with 25(OH)D concentration in non-pregnant adults. It is biologically plausible for *CYP27B1* to influence 25(OH)D concentrations as it converts 25(OH)D to its active hormone, $1,25(OH)_2D$. The lack of association in other studies may be due to small sample sizes and poor genotype coverage of candidate genes. Future studies using larger sample sizes and more gene coverage are needed to replicate these findings.

2.3.3 VDR gene

The *VDR* gene, which spans 63.49 kb on the reverse strand of chromosome 12q12-q14, is another candidate gene for vitamin D metabolism. Through *VDR*, $1,25(OH)_2D$ regulates up to 3% of the human genome which include functions in growth regulation, DNA repair, differentiation, apoptosis, membrane transport, metabolism, cell adhesion, and oxidative stress[118]. The gene is expressed in almost every human tissue; although at different concentrations[154]. The *VDR* is a nuclear receptor for steroid hormones and mediates its action via two main methods, one of which is at the gene expression level and the other a more rapid

non-genomic response involving signaling cascades[28]. The active metabolite, 1,25$(OH)_2$D, is transported by DBP to target cells that express vitamin D receptors. *VDR* gene encodes this nuclear hormone receptor for 1,25$(OH)_2$D. After 1,25$(OH)_2$D interacts with *VDR*, *RXR* attaches to the complex which forms a heterodimer. The heterodimer complex activates expression of genes involved in important biological mechanisms, like bone homeostasis and the inflammation process. These processes may be altered by variations in the *VDR* gene making up several important and common genotypes [118]. The list of diseases associated with *VDR* variants is extensive and includes multiple sclerosis, prostate cancer, breast cancer, and heart disease[155-158].

2.3.3.1 Association with 25(OH)D in non-pregnant adults

Three *VDR* variants were associated with serum 25(OH)D levels, including rs2228570 (previously known as rs10735810), rs2544037, and rs10783219 (Appendix Section: Table 3). The rs2228570 variant modifies the translation initiation site on *VDR*, changing the function of the receptor. The mutation changes the threonine into cysteine in the first initiation codon (ATG), introducing a FokI site to exon 2 and creating a restriction site. The presence of the restriction site creates a longer *VDR* protein isoform which has lower transcriptional activity than in the absence of the restriction site. Consistently, the variant carriers of Fok1 have higher serum levels compared to wild-type healthy individuals[32, 159].

Another important variant is rs2544037, which may be in linkage disequilibrium with a functional unknown variant in *VDR*. Wang et al conducted a Z-score weighted meta-analysis across several cohorts and reported an association for this variant; however the significance was lost in the replication cohort of 9366 participants[160]. The lack of replication may mean the variant is not the functional variant. Instead, it may be in linkage disequilibrium with the

functional variant. Carriers of a third variant, rs10783219, had a lower concentration of 25(OH)D compared to wild-type individuals. This association was also noted in a Hispanic subpopulation living in San Antonio (beta= -0.164; 95% CI -0.108, -0.272)[161]. However, the finding may be population specific since it was not replicated in two subpopulations in the same study.

One study on pregnant women assessed 4 *VDR* variants: ApaI, BsmI, FokI and TaqI. In the cohort of 354 women, 25(OH)D was also analyzed at 28-30 weeks of gestation[42]. The investigators did not find maternal 25(OH)D varied by genotype. However, this study was limited by small sample size for ApaI aa, Bsm1 bb and TaqI tt genotypes (n=2, 2, and 1, respectively) and by the small number of genotyped SNPs.

Overall, there was some evidence that *VDR* variant carriers may be associated with different levels of 25(OH)D concentration compared to wild-type individuals. However, two of the three variants might not be the functional variants underlying the association. Other studies did not find an association due to poor coverage of the *VDR* gene.

2.3.3.2 Variants and Pregnancy or Infant outcomes

Many studies attempted to detect a difference in risk of pregnancy and infant outcomes by *VDR* genotype, including one study on preeclampsia (Supplementary Section: Table 4). These pregnancy studies were limited by low sample size and poor coverage by genotyping only 1 to 4 polymorphisms in the *VDR* gene (Apa1, Bsm1, Fok1 or Taq1) [43, 162-164]. In addition, since the actions of $1,25(OH)_2D$ are influenced by *VDR*, it is important to measure and test vitamin D status as a confounder and possible effect modifier in these studies[155]. However, some of these studies did not measure 25(OH)D concentration [43, 165].

Several limitations in the study on preeclampsia may have limited power to report an association[43]. In Brazil, investigators conducted a study on women who were healthy (n=213),

diagnosed with gestational diabetes (n=154), and diagnosed with preeclampsia (n=164). Only three polymorphisms were genotyped: FokI, ApaI and BsmI. There was no difference by maternal genotype, allelic, and haplotype frequencies between the three groups of participants. The study may have missed functional variants important to the development of preeclampsia by limiting to three polymorphisms. In addition, this study did not measure 25(OH)D concentration levels. Overall, there was a lack of association perhaps due to several important limitations.

Unlike the preeclampsia study, a nested case-control study reported FokI variant carriers had increased odds for preterm birth for a population of Israeli women[165]. This study genotyped DNA for four SNPS (FokI, ApaI, TaqI, and BsmI) from 33 cases of spontaneous preterm birth (<35 weeks of gestation) and 98 controls. The frequency of TaqI and FokI varied by case status. However, only the FokI variant significantly increased risk for preterm birth in the additive approach. More specifically, using exact logistic regression, the maternal FokI variant carriers had a 3-fold increased odds of spontaneous preterm birth compared with wild-type carriers (OR=3.3 95% CI (1.1-9.6)). Furthermore, the study reported neonatal variation in VDR did not affect preterm birth risk; instead only maternal genotype impacted risk. The study did not measure 25(OH)D concentration; therefore, residual confounding may be present.

Offspring birth size was examined using 25(OH)D status measured at 28-30 weeks of gestation[162]. Despite a lack of association between genotype and vitamin D status(as previously discussed), investigators noted evidence of effect modification by infant FokI genotype. For this study, the effect of maternal vitamin D status on birth size was analyzed by FokI, ApaI, Bsm1, and Taq1 genotypes. In infants with FF/Ff genotype, vitamin D deficiency (25(OH)D <28 nmol/L) was associated with low birth weight compared with women with 25(OH)D >28 nmol/L (3296 g versus 3584 g; p=0.02). This finding was not repeated for ff

genotype. An important limitation for this study is the low sample size since only 1-2 infants carried ApaI aa, Bsm1 bb, or TaqI tt genotypes. Overall, this study reported differences of association of 25(OH)D status on birth size by *VDR* genotype. This may be evidence that 25(OH)D load is important when considering the effect of VDR variation on birth outcomes.

Several SNPs were genotyped in a nested case-control study on the risk for small-for-gestational age by maternal genotype[166]. Small-for-gestational age (SGA) was defined as a live-born infant with less than the 10^{th} percentile of birth weight based on gender and age of a reference population at Magee-Womens Hospital. In this study, 124 cases of SGA and 1074 controls were recruited less than 16 weeks of gestation. The investigators genotyped 70 tagging SNPs on and around the *VDR* gene. One of these variants, rs11168292, was associated with SGA in white participants. For every allele, the odds for SGA increased by 70% while controlling for 25(OH)D (OR=1.7; 95%CI 1.1, 2.8). In black participants, there were three variants that were significantly associated with SGA when controlling for 25(OH)D. These included rs11168287, rs2239179, and rs3782905. Carriers of minor alleles of rs11168287 had a 60% decreased odds of SGA compared to wild-type individuals, while rs2239179 carriers had a 70% decreased odds of SGA (OR 0.4 95%CI 0.2,0.8); OR 0.3 (0.1,0.7)). On the contrary, carriers of rs3782905 had almost a 6-fold increased odds of SGA (OR 5.3, 95%CI 1.8, 17.4). These variants may be important predictors for small-for-gestational age independent of 25(OH)D concentration; however, there are no other published studies on SGA and *VDR* variants.

There was a difference in low birth weight by maternal *VDR* genotype where low birth weight was defined as a live born infant weighing less than 2500 grams at birth[167]. Swamy et al examined the effects of 38 *VDR* polymorphisms on birth weight in a cohort of pregnant women recruited between 18 and 28 weeks of gestation. They selected haplotype tagging SNPs

for Yoruban and Caucasian populations of the HapMap Project. Using tagging SNPs from both populations increased representation for both blacks and whites in the study population. One variant was significantly associated with low birth weight in non-Hispanic blacks (rs7975232; p=0.0009). This association was not seen in non-Hispanic whites. The investigators did not measure 25(OH)D concentration; therefore, the association may have residual confounding.

2.3.3.3 Conclusion

Currently, there is no evidence that *VDR* SNPs impact 25(OH)D concentration in pregnancy. However, studies have reported *VDR* SNPs may have a role in preterm birth, low birth weight, SGA, and birth size. More research is needed to find if *VDR* genotypes are important in the development of preeclampsia. There are several limitations on these pregnancy studies such as limited number of SNPs, not controlling for 25(OH)D concentration, and small sample size. Replication is needed to determine if findings are repeatable and generalizable in other study populations.

2.3.4 Functional variants

There were three variants that have a functional significance in pregnancy, rs10877012 (*CYP27B1*), [148] [148] rs10735810 (commonly known as Fok1 in *VDR*), and rs7975232 (*VDR*). Variant rs10877012 is a promoter polymorphism which may change the conversion of 25(OH)D to 1,25 $(OH)_2D$, subsequently changing the amount of 1,25$(OH)_2D$. The Fok1 variant modifies the translation initiation site on *VDR*, changing the initiation site into a restriction site. The intron variant, rs7975232, is part of a haplotype that is associated with mRNA stability[167]. A change

in mRNA stability can affect the amount of protein produced. These variants may play an important role in vitamin D deficiency in pregnancy as well as pregnancy outcomes.

2.4 CONCLUSION

Overall, more research is needed to determine if risks of vitamin D deficiency and preeclampsia vary by risk allele. Past research has shown variants of *GC, CYP27B1* and *VDR* genes are variable between people which may affect 25(OH)D levels. Furthermore, there is increasing evidence that genetics may have an important role in several pregnancy and birth outcomes. However, there are several limitations including poor gene coverage, small sample sizes, and not controlling for 25(OH)D concentration. Our goal is to fill in these knowledge gaps by using samples and data from two large pregnancy studies.

Table 1: Risk variants on *GC* gene for vitamin D deficiency in non-pregnant adults

Variant	Type of study	Type of effect	Sample size	Population	Authors
rs7041	Case-control	By genotype	749 cases (prostate cancer) and 781 controls	USA	Ahn, 2009[125]
	GWAS		4501 cohort	USA	Ahn, 2010[126]
	Cross-sectional, family-based	Per allele	1190	USA	Engelman, 2008[107]
	Cohort	By genotype	9528	Norway	Jorde, 2012[131]
	Cohort	Per risk allele	741	Canada	Sinotte, 2009[168]
	GWAS		33996	USA	Wang, 2010[124]
rs4588	Cross-sectional, family-based	Per allele	1190	USA	Engelman, 2008[107]
	Cohort	By genotype	98		Fu, 2009[142]
	Case-control	By genotype	332 cases (grave's disease) and 185 controls	Poland	Kurylowicz, 2006[123]
	Cohort	Per risk allele	3210	Chinese Hans	Lu, 2012[141]
	Cohort	Per risk allele	741	Canada	Sinotte, 2009[168]
	Cohort	Per risk allele	2897	Shanghai	Zhang, 2013[135]
rs2282679	Case-control	By genotype	749 cases (prostate cancer) and 781 controls	USA	Ahn, 2009[125]
	Cohort	By genotype	9528	Norway	Jorde, 2012[131]
	Cohort	Per risk allele	3210	Chinese Hans	Lu, 2012[141]
	Cohort	By genotype	792	USA	Signorello, 2011[31]

Table 1 Continued

	GWAS		15 cohorts	USA	Wang, 2010[124]
rs1155563	GWAS		4501 cohort	USA	Ahn, 2010[126]
	Cohort	By genotype	9528	Norway	Jorde, 2012[131]
	Cohort	Per risk allele	3210	Chinese Hans	Lu, 2012[141]
rs12512631	Case-control	By genotype	749 cases (prostate cancer) and 781 controls	USA	Ahn, 2009[125]
rs17467825	Cohort	By genotype	9528	Norway	Jorde, 2012[131]
rs222020	Cohort	Per risk allele	156		Bu, 2010[128]
rs2298849	Cohort	By genotype	792	USA	Signorello, 2011[31]
rs3755967	Cohort	By genotype	9528	Norway	Jorde, 2012[131]
rs4588 and rs7041 haplotype	Case-control	By haplotype	1402 cases and 2608 controls	German	Abbas, 2008 [119]
	Cohort	By haplotype	1317	Netherlands	Fang, 2010[130]

Table 2: Risk variants on *CYP27B1* gene for vitamin D deficiency in non-pregnant adults

Variant	Type of study	Type of effect	Sample size	Population	Authors
rs10877012	Cohort	Per risk allele	7288	United Kingdom	Hypponen, 2009[150]
	Case-control	By genotype	222 cases, 104 controls	Germany	Ramos-Lopez, 2008[148]
	Cohort	By genotype	758	USA	Signorello, 2011[31]
rs4646536	Twin-study	Per risk allele	198	Canada	Orton, 2008[149]
rs7033841	Twin-study	Per risk allele	198	Canada	Orton, 2008[149]

Table 3: Risk variants for vitamin D deficiency on *VDR* in non-pregnant adults

Variant	Type of study	Type of effect	Sample size	Population	Authors
rs10783219	Family-study	Per risk allele	1190	USA	Engelman, 2008[129]
rs10735810	Case-control	By genotype	749 vitiligo cases and 763 controls	USA	Li, 2012[159]
	Cohort	Per risk allele	198	Canada	Orton, 2008[149]
	Case-control	By genotype	212 cases (multiple sclerosis)	The Netherlands	Smolders, 2009[32]

Table 4: Risk variants for infant and birth outcomes

Author	Outcome	Type of study	Sample size	Population	Significant variant	Type	Strength of Association
Aslani, 2011 [163]	gestational diabetes	cohort	303	Iran	*VDR*: rs10735810	N/A	N/A
Bodnar, 2010 [166]	small-for gestational age	case-control study	111 cases, 301 controls	USA	*VDR*: rs11168287 rs2239179 rs3782905 rs11168292	per allele	White women: Rs11168292 OR=1.70 CI (1.1,2.8) Black women: Rs11168287 OR=0.40 CI (0.2,0.8) Rs3782905 OR=5.30 CI (1.8,17.4) Rs2239179 OR=0.30 CI (0.1,0.7)
Fang, 2010 [164]	premature rupture of membranes	case-control	206 cases, 287 controls	China	*VDR*: C352T	By genotype	TT genotype OR 6.08 CI (1.39–26.60)
Manzon, 2014 [165]	preterm birth	case-control	33 cases, 98 controls	Israel	*VDR*: rs10735810	By variant	OR=3.32 CI (1.14-9.63)
Morley, 2009 [42]	birth weight, subscapular thickness, and suprailiac skinfolds; 25(OH)D	Cohort		Australia	*VDR*: rs10735810	By genotype	Subscapular skinfolds: ff versus Ff or F 5.7 versus 6.4mm CI for difference (-1.2-0.3) Suprailiac skinfolds: Ff versus Ff or FF 6.3 versus 7.0 mm CI of difference(-1.2, -0.2)

Table 5. Risk variants for pregnancy and infant health outcomes.

Author	Outcome	Type of study	Sample size	Population	Significant variant	Type	Strength of Association
Rezende, 2012[43]	Preeclampsia	case-control	162 cases, 213 controls	Brazil	NA	N/A	N/A
Swamy, 2011[167]	birth weight	Cohort	615	USA	VDR: rs7975232		Covariate adjusted p-value for non-Hispanic blacks: 0.0009
Ramos-Lopez, 2008[148]	25(OH)D in gestational diabetes patients	cohort	222	USA	CYP27B1: Rs10877012	By genotype and per allele	<50nmol/L CC 68 (47.9) [n(%)] CA 62 (43.7) AA 12 (8.5) >50 nmol/L CC 31 (38.8) CA 31 (38.8) AA 18 (22.5) P=0.0127 <50nmol/L C allele=198(63.3) A allele=86(30.3) >50 nmol/L C allele=93(58.1) A allele=67(41.9)

3 . METHODS

3.0 STUDY POPULATION

We will use two study populations to achieve our research goals: EVITA and CPP. The first aim will use a case-cohort study called EVITA. The genetic aims (aims 2 and 3) will use an additional study called Collaborative Perinatal Project (CPP). Candidate gene association studies are traditionally case-control studies using healthy controls [169]. Therefore, we will only use healthy controls for our genetic aims following a case-control design. We will combine the two studies (EVITA and CPP) through a meta-analysis. Separately, the two study populations have a potential for false positives and false negatives due to small sample size. By using a meta-analysis approach on the two study populations, we will increase power to find an association with risk variants. In addition, by using a diverse cohort (CPP), there will be more power to detect an association for variants commonly found in mothers with African ancestry.

EVITA uses data and banked serum samples from women who had aneuploidy screening at ≤20 weeks gestation and delivered at Magee-Womens Hospital of UPMC in Pittsburgh, Pennsylvania. These samples were collected and stored in -80◦C freezers. Pregnancies that were eligible for this study were identified after merging a validated perinatal database that contains clinical information on all hospital deliveries with a database of all pregnancies that received screening at the hospital's Center for Medical Genetics and Genomics. This case-cohort study uses 650 preeclampsia cases, as well as 2327 randomly selected subcohort. The subcohort was selected from 65,867 singleton live-born infants. The subcohort was predominately non-Hispanic

white, college graduates, married, non-smokers, and covered by private insurance at the start of the index pregnancy (Table 1). About half of the subcohort was normal weight and were nulliparous before pregnancy. Compared with the subcohort, preeclampsia cases were older, and more likely to be Black, nulliparous, and covered by Medicaid. They also completed fewer years of education and had a higher prepregnancy BMI.

The collaborative perinatal project (CPP) uses data and samples collected $\leq$26 weeks from 1959-65 at 12 U.S medical centers for a total of 55,908 participants. Data were collected by an in-person interview (sociodemographic factors, medical history, and obstetric history) at first prenatal visit, and obstetric and medical events were recorded every 8 weeks after initial visit. In addition, blood samples were collected and random urine samples were tested for albumin every 8 weeks. The case-cohort uses n=717 preeclamptic cases and 3068 in the subcohort. 5.5% of the total cohort was selected (Table 1). About fifty percent of the CPP subcohort was white, 50% smoked, and 60% were between 20-29 years old. The subcohort was predominately multiparous, had normal BMI, and married. Compared to the subcohort, the cases were more likely to be white, unmarried, had higher BMI, and completed less education.

There are some differences between the two cohorts that should be noted. A greater proportion of CPP is black, multiparous, and younger, had lower BMI, completed less education, and had higher rates of smoking. One of the biggest advantages of using the CPP cohort in the genetics analysis is the high proportion of blacks in the study population. However, it is possible the two study populations will have different risk variants by chance. The meta-analysis approach will reduce any of these spurious associations. High rates of smoking in CPP may hide the true effects of risk variants on preeclampsia since smoking decreases risk of

preeclampsia[170]. Adjusting for smoking might not fully correct for this strong effect causing residual confounding.

3.1 PREECLAMPSIA DEFINITIONS

In the EVITA cohort, ICD-9 codes were used to define preeclampsia (642.4-642.7). Severe (642.5 and 642.6) and mild subtypes of preeclampsia (642.4) were also identified. Gestational age at delivery was determined by best obstetric estimate comparing menstrual dating and ultrasound and was ascertained from the perinatal database. Our reliance on correct assignment of ICD-9 codes for preeclampsia may pose as a potential source of misclassification of preeclampsia[171].

In the CPP cohort, medical records were used to define preeclampsia as gestational hypertension and proteinuria, and return of abnormalities to normal in the postpartum period. More specifically, gestational hypertension was defined new onset (>24 weeks) of two or more measurements of systolic blood pressure ≥140 mm Hg and/or diastolic blood pressure ≥90 mm Hg. Proteinuria was determined by two random urine dipsticks of 1+ protein or one dipstick of 2+ protein. We would expect preeclampsia to be classified more accurately for CPP than the EVITA cohort since the CPP cohort was followed prospectively and we used the current definition of preeclampsia to define cases.

3.2 VITAMIN D

In EVITA, maternal aneuploidy screening samples were held at -20°C for one month before transfer to -80°C for long-term storage. Only samples drawn ≤ 20 weeks of gestation were used for this study because this generally precedes the clinical onset of preeclampsia. An aliquot of each sample that was previously thawed once was sent to a Vitamin D External Quality Assessment Scheme–proficient (DEQAS, London, United Kingdom) and Clinical Laboratory Improvement Amendments–certified (Centers for Disease Control and Prevention, Atlanta, Georgia) laboratory. The laboratory assayed total 25-hydroxyvitamin D (25(OH)D) (calculated as $25(OH)D_2 + 25(OH)D_3$) using liquid chromatography–tandem mass spectrometry[172]. Intra-assay variation was 9.6% and the inter-assay variation was 10.9%. We will test multiple categories of serum 25(OH)D concentration , as well as use cubic splines, quadratic, quintiles[173]. In addition, we will test multiple knots and knot placements, and use two criteria to select the final specification (BIC, F)[174]. We will also evaluate the sensitivity of our results by excluding n=2 observations with 25(OH)D >175 nmol/L.

For CPP, maternal serum samples were stored at −20°C with no recorded thaws. Random samples were selected from samples drawn at ≤26 weeks. Sixty-four percent of the samples were drawn 20-26 weeks of gestation, while 28% and 8% were drawn at 13-20 weeks and <13 weeks. Like EVITA, serum was shipped to the laboratory of Dr. Michael Holick and assayed for total 25-hydroxyvitamin D (25(OH)D) [$25(OH)D_2 + 25(OH)D_3$] using liquid-chromatography-tandem mass spectrometry based on National Institute of Standards and Technology (NIST) standards. The assay had a coefficient of variation of 6.0%.

There are differences in freezer temperatures and length of frozen time between the two studies. However, a previous study on 25(OH)D concentration in the CPP cohort reported no evidence of 25(OH)D degradation [175]. Long-term storage in 20◦C freezer will not cause a loss of 25(OH)D[176]; therefore storage differences between the two studies should not cause a significant difference in 25(OH)D concentration.

3.3 GENETIC VARIANTS

Candidate vitamin D pathway genes were selected from vitamin D literature and two recent GWAS [31, 32, 119, 123, 129, 141, 153, 159-161, 164-166, 177]. By focusing analysis on candidate genes, there is increased probability of finding associations. However, there is always a possibility of missing important variants in other genes. Potentially Functional SNPs within these candidate genes (*GC*, *CYP27B1*, and *VDR*) were selected for inclusion and additional tagging SNPs were identified using HapMap [178]. Using haplotype tagging minimizes redundancy data from the same haplotype block. Tagging SNPs were selected for CEU (Utah residents with ancestry from northern and western Europe) population first, and supplemented with additional tagging SNPs from ASW (Americans of African Ancestry in Southwest USA) population in order to represent the black and white populations in the two cohorts. By selecting the variants by race, there is less risk of excluding race/ethnic-specific variants. We specified for minor allele frequency (MAF) >10% since we are interested in common variants. In addition loci with a low MAF (<10%) have a significantly lower power to detect genotypic risk ratios than loci with higher MAF; therefore MAF <10% may result in false findings[179]. In total, 499

SNPs were selected: 128 markers for Genetic Ancestry, 39 for *CYP27B1*, 126 were for *GC*, and 206 were for *VDR*.

Frozen serum from EVITA and CPP were thawed, cleaned, amplified, and genotyped using these selected SNPs. By genotyping both studies on each open array plate, the chance that any genotyping differences by study will be caused by genotyping error is reduced. Genotyping was conducted at the Graduate School of Public Health (University of Pittsburgh) using the QuantStudio 12K Flex platform. All SNPs were genotyped for every sample. Genotypes were analyzed using the TaqMan Genotyper software (version 1.3.1) and visual assessment of the data was used for confirmation.

3.4 SAMPLE SIZE

Aim 1

Power estimates were calculated for preeclampsia and subtypes of preeclampsia (mild, severe, early-onset, and late-onset preeclampsia) in the EVITA cohort. STATA software[180] was used to estimate power for odds ratios=1.5, 1.8, 2, and 2.5 using the known sample sizes (Table 2). Since the outcome rate is low, the case-control approach can be used to estimate power[181]. Overall, power is > 80% to detect an association between 25(OH)D <25 nmol/L and overall preeclampsia, as well as late-onset, severe and mild forms of preeclampsia. We will not have power to report an association for early-onset preeclampsia.

Aim 2 and 3

For the genetic study, power was calculated using the total 3576 and 3111 samples for EVITA and CPP, respectively (Table 3). There was a control to case ratio of 4:1 and 2:1 for

whites and blacks in EVITA and 3.6:1 and 2:1 ratio for whites and blacks in CPP. Using Quanto version 1.2[182], power was calculated for each study by race for three inheritance modes: recessive, dominant, and log additive (Table 4). Calculations assumed loss of 20% of the individual samples and 10% of SNPs due to poor genotyping, sample quality, or failure of Hardy-Weinberg equilibrium. The calculations were based on allelic frequencies of 0.1, 0.2, and 0.3. Due to multiple comparisons and the assumption that 10% of SNPs will fail, we used a p-value of 0.00015 (p-value 0.05/340 SNPs). This is a conservative p-value for the calculations. Based on our assumptions, we will have power (>80%) to detect an odds ratio of 3 for whites and blacks in CPP and EVITA for variants with allele frequencies ≥ 0.1 and have either a log additive or a dominant mode of inheritance (Table 4). In addition, there is power to report an odds ratio of 2 for dominant and log additive modes of inheritance for CPP cohort (white and black) and EVITA (white). There is limited power to detect an association for recessive variants.

3.5 CONFOUNDERS

Potential confounder variables for EVITA and CPP are from the perinatal database and records from in-person interviews, respectively. Both sources provided maternal self-reported race/ethnicity, pre-pregnancy BMI (<18.5, 18.5-24.9, 25-29.9, ≥30), diabetes status(yes, no), education(<12 years, 12 years, and >12 years), insurance (private, Medicaid), marital status(single, married), maternal age (<20, 20-29, ≥30), gestational age at delivery(continuous), parity, and season of blood draw (winter (December–February), spring (March–May), summer (June–August), or fall (September–November). Gestational age at delivery in CPP was based on the mother's report of the first day of her last menstrual period. In comparison, gestational age at

delivery in EVITA was determined by best obstetric estimate comparing menstrual dating and ultrasound and was ascertained from the perinatal database. EVITA also contains additional information on provider type (hospital outpatient resident clinic, hospital-affiliated private practice). A composite socioeconomic status score was also available for CPP, which combines education, occupation, and family income data.

3.6 ANALYSIS

Aim 1

In the EVITA case-cohort, a total of 950 pregnancies are missing maternal height (which is needed to calculate prepregnancy BMI) because the perinatal database did not collect height until 2003. Women are also missing prepregnany weight (n=12), maternal education (n=379), diabetes (n=136), smoking (n=4), and parity (n=2). We will use multiple imputation to address missing covariate data. The data will be imputed to create imputed datasets that assumes a multivariable normal distribution with a Markov Chain Monte Carlo approach[183, 184]. In this approach, missing values are filled in with simulated values. This is preferable to deleting samples with missing data which would cause loss of power. A disadvantage of this approach is the possibility of false values; however this is unlikely when missing is at random. We will impute the following log-transformed variables: prepregnany weight, height, diabetes status, smoking status (n=4 missing), education, and parity by including preeclampsia, 25(OH)D concentration, gestational age at blood draw, season and year of blood draw, batch number, maternal age, maternal race/ethnicity, type of insurance, provider type, marital status, gestational

age at delivery, and sample weight in the imputation model. We will also run all analyses among women with complete data (n=1494).

We will use multivariable log-binomial regression to calculate unadjusted and adjusted risk ratios and 95% confidence intervals for associations between maternal 25(OH)D and preeclampsia as well as its subtypes. Log-binomial regression assumes a binomial distribution of the outcomes. Cases are oversampled from the cohort; therefore we will include robust standard errors and weights[185]. The subjects will be weighted by the inverse of sampling probabilities. We will develop a parsimonious model[186] for overall preeclampsia by including all potential confounders (year of blood draw, batch, prepregnancy BMI, smoking status, gestational age of sample, insurance, parity, marriage status, maternal age, maternal race/ethnicity, maternal education, and season of blood draw) and removing variables whose exclusion from the model will not change the main exposure point estimate by at least 10%. A parsimonious model will reduce the risk of having an over-fit model which will hide the underlying model. We will include the same confounders for all preeclampsia subtype models. Mild preeclampsia cases will be excluded from the severe preeclampsia model. The same will be done for severe preeclampsia cases in the mild preeclampsia model, as well as early-onset cases in the late-onset preeclampsia model. In the models where 25(OH)D will be specified as restricted cubic spline terms, we will use linear combinations to calculate estimates of association for selected values of 25(OH)D concentrations relative to 25(OH)D = 75 nmol/L.

In separate sensitivity analyses, we will exclude women with chronic hypertension or diabetes, as well as women with 25(OH)D >175 nmol/L. Women with these chronic conditions will be included in the original models. Chronic hypertension and diabetes may share a causal pathway with preeclampsia; therefore removing these mothers might hide the underlying

association. We will use the synergy index [187] to test for effect modification on the additive scale between vitamin D and race/ethnicity (Black vs. non-Black), parity (nulliparous vs. multiparity), and gestational age of blood sampling (< 15 vs. > 15 weeks gestation). Synergy index creates a ratio between combined effect and individual effects which tests for interaction on the additive scale versus the multiplicative scale. Assessing interaction in the additive scale may be more appropriate for biological interactions[188, 189].

Aims 2 And 3

Quality Control

Before analysis, multiple data-quality control steps must be performed for all genotype data using Plink software [190, 191]. First, we will assess genotype calling differences by plate to detect any problems in genotyping. If there are differences in call rates by plate, further investigation will assess if there was a laboratory error.

Allelic frequencies can vary greatly between ancestral populations which may create spurious associations in gene association studies[192]. We will address this problem by using population stratification based on genetic ancestry using ancestry informative markers (AIMs) [193]. AIMs are superior to self-report since many people a composite of multiple ancestries. Individual ancestral proportions will be calculated using the Bayesian Markov chain Monte Carlo clustering algorithm [194-197] implemented in STRUCTURE 2.3 assuming K = 1 to 10. The number of subgroups will be determined by using an ad hoc statistic which is based on the rate change in the log probability of data between clusters (K) [198].

Using the adjusted race data, we will split the data by study and ancestral cluster (K) for further quality control steps. We will assess data quality to remove any sub-standard markers and samples from subsequent association analysis. First, we will carefully scrutinize the distribution

of missing genotype rates by race, study, and between cases and controls to establish the most appropriate threshold. Typically, markers with a genotype call rate less than 95% are removed [199, 200]. There is a risk of including poor quality genotypes if this threshold is set too low. Extreme differential call rates between cases and controls may result in a bias. After removing poor quality markers, we will identify SNPs with a significant deviation form Hardy-Weinberg equilibrium (p-value >0.00001) [201]and remove of all markers with a very low minor allele frequency (<5%)[202]. These additional quality steps will identify extreme genotyping errors. It is important to assess HWE and MAF by race since some variants are ancestry-specific. Other quality control measures include creating scatter plots by race for controls only, by gene, and overall. We will consider including all variants with known functional effects.

To determine population stratification, we will compare allele frequencies by race using Chi-square tests. If allele frequencies differ by race, and we ignore population stratification, there will be an increased chance of a false positive significance test[203]. We will divide the analysis for each study by race if there is evidence of population stratification. Another approach is to estimate ancestry and use this estimate in the models to remove the effect from the association test statistic[204, 205].

Analysis

Analysis will follow a case-control design. Due to differences in the study populations, separate analyses will be conducted for CPP and EVITA and results will be combined together using meta-analysis. Depending on the model (preeclampsia versus 25(OH)D), the outcome will either be preeclampsia (0=no, 1=yes) or 25(OH)D concentration (continuous). We will also consider dichotomous 25(OH)D (≤ 25 nmol, ≤ 30 nmol/l, or ≤ 50 nmol/l).

Multiple inheritance patterns must be considered in our analysis. We will test for allelic (D allele versus d allele), dominant (DD+ Dd versus dd), recessive (DD versus Dd + dd), and genotypic (DD versus Dd versus dd) inheritance modes. To test whether there was a linear dose effect of the variant alleles (log-additive genetic model for trend test), variants will be coded as 0, 1, and 2. This approach assumes additive risk while genotypic model does not provide any sense of ordering across the genotype. In addition, risk scores will be calculated where the score will be equal to the sum of the number of risk alleles across the SNPs [124]. This will allow us to determine if carriers of multiple risk variants have an increased risk of low 25(OH)D. In addition, we will determine mean 25(OH)D and preeclampsia incidence within each group of homozygous referent, heterozygous, and homozygous variant for each SNP.

Parsimonious multivariable logistic regression model will be used for categorical outcomes and multivariable linear regression models for continuous outcomes. We will adjust for batch number (in EVITA only), year drawn, and 25(OH)D concentration (for preeclampsia models). Other possible covariates to include are gestational age at blood draw, season of blood draw, prepregnancy BMI, mother's age, parity, smoking status, insurance, mother's education, and site. Covariates will be selected using 10 % change-in-estimate rule[206]. We will compare R^2, log-likelihood, AIC, BIC to help find best fit model. In addition, interactions will be examined in preeclampsia models to assess if there is effect modification of 25(OH)D concentration in the preeclampsia models.

Due to the large number of SNPs, it is important to adjust for multiple testing. Failure to adjust may produce excessive false positives or true positives may be overlooked. A conservative method to consider is the Bonferroni method (p-value/# of SNPs). If SNPs are in linkage

disequilibrium, other less conservative methods are to be considered including Holm, Sidak or FDR [207-209].

Past literature on variants in *GC* and *VDR* genes have found several common haplotypes affect disease risk. Haplotype structure will be determined by using the method of Gabriel and colleagues[210]. Each haplotype will be tested in the regression model in comparison with all other haplotypes.

To combine studies, we will use fixed- and random-effects meta-analysis to calculate effect size for each variant weighted by study size[211]. Composite odd ratios will be calculated by weighting individual odd ratios by the inverse of their variance. For a more balanced weight for each study, a random effects model will be used since it assumes that there is a distribution of true effect sizes rather than one true effect. We will use the Higginś test (I^2) to calculate heterozygosity which measures the degree of inconsistency in the results of the studies[212].

Table 6. Descriptive characteristics of EVITA and CPP.

	EVITA		CPP	
	Subcohort N=2327	Preeclampsia cases n=650	Subcohort N=3068	Preeclampsia cases n=717
Maternal race				
White	74	65	49	38
Black	19	30	44	53
Parity				
0	68	66	33	52
≥1	32	34	67	48
Maternal age, years				
<20	6.5	9.4	23	37
20-29	42	43	61	43
≥30	52	52	16	20
Prepregnancy BMI				
<18.5	3	1	10	9
18.5-24.9	49	32	72	64
25-29.9	28	32	14	18
≥30	21	35	4	9
Marital status				
Married	64	53	81	67
Unmarried	36	47	19	33
Maternal education				
<12	7	9	53	62
High school	23	30	33	30
>12	70	62	14	8
Smoking during pregnancy				
Yes	11	9	53	35
No	89	91	47	65
Gestational age of blood sample				
<13 weeks	32	40	13	8
13 to <20 weeks	67	59	33	28
20 to 26	1	1	54	64
Season of blood sample				
Winter	21	22	23	26
Spring	28	26	26	25
Summer	25	28	26	25
Fall	26	25	25	24

Table 7. Power calculations for 25-hydroxyvitamin D< 25nmol/L and preeclampsia risk.

Odds ratio	Overall Preeclampsia Power (%) (n=650)	Mild Preeclampsia Power (%) (n=449)	Severe preeclampsia Power (%) (n=201)	Early-onset preeclampsia Power (%) (n=73)	Late-onset preeclampsia Power (%) (n=577)
1.5	47.2	38.8	25.0	15.1	44.5
1.8	79.3	68.8	46.2	26.6	76.3
2.0	91.0	82.8	59.4	34.7	88.8
2.5	99.4	97.3	83.1	53.9	99.0

Table 8. Sample size for EVITA and CPP by maternal race.

	Preeclampsia Cases (n)	Healthy Controls (n)	25(OH)D < 50 nmol/l (n)	25(OH)D ≥ 50 nmol/l (n)
EVITA				
White	384	1570	228	1977
Black	181	379	392	247
CPP				
White	270	958	664	1019
Black	380	763	1223	516

Table 9. Power calculations by mode of inheritance and allele frequency for EVITA and CPP by maternal race.

	Mode of inheritance					
	Recessive		Dominant		Log additive	
EVITA	Odds Ratio 2					
Allele frequency:	White	Black	White	Black	White	Black
0.1	6%	0%	83%	19%	93%	28%
0.2	10%	1.2%	95%	33%	100%	63%
0.3	43%	5.6%	94%	32%	100%	77%
	Odds Ratio 3					
Allele frequency:	White	Black	White	Black	White	Black
0.1	8.2%	10%	100%	90%	100%	97%
0.2	79%	17%	100%	97%	100%	100%
0.3	100%	57%	100%	95%	100%	100%
CPP	Odds Ratio 2					
Allele frequency:	White	Black	White	Black	White	Black
0.1	0.0%	0.0%	53%	66%	71%	82%
0.2	4.3%	6.1%	73%	85%	96%	99%
0.3	20%	28%	70%	83%	99%	100%
	Odds Ratio 3					
Allele frequency:	White	Black	White	Black	White	Black
0.1	3.0%	4.0%	100%	100%	100%	100%
0.2	43%	53%	100%	100%	100%	100%
0.3	89%	95%	100%	100%	100%	100%

4 . LOW MATERNAL 25-HYDROXYVITAMIN D CONCENTRATION INCREASES THE RISK OF SEVERE AND MILD PREECLAMPSIA

4.0 ABSTRACT

Objective: The objective of this case-cohort study was to evaluate the relationship between maternal 25-hydroxyvitamin D (25(OH)D) concentration and preeclampsia overall and by severity.

Methods: From an eligible cohort of 12,861 women who had serum banked from aneuploidy screening, we randomly sampled a subcohort of 2327 pregnancies and all remaining preeclampsia cases (n=650 cases). Preeclampsia (defined as new-onset hypertension and proteinuria) and its mild and severe forms were identified using ICD-9 codes. Maternal serum collected at ≤20 weeks gestation was measured for 25(OH)D. We used log-binomial regression with restricted cubic splines to estimate the association between 25(OH)D and preeclampsia after adjusting for confounders.

Results: Approximately 21% of the randomly-selected sample had 25(OH)D <50 nmol/L. We found that the adjusted risk of preeclampsia declined as serum 25(OH)D increased to 50 nmol/L, and then plateaued (test of non-linearity $p<0.05$). The adjusted preeclampsia risk ratios (95% confidence intervals) for 25(OH)D at <25 nmol/L, 25-≤50 nmol/L, and 50-≤ 75 nmol/L were 2.4 (1.2, 4.8), 1.9 (0.68, 1.7), and 1.2 (0.87, 1.7), respectively, compared with those with 25(OH)D

≥75 nmol/L. Similar dose-response associations were observed with severe and mild preeclampsia.

Conclusion: Randomized trials of vitamin D supplements are needed to test the causality of this association.

Keywords: Pregnancy; Preeclampsia; Hypertension; Vitamin D deficiency; 25-Hydroxyvitamin D

List of abbreviations and acronyms

25(OH)D, maternal 25-hydroxyvitamin D; BMI, body-mass index; ICD-9, International Classification of Diseases; RR, Risk ratio; CI, confidence interval; S, Synergy Index

4.1 INTRODUCTION

Preeclampsia is a pregnancy-specific, multi-systemic condition that is defined by new-onset hypertension and either proteinuria or end-organ dysfunction at 20 weeks of gestation or later. Complicating 3–5% of pregnancies in the U.S.[91, 213], preeclampsia is a leading cause of maternal and infant morbidity and mortality [214]. The origins of preeclampsia likely lie in abnormal placental development, which induces oxidative stress and maternal systemic inflammation that lead to the clinical symptoms seen in preeclampsia [215, 216].

Vitamin D may play a role in the etiology of preeclampsia by regulating the transcription and function of genes associated with placental function, including placental invasion, normal implantation, and angiogenesis [38, 217]. Vitamin D also modulates immune function and inflammatory response [39]. Many [17, 18, 102, 103, 218-223] but not all [224-226] observational studies suggest vitamin D deficiency before disease onset is a risk factor for

preeclampsia. Some of the uncertainty in the literature may be because preeclampsia has not been studied separately by subtype [227, 228]. Classifying preeclampsia cases into more homogenous subgroups based on severity may enhance our understanding of specific exposures in the pathogenesis of preeclampsia[229]. Our objective was to evaluate the relationship between maternal 25-hydroxyvitamin D (25(OH)D) concentration at ≤20 weeks gestation and the risk of preeclampsia and subtypes based on symptom severity.

4.2 METHODS

EVITA is a case-cohort study of vitamin D and adverse pregnancy outcomes that was approved by our institution review board. The study has been described in detail previously [230]. Briefly, EVITA used data and banked serum samples from women who had aneuploidy screening at ≤20 weeks gestation and who subsequently delivered at Magee-Womens Hospital of UPMC in Pittsburgh, Pennsylvania. We merged clinical data from a perinatal database with a database of all screened samples to identify a cohort of 12,861 eligible pregnancies. We randomly selected 2327 of these pregnancies, and augmented this subcohort with all remaining preeclampsia cases (n=650 total cases).

Preeclampsia was defined as new-onset hypertension and proteinuria for the first time after 20 weeks gestation on the basis of International Classification of Diseases-9 codes (ICD-9) (642.4-642.6). Mild preeclampsia without mention of preexisting hypertension was identified (ICD-9 code 642.4). Severe preeclampsia was defined as severe preeclampsia or eclampsia with no pre-existing hypertension (ICD-9 codes 642.5-642.6). Maternal aneuploidy screening samples drawn at ≤20 weeks gestation were used because this time period generally precedes the clinical

onset of preeclampsia. Samples were stored at -80°C for up to 12 years. Serum was assayed for total 25-hydroxyvitamin D (25(OH)D) ($25(OH)D_2$+$25(OH)D_3$) using liquid chromatography–tandem mass-spectrometry [172]. We categorized serum 25(OH)D concentration as <25 nmol/L, 25 to <50 nmol/L, 50 to <75 nmol/L, and ≥75 nmol/L[173]. To model flexible non-linear relations between 25(OH)D and preeclampsia, we used restricted cubic splines with three knots [174, 231, 232].

The perinatal database provided information on potential confounders: maternal self-reported race/ethnicity, prepregnancy body mass index (BMI) [self-reported weight (kg) divided by height (m^2)], prepregnancy diabetes, education, marital status, smoking, insurance, provider type, parity, and season of blood draw.

We used multiple imputation to address missing data on height (n=950 missing, because the perinatal database did not collect height data until 2003), prepregnancy weight (n=12), education (n=379), diabetes (n=136), smoking (n=4), or parity (n=2). The data were imputed to create 22 imputed datasets that assumed a multivariable normal distribution with a Markov Chain Monte Carlo approach [183, 184]. The number of imputed datasets was based on variance and unrestricted fraction of missing information [233]. A previous publication detailing the multiple imputation methods has been described elsewhere [230].

We used multivariable log-binomial regression to calculate risk ratios (RR) and 95% confidence intervals (CI) for associations between maternal 25(OH)D and preeclampsia as well as its subtypes. Cases were oversampled from the cohort; therefore, we used robust standard errors [185]. Subjects in the subcohort were weighted by the inverse of their sampling probability. We developed a parsimonious model by removing potential confounders from a full model if their exclusion did not change the main exposure point estimate by ≥10%. Only year at

blood drawn, batch number, and gestational age of sample collection met our definition of confounding. We included race/ethnicity, parity, smoking status, prepregnancy BMI, and season out of convention. We used the synergy index (S) [187] to test for effect modification on the additive scale by race/ethnicity, parity, and gestational age of blood sampling.

4.3 RESULTS

The subcohort was predominately non-Hispanic white, college graduates, married, normal weight, nulliparous, non-smokers, and had private health insurance (Table 10). Compared with the subcohort, preeclampsia cases were older, and more likely to be Black, nulliparous, and recipients of Medicaid. They also completed fewer years of education and had a higher prepregnancy BMI. When separated into mild and severe preeclampsia cases, a greater proportion of severe cases were Black, >30 years of age, obese, unmarried, recipients of Medicaid, chronic hypertensives and multiparous as compared with mild cases.

The geometric mean 25(OH)D was 64.6 (95% CI 64.4, 64.8) nmol/L in the subcohort and 57.8 (95% CI 57.3, 58.3) nmol/L among the cases. In the subcohort, 3.3%, 18.0%, 36.8%, and 41.9% of women had 25(OH)D concentration <25 nmol/L, 25 to < 50 nmol/L, 50 to < 75 nmol/L, and ≥ 75 nmol/L, respectively. The weighted incidence of preeclampsia in the cohort was 5.0%. The unadjusted weighted incidence of preeclampsia was 9.1%, 5.8%, 5.4%, 3.9% among women with 25(OH)D <25, 25 to <50, 50 to <75, and ≥75 nmol/L, respectively.

After adjusting for confounders, there was a dose-response association between maternal vitamin D status and the risk of preeclampsia (Figure 1). As serum 25(OH)D concentration increased, preeclampsia risk declined, then plateaued at approximately 50 nmol/L. Relative to

25(OH)D of 75 nmol/L, women with 25(OH)D of 20, 30, or 40 nmol/L had 2.3-, 1.7-, and 1.3-fold increases in risk of preeclampsia, respectively (Table 11). After excluding women with prepregnancy hypertension or diabetes, the associations were attenuated. Adjusted RR (95% CI) for 25(OH)D of 20, 30, or 40 nmol/L were 1.9 (1.0, 3.5), 1.5 (0.99, 2.3), and 1.2 (0.94, 1.6) compared with 75 nmol/L. Similar results were found in the categorical analysis (Table 12). For example, after confounder adjustment, there was a 2.4-fold increased risk of preeclampsia for women with 25(OH)D <25 nmol/L compared to those with 25(OH)D >75 nmol/L.

The weighted incidences of severe and mild preeclampsia were 1.5% and 3.5%, respectively, and the incidence of each subtype increased as 25(OH)D category increased (Table 12). Confounder-adjusted associations between categories of 25(OH)D and risk of severe and mild preeclampsia were similar to results observed for overall preeclampsia. The dose-response associations indicating high risk with low 25(OH)D and a plateau at approximately 50 nmol/L were also similar for mild and severe disease (Appendix Figures 1-2).

None of these results varied by race/ethnicity [S=8.5, p=0.897], parity [S=3.4, p=0.249], or gestational age of blood draw [S=17, p=0.102].

4.4 DISCUSSION

In this large contemporary cohort of pregnancies electing aneuploidy screening, we observed a dose-response association between maternal 25(OH)D at ≤20 weeks of gestation and risk of preeclampsia. Risk declined sharply from 20 to 40 nmol/L, and then plateaued at 50 nmol/L. We observed similar findings when preeclampsia cases were separated into severe and mild subtypes. Association remained after adjusting for confounders.

Our results are consistent with four meta-analyses of observational studies [17, 18, 234, 235] despite individual observational studies having mixed results. All four meta-analyses revealed women with 25(OH)D < 50 nmol/L in pregnancy experienced an increased risk of preeclampsia compared to women with 25(OH)D ≥75nmol/L, with pooled odds ratios ranging from 1.6 to 2.3 [17, 18, 234, 235]. Our observation of a threshold at 50 nmol/L 25(OH)D supports these findings. However, the numbers of women with 25(OH)D <37.5 or <25 nmol/L were too small for these meta-analyses to report meaningful estimates that we could compare with ours.

Past literature also reported that low 25(OH)D was associated with increased risk of severe preeclampsia [15, 20, 24, 105, 236]. In one of the largest studies of maternal 25(OH)D and preeclampsia (n=560 mild and n=157 severe cases), investigators observed mothers with 25(OH)D ≥ 50 nmol/L at ≤ 26 weeks gestation had a 40% reduction in severe preeclampsia risk compared with those with 25(OH)D <50 nmol/L (RR 0.65 95% CI 0.43, 0.98) after adjusting for confounders [15]. Contrary to our findings, they found no association for mild preeclampsia, perhaps due to differences in case definition, age of serum samples, or in population characteristics. All other studies only examined severe forms of preeclampsia [24, 105, 236].

Our results suggest that vitamin D deficiency may impact pathophysiologic changes found in both severe and mild subtypes of preeclampsia, including an inappropriate inflammatory response, endothelial dysfunction, and high blood pressure [237]. The active form of vitamin D has a role in maintaining an appropriate inflammatory response in the maternal-fetal interface [97]. Endothelial function is maintained via vitamin D by improving proliferation, migration and tubule formation [238, 239]. Furthermore, there is evidence that vitamin D metabolites protect endothelial cells from oxidative stress and minimize the effects of exposure

to preeclampsia-related factors [240, 241]. In addition, active vitamin D influences the renin-angiotensin-aldosterone system, including the regulation of blood pressure [40].

Our reliance on correct assignment of ICD-9 codes for preeclampsia may contribute to some misclassification. Since preeclampsia might manifest before proteinuria is induced, clinicians may be categorizing potential preeclamptic cases as gestational hypertension. Therefore, experts at American Congress of Obstetricians and Gynecologists recommend gestational hypertension and preeclampsia be classified with preeclampsia with or without severe characteristics [242]. We were unable to test the relationship between 25(OH)D and gestational hypertension since our perinatal database used ICD-9 codes which lack this detailed information.

Women who elect prenatal aneuploidy screening may be different than those who do not elect screening. However, we have previously demonstrated that there are no major differences in the eligible EVITA subcohort compared with the full cohort [230]. As with any observational study, there is potential for unmeasured confounding; lack of data on socioeconomic status, physical activity, or genetics may have biased our results. Our study population may have limited generalizability to more diverse populations. In addition, our case-cohort contained few women with 25(OH)D <25 nmol/L which led to imprecise estimates. We also did not have other biomarkers of vitamin D, such as the vitamin D binding protein (VDBP), which controls the bioavailability of free 25(OH)D [22].

Major strengths of this study were the large number of preeclamptic cases and our ability phenotype preeclampsia based on severity. In addition, we used serum samples collected before onset of symptoms, which is important to establish temporality. Lastly, we were able to successfully confirm findings from meta-analyses and reviews in overall preeclampsia and

severity subtypes. These findings further develop our understanding of the role of vitamin D in preeclampsia.

If others confirm our findings of a threshold effect for preeclampsia at approximately 25(OH)D of 50 nmol/L, these results have important implications. The Institute of Medicine recommendation of achieving 25(OH)D <50 nmol/L to reduce skeletal outcomes may be enough to prevent preeclampsia. This may imply low doses of vitamin D supplementation are required to reduce preeclampsia risk.

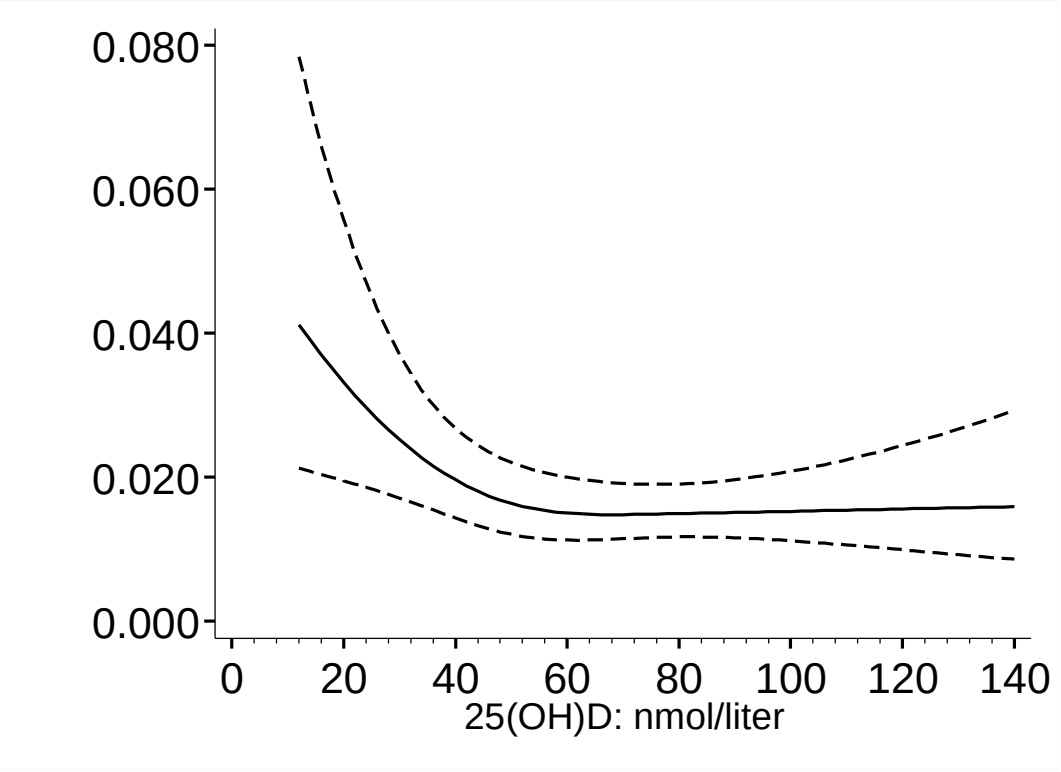

Figure 1. Risk of preeclampsia with increasing 25(OH)D concentration (nmol/L) using restricted cubic spline analysis with 3 knots.

Table 10. Characteristics of the randomly-selected subcohort and preeclampsia cases**.**

	Sub-cohort (n=2327) %	Preeclampsia Cases (n=650) %	Severe Cases(n=201) %	Mild Cases(n=449)%
Maternal race/ethnicity				
White	74	65	61	67
Black	19	30	35	27
Other race	7	5	3	6
Parity				
0	68	66	60	68
1 or more	32	34	40	32
Maternal age(years)				
<20	6	9	7	10
20-29	41	43	37	46
30-34	32	24	24	23
>35	21	24	31	21
Prepregnancy BMI (kg/m^2)				
<18.5	3	1	1	2
18.5-24.9	49	32	27	33
25-29.9	28	32	33	32
≥30	21	35	39	33
Marital status				
Unmarried	36	47	49	45
Married	64	53	51	55
Maternal education				
Less than high school	7	9	8	9
High school	23	30	26	30
Some college	23	22	26	20
College	47	40	39	41
Smoking during pregnancy				
Yes	11	9	8	9
No	89	91	92	91
Gestational age of sample				
<15 weeks	44	53	59	50
≥15 weeks	56	47	41	50
Season of blood sampling				
Winter	21	22	18	23
Spring	28	26	28	24
Summer	25	28	26	28
Fall	26	25	28	24
Type of provider				
Clinic	21	25	20	27
Private	79	75	80	72
Type of insurance				
Medicaid	37	41	44	40
Insured/self-pay	63	59	56	60

Table 11. Association between maternal serum 25(OH)D concentration and the risk of preeclampsia.

	Serum 25(OH)D(nmol/L) at ≤20 weeks gestation	Number of cases	Unadjusted Incidence[1]	Unadjusted relative risk (95%CI)	Adjusted relative risk[2] (95%CI)
Preeclampsia	Categories				
	< 25	42	0.091*	2.5 (1.7, 3.8)	2.4 (1.2, 4.7)
	25-50	138	0.058	1.5 (1.2, 2.0)	1.09 (0.68, 1.7)
	50-75	261	0.058	1.4 (1.2, 1.8)	1.2 (0.87, 1.7)
	≥75	209	0.039	Ref	Ref
	Splines[3]				
	20			2.0 (1.5, 2.7)	2.3 (1.3, 4.0)
	30			1.6 (1.3, 2.0)	1.7 (1.2, 2.5)
	40			1.3 (1.2, 1.5)	1.3 (1.0, 1.7)
	50			1.2 (1.0, 1.3)	1.1 (0.92, 1.3)
	75			Ref	Ref
	90			0.96 (0.88, 1.1)	1.0 (0.81, 1.3)
	100			0.94 (0.81, 1.1)	1.0 (0.75, 1.4)
Severe Preeclampsia	< 25	15	0.035*	3.2 (1.7, 5.9)	3.3 (1.4, 7.8)
	25-50	49	0.021	1.9 (1.3, 2.9)	1.3 (0.69, 2.4)
	50-75	78	0.017	1.5 (1.1, 2.2)	1.3 (0.78, 2.1)
	≥75	59	0.011	Ref	Ref
Mild Preeclampsia	< 25	27	0.061*	2.2 (1.4, 3.6)	2.4 (1.1, 5.1)
	25-50	89	0.038	1.4 (1.0, 1.8)	1.0 (0.61, 1.8)
	50-75	183	0.039	1.4 (1.1, 1.8)	1.2 (0.81, 1.8)
	≥75	150	0.028	Ref	Ref

[1] Based on weighted samples

[2] Adjusted for year of delivery, laboratory batch number, race/ethnicity, season, smoking status, prepregnancy BMI, insurance, parity and gestational age at blood draw.

[3]Restricted cubic spline model with 3 knots

* P-value < 0.05

5. VARIATION OF *CYP27B1*, *GC*, AND *VDR* GENES AND PREECLAMPSIA RISK USING TWO MULTI-ETHNIC PREGNANCY COHORTS

5.0 ABSTRACT

Objective: The objective of this candidate gene association study was to evaluate the relationship between maternal genetic variation of 3 vitamin D metabolism genes and preeclampsia risk.

Methods: Using two pregnancy cohorts, we studied 471 Black and 546 White preeclampsia cases and 1160 Black and 2059 White controls. For every sample, we genotyped 39 *CYP27B1* (Chromosome (chr.)12), 126 *GC* (chr. 4), and 206 *VDR* (chr.12) tagging SNPs. Using multivariable logistic regression models, we estimated the association between SNPs and preeclampsia risk using the allelic approach. Meta-analyses were conducted to calculate estimates of association between and within case-control studies.

Results: After adjusting for multiple comparisons, only trends of associations were observed for *GC* and *VDR* SNPs. For Black mother, those who carried the minor allele for three SNPs (rs11732451 [*GC*], rs4340112 [*VDR*], rs10459217 [*VDR*]) had 3.8-, 7.7-, and 8.4-fold increased odds of preeclampsia. In contrast, the odds of preeclampsia was 80-75% lowered for Black mothers who carried the minor allele for 3 SNPs (rs1099028 [*GC*], rs757344 [*VDR*], rs12721364 [*VDR*]). No *CYP27B1* variant was associated with preeclampsia risk. In the meta-analysis, two *VDR* intron variants (rs886441 and rs2853561) were associated with a decreased odds of

preeclampsia for all Black mothers. We did not observe any trends for White mothers in the meta-analysis.

Conclusion: Our research observed possible trends of association between several *GC* and *VDR* SNPs and preeclampsia risk. If these findings are replicated in only Black mothers, the effect of genetic variation on vitamin D and preeclampsia in observational studies may be minimized by controlling for race/ethnicity.

5.1 INTRODUCTION

Preeclampsia is a pregnancy-specific, multi-systemic condition that complicates 3-5% of all pregnancies [171, 243]. It is defined by new-onset hypertension and proteinuria, and/or any signs or symptoms of end-organ dysfunction (renal insufficiency, impaired liver function, pulmonary edema, thrombocytopenia) at 20 weeks of gestation or later [242]. In addition to contributing to 18% of maternal mortality in the United States [244], preeclampsia is one of the leading causes of maternal morbidity [245] and infant mortality and morbidity [91], and its complications are more likely to affect Black than White women [246]. Etiology of preeclampsia is thought to involve defective invasion of extravillous trophoblasts, which causes a failure in maternal spiral artery remodeling and impairment in blood flow [215]. This abnormal placentation leads to oxidative stress [216] and eventually maternal systemic inflammatory response and maternal symptoms.

Vitamin D may have an important role in the development of preeclampsia. Vitamin D regulates the transcription and function of genes associated with placental invasion and normal implantation [38, 217] and modulates the activities of antioxidant enzymes, inflammatory

response, and hypertension [40, 238]. Observational studies on vitamin D deficiency and preeclampsia risk have produced mixed results [17, 222]. Some of this heterogeneity in findings may be explained by a genetic effect or gene-environment interaction.

Vitamin D from sunlight, foods, or supplements is transported by chylomicrons and the vitamin D binding protein (DBP, encoded by *GC* gene) to the liver [29, 48, 49]. Enzyme-driven hydroxylating reactions convert vitamin D to 25(OH)D (25-hydroxyvitamin D) [29]. 25(OH)D is further hydroxylated by 1 alpha-hydroxylase (encoded by *CYP27B1*) to produce the active form of vitamin D, $1,25(OH)_2D$ (1,25-dihydroxyvitamin D) [27]. $1,25(OH)_2D$ attaches to the vitamin D receptor (*VDR*) expressed in target organs and cells [54]. The *GC, CYP27B1,* and *VDR* genes are integral to vitamin D metabolism and comprise of common variations that may increase risk of preeclampsia. Our objective was to determine the associations between allelic sequence variation in *GC*, *CYP27B1,* and *VDR* genes and preeclampsia risk.

5.2 METHODS

This candidate gene association study used existing data and stored blood samples from two pregnancy populations have been described in detail previously, the Collaborative Perinatal Project (CPP) and EVITA [227, 230]. CPP enrolled more than 55,000 women at 12 U.S medical centers from 1959-1965 [247]. From the subsample of 28,429 women with singleton pregnancies, with no preexisting conditions (pregestational diabetes, hypertension, or cardiovascular disease) and a banked serum sample at <26 weeks, we selected all preeclampsia cases (n=173 White, n=283 Black) and 2986 randomly-selected non-preeclamptic controls among Black or White women.

EVITA studied 12,861 women with singleton pregnancies who had an available banked serum from aneuploidy screening at ≤20 weeks and subsequent delivery at Magee-Womens Hospital of UPMC in Pittsburgh, Pennsylvania. From this sample, we selected all preeclampsia cases (n=313 White, n=173 Black) and 2327 randomly-selected non-preeclamptic controls also among Black and White women.

Preeclampsia was defined in the CPP as gestational hypertension and proteinuria for the first time after 20 weeks, and return of abnormalities to normal in the postpartum period. Gestational hypertension was classified based on two or more measurements of systolic blood pressure ≥ 140 mm Hg and/or diastolic blood pressure ≥ 90 mm Hg. Proteinuria was determined by two random urine dipsticks of 1+ protein or one dipstick of 2+ protein which were tested for albumin every 8 weeks. At each visit, medical and obstetric events were recorded and random urine samples were tested for albumin. A validation study showed a high degree of accuracy between blood pressure and urinary albumin and original medical records [248]. In EVITA, preeclampsia was defined by ICD-9 codes of 642.4-642.6).

We selected three candidate pathway genes (*GC*, *CYP27B1*, and *VDR*) from vitamin D literature [31, 119, 122, 124, 126-135]. For each gene, functional candidate SNPs were selected and then augmented with additional tagging SNPs using HapMap phase 3 data [178]. Tagging SNPs were selected for the CEU (Utah residents with ancestry from northern and western Europe) population first after specifying minor allele frequency >10% and 50 kilobases up and down stream of the gene. This SNP selection was then supplemented with additional tagging SNPs from the ASW (Americans of African Ancestry in Southwest USA) population in order to represent the Black and White populations in the two case-control studies. In total, 499 SNPs were selected: 39 for *CYP27B1*, 126 for *GC*, 206 for *VDR*, and 128 markers for genetic ancestry.

Serum for CPP and EVITA were frozen for 50 and 12 years, respectively. Serum was thawed and cleaned using Qiagen kits (QIAGEN, Valencia, CA) to reduce contaminants and inhibitors. We used REPLi-g midi kit (QIAGEN, Valencia, CA) to amplify the resulting product. QuantStudio 12K Flex platform (Life Technologies, Carlsbad, CA) was used to genotype after amplification for the selected SNPs. All SNPs were genotyped for every sample. Genotypes were analyzed using the TaqMan Genotyper software (Version 1.3.1, Grand Island, NY) and visual assessment of the data was used for confirmation. All plates included controls and 10% of samples were genotyped in duplicate as internal controls.

Potential confounding variables came from the perinatal database in EVITA and in-person interviews for CPP. Gestational age was based on best obstetric estimate comparing menstrual dating and ultrasound estimates was used in EVITA and on the mother's report of the first day of her last menstrual period in CPP (ultrasound was not available in the 1960s). Both studies included data on pre-pregnancy BMI (<18.5, 18.5-24.9, 25-29.9, ≥30), preexisting diabetes (yes, no), maternal education (<12 years, 12 years, and >12 years), marital status (single, married), maternal age (<20, 20-29, ≥30), parity, and season of blood draw (winter (December–February), spring (March–May), summer (June–August), or fall (September–November)). Data on provider type (hospital outpatient resident clinic, hospital-affiliated private practice) were available for women in EVITA. A composite socioeconomic status score was available for CPP, which combines education, occupation, and family income data [247].

Statistical methods

Multiple data-quality control steps were performed for all genotype data using PLINK software (Version 1.07, Boston, MA) [190, 191] (Figure 1). Samples and markers with sample genotyping call rate or marker genotyping call rate <50% were omitted. All SNPs included in the

analysis had call rate >50%, MAF >0.05 and were in Hardy–Weinberg equilibrium (P>0.00001). Beagle software (Version 4.0, Seattle, WA) was used to impute missing genotypes for SNPs and samples that passed the quality control measures[249]. Missing calls with >0.85 r^2 were replaced by imputed calls for further analysis (9.7% of CPP Black mothers, 13% of CPP White mothers, 8.4% EVITA White mothers, and 11% EVITA Black mothers).

Population stratification to categorize mothers as Black or White was based on genetic ancestry and self-reported race [193]. First, individual ancestral proportions were calculated and implemented in STRUCTURE 2.3 (Stanford, CA) assuming K = 1 to 4 [194-197]. The number of subgroups (K=2) was determined by using an ad hoc statistic based on the rate change in the log probability of data between clusters (K) [198]. Using this approach, each individual was assigned into one of two ancestral groups—White or Black—the genetic ancestral probability of belonging to CEU or ASW was >85%, respectively. Race was based on self-report when the probability was $\leq$85%.

Separate analyses were conducted by case-control study (CPP or EVITA) and race (Black or White). Allelic frequencies (carriers of minor allele versus major allele) by cases status were assessed for each SNP using a chi-square test. SNPs were further examined in univariate logistic regression models to estimate preeclampsia odds ratios (OR) and 95% confidence intervals (95%CI) using the allelic approach (minor allele vs. major allele). We first modeled functional SNPs in parsimonious models by removing potential confounders from the model if their exclusion did not change the main exposure point estimate by ≥10%. By using this method, we adjusted for batch number (in EVITA only), year drawn (in EVITA only), site (CPP), season of blood draw, and mother's age. Other variables did not change the estimate. Season of blood draw and mother's age met our definition of confounding in all models. In EVITA models, we also

included batch number and year, and in CPP models, we additionally included study site. For consistency, we included the same covariates for the tagging SNP models.

We used random-effects meta-analysis to summarize race-specific beta coefficients across the two studies by weighting individual coefficients by the inverse of their variance. The Higgins test (I^2) was used to calculate heterogeneity as a measure of inconsistencies in the results of the studies.

5.3 RESULTS

After quality control steps, the analysis included 992 preeclampsia cases (506 CPP, 486 EVITA), and 3,161 controls (1619 CPP, 1542 EVITA). The basic characteristics of CPP and EVITA studies are presented in Tables 12 and 13. The mean age of CPP was 24 (+/- 5.6) years and of EVITA was 29 (+/- 6.3) years. Compared to Black controls in CPP, Black cases completed less years of education, smoked less, and had higher prepregnancy BMI. The same differences were observed for White controls and cases. Compared to Black controls in EVITA, Black cases were more likely to be nulliparous and non-smokers, and more likely to have completed less years of education and have higher prepregnancy BMI. There were similar differences for White mothers in EVITA, except there was no significant difference in smoking status.

For both studies and maternal race, geometric means of 25(OH)D were lower for cases compared with controls (Table 14). Additionally, despite case-status and study, Black mothers had lower geometric mean of 25(OH)D compared with White mothers.

Black mothers from EVITA had the most gene coverage with 44, 39, and 36% of each *CYP27B1*, *GC* SNPs, and *VDR* SNPs successfully passing all quality control steps (Figure 1).

For White mothers in EVITA, 10, 8, and 6% of each *CYP27B1*, *GC*, and *VDR* SNPs passed the same quality measures. Black mothers in CPP also had lower genotype coverage, with 5 % of *CYP27B1*, 5% of *GC*, and 4% *VDR* SNPs passing quality control, while 3% of *CYP27B1*, 12% of *GC* and 9% of *VDR* SNPs passed for White mothers.

Using only SNPs that passed quality control steps, preeclampsia risk differed by allelic frequency for 1 *CYP27B1* SNP, 3 *GC* SNPs and 3 *VDR* SNPs (Table 15). In Black mothers in EVITA, the minor alleles of rs238522 (CYP27B1) and rs16847105 (GC) were 30% and 15% more commonly carried by controls than cases, respectively. Also in Black mothers from EVITA, the GC SNP rs843010 and VDR SNP rs7975232 minor allele were 29-47% more likely carried by cases compared with controls, while the minor alleles of two VDR SNPs (rs10459217 and rs1544410) were 21-26% more likely carried by controls than cases. Lastly, for White mothers in CPP, the minor allele of GC SNP rs1526692 was 18% more likely carried by cases compared with controls.

After adjusting for confounders, we observed associations between 11 SNPs (5 *GC* SNPs and 6 *VDR* SNPs) and preeclampsia risk (Table 16). Further adjusting for multiple comparisons and linkage disequilibrium, no SNP remained statistically significant; however 2 *GC* SNPs (rs11732451 and rs1099028) and 4 *VDR* SNPs (rs757344, rs12721364, rs4340112, and rs10459217) had trends of associations with preeclampsia risk ($p<0.01$) for only Black mothers in EVITA. Compared with major allele carriers of rs11732451 for Black mothers in EVITA, those who carried the minor allele had 3.6-fold increased odds of preeclampsia. In contrast, the odds of preeclampsia was 75% lower for minor allele carriers of rs1099028 (intron variant) compared with major allele carriers. Adjusted odds ratios and 95% confidence intervals for minor allele carriers of 4 *VDR* SNPs (rs757344 (intron), rs12721364, rs4340112, rs10459217)

were 0.24 (0.10, 0.60), 0.20 (0.06, 0.65), 7.7 (1.6, 36), and 8.4 (1.9, 37) compared with major allele carriers, respectively. No *CYP27B1* variant was associated with preeclampsia risk.

Only three functional variants successfully passed quality control steps. Minor allele carriers of missense mutations in *CYP27B1* (rs2229103) (OR 0.79 95%CI 0.45, 1.4; p=0.42), *VDR* (rs11574115) (OR 1.4 95%CI 0.28, 7.1; p=0.68), and in the *GC* gene (rs4588) (OR 0.65 95%CI 0.26, 1.6; p=0.340) were not associated with preeclampsia risk.

A meta-analysis was conducted to summarize associations by race/ethnicity between CPP and EVITA, as well as to summarize associations between Black and White mothers. In the meta-analysis, we identified two *VDR* intron variants (rs886441 and rs2853561) with trends of decreased odds of preeclampsia for Black mothers in CPP and EVITA (Figure 2). There was a reduced odds of preeclampsia for minor allele carriers of rs886441 (Meta OR 0.53 95% CI 0.73, 1.0; p=0.05) and rs2853561 (OR 0.76 95%CI 0.58, 0.99 p=0.043). There were no significant differences in the estimates by study (I^2 0% p=0.34; I^2 0% p=0.89).

5.4 DISCUSSION

In a study of two racially diverse populations, there were trends towards association for Black mothers in those who carried minor alleles of 2 *GC* and 4 *VDR* SNPs with preeclampsia risk. By using the meta-analysis approach, we identified two other *VDR* variants with trends of associations for all Black mothers. However, the analysis could not detect associations for White mothers, perhaps due to gene coverage differences. Due to the lack of statistical significance, it is not possible to make conclusions without improving our sample size and gene coverage.

Our study has the largest gene coverage compared with other gene association studies on preeclampsia risk. Studies relating allelic variation in vitamin D metabolizing genes and preeclampsia risk are sparse. In a cohort study of 164 preeclamptics, 154 gestational diabetics, and 213 pregnant women without either condition, there were no differences in three *VDR* SNPs by maternal genotype, allelic, and haplotype frequencies [43]. However, limiting the genotyping to three SNPs may have missed functional variants important to the development of poor pregnancy outcomes. In a Chinese population, minor allele carriers of rs1544410 SNP were more likely to develop preeclampsia (OR=1.41, 95% CI 1.11–1.61) [250]. Our study did not observe an association with the same SNP for Black mothers from EVITA (OR=0.64 95%CI 0.23, 1.8), perhaps due to differences in ancestral backgrounds.

Genetic variations in these two vitamin D metabolizing genes may impact maternal 25(OH)D concentrations through several mechanisms. Past studies of non-pregnant adults observed common variations in *GC* gene modulate DBP levels and affect affinity of DBP to 25(OH)D [116]. Poor GC-25(OH)D binding may reduce the 25(OH)D concentration and other vitamin D metabolites in serum [117]. Adequate 25(OH)D and $1,25(OH)_2D$ are important for mounting an appropriate inflammatory response in the maternal-fetal interface[97], maintaining endothelial function[238], and protecting endothelial cells from oxidative stress and minimizing the effects of exposure to preeclampsia-related factors[240, 241].

Genetic variation in *VDR* may impact the activity of *VDR*, increasing a mother's susceptibility to preeclampsia. Variations in the *VDR* gene may enhance or reduce expression or binding activity of *VDR* [118] and even determine its own expression[251]. Several cell culture studies have observed possible functions of *VDR* that may influence the development of preeclampsia, including maintaining an inflammatory response [97] and endothelial repair [238].

There is also evidence from cell culture that, through a *VDR*-mediated mechanism,1,25$(OH)_2D_3$ suppresses renin transcription, which is important in the regulation of blood pressure [40].

Our study observed 2 intron variants in *GC* gene and 6 tagging SNPs in *VDR* may be associated with preeclampsia risk, while there was no evidence of associations for functional SNPs. These intron SNPs and tagging SNPs may be in LD with functional variants that increase risk of preeclampsia, or the associations may have occurred by chance. However, we took measures to increase our ability of finding true associations by controlling for multiple comparisons. Additionally, minor allele carriers of these SNPs increased or decreased risk of preeclampsia in Black mothers which may be causing some of the heterogeneity observed in observational studies. However, these trends may have only occurred in Black mothers due to the lower genotyping rates for samples from White mothers.

We had a limited ability to conduct meta-analyses within racial/ethnic groups as well as between the two studies due to genotyping coverage differences between these groups. Our study had a large number of samples that failed quality control steps, therefore we could not adequately compare associations by race/ethnicity. For the same reasons, our analysis was limited to the allelic approach. DBP protein controls the bioavailability of free 25(OH)D[22], therefore DBP may be a better predictor of preeclampsia. However, we genotyped common SNPs in the *GC* gene that encodes DBP which may be serving as a proxy for DBP protein levels. Additionally, there may be unmeasured confounders between CPP and EVITA study populations that would reduce our ability to compare results; yet by using the meta-analysis approach, we were able to combine results despite this limitation. There are some unmeasured confounders to consider, including vitamin D supplement-use, diet, and sun exposure, which may have led to biased results. Instead, we controlled for season of blood draw to minimize the effect of vitamin D

exposure differences. The relationship between SNPs and preeclampsia risk may depend on 25(OH)D concentration; however, we were limited in power to test this interaction. Past literature on variants in *GC* and *VDR* genes have found several common haplotypes affect disease risk; however we did not have the statistical power to address this analysis. Our findings are not generalizable to other racial groups due to differences in allelic frequencies by race/ethnicity.

Our study had several strengths, including a large sample of racially diverse pregnant women consisting of a large number of preeclamptic cases. Our use of two pregnancy populations in the genetic analysis improved our power for detecting associations. All of our subjects were genotyped for ancestral markers, therefore we did not rely on self-reported race. Since many variants are ancestral-specific, this strategy decreased spurious associations. In addition, by using a tagging SNP approach, we reduced redundancy while improving our statistical power for detecting associations.

Our study observed several *GC* and *VDR* SNPs may be associated with preeclampsia risk. Avoiding genetic variability between individuals in observational studies may cause an underestimation or overestimation in the association between 25(OH)D and preeclampsia risk. If associations are ancestry specific, stratifying or controlling by race may decrease the effect of genetics in observational studies. However, a replication study is needed to confirm these findings in another racially diverse pregnancy population.

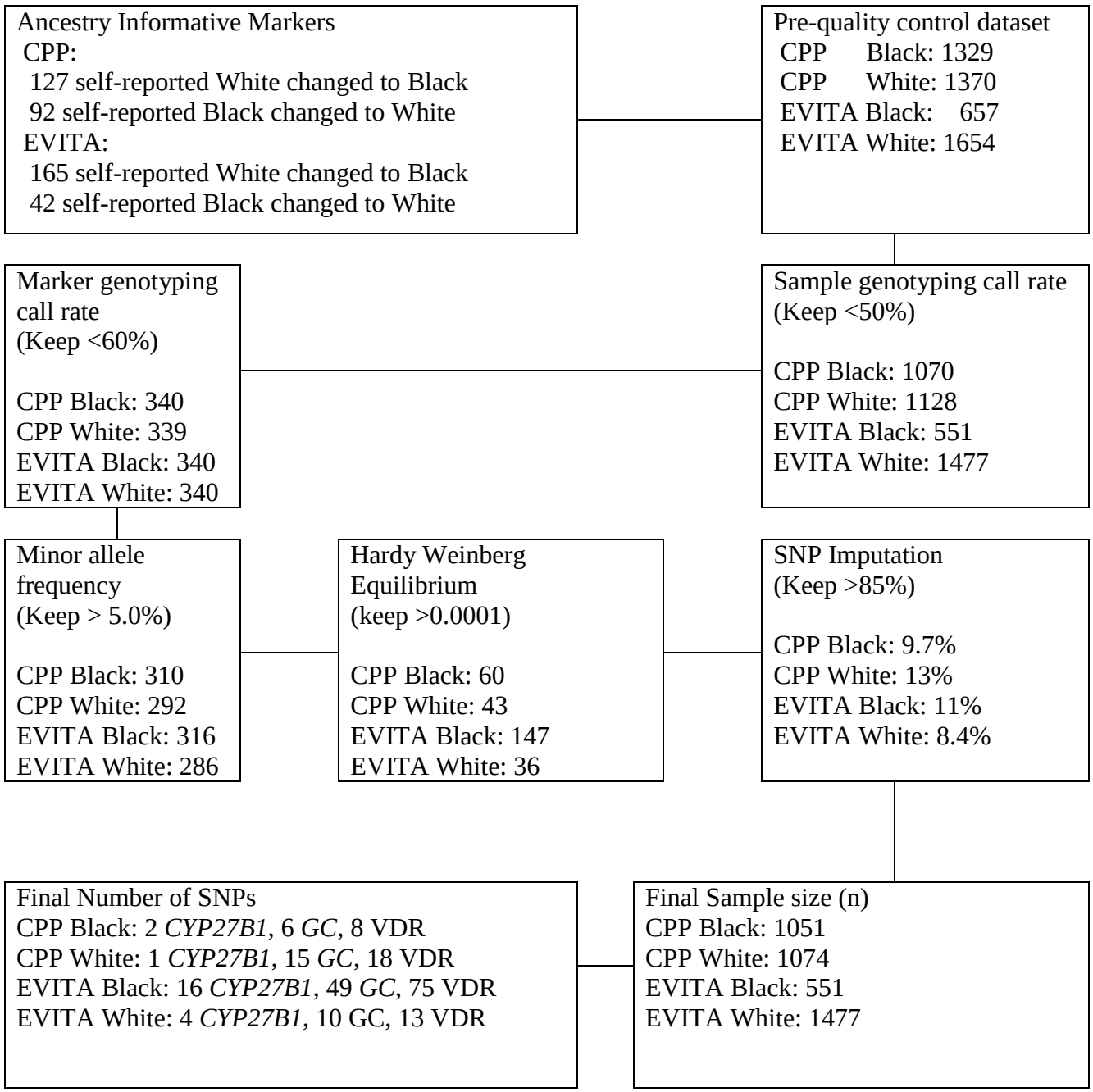

Figure 2. Quality control measures for EVITA and CPP

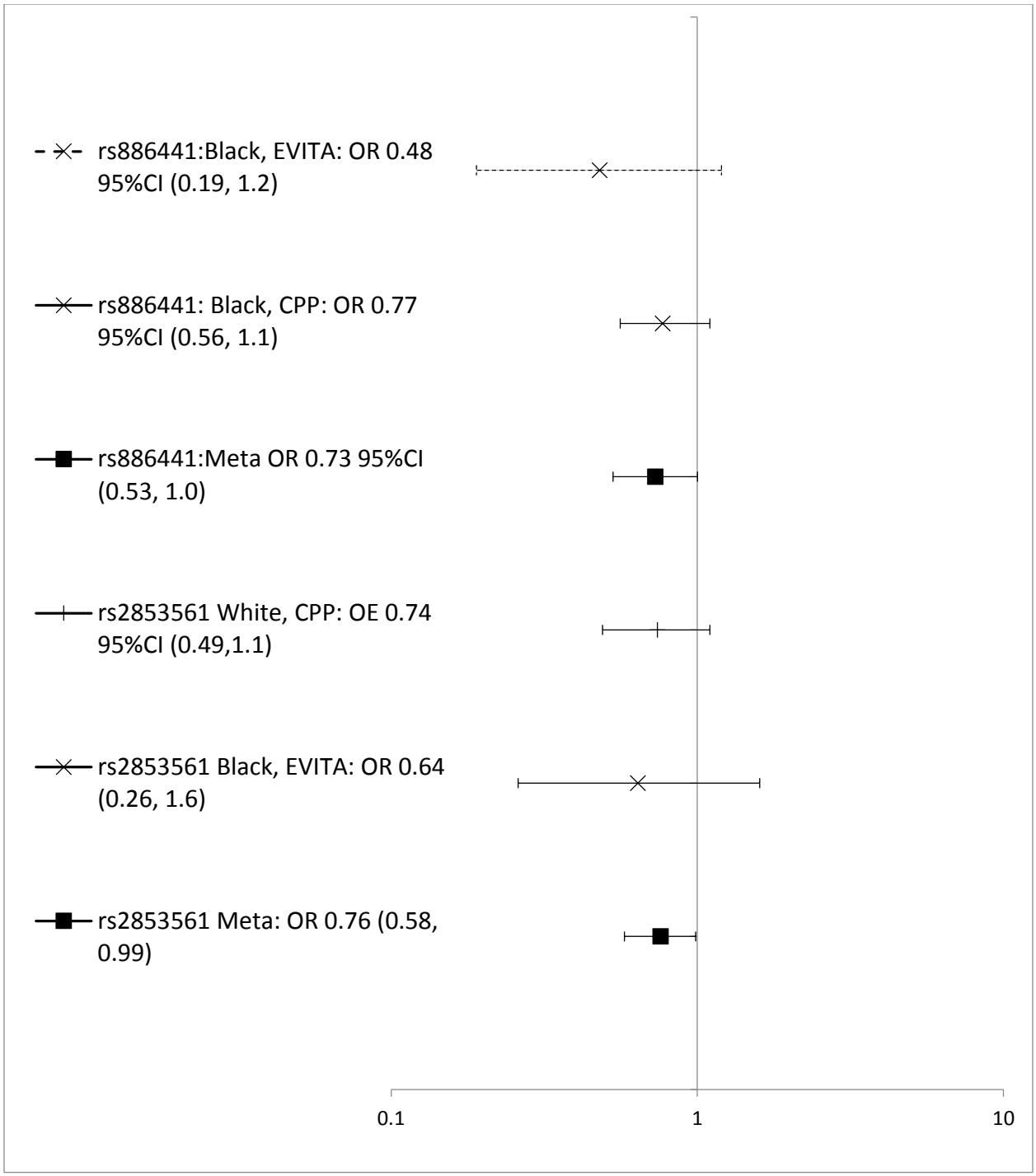

Figure 3. Meta-analysis between and with EVITA and CPP. Covariate adjusted (site (CPP), batch (EVITA), year drawn (EVITA), BMI, mother's age, and season of blood draw) OR and 95% CI.

Table 12: Population characteristics for controls and cases in CPP by race and preeclampsia status.

	CPP: Black cases n=288 Number (%)	CPP: Black controls n=782 Number (%)	CPP: White cases n=233 Number (%)	CPP: White controls n=895 Number (%)
25(OH)D				
<25 nmol/L	71 (25)*	182 (23)*	21 (9)*	62 (7)*
25-50 nmol/L	130 (45)	362 (46)	81 (35)	307 (32)
50-75 nmol/L	53 (18)	160 (20)	70 (30)	299 (32)
>=75	34 (12)	78 (10)	61 (26)	277 (29)
Season				
Winter	68 (24)	172 (22)	65 (28)	207 (23)
Spring	70 (24)	202 (26)	54 (23)	247 (27)
Summer	69 (24)	220 (28)	62 (27)	217 (24)
Fall	81 (28)	188 (24)	52 (22)	228 (25)
BMI				
<18.5	24 (9)*	67 (9)*	11 (5)*	76 (9)*
18.5-24.9	175 (63)	522 (69)	151 (69)	632 (78)
25-29.9	52 (19)	128 (17)	36 (17)	84 (10)
≥30	27 (10)	37 (5)	20 (9)	19 (2)
Socioeconomic status score				
0-20	51 (18)*	95 (12)*	9 (4)*	22 (3)*
20-39	107 (38)	271 (35)	57 (25)	139 (16)
40-59	89 (32)	276 (36)	71 (31)	262 (30)
60-79	26 (9)	101 (13)	56 (24)	245 (28)
80-100	5 (2)	25 (3)	36 (16)	208 (24)
Smoker				
Yes	96 (34)*	348 (45)*	102 (44)*	473 (53)*
No	189 (66)	432 (55)	130 (56)	415 (47)
Nulliparity				
Yes	130 (45)*	539 (69)*	104 (45)*	305 (34)*
No	158 (55)	242 (31)	129 (55)	588 (66)

***** P-value <0.05 for t-test

Table 13. Population characteristics for controls and cases in EVITA by maternal race and preeclampsia status.

	EVITA: Black cases n=173 Number (%)	EVITA: Black controls n=378 Number (%)	EVITA: White cases n=313 Number (%)	EVITA: White controls n=1164 Number (%)
25(OH)D				
<25 nmol/L	25 (14)*	34 (9)*	7 (2)*	12 (1)*
25-50 nmol/L	64 (37)	143 (38)	40 (13)	105 (9)
50-75 nmol/L	59 (34)	124 (33)	133 (42)	450 (39)
>=75 nmol/L	25 (14)	77 (20)	133 (42)	597 (52)
Season				
Winter	41 (24)	88 (23)	63 (20)	237 (20)
Spring	42 (24)	99 (26)	85 (27)	350 (30)
Summer	54 (31)	109 (29)	81 (26)	276 (24)
Fall	36 (21)	82 (22)	84 (27)	301(26)
BMI				
<18.5	1 (1)*	3 (1)*	5 (2)*	13 (1)*
18.5-24.9	28 (16)	66 (17)	52 (17)	281 (24)
25-29.9	33 (19)	63 (17)	65 (21)	174 (15)
≥30	111 (64)	246 (65)	191 (61)	696 (60)
Education				
Some high school	22 (15)*	45 (14)*	12 (5)*	48 (5)
High school	54 (36)	112 (34)	65 (25)	185 (18)
Some college	42 (28)	94 (28)	51 (19)	197 (19)
College	30 (20)	81 (24)	135 (51)	590 (58)
Smoker				
Yes	19 (11)*	51 (14)*	27 (9)*	122 (11)*
No	154 (89)	326 (86)	286 (91)	1039 (89)
Nulliparity				
Yes	112 (65)*	177 (47)*	108 (35)*	593 (51)*
No	61 (35)	201 (53)	205 (65)	570 (49)

P-value <0.05 for t-test

Table 14. Geometric means and 95% confidence intervals of 25-hydroxvitamin D by study, maternal race, and case-status.

Study	Maternal race	Geometric Mean (95%CI) Cases	Geometric Mean (95%CI) Controls
EVITA	Black	44 (43, 45) nmol/L	50 (49, 50) nmol/L
	White	69 (68,70) nmol/L	72 (72,72) nmol/L
CPP	Black	44 (43,45) nmol/L	50 (49, 50) nmol/L
	White	51 (50, 51) nmol/L	52 (52, 53) nmol/L

Table 15. Allelic frequency for SNPs in *CYP27B1*, *GC*, and *VDR* by preeclampsia case-status and maternal race in CPP and EVITA.

						allele frequency by allele (1 or 2) and case status				
						Cases		Controls		
Gene	Race	Study	SNP	Allele 1 /Allele 2	Cases/ Control (n)	1	2	1	2	p-value
CYP27B1	Black	EVITA	rs238522	C/T	155/333	0.73	0.27	0.61	0.39	<0.01*
GC	Black	CPP	rs16847105	T/A	243/662	0.72	0.28	0.67	0.33	0.04*
			rs843010	C/T	154/333	0.85	0.15	0.92	0.08	<0.01*
	White	CPP	rs1526692	C/A	196/424	0.67	0.33	0.73	0.27	0.02*
VDR	Black	EVITA	rs10459217	A/C	134/289	0.78	0.22	0.72	0.28	0.04*
			rs1544410	G/A	143/313	0.66	0.34	0.57	0.43	0.01*
			rs7975232	C/T	137/282	0.76	0.24	0.83	0.17	0.02*

*p-value<0.05

Table 16. Association between selected *CYP27B1*, *GC*, and *VDR* single nucleotide polymorphisms and the risk of preeclampsia by maternal race and study.

Gene	Race	Study	SNP		OR (95%CI)	p-value	Adj OR[1]	p-value
CYP27B1								
	Black	EVITA	rs238522	C/T	0.58 (0.39,0.85)	0.01*	0.18 (0.03, 1.1)	0.06
GC								
	Black	CPP	rs16847105	T/A	0.74 (0.55,1.0)	0.05	0.84 (0.61,1.1)	0.27
			rs842881	CT	0.70 (0.49,0.98)	0.04	0.73 (0.51,1.0)	0.08
		EVITA	rs60696209	G/C	0.66 (0.45,0.98)	0.04	1.0 (0.39,2.6)	0.97
			rs842881	AG	0.70(0.44,1.1)	0.12	0.32 (0.11,0.96)	0.04
			rs11732451	AG	1.4 (0.92,2.0)	0.13	3.8 (1.6,9.5)	0.00*
			rs1099028	AG	0.94 (0.55,1.6)	0.81	0.20 (0.06,0.61)	0.01*
			rs843010	C/T	1.7 (1.1,2.8)	0.03	0.80 (0.32,2.0)	0.64
	White	CPP	rs1526692	C/T	1.4 (1.0,1.9)	0.04	1.5 (1.1,2.1)	0.02
VDR								
	Black	EVITA	rs757344	AG	0.68 (0.45,1.0)	0.07	0.24 (0.10,0.60)	<0.01*
			rs12721364	C/T	0.86 (0.54,1.4)	0.54	0.19 (0.06,0.66)	<0.01*
			rs4340112	C/T	1.1 (0.64,1.9)	0.73	7.7 (1.6,36)	0.01*
:			rs10459217	A/C	1.2 (0.61,2.2)	0.65	8.4 (1.9,37)	<0.01*
			rs11168319	GA	0.57(0.36,0.91)	0.02	0.48 (0.13,1.8)	0.28
			rs1544410	GA	0.65 (0.43,0.98)	0.04	0.64 (0.23,1.8))	0.37
			rs7975232	C/T	1.7(1.1,2.7)	0.01	1.5 (0.52,4.3)	0.46
		CPP	rs11835083	C/G	0.69 (0.48,0.99)	0.04	0.75 (0.51,1.1)	0.13
	White	EVITA	rs2544037	AG	0.86 (0.65,1.1)	0.30	0.62 (0.39,0.99)	0.04

* P<0.01

[1] Adjusted for maternal age, site (CPP only), season of blood draw, year of blood draw (EVITA only), and batch (EVITA only)

6. VARIATION IN *CYP27B1, GC, VDR* GENES AND VITAMIN D STATUS IN TWO MULTI-ETHNIC PREGNANCY COHORTS.

6.0 ABSTRACT

Objective: The objective of this candidate gene association study was to evaluate the relationship between maternal genetic variation in 3 vitamin D metabolism genes and 25-hydroxyvitamin D concentration.

Methods: From two pregnancy cohorts, we genotyped 39 *CYP27B1* [chromosome (Chr.) 12] single nucleotide polymorphisms (SNPs), 126 *GC* (Chr.4) SNPs, and 206 *VDR* (Chr.12) SNPs. These cohorts consisted of 768 Black and 851 White pregnant mothers in the Collaborative Perinatal Project and 378 Black and 1164 White pregnant mothers in the EVITA cohort. Maternal serum collected at ≤26 weeks gestation was measured for 25(OH)D. We used multivariable linear regression to estimate the association between genotype and log-transformed 25(OH)D after adjusting for confounders. Meta-analyses were conducted to calculate estimates of association for minor allele genotype carriers compared with major allele carriers between and within cohorts.

Results: Approximately 28 and 54% of each cohort had 25(OH)D <50 nmol/L. For Black mothers in EVITA only, compared with major allele genotypes, minor allele genotypes of rs1844885 (*GC*) and rs11168275 (*VDR*) had 32-51% increases in 25(OH)D and those of

rs11732451 (*GC*) had a 23% decrease in 25(OH)D on average after controlling for multiple comparisons and adjusting for multiple comparisons. For all Black mothers, there was a significant increase in 25(OH)D for rs1844885 and a trend of decreased 25(OH)D for *CYP27B1* rs10877016 SNP. Within EVITA, the meta-analysis observed trends of decreased 25(OH)D (*GC*: rs962227) and increased 25(OH)D (*VDR*: rs2238135).

Conclusions: Our findings support a role for common genetic variants in the regulation of maternal serum 25(OH)D concentration. A replication study using a multi-ethnic pregnancy population is needed to confirm these findings.

6.1 INTRODUCTION

One in four pregnant women in the U.S. has suboptimal vitamin D status [73-75], as defined by serum 25-hydroxyvitamin D (25(OH)D) concentrations <50 nmol/L [252]. Non-Hispanic Black women are 3-5 times as likely as non-Hispanic White women to be deficient in vitamin D [70, 253]. The widespread deficiency and racial/ethnic gap in its prevalence is concerning because of the association between low 25(OH)D and the increase risk of offspring rickets [254], preterm birth [255], preeclampsia [18], small-for-gestational age birth [234], childhood asthma [256], and type 1 diabetes mellitus [257]. Understanding factors that influence vitamin D status and whether they differ by maternal race/ethnicity is of major public health importance.

Single nucleotide polymorphisms (SNPs) in several key genes in the vitamin D metabolic pathway may partially determine 25(OH)D status [124]. Vitamin D binding protein [DBP, encoded by the vitamin D-binding (*GC*) gene] is critical for transporting vitamin D in serum to

the liver [29, 48, 49], where it is hydroxylated to 25(OH)D [29]. Allelic variants in the *GC* gene modulate DBP levels and affinity for 25(OH)D [116]. After 25(OH)D is returned to circulation, it is further hydroxylated by 1 alpha-hydroxylase (encoded by *CYP27B1* also known as cytochrome p450 27B1) to the active form of vitamin D, $1,25(OH)_2D$ (1,25-dihydroxyvitamin D), in the kidney and other cells and organs in the body [27]. $1,25(OH)_2D$ attaches to the vitamin D receptor (*VDR*) expressed in target organs and cells to perform genomic roles [54]. SNPs in *CYP27B1* and *VDR* influence serum 25(OH)D by changing the rate at which 25(OH)D is hydroxylated [143] either directly or through *VDR* triggering a negative feedback loop.

While there have been many studies exploring the contribution of genetic variation in vitamin D status among non-pregnant adults separately by race/ethnicity [31, 124, 258], there is a paucity of data among pregnant women. Pregnancy is accompanied by important alterations in vitamin D metabolism, including increasing circulation of DBP[259], enhanced hydroxylation of 25(OH)D to $1,25(OH)_2D$ in kidneys and other cells and organs in the human body[260], and transportation of 25(OH)D into the placenta [58]. Our objective was to evaluate the relationship between common genotypes in *GC*, *CYP27B1*, and *VDR* genes and maternal 25(OH)D concentrations in two large pregnancy cohorts of Black and White mothers.

6.2 METHODS

We used data and samples from two pregnancy cohort studies on vitamin D and adverse birth outcomes that were approved by the University of Pittsburgh Institutional Review Board and have been described in detail elsewhere [227, 230]. The Collaborative Perinatal Project (CPP) enrolled >55,000 women at 12 U.S medical centers in 1959-65. The cohort of eligible

women for the parent study included women with singleton pregnancies, no preexisting conditions, and a banked serum sample at ≤ 26 weeks (n= 28,429). We then randomly selected a subcohort of 2986 eligible women for 25(OH)D assay. From this group, we selected 1370 White and 1329 Black women with non-preeclamptic term deliveries for the present analysis.

EVITA studied an eligible cohort of 12,861 singleton pregnancies receiving aneuploidy screening at ≤20 weeks gestation and subsequent delivery at Magee-Womens Hospital of UPMC in Pittsburgh, Pennsylvania. A subcohort of 2327 eligible women were selected for 25(OH)D assessment, and from this group we chose all term pregnancies without preeclampsia of non-Hispanic White (n=1654) and non-Hispanic Black (n=657) women. In both cohorts, racial/ethnic groups other than White and Black were excluded from genotyping because of small samples.

Maternal serum samples in CPP were collected at ≤26 weeks of gestation and stored for 40 years at −20°C with no recorded thaws. EVITA samples were collected at ≤20 weeks of gestation and were stored at -80°C for 2 to 12 years. Sera from both cohorts were sent to the same Vitamin D External Quality Assessment Scheme certified laboratory. Samples were assayed for total 25-hydroxyvitamin D (25(OH)D) [$25(OH)D_2$ + $25(OH)D_3$] using liquid-chromatography-tandem mass spectrometry based on National Institute of Standards and Technology (NIST) standards[172]. Intra- and the inter-assay variations were ≤9.6% and ≤10.9%, respectively. We defined vitamin D deficiency as 25(OH)D < 50 nmol/L[252].

We selected functional candidate SNPs from three candidate pathway genes (*CYP27B1*, *GC*, *VDR*) using vitamin D literature and 2 genome-wide association studies [31, 119, 122, 124, 126-135]. In addition, we included tagging SNPs using HapMap phase 3 data [178].Tagging SNPs were first selected from Utah residents with ancestry from northern and western Europe (CEU) population and then supplemented with additional tagging SNPs from Americans of

African Ancestry in Southwest USA (ASW). We specified SNPs with a minor allele frequency >10% and 50 kilobases up- and downstream of each gene. In total, 499 SNPs were selected: 39 for *CYP27B1*, 126 for *GC*, 206 for *VDR*, and 128 markers for genetic ancestry.

Before amplifying and genotyping the DNA, serum was thawed and cleaned using Qiagen kits (QIAGEN, Valencia, CA) to reduce impurities. The REPLi-g midi kit (QIAGEN, Valencia, CA) was then used for whole genome amplification. After amplification, all SNPs were genotyped for every sample using QuantStudio 12K Flex platform (Life Technologies, Carlsbad, CA). The quality of the genotype calls were analyzed using the TaqMan Genotyper software (Version 1.3.1, Grand Island, NY) and was confirmed by visual assessment of the data.

Gestational age in EVITA was based on best obstetric estimate, while in CPP it was based on the mother's report of the first day of her last menstrual period. Maternal self-report and medical records were used to ascertain pre-pregnancy body-mass index (BMI) (<18.5, 18.5-24.9, 25-29.9, ≥30), preexisting diabetes status (yes, no), marital status (single, married), maternal age (<20, 20-29, ≥30), parity, and season of blood draw [winter (December–February), spring (March–May), summer (June–August), or fall (September–November)]. Data on provider type (hospital outpatient resident clinic, hospital-affiliated private practice), maternal education (<12 years, 12 years, and >12 years), and insurance (private and Medicaid) were available for the EVITA cohort. A composite socioeconomic status (SES) score, which combines education, occupation, and family income data [247], was available for CPP.

Statistical methods

Genetic ancestry was calculated and implemented using STRUCTURE 2.3 (Stanford, CA) assuming K = 1 to 4 [194-197]. The number of subgroups (K=2) was determined by using an ad hoc statistic based on the rate change in the log probability of data between clusters (K)

[198]. Using this genetic ancestry approach, each study sample was classified as coming from a White or Black mother. If the probability of belonging to CEU or ASW was <85%, samples were assigned to self-reported race.

Samples with a genotyping call rate <50% were omitted from further quality control steps. All SNPs included in the analysis had call rates >50%, minor allele frequency (MAF) >0.05 and were in Hardy–Weinberg equilibrium (P>0.0001). These data-quality control steps were performed using PLINK software (Version 1.07, Boston, MA) [190, 191] (Figure 1). Beagle software (Version 4.0, Seattle, WA) [249] was used on the remaining data after these quality control steps. The software was used for conducting multiple imputations on missing genotype data to create probabilities of genotypes (AA, AB, and BB). Data with an imputation probability of >0.85 replaced the missing data, while the remaining data (≤ 0.85) remained missing.

Using Stata software (version 12.1, College Station, Texas), we also used multiple imputation to address missing covariate data for EVITA [height (n=645 missing, because the perinatal database for EVITA did not collect height data until 2003), prepregnancy weight (n=10), education (n=265), diabetes (n=95), and smoking (n=4)] and CPP [BMI (n=137), SES (n=47), diabetes (n=7), smoking (n=13), and parity (n=3)]. The data were imputed separately by study and race to create 10 imputed datasets each that assumed a multivariable normal distribution with a Markov Chain Monte Carlo approach [183, 184]. A previous publication detailing the multiple imputation methods has been described elsewhere [230].

Analyses were stratified by cohort (CPP and EVITA) and race/ethnicity (Black or White). For each cohort and racial group, we calculated the geometric mean of 25(OH)D (nmol/L) by genotype and tested for differences using nonparametric trend tests. SNPs

associated with 25(OH)D concentration were further examined in multivariable linear regression models. 25(OH)D was log-transformed to reduce skewness of the data. These models were used to estimate associations between 25(OH)D and genotype. Genotype of the major allele was set as the reference genotype (AA, AB, BB) for each SNP. We first modeled functional SNPs in parsimonious models by removing potential confounders from the model if their exclusion did not change the main exposure point estimate by ≥10%. We adjusted for batch number (in EVITA only), year drawn (in EVITA only), site (CPP), season of blood draw, and mother's age. Other variables did not change the estimate >10% (education, insurance, smoking status, diabetes status, and parity). For comparability, we used the same model for tagging SNPs. All associations were adjusted for multiple comparisons and linkage disequilibrium (LD). LD of two SNPs characterizes dependent heritability which was measured using PLINK software.

We used random-effects meta-analysis to summarize race-specific beta coefficients across the two studies by weighting individual coefficients by the inverse of their variance [211]. The Higginś test (I^2) was used to calculate heterogeneity to measure the degree of inconsistency in the results of the studies[212].

6.3 RESULTS

Based on ancestral informative markers, we reclassified the racial group of 8.1% and 9.0% of women in CPP and EVITA, respectively (Figure 1). Of the 4226 samples that passed QC steps, up to 13% of the genotyped data was imputed. After imputation, 1065 preeclampsia and preterm birth cases were excluded, resulting in a total of 1619 women in CPP and 1542 women in EVITA.

Compared with Black mothers in EVITA, greater proportions of Black mothers in CPP completed more years of education, smoked, had a lower prepregnancy BMI, and were nulliparous (Table 17). Similarly, compared with White mothers in EVITA, White mothers in CPP had a lower prepregnancy BMI, and a higher proportion of smokers and nulliparity. Unlike Black mothers, White mothers in CPP completed less years of education compared with White mothers in EVITA. In addition, the prevalence of women with 25(OH)D <50 nmol/L was higher in Black mothers of CPP than EVITA (78% vs 47%) and higher in White mothers of CPP compared with White mothers in EVITA (43% vs 10%). The geometric means and 95% confidence intervals of 25(OH)D in Black mothers in CPP and EVITA were 44 nmol/L (43, 45 nmol/L**)** and 66 nmol/L (64, 67 nmol/L), respectively. These means and 95%CI of 25(OH)D in White mothers in CPP and EVITA were 52 nmol/L (50, 53 nmol/L) and 71 nmol/L (70, 76 nmol/L), respectively.

The highest success genotyping rate was in Black mothers of EVITA, with 33% of *CYP27B1*, 24% of *GC*, and 21% of *VDR* SNPs successfully genotyped (Figure 1). The success rate was lower for White mothers in EVITA with 8% of *CYP27B1*, 6% of *GC*, and 5% of *VDR* SNPs genotyped. Overall, few SNPs passed quality control steps in CPP. For each *CYP27B1*, *GC*, and *VDR* SNPs, 5 % passed for Black mothers, while 3% of *CYP27B1*, 8% of *GC* and 8% of *VDR* SNPs passed for White mothers.

The test of trend for 25(OH)D concentration by genotype (AA, AB, BB) was significant ($p<0.05$) for two *CYP27B1* SNPs (rs10877016 and rs10877012), one *GC* SNP (rs11732451), and 4 *VDR* SNPs (rs11168261, rs11168275, rs7974708, and rs2238135) (Table 18). Compared with the univariate results, associations were slightly attenuated after adjusting for season of blood draw, mother's age, BMI, site (for CPP), batch (for EVITA), and year drawn (for EVITA) (Table

19). After adjusting for multiple comparisons and linkage disequilibrium, only estimates for 2 *GC* SNPs (rs1844885, rs11732451) and 1*VDR* SNP (rs11168275) remained statistically significant for only Black mothers in EVITA. Mothers who carried the minor allele genotypes of these SNPs had 32-51% increases (rs1844885, rs11168275) or a 23% decrease (rs11732451) in 25(OH)D concentration compared with carriers of major allele genotypes. Heterozygous carriers for 4 SNPs were also associated with trends of lowered (rs11732451) and increased (rs7974708, rs2238135, rs11168275) 25(OH)D concentrations. Only one functional SNP passed quality control steps: a missense mutation in CYP27B1 (rs2229103). In EVITA, White women with the minor allele genotype were not associated with a significant change in 25(OH)D (beta=-0.13 95%CI 0.33, 0.06; p=0.17).

Using a meta-analysis, we summarized associations within all Black mothers and all White mothers, as well as between cohorts. Only *GC* SNP rs1844885 SNP had a significant association with 25(OH)D in the meta-analysis after adjusting for multiple comparisons (Meta-beta 0.27 95%CI 0.14, 0.40; p<0.00001) with no significant variation between studies (I^2=0% p=0.454). The meta-analysis on *CYP27B1* SNP rs10877016 for White mothers in EVITA and CPP observed a trend of decreased 25(OH)D (Meta beta -16, 95%CI -25, -5.9 p= 0.003; I^2 =0% p=0.680). Further meta-analyses on 2 SNPs between Black and White mothers only in EVITA also had trends of associations with no significant variation between race/ethnicities (*GC* SNP rs962227, meta-beta -0.17, 95%CI -0.297, -0.043 p=0.009; I^2 0% p=0.444, *VDR* SNP rs2238135 minor allele carriers meta-beta 0.10 95%CI 0.02, 0.18 p= 0.028 I^2=0.0% p=0.632).

6.4 DISCUSSION

Using two large and racially diverse pregnancy cohorts, we observed allelic variation of genes with roles in vitamin D transport (*GC*), hydroxylation (*CYP27B1*) and gene regulation (*VDR*) were associated with 25(OH)D concentration in pregnant women who delivered to term. If these findings are confirmed in a replication study, our findings support a role for common genetic variants in the regulation of maternal circulating 25(OH)D concentration in pregnancy.

To our knowledge, ours is the most complete candidate gene analysis of vitamin D in pregnant women to date. Two previous candidate gene association studies have analyzed the relationship between maternal 25(OH)D concentration and maternal genetic variation of *CYP27B1* and *VDR* genes [42, 177], while ours is the first to have analyzed maternal SNPs in the *GC* gene. Using a cohort of 222 pregnant White women with gestational diabetes, researchers observed the minor allele genotype of rs10877012 was more common in women with 25(OH)D $\geq$ 50 nmol/L at 24 weeks of gestation compared with the major allele genotype (p=0.01) [177]. This is the only analyzed *CYP27B1* SNP in the study and they did not adjust for covariates. Our study only observed a trend of increased 25(OH)D for the same SNP (beta=0.06 95% CI 0.00, 0.13; p=0.046) for White mothers of EVITA. In a separate study of 354 White pregnant women, maternal 25(OH)D did not vary by *VDR* genotype. The study included 4 SNPs, none of which were included in our analysis. However, their study was limited by small sample size for the minor alleles (n=1 to 2) [42].

Genetic variations in these three vitamin D metabolizing genes may impact maternal 25(OH)D concentrations through several mechanisms. Variations in the *CYP27B1* gene may influence serum 25(OH)D and $1,25(OH)_2D$ concentrations by changing the rate of hydroxylation [143]. Common genetic variations in *GC* gene are associated with varying 25(OH)D levels, DBP

levels, and DBP affinity for 25(OH)D [116]. DBP protein controls the bioavailability of free 25(OH)D [22], therefore DBP levels may be a predictor of 25(OH)D concentration. We genotyped common SNPs in the *GC* gene that encode DBP, which may be serving as a proxy for DBP protein levels. Additionally, studies suggest 25(OH)D concentrations may be limited via activated *VDR* by downregulating 25(OH)D hydroxylation and the formation of 25(OH)D [261]. Our study observed the minor allele genotype for intron variants and tagging SNPs were associated with 25(OH)D concentration. These SNPs may be in LD with functional variants that, in turn, reduce 25(OH)D concentration.

There are several important limitations to consider in our study. We had a limited ability to conduct meta-analyses within racial/ethnic groups as well as between cohorts due to genotyping coverage differences between groups. Our study had a large number of samples that failed quality control steps. Therefore, we could not adequately compare associations by race/ethnicity. There may be unmeasured confounders that differ between CPP and EVITA study populations; however, by using the meta-analysis approach, we were able to combine results despite this limitation. Our study could not determine causality between genetic variation and 25(OH)D concentration. Instead, it is possible that the associations we observed may be due to shared factors. We were also limited from conducting gene-gene interaction and haplotype analyses. Lastly, due to differences in allelic frequencies by race/ethnicity, our findings might not be generalizable to other racial groups.

Our study has several strengths, including a large sample of racially diverse pregnant women. All of our subjects were genotyped for ancestral markers. Therefore, we did not rely on self-reported race. Since many variants are ancestral-specific, this strategy will decrease spurious associations in our genetic analysis. Circulating 25(OH)D concentrations were measured using

the gold standard approach at one laboratory, in as few batches and in as short a time frame as possible, thus attenuating any potential technical errors of measurement. In addition, by using tagging SNP approach, we reduced redundancy while improving our statistical power for detecting associations.

Maternal 25(OH)D concentration is largely dependent on the sun exposure and diet [60], but our research observed genetic differences may also be contributing. The rate at which vitamin D is metabolized may differ by genotype after controlling for confounders. Therefore, to improve maternal vitamin D status in pregnancy, vitamin D supplementation may need to consider variability due to genetic variation in vitamin D metabolism genes. However, these findings need to be replicated in a different pregnancy study before making these conclusions.

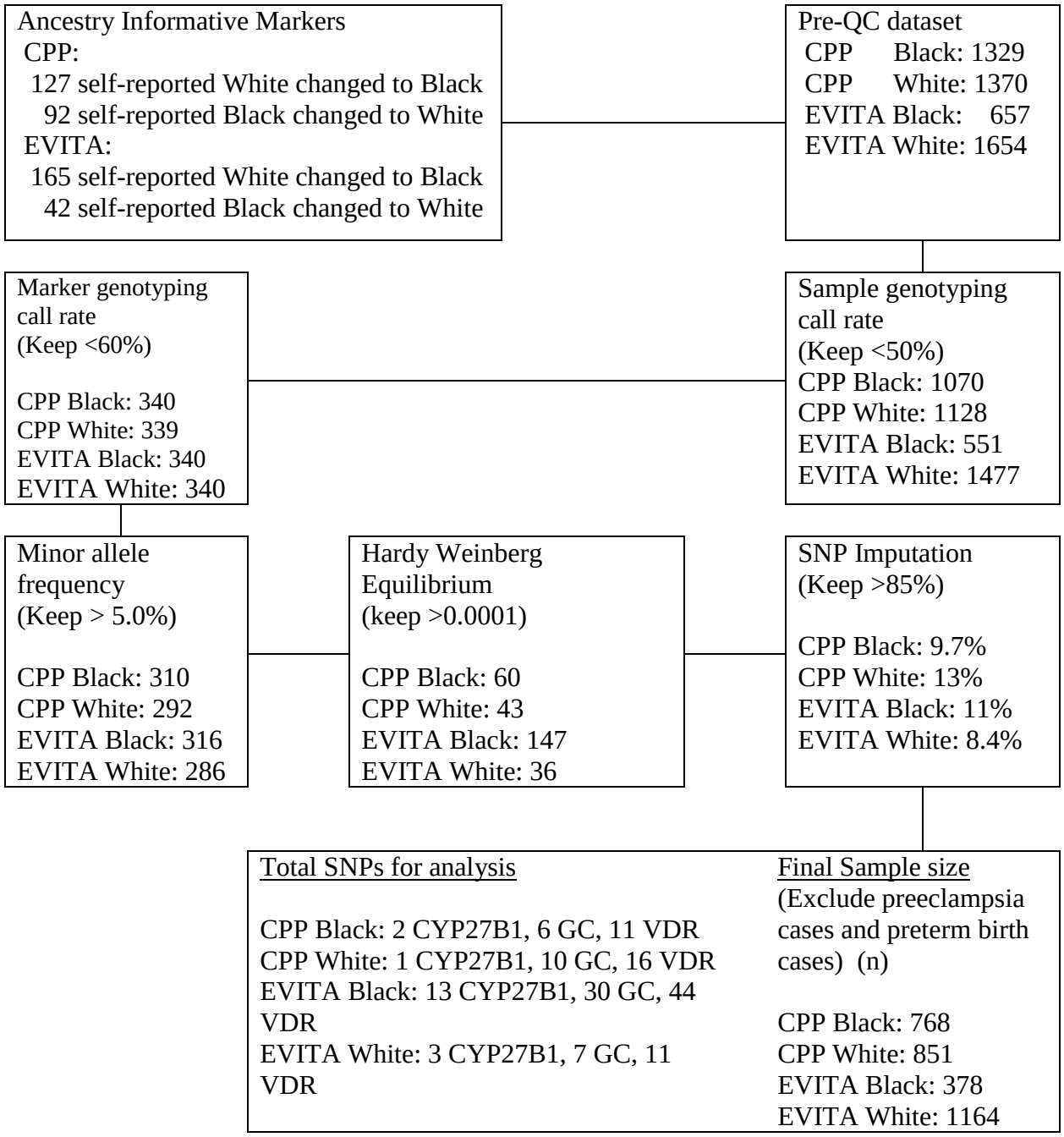

Figure 4. Quality control measures for EVITA and CPP.

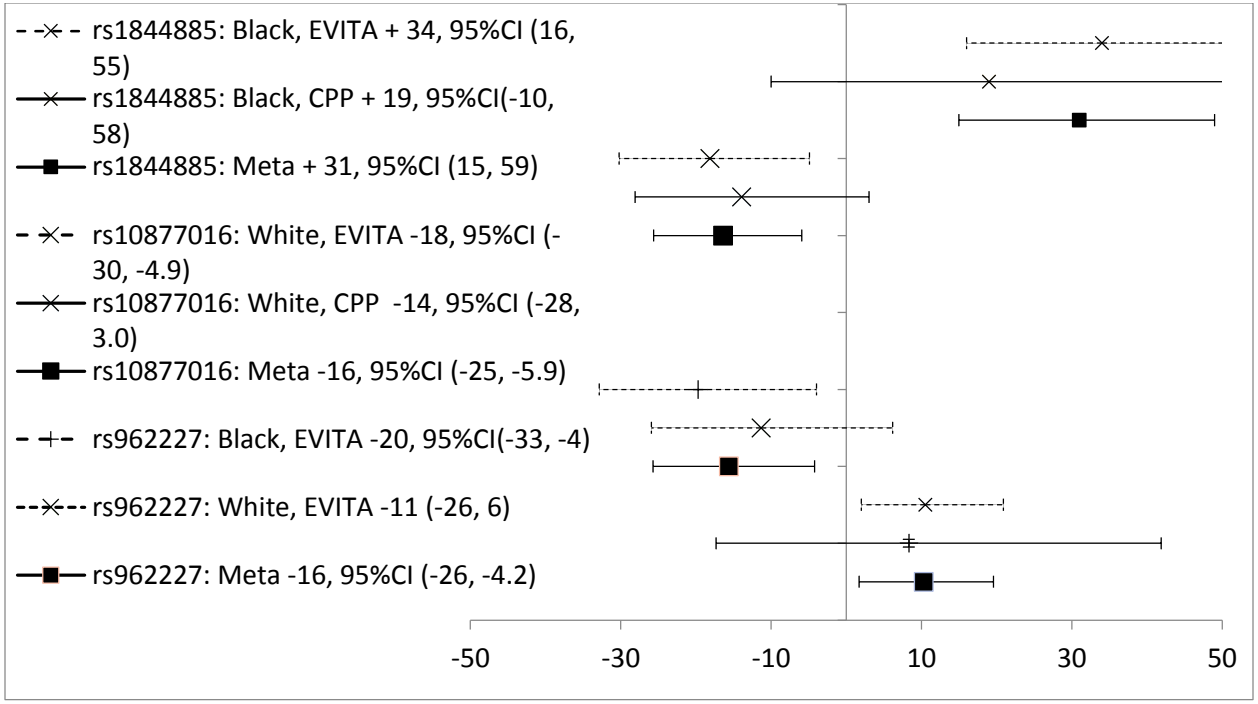

Figure 5. Meta-analysis between and within EVITA and CPP. Covariate adjusted (site (CPP), batch (EVITA), year drawn (EVITA), BMI, and season of blood draw) and 95%CI.

Table 17. Descriptive characteristics of CPP and EVITA by maternal race.

	CPP	EVITA	CPP	EVITA
	Black maternal race: n=768 Number (%)	Black maternal race: 378 Number (%)	White maternal race: 851 Number (%)	White maternal race: 1164 Number (%)
25(OH)D (nmol/L)				
<25	184 (24)*	34 (9)*	60 (7)*	12 (1)*
25-50	361 (47)	143 (38)	306 (36)	105 (9)
50-75	161 (21)	125 (33)	298 (35)	454 (39)
>=75	62 (8)	76 (20)	187 (22)	593 (51)
Season				
Winter	169 (22)	87 (23)	196 (23)	233 (20)
Spring	200 (26)	98 (26)	230 (27)	349 (30)
Summer	215 (28)	110 (29)	204 (24)	279 (24)
Fall	184 (24)	83 (22)	221 (26)	303 (26)
BMI				
<18.5	69 (9)*	15 (4)*	85 (10)*	23 (2)*
18.5-24.9	530 (69)	136 (36)	655 (77)	594 (51)
25-29.9	131 (17)	110 (29)	94 (11)	326 (28)
≥30	38 (5)	117 (31)	17 (2)	233 (20)
Education				
Less than high school	54 (7)*	60 (16)*	51 (6)*	58 (5)*
High school	61 (8)	129 (34)	298 (35)	221 (19)
Some college	361 (47)	102 (27)	306 (36)	244 (21)
College	284 (37)	87 (23)	196 (23)	641 (55)
Smoker				
Yes	422 (55)*	49 (13)*	451 (53)*	116 (10)*
No	346 (45)	329 (87)	400 (47)	1048 (90)
Nulliparity				
Yes	530 (69)*	178 (47)*	545 (64)*	570 (49)*
No	238 (31)	200 (53)	306 (36)	594 (51)

* $p<0.05$ for t-test comparing White and Black mothers between each cohort.

Table 18. Single nucleotide polymorphisms with significant differences in mean 25-hydroxy vitamin D by genotype.

SNP	Maternal race	Study	SNP Genotype	n	mean(nmol/L)	Minor allele frequency
CYP27B1	Black	EVITA	rs701008			10%
			AA	294	49 (47, 53)	
			AT	57	48 (43, 53)	
			TT	8	68 (56, 83)	
	White	EVITA	rs10877016*			12%
			GG	795	72 (71, 74)	
			GA	207	71 (68, 74)	
			AA	23	58 (49, 70)	
	White	EVITA	rs10877012*			27%
			CC	578	71 (69, 73)	
			CT	352	71 (69, 74)	
			TT	96	77 (73, 82)	
GC	Black	EVITA	rs1844885			15%
			GG	254	49 (46, 52)	
			GA	72	47 (42, 53)	
			AA	17	68 (59, 80)	
	Black	EVITA	rs11732451*			34%
			AA	156	58 (53, 62)	
			AG	125	45 (41, 49)	
			GG	52	43 (38, 48)	
	Black	EVITA	rs962227			21%
			CC	200	51 (47, 54)	
			CT	81	46 (41, 52)	
			TT	25	43 (36, 52)	
VDR	Black	EVITA	rs11168261*			30%
			CC	174	53 (49, 57)	
			CT	116	49 (45, 53)	
			TT	40	40 (34, 46)	
	Black	EVITA	rs11168275*			11%
			CC	276	46 (44, 49)	
			CT	56	59 (53, 65)	
			TT	11	79 (65, 94)	
	White	EVITA	rs7974708*			17%
			GG	716	70 (68, 72)	
			GT	252	75 (73, 78)	
			TT	46	74 (68, 80)	
	White	EVITA	rs2238135*			22%
			GG	585	70 (68, 72)	
			GT	368	75 (72, 77)	
			TT	39	79 (73, 86)	

*p-value< 0.05, test of trend

Table 19. Selected *CYP27B1*, *GC*, and *VDR* SNPs: unadjusted and adjusted coefficients and 95%CI for maternal log-25(OH)D[1].

Gene	Race/ ethnicity	Study	SNP and genotype	Coefficient (95%CI)	p-value	Adjusted coefficient[2] (95%CI)	p-value
CYP27B1	Black	EVITA	rs701008				
			AA	ref		ref	
			AT	-0.04 (-0.16, 0.08)	0.53	-0.08 (-0.20, 0.04)	0.18
			TT	0.32 (0.16, 0.49)	<0.001	0.20 (0.05,0.33)	<0.001
	White	EVITA	rs10877016				
			GG	ref		ref	
			GA	-0.01 (-0.06, 0.04)	0.58	-0.02(-0.07, 0.02)	0.37
			AA	-0.21 (-0.38, -0.05)	0.01	-0.20 (-0.34, -0.04)	0.01
GC	Black	EVITA	rs1844885				
|			GG	ref		ref	
			GA	-0.04 (-0.17, 0.08)	0.49	-0.02 (-0.13, 0.09)	0.74
			AA	0.33 (0.18, 0.48)	<0.001	0.28 (0.14,0.41)	<0.001*
	Black	EVITA	rs11732451				
			AA	ref		ref	
			AG	-0.24 (-0.36,-0.13)	<0.001	-0.18 (-0.29,-0.08)	<0.001
			GG	-0.30 (-0.44,-0.16)	<0.001	-0.26 (-0.40, -0.12)	<0.001*
VDR	Black	EVITA	rs11168261				
			CC	ref		ref	
			CT	-0.08 (-0.19, 0.04)	0.18	-0.09 (-0.20,-0.01)	0.08
			TT	-0.28 (-0.44,-0.12)	<0.001	-0.26 (-0.41,-0.1)	<0.001
	Black	EVITA	rs11168275				
			CC	ref		ref |	
			CT	0.24 (0.13, 0.36*	<0.001	0.15 (0.04,0.26)	0.01
			TT	0.52 (0.35,0.69)	<0.001	0.41 (0.24,0.58)	<0.001*
	White	EVITA	rs7974708				
			GG	ref		ref	
			GT	0.08 (0.03, 0.12)	<0.001	0.06 (0.02, 0.11)	<0.001*
			TT	0.05 (-0.03, 0.13)	0.19	0.04 (-0.04,0.12)	0.35
	White	EVITA	rs2238135				
			GG	ref		ref	
			GT	0.06 (0.02, 0.10)	0.01	0.05 (0.01, 0.10)	0.09
			TT	0.12(0.04, 0.20)	<0.001	0.12 (0.03,0.20)	0.01

*significant p-values after adjusting for multiple comparisons and linkage disequilibrium

[1] No associations in CPP were significant after adjusting for multiple imputations and linkage disequilibrium.

[2] Adjusted for maternal age, BMI, year of blood draw, and batch

7. SYNTHESIS

7.0 OVERVIEW OF RESEARCH FINDINGS

This dissertation used data from two pregnancy cohorts to study the relationship between 25(OH)D and preeclampsia risk, and the associations of each 25(OH)D concentration and preeclampsia risk with maternal genetic variation of three vitamin D metabolizing genes (*GC*, *CYP27B1*, and *VDR*). Our research observed 25(OH)D <50 nmol/L increased risk of preeclampsia, as well as mild and severe forms of preeclampsia. In the candidate gene association studies, we observed trends of associations between SNPs of *GC* and *VDR* and risk of preeclampsia. Additionally, we observed allelic variations of *GC*, *CYP27B1*, and *VDR* were associated with 25(OH)D concentration in pregnant women who delivered to term. The following summarizes our findings for this dissertation.

1. Determine the effect of vitamin D deficiency on risk of preeclampsia.

We used a randomly sampled subcohort of 2327 pregnancies and all remaining preeclampsia cases (n=650 cases) from EVITA. The association between 25(OH)D and preeclampsia was estimated by using log-binomial regression with restricted cubic splines. Adjusting for confounders, risk declined sharply from 20 to 40 nmol/L, and then plateaued at 50 nmol/L (test of non-linearity $p<0.05$). Similar dose-response associations were observed with severe and mild preeclampsia. The adjusted preeclampsia risk ratios (95% confidence intervals)

for 25(OH)D at <25 nmol/L, 25-≤50 nmol/L, and 50-≤ 75 nmol/L were 2.4 (1.2, 4.8), 1.9 (0.68, 1.7), and 1.2 (0.87, 1.7), respectively, compared with those with 25(OH)D ≥75 nmol/L. Our results are consistent with four meta-analyses of observational studies [17, 18, 234, 235], and with past literature on low 25(OH)D and severe preeclampsia[15, 20, 24, 105, 236]. Contrary to a previous study, we observed an association for mild preeclampsia[15].

2. Determine the relationship between maternal genetic variation in *GC*, *CYP27B1*, and *VDR* on risk of preeclampsia.

Using a candidate gene approach, we studied 283 Black and 223 White preeclampsia cases, as well as 768 Black and 851 White controls from CPP. From EVITA, we used 173 Black and 313 White cases, and 378 Black and 1164 White controls. We genotyped 39 *CYP27B1*, 126 *GC*, and 206 *VDR* tagging SNPs for every sample. Using multivariable logistic regression models, we estimated the association between SNPs and preeclampsia risk using the allelic approach. Meta-analyses were conducted to calculate estimates of association for minor allele carriers compared with major allele carriers between and within cohorts. In summary, only trends of association were observed in this study. Black mothers in EVITA who carried the minor alleles for three SNPs (rs11732451 [*GC*], rs4340112 [*VDR*], rs10459217 [*VDR*]) had 3.8-, 7.7-, and 8.4-fold increased odds of preeclampsia. In contrast, the odds of preeclampsia was 80-75% lowered for mothers who carried the minor alleles for 3 SNPs (rs1099028 [*GC*], rs757344 [*VDR*], rs12721364 [*VDR*]). No *CYP27B1* variant was associated with preeclampsia risk. In the meta-analysis, two *VDR* intron variants (rs886441 and rs2853561) were associated with decreased odds of preeclampsia for all Black mothers (CPP and EVITA). We did not observe any trends for White mothers in the meta-analysis.

In summary, our study observed 2 intron variants in *GC* and 6 tagging SNPs in *VDR* may be associated with preeclampsia risk. Due to the lack of statistical significance, it is not possible to make conclusions without improving our sample size and gene coverage. Significant associations between SNPs and preeclampsia risk would support evidence of a genetic influence in the relationship between vitamin D and preeclampsia risk.

3. Determine the effect of *GC*, *CYP27B1*, and *VDR* gene on vitamin D deficiency.

From CPP and EVITA, we genotyped 39 *CYP27B1,* 126 *GC,* and 206 *VDR* SNPs using the tagging SNP approach. These cohorts consisted of 768 Black and 851 White pregnant mothers in CPP and 378 Black and 1164 White pregnant mothers in the EVITA cohort. After excluding preterm and preeclampsia cases from the analysis, we used multivariable linear regression to estimate the association between genotype and log-transformed 25(OH)D after adjusting for confounders. Meta-analyses were conducted to calculate estimates of association for minor allele genotype carriers compared with major allele carriers between and within cohorts. Compared with major allele genotypes, Black women in EVITA with minor allele genotypes of rs1844885 (*GC*) and rs11168275 (*VDR*) had 32-51% increases in 25(OH)D and those of rs11732451 (*GC*) had a 23% decrease in 25(OH)D after controlling for multiple comparisons and covariates. For all Black mothers (EVITA and CPP), there was a significant increase in 25(OH)D for rs1844885 and a trend of decreased 25(OH)D for rs10877016 *CYP27B1* SNP. Within EVITA, the meta-analysis observed trends of decreased 25(OH)D (*GC*: rs962227) and increased 25(OH)D (*VDR*: rs2238135).

Overall, common SNPs in vitamin D metabolism genes were associated with maternal circulating 25(OH)D concentration in pregnancy. To our knowledge, our study is the most complete candidate gene analysis of vitamin D in pregnant women to date. Due to the limited

literature, our study could not verify previous findings on maternal genetic variations of *GC*, *VDR*, and *CYP27B1* genes and maternal 25(OH)D concentration.

7.1 STRENGTHS AND LIMITATIONS

Several limitations of this research need to be considered. We had a limited ability to conduct meta-analyses within racial/ethnic groups as well as between cohorts due to genotyping coverage differences between these groups. Our study had a large number of samples that failed quality control steps, therefore we could not adequately compare associations by race/ethnicity. Also, our analysis for the gene candidate study on preeclampsia risk was limited to the allelic approach due to small sample size and poor genotype quality. We did not have enough power to study gene-by-gene interactions. Several common haplotypes affect disease risk; however we did not have the statistical power to address this in our analyses. In addition, the relationship between SNPs and preeclampsia risk may depend on 25(OH)D status which we could not measure.

Our study could not determine causality between 25(OH)D and preeclampsia risk, as well as genetic variation and 25(OH)D concentration. It is possible that the associations we observed may be due to shared factors. DBP controls the bioavailability of free 25(OH)D [22], therefore DBP levels may be a better predictor of preeclampsia risk, as well as 25(OH)D concentration. However, for our genetic studies, we genotyped common SNPs in the *GC* gene that encodes DBP. These SNPs may be serving as a proxy for DBP protein levels. Additionally, there may be unmeasured confounders between CPP and EVITA study populations that would reduce our ability to compare results; yet by using the meta-analysis approach, we were able to combine

results despite this limitation. Our data did not include information on vitamin D supplement use, diet, or sun-exposure. Therefore, our results may be biased by unmeasured confounders. Multiple measurements of 25(OH)D could have resulted in a better estimate of longer-term exposure and reduce the degree of non-differential measurement error; however, our data was limited to one measurement. Our genetic analysis may not be generalizable to other racial groups due to differences in allelic frequencies by race/ethnicity. Lastly, for our case-cohort, few women had 25(OH)D <25 nmol/L which led to imprecise estimates.

Our study has several strengths, including a large sample of racially diverse pregnant women consisting of a large number of preeclamptic cases. Our use of two pregnancy cohorts in the genetic analysis improved our power for detecting associations when dividing the data by race/ethnicity. All of our subjects in the genetic studies were genotyped for ancestral markers, therefore we did not rely on self-reported race. Since many variants are ancestral-specific, this strategy decreased spurious associations. Circulating 25(OH)D concentrations were measured at one laboratory, in as few batches and in as short a time frame as possible, thus attenuating any potential technical errors of measurement. In addition, by using tagging SNP approach, we reduced redundancy while improving our statistical power for detecting associations. Lastly, we used serum samples collected before onset of preeclampsia symptoms, which is important in establishing temporality in our first study.

7.2 PUBLIC HEALTH SIGNIFICANCE

This dissertation makes a significant contribution to public health. Our first paper implies low doses of vitamin D supplementation may reduce preeclampsia risk. It also implies that the

dose of vitamin D (400IU) in prenatal supplements may be enough to reduce preeclampsia risk. Therefore, measures to increase prenatal vitamin D uptake may be more important than recommending additional vitamin D supplementation during pregnancy. This is significant since the concentration of 25(OH)D to reduce preeclampsia may be the same level recommended by the Institute of Medicine to reduce skeletal outcomes. If our findings are supported by randomized clinical trials, prenatal supplement uptake may be an important means of preventing preeclampsia.

Due to a dose-response between 25(OH)D and severe and mild preeclampsia, we suggest vitamin D deficiency may impact pathophysiologic changes found in both severe and mild subtypes of preeclampsia. These changes include an inappropriate inflammatory response, endothelial dysfunction, and high blood pressure [237] which may be mediated by *VDR* and vitamin D metabolites [40, 97, 238]. It is possible that common genetic variations in *VDR* impact these pathophysiologic mechanisms and increase a mother's risk of preeclampsia.

Our research also observed possible trends of association between several *GC* and *VDR* SNPs and preeclampsia risk. Despite the poor genotype coverage, one of these *GC* SNPs (rs11732451) was associated with a trend of increased risk of preeclampsia, as well as significantly decreased 25(OH)D in Black mothers. If these trends prove to be significant, this research will be supporting evidence of a genetic influence in the relationship between 25(OH)D and preeclampsia risk. Furthermore, not controlling for this genetic variability between individuals may have overestimated the association observed between 25(OH)D and preeclampsia risk in EVITA. However, if associations are ancestry-specific, controlling by race may have decreased the effect of genetics in EVITA.

Our study observed common genetic variations in three vitamin D metabolizing genes are associated with maternal 25(OH)D. 25(OH)D status in pregnancy is largely dependent on sun exposure and diet, but our research observed genetic differences may also be contributing. Due to the heritability of 25(OH)D, increasing vitamin D supplements can reduce the environmental component of the variability in 25(OH)D between individuals (supplements and season), which may expose the inherited inter-individual differences in vitamin D metabolism directed by different genotypes. In other words, the rate at which a dose of vitamin D (sunlight or supplement) is metabolized may differ by genotype after controlling for confounders, leading to different 25(OH)D concentrations by genotype. Therefore, to improve maternal 25(OH)D status in pregnancy, vitamin D supplementation may need to consider the influence of genetic variation if these findings are confirmed in a replication study.

7.3 FUTURE DIRECTIONS

Randomized control trials are needed to determine causality between vitamin D deficiency and preeclampsia risk. Previous randomized control trials of vitamin D supplements in pregnancy may have started supplementation too late in pregnancy to observe an effect on reducing risk of preeclampsia [102, 103], and small sample sizes may have reduced power to find a significant difference in preeclampsia cases by vitamin D dose. If our research findings are correct, then the control dose (400-1000 units/day) may be enough to increase 25(OH)D to 50 nmol/L and reduce risk of preeclampsia. Therefore, past randomized control trials may have been limited in power to detect a difference in preeclampsia risk by intervention group.

As with any other genetic study, replication is needed to confirm results. Since common SNPs in these genes have race-specific allelic variations, future work on variation in vitamin D metabolism genes and preeclampsia risk should continue to measure associations by race/ethnicity. Additionally, the relationship between genetic variation and preeclampsia risk may depend on 25(OH)D status. Therefore, additional analyses should test the significance of an interaction between vitamin D deficiency and risk allele. Future studies should measure DBP levels to provide some insight in the effects of genetic variation in the *GC* gene on 25(OH)D. Other genes are involved in vitamin D metabolism that should be explored in future studies. These genes are involved in converting vitamin D to 25(OH)D (*CYP2R1*), inactivating 25(OH)D and 1,25$(OH)_2$D (*CYP24A1*).

Assuming these genetic findings are causal, it will be important to find if vitamin D supplementation improves 25(OH)D status, as well as reduce their risk of preeclampsia, for those with risk variants. To achieve this, the change in 25(OH)D (before and after supplementation) by genotype should be assessed to find if rate of vitamin D metabolism differs by genotype. If not, then it might not be worthwhile to screen women for genetic factors before supplementation.

APPENDIX: TABLES AND FIGURES

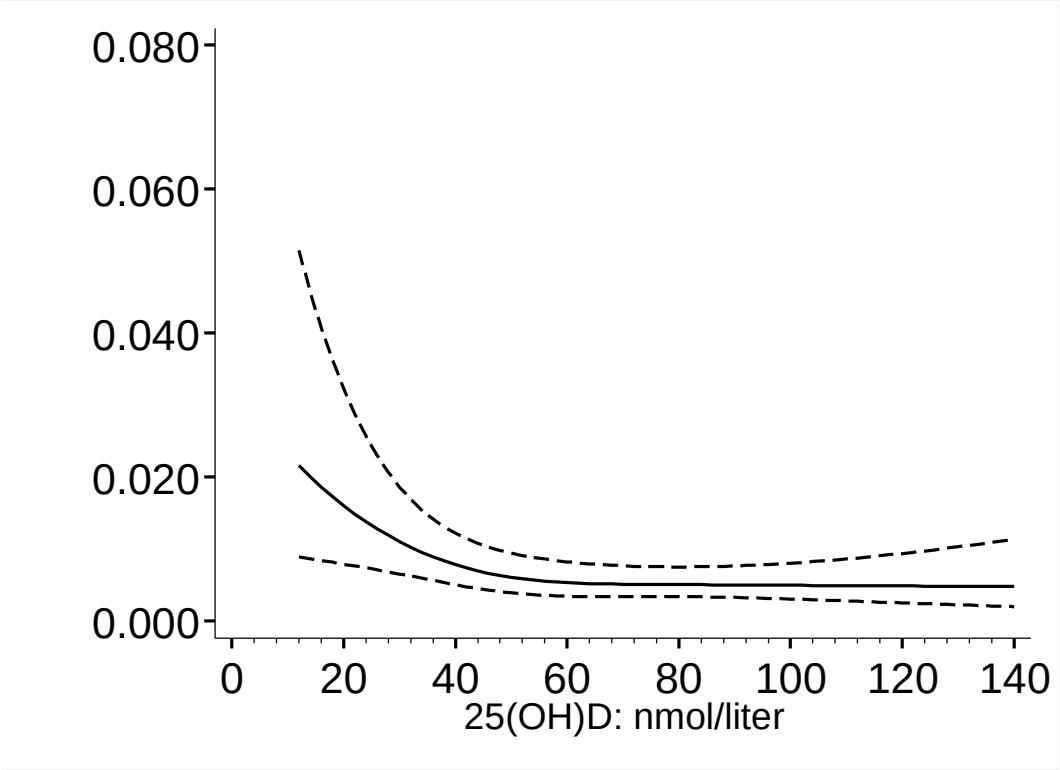

Figure 6. Risk of severe preeclampsia with increasing 25(OH)D concentration (nmol/L) using restricted cubic spline analysis with 3 knots.

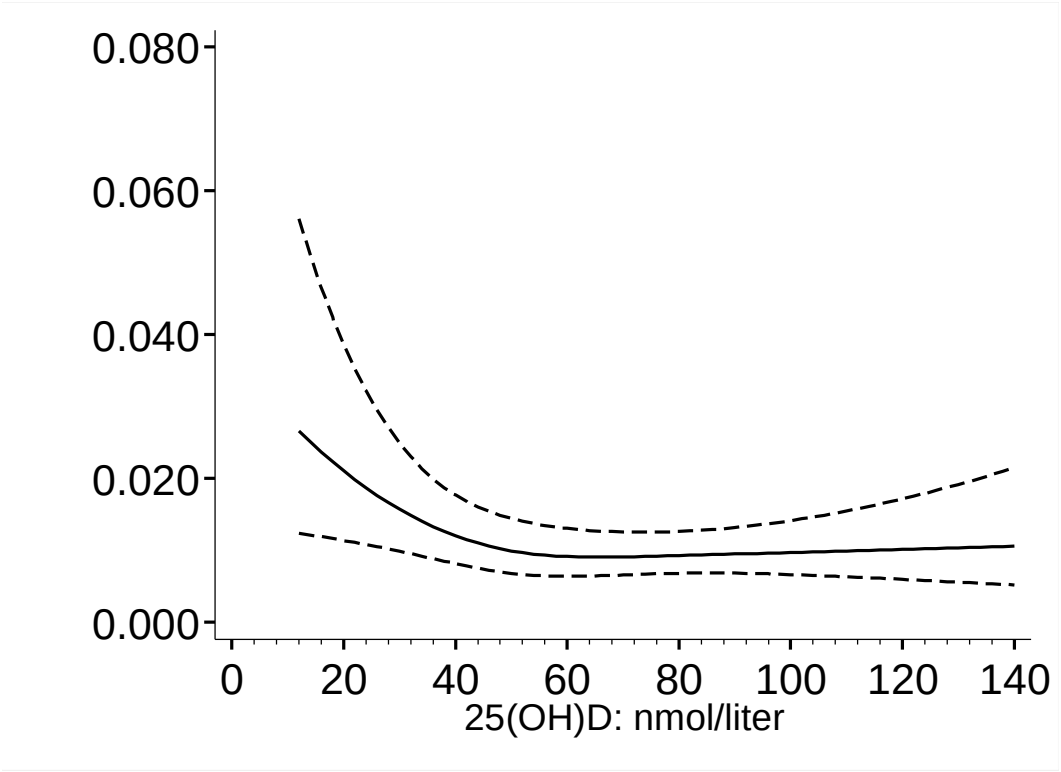

Figure 7. Risk of mild preeclampsia with increasing 25(OH)D concentration (nmol/L) using restricted cubic spine analysis with 3 knots.

BIBLIOGRAPHY

1. Holick, M.F., *Resurrection of vitamin D deficiency and rickets.* The Journal of Clinical Investigation, 2006. **116**(8): p. 2062-2072.
2. Li, Y.C., *Vitamin D regulation of the renin-angiotensin system.* J Cell Biochem, 2003. **88**(2): p. 327-31.
3. Dimeloe, S., et al., *Regulatory T cells, inflammation and the allergic response-The role of glucocorticoids and Vitamin D.* J Steroid Biochem Mol Biol, 2010. **120**(2-3): p. 86-95.
4. Holick, M.F., *Vitamin D and bone health.* J Nutr, 1996. **126**(4 Suppl): p. 1159S-64S.
5. Garland, C.F., et al., *The role of vitamin D in cancer prevention.* Am J Public Health, 2006. **96**(2): p. 252-61.
6. Kovacs, C.S., *Vitamin D in pregnancy and lactation: maternal, fetal, and neonatal outcomes from human and animal studies.* Am J Clin Nutr, 2008. **88**(2): p. 520S-528S.
7. Wang, T.J., et al., *Vitamin D deficiency and risk of cardiovascular disease.* Circulation, 2008. **117**(4): p. 503-11.
8. Pittas, A.G., et al., *The role of vitamin D and calcium in type 2 diabetes. A systematic review and meta-analysis.* J Clin Endocrinol Metab, 2007. **92**(6): p. 2017-29.
9. Chalmers, J.D., et al., *Vitamin-D deficiency is associated with chronic bacterial colonisation and disease severity in bronchiectasis.* Thorax, 2013. **68**(1): p. 39-47.
10. Lim, S., et al., *Association of vitamin D deficiency with incidence of type 2 diabetes in high-risk Asian subjects.* The American Journal of Clinical Nutrition, 2013. **97**(3): p. 524-530.
11. Anglin, R.E.S., et al., *Vitamin D deficiency and depression in adults: systematic review and meta-analysis.* The British Journal of Psychiatry, 2013. **202**(2): p. 100-107.
12. Norman, A.W., *From vitamin D to hormone D: fundamentals of the vitamin D endocrine system essential for good health.* The American Journal of Clinical Nutrition, 2008. **88**(2): p. 491S-499S.
13. Shand, A.W., et al., *Maternal vitamin D status in pregnancy and adverse pregnancy outcomes in a group at high risk for pre-eclampsia.* BJOG: An International Journal of Obstetrics & Gynaecology, 2010. **117**(13): p. 1593-1598.
14. Bodnar, L.M., et al., *Maternal Vitamin D Deficiency Increases the Risk of Preeclampsia.* The Journal of Clinical Endocrinology & Metabolism, 2007. **92**(9): p. 3517-3522.
15. Bodnar, L.M., et al., *Maternal vitamin D status and the risk of mild and severe preeclampsia.* Epidemiology, 2014. **25**(2): p. 207-214.
16. Wetta, L.A., et al., *Is midtrimester vitamin D status associated with spontaneous preterm birth and preeclampsia?* Am J Perinatol, 2014. **31**(6): p. 541-6.

17. Tabesh, M., A. Salehi-Abargouei, and A. Esmaillzadeh, *Maternal vitamin D status and risk of pre-eclampsia: a systematic review and meta-analysis.* J Clin Endocrinol Metab, 2013. **98**(8): p. 3165-73.
18. Hypponen, E., et al., *Vitamin D and pre-eclampsia: original data, systematic review and meta-analysis.* Ann Nutr Metab, 2013. **63**(4): p. 331-40.
19. Scholl, T.O., X. Chen, and T.P. Stein, *Vitamin D, secondary hyperparathyroidism, and preeclampsia.* The American Journal of Clinical Nutrition, 2013. **98**(3): p. 787-793.
20. Wei, S.Q., et al., *Longitudinal vitamin D status in pregnancy and the risk of pre-eclampsia.* BJOG: An International Journal of Obstetrics & Gynaecology, 2012. **119**(7): p. 832-839.
21. Azar, M., et al., *Serum Carotenoids and Fat-Soluble Vitamins in Women With Type 1 Diabetes and Preeclampsia: A longitudinal study.* Diabetes Care, 2011. **34**(6): p. 1258-1264.
22. Powe, C.E., et al., *Vitamin D–Binding Protein and Vitamin D Status of Black Americans and White Americans.* New England Journal of Medicine, 2013. **369**(21): p. 1991-2000.
23. Powe, C.E., et al., *First trimester vitamin D, vitamin D binding protein, and subsequent preeclampsia.* Hypertension, 2010. **56**(4): p. 758-763.
24. Baker, A.M., et al., *A nested case-control study of midgestation vitamin D deficiency and risk of severe preeclampsia.* The Journal of Clinical Endocrinology & Metabolism, 2010. **95**(11): p. 5105-5109.
25. Achkar, M., et al., *Vitamin D status in early pregnancy and risk of preeclampsia.* American journal of obstetrics and gynecology, 2014.
26. White, P. and N. Cooke, *The Multifunctional Properties and Characteristics of Vitamin D-binding Protein.* Trends in Endocrinology & Metabolism, 2000. **11**(8): p. 320-327.
27. MF, H., *The vitamin D deficiency pandemic: a forgotten hormone important for health.* Public Health Reviews, 2010. **32**: p. 267-283.
28. Haussler, M.R., et al., *Vitamin D receptor (VDR)-mediated actions of 1α,25(OH)2vitamin D3: Genomic and non-genomic mechanisms.* Best Practice & Research Clinical Endocrinology & Metabolism, 2011. **25**(4): p. 543-559.
29. Omdahl, J.L., H.A. Morris, and B.K. May, *Hydroxylase enzymes of the vitamin D pathway: expression, function, and regulation.* Annual review of nutrition, 2002. **22**(1): p. 139-166.
30. Malik, S., et al., *Common variants of the vitamin D binding protein gene and adverse health outcomes**. Critical reviews in clinical laboratory sciences, 2013. **50**(1): p. 1-22.
31. Signorello, L.B., et al., *Common variation in vitamin D pathway genes predicts circulating 25-hydroxyvitamin D levels among African Americans.* PLoS One, 2011. **6**(12): p. e28623.
32. Smolders, J., et al., *Fok-I vitamin D receptor gene polymorphism (rs10735810) and vitamin D metabolism in multiple sclerosis.* Journal of neuroimmunology, 2009. **207**(1): p. 117-121.
33. Blazer, D.G., et al., *Vitamin D receptor polymorphisms and prostate cancer.* Molecular carcinogenesis, 2000. **27**(1): p. 18-23.
34. McKay, J.D., et al., *Vitamin D receptor polymorphisms and breast cancer risk: results from the National Cancer Institute Breast and Prostate Cancer Cohort Consortium.* Cancer Epidemiology Biomarkers & Prevention, 2009. **18**(1): p. 297-305.

35. Ortlepp, J., et al., *The vitamin D receptor gene variant is associated with the prevalence of type 2 diabetes mellitus and coronary artery disease.* Diabetic medicine, 2001. **18**(10): p. 842-845.
36. Aghajafari, F., et al., *Association between maternal serum 25-hydroxyvitamin D level and pregnancy and neonatal outcomes: systematic review and meta-analysis of observational studies.* Bmj, 2013. **346**: p. f1169.
37. Ponsonby, A.-L., et al., *Vitamin D status during Pregnancy and Aspects of Offspring Health.* Nutrients, 2010. **2**(3): p. 389-407.
38. Evans, K.N., et al., *Vitamin D and placental-decidual function.* J Soc Gynecol Investig, 2004. **11**(5): p. 263-71.
39. Hewison, M., *Vitamin D and the immune system: new perspectives on an old theme.* Endocrinol Metab Clin North Am, 2010. **39**(2): p. 365-79.
40. Li, Y.C., et al., *1,25-Dihydroxyvitamin D(3) is a negative endocrine regulator of the renin-angiotensin system.* J Clin Invest, 2002. **110**(2): p. 229-38.
41. Uzan, J., et al., *Pre-eclampsia: pathophysiology, diagnosis, and management.* Vascular health and risk management, 2011. **7**: p. 467.
42. Morley, R., et al., *Maternal 25-hydroxyvitamin D concentration and offspring birth size: effect modification by infant VDR genotype.* Eur J Clin Nutr, 2009. **63**(6): p. 802-4.
43. Rezende, V., et al., *Vitamin D receptor polymorphisms in hypertensive disorders of pregnancy.* Molecular Biology Reports, 2012. **39**(12): p. 10903-10906.
44. Calvo, M.S. and S.J. Whiting, *Survey of current vitamin D food fortification practices in the United States and Canada.* J Steroid Biochem Mol Biol, 2013. **136**: p. 211-3.
45. Macdonald, H.M., *Contributions of sunlight and diet to vitamin D status.* Calcif Tissue Int, 2013. **92**(2): p. 163-76.
46. O'Neil, C.E., et al., *Food sources of energy and nutrients among adults in the US: NHANES 2003-2006.* Nutrients, 2012. **4**(12): p. 2097-120.
47. Japelt, R.B. and J. Jakobsen, *Vitamin D in plants: a review of occurrence, analysis, and biosynthesis.* Front Plant Sci, 2013. **4**(136).
48. Jones, G., S.A. Strugnell, and H.F. DeLuca, *Current understanding of the molecular actions of vitamin D.* Physiol Rev, 1998. **78**(4): p. 1193-231.
49. Holick, M.F., et al., *Photosynthesis of previtamin D3 in human skin and the physiologic consequences.* Science, 1980. **210**(4466): p. 203-205.
50. Holick, M.F., *The use and interpretation of assays for vitamin D and its metabolites.* J Nutr, 1990. **11**: p. 1464-9.
51. Bikle, D.D., et al., *Assessment of the free fraction of 25-hydroxyvitamin D in serum and its regulation by albumin and the vitamin D-binding protein.* J Clin Endocrinol Metab, 1986. **63**(4): p. 954-9.
52. Safadi, F.F., et al., *Osteopathy and resistance to vitamin D toxicity in mice null for vitamin D binding protein.* J Clin Invest, 1999. **103**(2): p. 239-51.
53. BIKLE, D.D. and E. GEE, *Free, and Not Total, 1,25-Dihydroxyvitamin D Regulates 25-Hydroxyvitamin D Metabolism by Keratinocytes.* Endocrinology, 1989. **124**(2): p. 649-654.
54. Haussler, M.R., et al., *The nuclear vitamin D receptor controls the expression of genes encoding factors which feed the "Fountain of Youth" to mediate healthful aging.* J Steroid Biochem Mol Biol, 2010. **121**(1-2): p. 88-97.

55. Papapetrou, P.D., *The interrelationship of serum 1,25-dihydroxyvitamin D, 25-hydroxyvitamin D and 24,25-dihydroxyvitamin D in pregnancy at term: a meta-analysis.* Hormones (Athens), 2010. **9**(2): p. 136-44.
56. Holick, M.F., et al., *Evaluation, treatment, and prevention of vitamin D deficiency: an Endocrine Society clinical practice guideline.* Journal of Clinical Endocrinology & Metabolism, 2011. **96**(7): p. 1911-1930.
57. Brannon, P.M. and M.F. Picciano, *Vitamin D in Pregnancy and Lactation in Humans.* Annual Review of Nutrition, 2011. **31**(1): p. 89-115.
58. Shin, J.S., et al., *Vitamin D Effects on Pregnancy and the Placenta.* Placenta, 2010. **31**(12): p. 1027-1034.
59. Brannon, P.M. and M.F. Picciano, *Vitamin D in pregnancy and lactation in humans.* Annual review of nutrition, 2011. **31**: p. 89-115.
60. Holick, M.F., *Vitamin D deficiency.* New England Journal of Medicine, 2007. **357**(3): p. 266-281.
61. Mithal, A., et al., *Global vitamin D status and determinants of hypovitaminosis D.* Osteoporosis international, 2009. **20**(11): p. 1807-1820.
62. Snellman, G., et al., *Seasonal genetic influence on serum 25-hydroxyvitamin D levels: a twin study.* PLoS One, 2009. **4**(11): p. e7747.
63. Zhang, R. and D.P. Naughton, *Vitamin D in health and disease: current perspectives.* Nutr J, 2010. **9**: p. 65.
64. Ross, A.C., et al., *The 2011 report on dietary reference intakes for calcium and vitamin D from the Institute of Medicine: what clinicians need to know*. The Journal of Clinical Endocrinology & Metabolism, 2011. **96**(1): p. 53-58.
65. Tsiaras, W.G. and M.A. Weinstock, *Factors influencing vitamin D status.* Acta dermato-venereologica, 2011. **91**(2): p. 115-124.
66. IOM, *Dietary Reference Intakes for Calcium and Vitamin D*, E. National Academy Press, Editor. 2010: D.C. Washington.
67. Holick, M.F., et al., *Evaluation, treatment, and prevention of vitamin D deficiency: an Endocrine Society clinical practice guideline.* J Clin Endocrinol Metab, 2011. **96**(7): p. 1911-30.
68. Bischoff-Ferrari, H.A., *Optimal serum 25-hydroxyvitamin D levels for multiple health outcomes.* Adv Exp Med Biol, 2014. **810**: p. 500-25.
69. Dawson-Hughes, B., et al., *Estimates of optimal vitamin D status*. Osteoporos Int. 2005 Jul;16(7):713-6. Epub 2005 Mar 18.
70. Forrest, K.Y. and W.L. Stuhldreher, *Prevalence and correlates of vitamin D deficiency in US adults.* Nutr Res, 2011. **31**(1): p. 48-54.
71. Holick, M.F., *High Prevalence of Vitamin D Inadequacy and Implications for Health.* Mayo Clinic Proceedings, 2006. **81**(3): p. 353-373.
72. Ginde, A.A., M.C. Liu, and C.A. Camargo, Jr., *Demographic differences and trends of vitamin D insufficiency in the US population, 1988-2004.* Arch Intern Med, 2009. **169**(6): p. 626-32.
73. Looker, A.C., et al., *Vitamin D status: United States, 2001-2006.* NCHS Data Brief, 2011. **59**: p. 1-8.
74. Johnson, D.D., et al., *Vitamin D deficiency and insufficiency is common during pregnancy.* Am J Perinatol, 2011. **28**(1): p. 7-12.

75. Bodnar, L.M., et al., *High prevalence of vitamin D insufficiency in black and white pregnant women residing in the northern United States and their neonates.* J Nutr, 2007. **137**(2): p. 447-52.
76. DeLuca, H.F., *The vitamin D story: a collaborative effort of basic science and clinical medicine.* Faseb J, 1988. **2**(3): p. 224-36.
77. Pols, H.A., et al., *Vitamin D: a modulator of cell proliferation and differentiation.* J Steroid Biochem Mol Biol, 1990. **37**(6): p. 873-6.
78. Steegers, E.A., et al., *Pre-eclampsia.* Lancet, 2010. **376**(9741): p. 631-44.
79. Wallis, A.B., et al., *Secular trends in the rates of preeclampsia, eclampsia, and gestational hypertension, United States, 1987-2004.* Am J Hypertens, 2008. **21**(5): p. 521-6.
80. Ananth, C.V., K.M. Keyes, and R.J. Wapner, *Pre-eclampsia rates in the United States, 1980-2010: age-period-cohort analysis.* Bmj, 2013. **347**: p. f6564.
81. Lisonkova, S. and K.S. Joseph, *Incidence of preeclampsia: risk factors and outcomes associated with early- versus late-onset disease.* American journal of obstetrics and gynecology, 2013. **209**(6): p. 544.e1-544.e12.
82. Mbah, A.K., et al., *Pre-eclampsia in the first pregnancy and subsequent risk of stillbirth in black and white gravidas.* European Journal of Obstetrics & Gynecology and Reproductive Biology, 2010. **149**(2): p. 165-169.
83. Mostello, D., et al., *Recurrence of preeclampsia: effects of gestational age at delivery of the first pregnancy, body mass index, paternity, and interval between births.* Am J Obstet Gynecol, 2008. **199**(1): p. 15.
84. Xiong, X., et al., *Association of preeclampsia with high birth weight for age.* Am J Obstet Gynecol, 2000. **183**(1): p. 148-55.
85. Rasmussen, S. and L.M. Irgens, *Fetal growth and body proportion in preeclampsia.* Obstet Gynecol, 2003. **101**(3): p. 575-83.
86. MacKay, A.P., C.J. Berg, and H.K. Atrash, *Pregnancy-related mortality from preeclampsia and eclampsia.* Obstetrics & Gynecology, 2001. **97**(4): p. 533-538.
87. Raymond, D. and E. Peterson, *A Critical Review of Early-Onset and Late-Onset Preeclampsia.* Obstetrical & Gynecological Survey, 2011. **66**(8).
88. Roberts, J.M. and M.J. Bell, *If we know so much about preeclampsia, why haven't we cured the disease?* J Reprod Immunol, 2013. **99**(1-2): p. 1-9.
89. Coomarasamy, A., et al., *Aspirin for prevention of preeclampsia in women with historical risk factors: a systematic review.* Obstetrics & Gynecology, 2003. **101**(6): p. 1319-1332.
90. Bucher, H.C., et al., *Effect of calcium supplementation on pregnancy-induced hypertension and preeclampsia: A meta-analysis of randomized controlled trials.* JAMA, 1996. **275**(14): p. 1113-1117.
91. Duley, L., *The global impact of pre-eclampsia and eclampsia.* Semin Perinatol, 2009. **33**(3): p. 130-7.
92. Redman, C.W. and I.L. Sargent, *Latest advances in understanding preeclampsia.* Science, 2005. **308**(5728): p. 1592-4.
93. Goldman-Wohl, D. and S. Yagel, *Regulation of trophoblast invasion: from normal implantation to pre-eclampsia.* Mol Cell Endocrinol, 2002. **187**(1-2): p. 233-8.
94. Redman, C.W. and I.L. Sargent, *Pre-eclampsia, the placenta and the maternal systemic inflammatory response--a review.* Placenta, 2003. **24**(7): p. S21-7.

95. Zehnder, D., et al., *Synthesis of 1,25-dihydroxyvitamin D(3) by human endothelial cells is regulated by inflammatory cytokines: a novel autocrine determinant of vascular cell adhesion.* J Am Soc Nephrol, 2002. **13**(3): p. 621-9.
96. Forman, J.P., J.S. Williams, and N.D.L. Fisher, *Plasma 25-Hydroxyvitamin D and Regulation of the Renin-Angiotensin System in Humans.* Hypertension, 2010. **55**(5): p. 1283-1288.
97. Tamblyn, J.A., et al., *Immunological role of vitamin D at the maternal-fetal interface.* J Endocrinol, 2015. **224**(3): p. R107-21.
98. Piccinni, M.P., et al., *Role of hormone-controlled Th1- and Th2-type cytokines in successful pregnancy.* J Neuroimmunol, 2000. **109**(1): p. 30-3.
99. Cardus, A., et al., *1,25-dihydroxyvitamin D3 regulates VEGF production through a vitamin D response element in the VEGF promoter.* Atherosclerosis, 2009. **204**(1): p. 85-9.
100. Cardus, A., et al., *1,25-Dihydroxyvitamin D3 regulates VEGF production through a vitamin D response element in the VEGF promoter.* Atherosclerosis, 2009. **204**(1): p. 85-89.
101. Carrara, D., et al., *Cholecalciferol administration blunts the systemic renin–angiotensin system in essential hypertensives with hypovitaminosis D.* Journal of Renin-Angiotensin-Aldosterone System, 2013.
102. Hollis, B.W., et al., *Vitamin D supplementation during pregnancy: double-blind, randomized clinical trial of safety and effectiveness.* J Bone Miner Res, 2011. **26**(10): p. 2341-57.
103. Wagner, C.L., et al., *A randomized trial of vitamin D supplementation in 2 community health center networks in South Carolina.* Am J Obstet Gynecol, 2013. **208**(2): p. 137 e1-13.
104. Fernández-Alonso, A.M., et al., *First-trimester maternal serum 25-hydroxyvitamin D3 status and pregnancy outcome.* International Journal of Gynecology & Obstetrics, 2012. **116**(1): p. 6-9.
105. Robinson, C.J., et al., *Plasma 25-OH-Vitamin D Levels in Early Onset, Severe Preeclampsia.* American journal of obstetrics and gynecology, 2010. **203**(4): p. 366.e1-366.e6.
106. Wjst, M., et al., *A genome-wide linkage scan for 25-OH-D 3 and 1, 25-(OH) 2-D 3 serum levels in asthma families.* The Journal of steroid biochemistry and molecular biology, 2007. **103**(3): p. 799-802.
107. Engelman, C.D., et al., *Genetic and environmental determinants of 25-hydroxyvitamin D and 1, 25-dihydroxyvitamin D levels in Hispanic and African Americans.* The Journal of Clinical Endocrinology & Metabolism, 2008. **93**(9): p. 3381-3388.
108. Chun, R.F., et al., *Vitamin D and DBP: The free hormone hypothesis revisited.* The Journal of steroid biochemistry and molecular biology, 2014. **144, Part A**(0): p. 132-137.
109. Chun, R.F., et al., *Vitamin D and DBP: the free hormone hypothesis revisited.* J Steroid Biochem Mol Biol, 2014. **144 Pt A**: p. 132-7.
110. Bikle, D.D., et al., *Free 1,25-dihydroxyvitamin D levels in serum from normal subjects, pregnant subjects, and subjects with liver disease.* Journal of Clinical Investigation, 1984. **74**(6): p. 1966-1971.

111. Vieth, R., *Simple method for determining specific binding capacity of vitamin D-binding protein and its use to calculate the concentration of "free" 1,25-dihydroxyvitamin D.* Clin Chem, 1994. **40**(3): p. 435-41.
112. Nykjaer, A., et al., *An endocytic pathway essential for renal uptake and activation of the steroid 25-(OH) vitamin D3.* Cell, 1999. **96**(4): p. 507-15.
113. Rowling, M.J., et al., *Megalin-mediated endocytosis of vitamin D binding protein correlates with 25-hydroxycholecalciferol actions in human mammary cells.* J Nutr, 2006. **136**(11): p. 2754-9.
114. Abboud, M., et al., *Evidence for a specific uptake and retention mechanism for 25-hydroxyvitamin D (25OHD) in skeletal muscle cells.* Endocrinology, 2013. **154**(9): p. 3022-30.
115. Ponda, M.P., D. McGee, and J.L. Breslow, *Vitamin D–Binding Protein Levels Do Not Influence The Effect of Vitamin D Repletion on Serum PTH and Calcium: Data From a Randomized, Controlled Trial.* The Journal of Clinical Endocrinology & Metabolism, 2014. **99**(7): p. 2494-2499.
116. Speeckaert, M., et al., *Biological and clinical aspects of the vitamin D binding protein (Gc-globulin) and its polymorphism.* Clinica Chimica Acta, 2006. **372**(1–2): p. 33-42.
117. Bouillon, R., *The vitamin D binding protein DBP.* Vitamin D, 2011. **1**: p. 57-72.
118. Bouillon, R., et al., *Vitamin D and Human Health: Lessons from Vitamin D Receptor Null Mice.* Endocrine Reviews, 2008. **29**(6): p. 726-776.
119. Abbas, S., et al., *The Gc2 allele of the vitamin D binding protein is associated with a decreased postmenopausal breast cancer risk, independent of the vitamin D status.* Cancer Epidemiology Biomarkers & Prevention, 2008. **17**(6): p. 1339-1343.
120. Weinstein, S.J., et al., *Circulating 25-hydroxyvitamin D, vitamin D-binding protein and risk of prostate cancer.* International Journal of Cancer, 2013. **132**(12): p. 2940-2947.
121. Malik, S., et al., *Common variants of the vitamin D binding protein gene and adverse health outcomes.* Critical reviews in clinical laboratory sciences, 2013. **50**(1): p. 1-22.
122. Janssens, W., et al., *Vitamin D deficiency is highly prevalent in COPD and correlates with variants in the vitamin D-binding gene.* Thorax, 2010. **65**(3): p. 215-220.
123. Kurylowicz, A., et al., *Vitamin D-binding protein (DBP) gene polymorphism is associated with Graves' disease and the vitamin D status in a Polish population study.* Experimental and clinical endocrinology & diabetes : official journal, German Society of Endocrinology [and] German Diabetes Association, 2006. **114**(6): p. 329-335.
124. Wang, T.J., et al., *Common genetic determinants of vitamin D insufficiency: a genome-wide association study.* Lancet, 2010. **376**(9736): p. 180-188.
125. Ahn, J., et al., *Vitamin D-related genes, serum vitamin D concentrations and prostate cancer risk.* Carcinogenesis, 2009. **30**(5): p. 769-76.
126. Ahn, J., et al., *Genome-wide association study of circulating vitamin D levels.* Hum Mol Genet, 2010. **19**(13): p. 2739-45.
127. Benjamin, E.J., et al., *Genome-wide association with select biomarker traits in the Framingham Heart Study.* BMC Med Genet, 2007. **8 Suppl 1**: p. S11.
128. Bu, F.X., et al., *Comprehensive association analysis of nine candidate genes with serum 25-hydroxy vitamin D levels among healthy Caucasian subjects.* Hum Genet, 2010. **128**(5): p. 549-56.

129. Engelman, C.D., et al., *Genome-wide association study of vitamin D concentrations in Hispanic Americans: the IRAS family study.* J Steroid Biochem Mol Biol, 2010. **122**(4): p. 186-92.
130. Fang, Y., et al., *Vitamin D binding protein genotype and osteoporosis.* Calcif Tissue Int, 2009. **85**(2): p. 85-93.
131. Jorde R, S.H., Wilsgaard T, Joakimsen RM, Mathiesen EB, Njølstad I, Løchen ML, Figenschau Y, Berg JP, Svartberg J, Grimnes G., *Polymorphisms related to the serum 25-hydroxyvitamin D level and risk of myocardial infarction, diabetes, cancer and mortality. The Tromsø Study.* PLoS One, 2012. **7**(5).
132. Lu, L., et al., *Associations between common variants in GC and DHCR7/NADSYN1 and vitamin D concentration in Chinese Hans.* Hum Genet, 2012. **131**(3): p. 505-12.
133. Sinotte, M., et al., *Genetic polymorphisms of the vitamin D binding protein and plasma concentrations of 25-hydroxyvitamin D in premenopausal women.* The American Journal of Clinical Nutrition, 2009. **89**(2): p. 634-640.
134. Wjst, M., et al., *Asthma families show transmission disequilibrium of gene variants in the vitamin D metabolism and signalling pathway.* Respir Res, 2006. **7**: p. 60.
135. Zhang, Z., et al., *An analysis of the association between the vitamin D pathway and serum 25-hydroxyvitamin D levels in a healthy Chinese population.* J Bone Miner Res, 2013.
136. Zhang, Y.Y., et al., *Estrogen receptor alpha and vitamin D receptor gene polymorphisms and bone mineral density: association study of healthy pre- and postmenopausal Chinese women.* Biochem Biophys Res Commun, 2003. **308**(4): p. 777-83.
137. Baker, A.R., et al., *Cloning and expression of full-length cDNA encoding human vitamin D receptor.* Proceedings of the National Academy of Sciences of the United States of America, 1988. **85**(10): p. 3294-3298.
138. Kurylowicz, A., et al., *Vitamin D-binding protein (DBP) gene polymorphism is associated with Graves' disease and the vitamin D status in a Polish population study.* Exp Clin Endocrinol Diabetes, 2006. **114**(6): p. 329-35.
139. Engelman, C.D., et al., *Genetic and Environmental Determinants of 25-Hydroxyvitamin D and 1,25-Dihydroxyvitamin D Levels in Hispanic and African Americans.* The Journal of Clinical Endocrinology & Metabolism, 2008. **93**(9): p. 3381-3388.
140. Cooper, J.D., et al., *Inherited variation in vitamin D genes is associated with predisposition to autoimmune disease type 1 diabetes.* Diabetes, 2011. **60**(5): p. 1624-31.
141. Lu, L., et al., *Associations between common variants in GC and DHCR7/NADSYN1 and vitamin D concentration in Chinese Hans.* Human Genetics, 2012. **131**(3): p. 505-512.
142. Fu, L., et al., *Common genetic variants of the vitamin D binding protein (DBP) predict differences in response of serum 25-hydroxyvitamin D [25(OH)D] to vitamin D supplementation.* Clin Biochem, 2009. **42**(10-11): p. 1174-7.
143. Wu, S., et al., *Splice Variants of the CYP27b1 Gene and the Regulation of 1,25-Dihydroxyvitamin D3 Production.* Endocrinology, 2007. **148**(7): p. 3410-3418.
144. Dardenne, O., et al., *Targeted Inactivation of the 25-Hydroxyvitamin D3-1α-Hydroxylase Gene (CYP27B1) Creates an Animal Model of Pseudovitamin D-Deficiency Rickets 1.* Endocrinology, 2001. **142**(7): p. 3135-3141.
145. Ramagopalan, S.V., et al., *Rare variants in the CYP27B1 gene are associated with multiple sclerosis.* Annals of neurology, 2011. **70**(6): p. 881-886.

146. Bailey, R., et al., *Association of the vitamin D metabolism gene CYP27B1 with type 1 diabetes.* Diabetes, 2007. **56**(10): p. 2616-2621.
147. Lopez, E.R., et al., *A promoter polymorphism of the CYP27B1 gene is associated with Addison's disease, Hashimoto's thyroiditis, Graves' disease and type 1 diabetes mellitus in Germans.* European Journal of Endocrinology, 2004. **151**(2): p. 193-197.
148. Ramos-Lopez, E., et al., *Gestational diabetes mellitus and vitamin D deficiency: genetic contribution of CYP27B1 and CYP2R1 polymorphisms.* Diabetes Obes Metab, 2008. **10**(8): p. 683-5.
149. Orton, S.M., et al., *Evidence for genetic regulation of vitamin D status in twins with multiple sclerosis.* Am J Clin Nutr, 2008. **88**(2): p. 441-7.
150. Hypponen, E., et al., *Serum 25-hydroxyvitamin D and IgE - a significant but nonlinear relationship.* Allergy, 2009. **64**(4): p. 613-620.
151. Hyppönen, E., et al., *Serum 25-hydroxyvitamin D and IgE – a significant but nonlinear relationship.* Allergy, 2009. **64**(4): p. 613-620.
152. Lange, C.M., et al., *Vitamin D deficiency and a CYP27B1-1260 promoter polymorphism are associated with chronic hepatitis C and poor response to interferon-alfa based therapy.* Journal of Hepatology, 2011. **54**(5): p. 887-893.
153. Orton, S.-M., et al., *Evidence for genetic regulation of vitamin D status in twins with multiple sclerosis.* The American Journal of Clinical Nutrition, 2008. **88**(2): p. 441-447.
154. Haussler, M.R., et al., *The Nuclear Vitamin D Receptor: Biological and Molecular Regulatory Properties Revealed.* Journal of Bone and Mineral Research, 1998. **13**(3): p. 325-349.
155. Haussler, M.R., et al., *The vitamin D hormone and its nuclear receptor: molecular actions and disease states.* J Endocrinol, 1997. **154**(73): p. S57-73.
156. Hayes, C.E., M.T. Cantorna, and H.F. DeLuca, *Vitamin D and multiple sclerosis.* Experimental Biology and Medicine, 1997. **216**(1): p. 21-27.
157. Mondul, A.M., et al., *Vitamin D-associated genetic variation and risk of breast cancer in the Breast and Prostate Cancer Cohort Consortium (BPC3).* Cancer Epidemiology Biomarkers & Prevention, 2014: p. cebp. 1127.2014.
158. Witham, M.D., et al., *Vitamin D Therapy to Reduce Blood Pressure and Left Ventricular Hypertrophy in Resistant Hypertension Randomized, Controlled Trial.* Hypertension, 2014. **63**(4): p. 706-712.
159. Li, K., et al., *The association of vitamin D receptor gene polymorphisms and serum 25-hydroxyvitamin D levels with generalized vitiligo.* Br J Dermatol, 2012. **167**(4): p. 815-21.
160. Wang, T.J., et al., *Common genetic determinants of vitamin D insufficiency: a genome-wide association study.* The Lancet, 2010. **376**(9736): p. 180-188.
161. Engelman, C.D., et al., *Genetic and Environmental Determinants of 25-Hydroxyvitamin D and 1,25-Dihydroxyvitamin D Levels in Hispanic and African Americans.* The Journal of Clinical Endocrinology and Metabolism, 2008. **93**(9): p. 3381-3388.
162. Morley, R., et al., *Maternal 25-Hydroxyvitamin D and Parathyroid Hormone Concentrations and Offspring Birth Size.* The Journal of Clinical Endocrinology & Metabolism, 2006. **91**(3): p. 906-912.
163. Aslani, S., et al., *VDR FokI polymorphism and its potential role in the pathogenesis of gestational diabetes mellitus and its complications.* Gynecol Endocrinol, 2011. **27**(12): p. 1055-60.

164. Fang, X., et al., *Polymorphisms in the MTHRF, VDR, MMP-9 and IL-β genes and the risk of premature rupture of membranes.* Gynecologic and obstetric investigation, 2009. **70**(3): p. 206-214.
165. Manzon, L., et al., *Vitamin D receptor polymorphism FokI is associated with spontaneous idiopathic preterm birth in an Israeli population.* Eur J Obstet Gynecol Reprod Biol, 2014. **177**: p. 84-8.
166. Bodnar, L.M., et al., *Maternal Serum 25-Hydroxyvitamin D Concentrations Are Associated with Small-for-Gestational Age Births in White Women.* The Journal of nutrition, 2010. **140**(5): p. 999-1006.
167. Swamy, G.K., et al., *Maternal vitamin D receptor genetic variation contributes to infant birthweight among black mothers.* American Journal of Medical Genetics Part A, 2011. **155**(6): p. 1264-1271.
168. Sinotte, M., et al., *Genetic polymorphisms of the vitamin D binding protein and plasma concentrations of 25-hydroxyvitamin D in premenopausal women.* Am J Clin Nutr, 2009. **89**(2): p. 634-40.
169. Daly, A.K. and C.P. Day, *Candidate gene case-control association studies: advantages and potential pitfalls.* Br J Clin Pharmacol, 2001. **52**(5): p. 489-99.
170. Conde-Agudelo, A., et al., *Cigarette smoking during pregnancy and risk of preeclampsia: a systematic review.* American journal of obstetrics and gynecology, 1999. **181**(4): p. 1026-1035.
171. Ananth, C.V., K.M. Keyes, and R.J. Wapner, *Pre-eclampsia rates in the United States, 1980-2010: age-period-cohort analysis*. Vol. 347. 2013.
172. Holick, M., et al., *Prevalence of vitamin D inadequacy among postmenopausal North American women receiving osteoporosis therapy.* J Clin Endocrinol Metab, 2005. **90**: p. 3215 - 3224.
173. Lips, P., *Which circulating level of 25-hydroxyvitamin D is appropriate?* The Journal of Steroid Biochemistry and Molecular Biology, 2004. **89–90**(0): p. 611-614.
174. Kuha, J., *AIC and BIC: Comparisons of Assumptions and Performance.* Sociological Methods & Research, 2004. **33**(2): p. 188-229.
175. Bodnar, L.M., et al., *Racial and seasonal differences in 25-hydroxyvitamin D detected in maternal sera frozen for over 40 years.* Br J Nutr, 2009. **101**(2): p. 278-84.
176. Zerwekh, J.E., *The measurement of vitamin D: analytical aspects.* Ann Clin Biochem, 2004. **41**(Pt 4): p. 272-81.
177. Ramos-Lopez, E., et al., *Gestational diabetes mellitus and vitamin D deficiency: genetic contribution of CYP27B1 and CYP2R1 polymorphisms.* Diabetes, Obesity and Metabolism, 2008. **10**(8): p. 683-685.
178. *The International HapMap Project.* Nature, 2003. **426**(6968): p. 789-796.
179. Ardlie, K.G., K.L. Lunetta, and M. Seielstad, *Testing for Population Subdivision and Association in Four Case-Control Studies.* American Journal of Human Genetics, 2002. **71**(2): p. 304-311.
180. StataCorp, *Stata Statistical Software: Release 12*, S. LP, Editor. 2011, College Station.
181. Kim, M.Y., X. Xue, and Y. Du, *Approaches for Calculating Power for Case–Cohort Studies.* Biometrics, 2006. **62**(3): p. 929-933.
182. Gauderman, W. and J. Morrison, *QUANTO version 1.2. 3: a computer program for power and sample size calculations for genetic-epidemiology studies*. 2007.

183. Royston, P., *Multiple imputation of missing values: Further update of ICE, with an emphasis on categorical variables. Stata Journal, 9, 466–477*. 2009.
184. Sterne, J.A., et al., *Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls.* Bmj, 2009. **338**.
185. Cologne, J., et al., *Conventional case–cohort design and analysis for studies of interaction.* International journal of epidemiology, 2012: p. dys102.
186. Bentler, P.M. and A. Mooijaart, *Choice of structural model via parsimony: a rationale based on precision.* Psychological bulletin, 1989. **106**(2): p. 315.
187. ROTHMAN, K.J., *Synergy and antagonism in cause-effect relationships.* American Journal of Epidemiology, 1974. **99**(6): p. 385-388.
188. Rothman, K., S. Greenland, and T. Lash, *Modern Epidemiology, Lippincott Williams and Wilkins, Philadelphia, PA.* 1998.
189. Thompson, W.D., *Effect modification and the limits of biological inference from epidemiologic data.* Journal of Clinical Epidemiology. **44**(3): p. 221-232.
190. Anderson, C.A., et al., *Data quality control in genetic case-control association studies.* Nat. Protocols, 2010. **5**(9): p. 1564-1573.
191. Purcell S, N.B., Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ & Sham PC, *PLINK: a toolset for whole-genome association and population-based linkage analysis.* American Journal of Human Genetics, 2007. **81**.
192. Enoch, M.-A., et al., *Using ancestry-informative markers to define populations and detect population stratification.* Journal of Psychopharmacology, 2006. **20**(4 suppl): p. 19-26.
193. Kosoy, R., et al., *Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America.* Hum Mutat, 2009. **30**(1): p. 69-78.
194. Pritchard, J.K., M. Stephens, and P. Donnelly, *Inference of population structure using multilocus genotype data.* Genetics, 2000. **155**(2): p. 945-959.
195. Falush, D., M. Stephens, and J.K. Pritchard, *Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies.* Genetics, 2003. **164**(4): p. 1567-1587.
196. Falush, D., M. Stephens, and J.K. Pritchard, *Inference of population structure using multilocus genotype data: dominant markers and null alleles.* Molecular ecology notes, 2007. **7**(4): p. 574-578.
197. Hubisz, M.J., et al., *Inferring weak population structure with the assistance of sample group information.* Molecular ecology resources, 2009. **9**(5): p. 1322-1332.
198. Evanno, G., S. Regnaut, and J. Goudet, *Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study.* Molecular ecology, 2005. **14**(8): p. 2611-2620.
199. Silverberg, M.S., et al., *Ulcerative colitis-risk loci on chromosomes 1p36 and 12q15 found by genome-wide association study.* Nat Genet, 2009. **41**(2): p. 216-20.
200. Fisher, S.A., et al., *Genetic determinants of ulcerative colitis include the ECM1 locus and five loci implicated in Crohn's disease.* Nat Genet, 2008. **40**(6): p. 710-2.
201. Teo, Y.Y., et al., *On the Usage of HWE for Identifying Genotyping Errors.* Annals of Human Genetics, 2007. **71**(5): p. 701-703.
202. Tabangin, M.E., J.G. Woo, and L.J. Martin, *The effect of minor allele frequency on the likelihood of obtaining false positives.* BMC Proceedings, 2009. **3**(Suppl 7): p. S41-S41.

203. Thomas, D.C. and J.S. Witte, *Point: Population Stratification: A Problem for Case-Control Studies of Candidate-Gene Associations?* Cancer Epidemiology Biomarkers & Prevention, 2002. **11**(6): p. 505-512.
204. Devlin, B. and K. Roeder, *Genomic control for association studies.* Biometrics, 1999. **55**(4): p. 997-1004.
205. Pritchard, J.K., et al., *Association Mapping in Structured Populations.* American Journal of Human Genetics, 2000. **67**(1): p. 170-181.
206. Weng, H.Y., et al., *Methods of covariate selection: directed acyclic graphs and the change-in-estimate procedure.* Am J Epidemiol, 2009. **169**(10): p. 1182-90.
207. Holm, S., *A simple sequentially rejective multiple test procedure.* Scandinavian journal of statistics, 1979: p. 65-70.
208. SIMES, R.J., *An improved Bonferroni procedure for multiple tests of significance.* Biometrika, 1986. **73**(3): p. 751-754.
209. Benjamini, Y. and Y. Hochberg, *Controlling the false discovery rate: a practical and powerful approach to multiple testing.* Journal of the Royal Statistical Society. Series B (Methodological), 1995: p. 289-300.
210. Gabriel, S.B., et al., *The Structure of Haplotype Blocks in the Human Genome.* Science, 2002. **296**(5576): p. 2225-2229.
211. Deeks, J.J., D.G. Altman, and M.J. Bradburn, *Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis.* Systematic Reviews in Health Care: Meta-Analysis in Context, Second Edition, 2008: p. 285-312.
212. Higgins, J. and S.G. Thompson, *Quantifying heterogeneity in a meta-analysis.* Statistics in medicine, 2002. **21**(11): p. 1539-1558.
213. Ananth, C.V., K.M. Keyes, and R.J. Wapner, *Pre-eclampsia rates in the United States, 1980-2010: age-period-cohort analysis.* BMJ, 2013. **347**.
214. Hutcheon, J.A., S. Lisonkova, and K.S. Joseph, *Epidemiology of pre-eclampsia and the other hypertensive disorders of pregnancy.* Best Practice & Research Clinical Obstetrics & Gynaecology, 2011. **25**(4): p. 391-403.
215. Ji, L., et al., *Placental trophoblast cell differentiation: Physiological regulation and pathological relevance to preeclampsia.* Molecular Aspects of Medicine, 2013. **34**(5): p. 981-1023.
216. Marseglia, L., et al., *Oxidative Stress-Mediated Aging during the Fetal and Perinatal Periods.* Oxidative Medicine and Cellular Longevity, 2014. **2014**: p. 358375.
217. Chan, S.Y., et al., *Vitamin D promotes human extravillous trophoblast invasion in vitro.* Placenta, 2015.
218. Wetta, L.A., et al., *Is Midtrimester Vitamin D Status Associated with Spontaneous Preterm Birth and Preeclampsia?* Am J Perinatol, 2013.
219. Bodnar, L.M., et al., *Maternal vitamin D deficiency increases the risk of preeclampsia.* Journal of Clinical Endocrinology & Metabolism, 2007. **92**(9): p. 3517-22.
220. Ullah, M.I., et al., *Vitamin d deficiency and the risk of preeclampsia and eclampsia in bangladesh.* Horm Metab Res, 2013. **45**(9): p. 682-7.
221. Xu, L., et al., *The Relationship of Hypovitaminosis D and IL-6 in Preeclampsia.* American Journal of Obstetrics and Gynecology, (0).
222. Wei, S.Q., et al., *Maternal plasma 25-hydroxyvitamin D levels, angiogenic factors, and preeclampsia.* Am J Obstet Gynecol, 2013. **208**(5): p. 390 e1-6.

223. Azar, M., et al., *Serum carotenoids and fat-soluble vitamins in women with type 1 diabetes and preeclampsia: a longitudinal study.* Diabetes Care, 2011. **34**(6): p. 1258-64.
224. Shand, A.W., et al., *Maternal vitamin D status in pregnancy and adverse pregnancy outcomes in a group at high risk for pre-eclampsia.* BJOG, 2010. **117**(13): p. 1593-8.
225. Powe, C.E., et al., *First trimester vitamin D, vitamin D binding protein, and subsequent preeclampsia.* Hypertension, 2010. **56**(4): p. 758-63.
226. Wetta, L.A., et al., *Is midtrimester vitamin D status associated with spontaneous preterm birth and preeclampsia?* American journal of perinatology, 2014. **31**(6): p. 541-546.
227. Bodnar, L.M., et al., *Maternal vitamin D status and the risk of mild and severe preeclampsia.* Epidemiology, 2014. **25**(2): p. 207-14.
228. Baker, A.M., et al., *A nested case-control study of midgestation vitamin D deficiency and risk of severe preeclampsia.* J Clin Endocrinol Metab, 2010. **95**(11): p. 5105-9.
229. Myatt, L., et al., *Strategy for Standardization of Preeclampsia Research Study Design.* Hypertension, 2014. **63**(6): p. 1293-1301.
230. Bodnar, L.M., R.W. Platt, and H.N. Simhan, *Early-pregnancy vitamin D deficiency and risk of preterm birth subtypes.* Obstet Gynecol, 2015. **125**(2): p. 439-47.
231. Rao, S.J., *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis.* Journal of the American Statistical Association, 2003. **98**(461): p. 257-258.
232. Harrell Jr, F.E., *Multivariable modeling strategies*, in *Regression Modeling Strategies*. 2001, Springer. p. 53-85.
233. StataCorp, *Stata Statistical Software: Release 14*, S. LP, Editor. 2015: College Station, TX.
234. Wei, S.Q., et al., *Maternal vitamin D status and adverse pregnancy outcomes: a systematic review and meta-analysis.* J Matern Fetal Neonatal Med, 2013. **26**(9): p. 889-99.
235. Aghajafari, F., et al., *Association between maternal serum 25-hydroxyvitamin D level and pregnancy and neonatal outcomes: systematic review and meta-analysis of observational studies.* Bmj, 2013. **346**: p. f1169.
236. Robinson, C.J., et al., *Association of Maternal Vitamin D and Placenta Growth Factor with the Diagnosis of Early Onset Severe Preeclampsia.* Am J Perinatol, 2012.
237. Cantorna, M.T., et al., *Vitamin D status, 1,25-dihydroxyvitamin D3, and the immune system.* The American Journal of Clinical Nutrition, 2004. **80**(6): p. 1717S-1720S.
238. von Versen-Hoynck, F., et al., *Vitamin D antagonizes negative effects of preeclampsia on fetal endothelial colony forming cell number and function.* PLoS One, 2014. **9**(6).
239. Merke, J., et al., *Identification and regulation of 1,25-dihydroxyvitamin D3 receptor activity and biosynthesis of 1,25-dihydroxyvitamin D3. Studies in cultured bovine aortic endothelial cells and human dermal capillaries.* Journal of Clinical Investigation, 1989. **83**(6): p. 1903-1915.
240. Uberti, F., et al., *Vitamin D Protects Human Endothelial Cells From Oxidative Stress Through the Autophagic and Survival Pathways.* The Journal of Clinical Endocrinology & Metabolism, 2014. **99**(4): p. 1367-1374.
241. Brodowski, L., et al., *Vitamin D prevents endothelial progenitor cell dysfunction induced by sera from women with preeclampsia or conditioned media from hypoxic placenta.* PLoS One, 2014. **9**(6): p. e98527.

242. *Hypertension in pregnancy. Report of the American College of Obstetricians and Gynecologists' Task Force on Hypertension in Pregnancy.* Obstet Gynecol, 2013. **122**(5): p. 1122-31.
243. Duley, L. *The global impact of pre-eclampsia and eclampsia*. in *Seminars in perinatology*. 2009. Elsevier.
244. Sibai, B., G. Dekker, and M. Kupferminc, *Pre-eclampsia.* The Lancet, 2005. **365**(9461): p. 785-799.
245. Ghulmiyyah, L. and B. Sibai, *Maternal Mortality From Preeclampsia/Eclampsia.* Seminars in Perinatology, 2012. **36**(1): p. 56-59.
246. Lo, J.O., J.F. Mission, and A.B. Caughey, *Hypertensive disease of pregnancy and maternal mortality.* Current Opinion in Obstetrics and Gynecology, 2013. **25**(2): p. 124-132.
247. Niswander, K. and M. Gordon, *The Collaborative Perinatal Study of the National Institute of Neurological Diseases and Stroke: The Women and Their Pregnancies. 1972.* Philadelphia, PA: WB Saunders.
248. Friedman, E.N., RK. , *Pregnancy hypertension: A systematic evaluation of clinical diagnostic criteria*. 1977: Littleton, MA.
249. Browning, B.L. and Z. Yu, *Simultaneous Genotype Calling and Haplotype Phasing Improves Genotype Accuracy and Reduces False-Positive Associations for Genome-wide Association Studies.* The American Journal of Human Genetics, 2009. **85**(6): p. 847-861.
250. Zhan, Y., et al., *Genetic variations in the vitamin-D receptor (VDR) gene in preeclampsia patients in the Chinese Han population.* Hypertens Res, 2015. **38**(7): p. 513-7.
251. Ogunkolade, B.-W., et al., *Vitamin D receptor (VDR) mRNA and VDR protein levels in relation to vitamin D status, insulin secretory capacity, and VDR genotype in Bangladeshi Asians.* Diabetes, 2002. **51**(7): p. 2294-2300.
252. Ross, A., et al., *Dietary reference intakes for calcium and vitamin D. Institute of Medicine (US) Committee to Review Dietary Reference Intakes for Vitamin D and Calcium*. 2011, National Academies Press (US), Washington (DC).
253. Taksler, G.B., et al., *Vitamin D deficiency in minority populations.* Public health nutrition, 2015. **18**(03): p. 379-391.
254. Elder, C.J. and N.J. Bishop, *Rickets.* Lancet, 2014. **383**(9929): p. 1665-76.
255. Wei, S.Q., *Vitamin D and pregnancy outcomes.* Curr Opin Obstet Gynecol, 2014. **26**(6): p. 438-47.
256. Pojsupap, S., et al., *Efficacy of high-dose vitamin D in pediatric asthma: a systematic review and meta-analysis*. J Asthma, 2015. **52**(4): p. 382-90.
257. Chakhtoura, M. and S.T. Azar, *The Role of Vitamin D Deficiency in the Incidence, Progression, and Complications of Type 1 Diabetes Mellitus.* International Journal of Endocrinology, 2013. **2013**: p. 148673.
258. Ahn, J., et al., *Genome-wide association study of circulating vitamin D levels.* Human Molecular Genetics, 2010.
259. Heijboer, A.C., et al., *Accuracy of 6 Routine 25-Hydroxyvitamin D Assays: Influence of Vitamin D Binding Protein Concentration.* Clinical Chemistry, 2012. **58**(3): p. 543-548.
260. Evans, K.N., et al., *Vitamin D and Placental-Decidual Function.* Journal of the Society for Gynecologic Investigation, 2004. **11**(5): p. 263-271.
261. Dusso, A.S., A.J. Brown, and E. Slatopolsky, *Vitamin D.* Am J Physiol Renal Physiol, 2005. **289**(1): p. F8-28.