**EPIDEMIOLOGY**

# A common *IGF1R* gene variant predicts later life breast cancer risk in women with preeclampsia

Mark Powell[1] · Sophia Fuller[2] · Erica Gunderson[3] · Christopher Benz[1]

## Abstract

**Purpose** Preeclampsia has been inconsistently associated with altered later life risk of cancer. This study utilizes the Nurses' Health Study 2 (NHS2) to determine if the future risk of breast and non-breast cancers in women who experience preeclampsia is modified by carrying a protective variant of rs2016347, a functional insulin-like growth factor receptor-1 (*IGF1R*) single nucleotide polymorphism.

**Methods** This retrospective cohort study completed within the NHS2 evaluated participants enrolled in 1989 and followed them through 2015, with a study population of 86,751 after exclusions. Cox proportional hazards models both with and without the impact of rs2016347 genotype were used to assess the risk of invasive breast cancer, hormone receptor-positive (HR+) breast cancer, and non-breast cancers.

**Results** Women with preeclampsia had no change in risk of all breast, HR+ breast, or non-breast cancers when not considering genotype. However, women carrying at least one T allele of rs2016347 had a lower risk of HR+ breast cancer, HR 0.67, 95% CI: 0.47–0.97, $P = 0.04$, with interaction term $P = 0.06$. For non-breast cancers as a group, women carrying a T allele had an HR 0.76, 95% CI: 0.53–1.08, $P = 0.12$, with interaction term $P = 0.26$.

**Conclusions** This retrospective cohort study found that women with preeclampsia who carry a T allele of *IGF1R* rs2016347 had a reduced future risk of developing HR+ breast cancer, and a reduced but not statistically significant decreased risk of non-breast cancers suggesting a possible role for the IGF-1 axis in the development of cancer in these women.

**Keywords** Preeclampsia · Breast cancer · IGF1R · Hormone receptor-positive · SNP · rs2016347

## Introduction

Preeclampsia develops in 5–12% of all pregnancies depending on race, ethnicity, and other risk factors within the population and is a leading cause of maternal/fetal perinatal mortality as well as being linked with later life maternal cardiovascular disease (CVD) [1]. Much less widely accepted, however, is that preeclampsia may protect against later life maternal breast cancer development, as the evidence is mixed depending on specific cohort characteristics and whether or not a woman's inherited genotype for the common insulin-like growth factor receptor-1 (*IGF1R*) gene variant rs2016347 is also considered [2–4]. Further validation and mechanistic insight into the breast cancer protective impact of preeclampsia as influenced by this functional *IGF1R* single nucleotide polymorphism (SNP) might immediately improve breast cancer risk assessment and personalized screening, as well as suggest new breast cancer prevention strategies.

Preeclampsia is characterized by the development of high blood pressure after the 20th week of gestation plus the presence of proteinuria or other end organ dysfunction, which distinguishes it from gestational hypertension, the other major component of the broader class of disorders known as hypertensive disorders of pregnancy (HDP) [5, 6]. The association of preeclampsia with increased maternal risk of future CVD and mortality is now well-established [7, 8]. An association of preeclampsia with lower breast cancer risk was initially reported decades ago, and many subsequent

✉ Mark Powell
markp@zerobreastcancer.org

1 Buck Institute for Research On Aging, 8001 Redwood Blvd, Novato, CA 94945, USA

2 Graduate Group in Biostatistics, School of Public Health, University of California, Berkeley, CA, USA

3 Division of Research, Kaiser Permanente Northern California, Oakland, CA, USA

large cohort studies have shown 10–20% lower relative risk, although these findings have not been universal [9–13]. The association of preeclampsia with later life risk for developing other cancers has also yielded inconsistent results overall and for specific cancer subtypes [14–18].

The systemwide sequelae of preeclampsia have long been noted and debated; and while preeclampsia is still primarily considered a placental disorder its exact pathogenesis remains unclear. Postulated systemic impacts include vascular ischemia, immune dysfunction, oxidative stress, and dysregulated angiogenesis, along with derangements of the insulin-like growth factor-1 (IGF-1) axis [5, 19]. With regard to the latter, several reports have described low IGF-1 blood levels as well as decreased placental expression of IGF-1 in women with preeclampsia, and this has been hypothesized to play a role in the future risk of breast and other cancers [20–25].

IGF-1 plays an essential role in normal breast development; yet upregulation of the IGF-1 axis and increased IGF1R expression appear to be associated with the development of breast cancer [26–29]. A pooled analysis of 17 studies demonstrated that higher levels of circulating IGF-1 are associated with increased risk of breast cancer, with this being especially true for hormone receptor-positive (HR+) breast cancer [30]. Similarly, elevated serum levels or increased expression of IGF-1 has been implicated in the development of various non-breast cancers [31–35]. Of mechanistic relevance to the present study, the normal breast tissue of women with a history of preeclampsia who also inherit the lower expressing T allele of *IGF1R* rs2016347 exhibits increased mammary gland involution, a histologic predictor of reduced breast cancer risk [36].

Prebil in 2014 first reported that women with HDP whose germlines possess the T allele of rs2016347 had significantly lower mammographic density, a clinically recognized proxy for breast cancer risk [37]. This was postulated to result from the fact that inheritance of the functional rs2016347 T allele had been shown to be associated with significantly lower IGF1R mRNA expression in various healthy organs including the breast [38]. A subsequent case–control analysis within the California Teachers Study (CTS) demonstrated a decreased risk of breast cancer of over 60% in women with preeclampsia carrying the rs2016347 T allele, with a larger protective effect seen against HR+ breast cancer development [2]. However, a more recent UK Biobank analysis of women with a history of HDP did not find a lower risk for developing breast cancer in those carrying the T allele, although this study did not include specific preeclampsia history or data on HR+ breast cancer [18]. Others have shown that the outcome impact of the rs2016347 T allele is greater in women diagnosed with HR+ breast cancers [39, 40]. When looking at non-breast cancers, the UK Biobank study did demonstrate that women with HDP carrying the

protective T allele had a statistically significant 41% lower risk of developing non-breast cancers, while there was no impact on cancer risk by HDP or the T allele alone [18].

Given the above inconsistent population-based evidence, we turned to the large and well-studied longitudinal Nurses' Health Study 2 (NHS2) with its specific data on preeclampsia, HR+ status, and non-breast cancer outcomes to confirm the potentially important modulating impact of rs2016347 T allele inheritance on a woman's future risk of developing cancer.

# Methods

## Study population and design

This retrospective cohort study was completed within the NHS2, which enrolled 116,430 female registered nurses in 1989 with ages at entry ranging from 25 to 42. The original detailed entry questionnaire collected extensive data as did subsequent follow-up biennial questionnaires (available at https://nurseshealthstudy.org/participants/questionnaires) including preeclampsia history and necessary covariates such as age at entry, ethnicity, Body Mass Index (BMI), parity, age at first birth, age at menarche, family history of breast cancer, diet, physical activity, and smoking history. Covariates were modeled as continuous except for ethnicity, age at menarche, smoking history, and family history of breast cancer which were categorical with nominal categories as described in Table 1. Many papers detail the NHS2 and its extensive contributions to medical research [41–43]. The retrospective cohort sample for this study excluded 19,969 participants who were nulliparous, 9493 who had missing covariates, and 217 who developed a breast or a non-breast cancer prior to their diagnosis of preeclampsia leaving a final study number of 86,751. A subset of participants provided blood and buccal DNA samples for potential genotyping resulting in available information for our *IGF1R* SNP of interest, rs2016347, on 6577 participants; details on the genotyping process are provided below. A flow chart of study numbers is presented in Fig. 1. The NHS2 study protocol was approved by the Institutional Review Board (IRB) of the Brigham and Women's Hospital, and the IRB allowed participants' completion of questionnaires as implied consent.

## Study exposures and outcomes

Assessment of the primary exposure of preeclampsia was by self-report. Participants were asked if they had experienced preeclampsia in any pregnancy on their entry questionnaire in 1989 and on subsequent biennial questionnaires through 2001. Preeclampsia was defined as hypertension in

**Table 1** Characteristics of NHS2 participants by preeclampsia status[a]

| Characteristic | Preeclampsia+ $N = 11,328$ | Preeclampsia − $N = 75,423$ |
|---|---|---|
| Age at entry | 34.3 (4.6) | 34.4 (4.7) |
| Race/ethnicity | | |
|   White non-Hispanic | 10,550 (93.1%) | 70,665 (93.7%) |
|   Other | 778 (6.9%) | 4,758 (6.3%) |
| BMI at entry[b] | 26.2 (5.9) | 23.6 (4.5) |
| Parity | 2.40 (1.05) | 2.41 (1.04) |
| Age at first birth | 26.0 (4.7) | 26.6 (4.8) |
| Age at menarche | | |
|   <= 11 | 3,380 (29.8%) | 17,348 (23.1%) |
|   12–13 | 6,257 (55.2%) | 43,965 (58.3%) |
|   14+ | 1,653 (15.0%) | 13,894 (18.4%) |
| Smoking history | | |
|   Never | 7,319 (64.6%) | 49,770 (66.0%) |
|   Past/current | 4,009 (35.4%) | 25,653 (34.0%) |
| Family history of breast cancer | 1,607 (14.2%) | 10,949 (14.5%) |
| Diet (HEI)[c] | 53.8 (10.3) | 55.2 (10.5) |
| Physical activity (METS/week)[d] | 20.5 (20.3) | 22.3 (20.8) |

[a]Mean (standard deviation); $n$ (%)

[b]*BMI* body mass index

[c]*HEI* healthy eating index

[d]*METS/week* metabolic equivalents per week

pregnancy plus proteinuria on the entry questionnaire and as "preeclampsia/toxemia" on the follow-up biennial questionnaires, and participants responding yes on any questionnaire were deemed to have had a positive history of preeclampsia.

Outcome diagnoses of invasive breast cancer and non-breast cancers (defined as all cancers except breast and non-melanoma skin cancers) were determined by self-report with subsequent validation by NHS2 staff who obtained consent to review medical records. HR (±) breast cancer status was determined from medical records and/or pathology reports. Participants with HR− breast cancer or unclear HR status were excluded entirely from the HR+ breast cancer analyses, and the low number of documented HR− cases precluded their analysis as a subgroup.

## Samples for genotyping

In order to obtain genetic data for the NHS2, blood samples were collected from 29,611 and buccal smears from 29,392 participants from which DNA was extracted using Qiagen PureGene DNA Isolation Kits (Gentra Systems, Minneapolis, MN). Initial genotyping results of rs2016347 by the NHS2 was available on 6434 (7.4%) of study participants utilizing at least one of five different DNA analysis platforms as previously reported [44, 45]. The initial genotyping was performed entirely in 6 case control studies and preeclampsia disease status was not a selection criterion in any of these studies. The diseases studied and resulting

genotyped participants included in the current study include breast cancer ($n = 2,560$), post-traumatic stress disorder ($n = 1,553$), endometriosis ($n = 1,461$), nephrolithiasis ($n = 452$), venous thromboembolism ($n = 295$), and ovarian cancer ($n = 113$). Of note, in the initial genotyping attempts were made to genotype all cases of breast cancer at that time.

In an effort to increase power for the genetic analyses, participants with preeclampsia who subsequently developed invasive breast cancer ($n = 50$) and had DNA available for genotyping but had not been initially genotyped had their DNA sent to the Brigham and Women's Hospital/Harvard Cohorts Biorepository for rs2016347 genotyping. Similarly, 138 participants with preeclampsia and non-breast cancers with available DNA had genotyping requested. Because of limitations in DNA quantity/quality, only 31 of the 50 breast cancer specimens and 112 of the 138 non-breast cancer specimens were successfully genotyped, resulting in a final total of 6577 participants with rs2016347 genotyping. Among women with preeclampsia, these included 163 for all breast cancers, 126 for HR+ breast cancers, and 154 for non-breast cancers. The non-breast cancers consisted of 25 cases each of uterine cancer and melanoma, 24 cases of thyroid cancer, 15 of colon cancer, 14 with ovarian cancer, with all other cancer types contributing less than 10 cases each. Genotyping of rs2016347 among study participants was in Hardy Weinberg Equilibrium ($P = 0.21$), with an overall T/G allele frequency of 0.52/0.48.
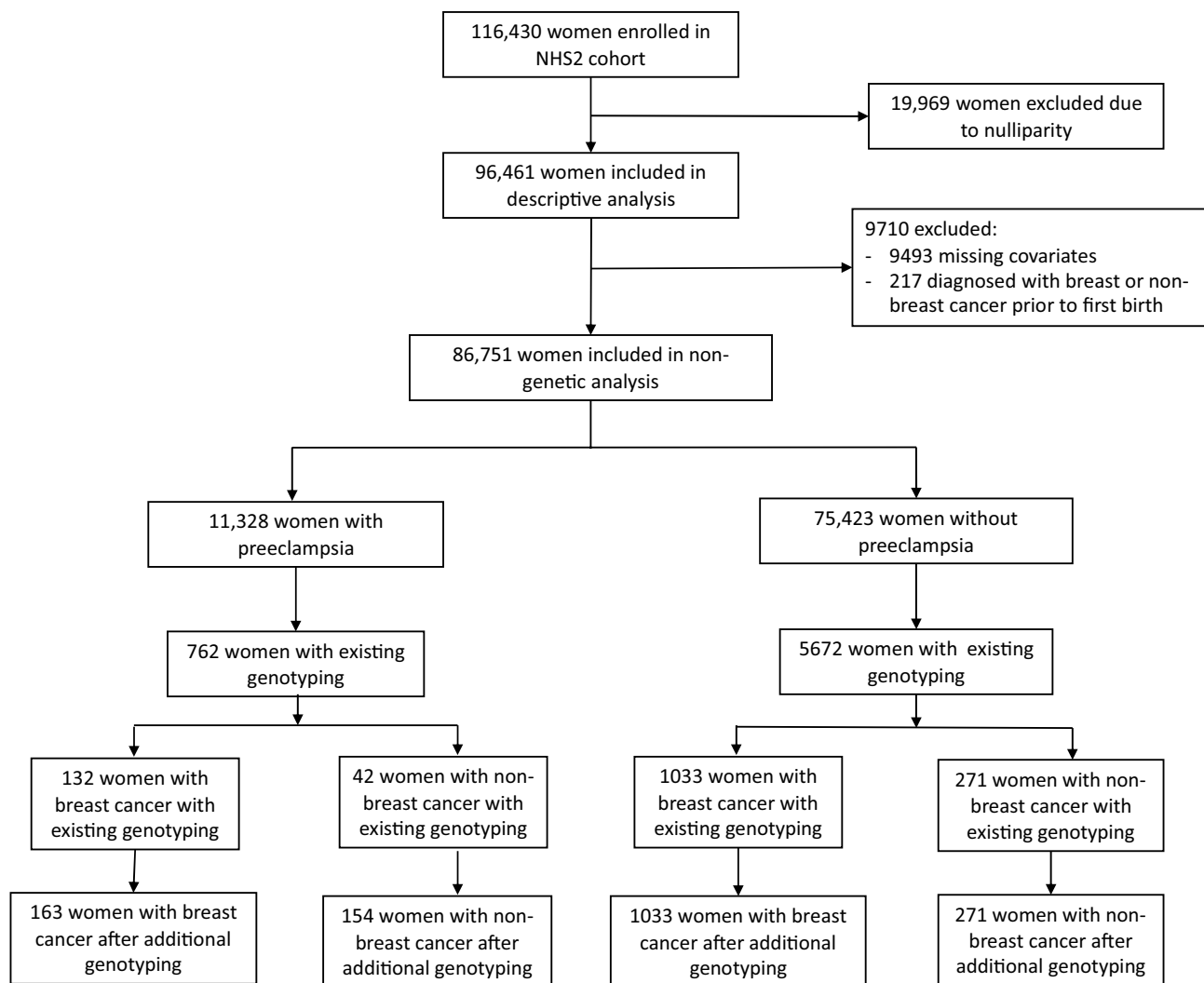
**Fig. 1** Study flow chart

## Genetic model selection

An a priori decision was made to focus on the dominant model in these analyses based on results of the impact of the rs2016347 genotype on previously studied outcomes, and with a desire to maximize statistical power while limiting potential multiple testing issues. The results of earlier published studies of the associations of rs2016347 with similar disease outcomes have demonstrated that the GT genotype tracks most closely with the TT genotype suggesting that only one T allele is necessary for risk modification [18, 37, 39, 40]. As our current analyses and hypothesis is focused on the potential "protective effect" of carrying a T allele, our results report the risk for those participants carrying at least one T allele (TT and GT combined) compared to the reference group of GG carriers.

## Statistical analysis

Hazard ratios (HR) and 95% confidence intervals (CI) for the association of preeclampsia with cancer outcomes were estimated from Cox proportional hazards models using person-years of time as the time scale. Person-years began accumulating for each of the years the participant remained in the study from the year of their first birth, and stopped accumulating upon the year of the final questionnaire (2015), year of their cancer diagnosis, loss to follow-up, or death, whichever occurred first. Participants that had a breast or non-breast cancer diagnosis after their first pregnancy but before study entry were included in the study as dates of cancer diagnosis were deemed adequately reliable to assure proper temporality. All participants with a breast or non-breast cancer diagnosis prior to preeclampsia history were excluded from

the study. Breast cancer models are adjusted for age, ethnicity, BMI, parity, age at first birth, smoking history, family history of breast cancer, physical activity and diet (Healthy Eating Index); non-breast cancers for age, ethnicity, BMI, smoking, and diet. Data for alcohol intake was inadequate for utilization as a covariate. Proportional hazards assumptions for the models were evaluated using standard model-based methods [46]. Cox proportional hazard models were also used to estimate HRs and 95% confidence intervals for the association of genotype and cancer outcomes stratified by preeclampsia history. Linear interaction was then evaluated comparing the risk of each outcome for GG to GT/TT genotype of rs2016347 among those with preeclampsia history. The genetic models were adjusted for the same covariates used in the non-genetic models. Analyses were conducted using R version 4.2.0 "Vigorous Calisthenics", and "survival" and "stats" packages were used to estimate hazard ratios and 95% confidence intervals [47–49]. Visualizations were created using "ggplot2", and "gtummary" was used to create tables and figures [50, 51]. All code needed to reproduce the analyses is available on Github at https://github.com/sfuller2/tox-analysis.

## Results

### Participant characteristics

Characteristics of study participants by preeclampsia exposure status are presented in Table 1. Women with preeclampsia were slightly less likely to be White (White non-Hispanic), 93.1% vs 93.7%, had a lower age at menarche, and were slightly more likely to have smoked. They had a younger mean age at first birth of 26.0 vs 26.6 years, a higher BMI of 26.2 vs 23.6, a lower level of physical activity of 20.8 METS per week vs 22.7, and ate a less healthy diet than participants without preeclampsia. Both obesity and non-White race are well-established risk factors for preeclampsia [52].
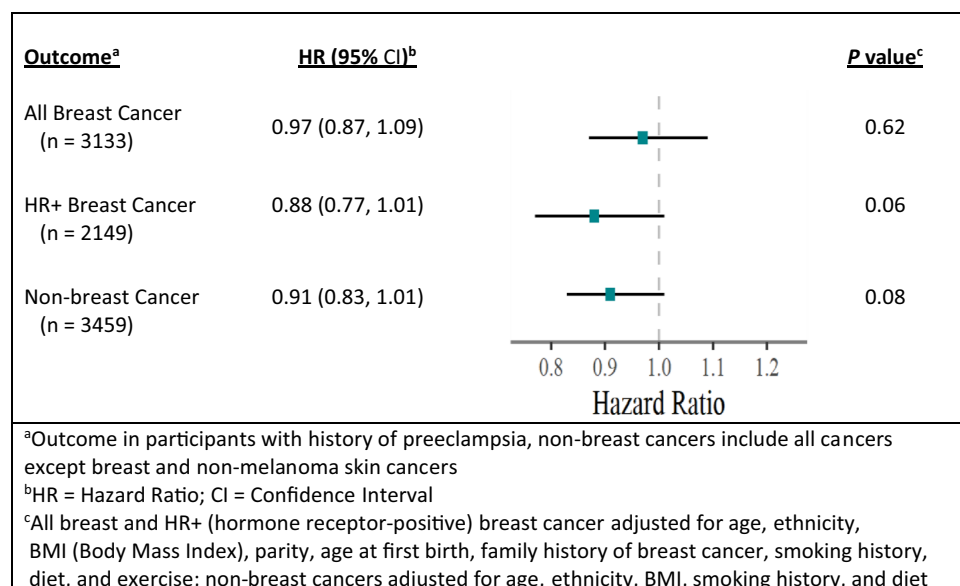
Of participants in the analyses, 3133 (3.6%) had a history of breast cancer, 2149 (2.5%) had a history of HR+ breast cancer, and 3459 (4.0%) had a history of a non-breast cancer.

Comparisons of the characteristics of participants with all breast cancers, HR+ breast cancer only, and non-breast cancers compared to participants without are provided in Supplemental Tables S1–S3, respectively. Participants with all breast or HR+ breast cancer were older at entry, had lower parity, higher age at first birth, stronger family history of breast cancer, increased smoking history, decreased physical activity, and a slightly better diet than those without breast cancer. Participants with non-breast cancers were older, more likely to be White, have a smoking history, a higher BMI, and did less physical activity than participants without.

### Risks for breast and non-breast cancers in women with preeclampsia

Women with a history of preeclampsia had no increase in risk of developing breast cancer when not considering genotype, HR 0.97; 95% CI, 0.87–1.09; $P = 0.62$. When looking at the subgroup of HR+ breast cancers, the risk decreased somewhat, HR 0.88; 95% CI, 0.77–1.01, $P = 0.06$. Assessment of the risk of developing a non-breast cancer revealed an HR 0.91; 95% CI, 0.83–1.01, $P = 0.08$. These results are summarized in Fig. 2.

**Fig. 2** Outcome risk for NHS2 participants with preeclampsia



| Outcome[a] | HR (95% CI)[b] | | P value[c] |
|---|---|---|---|
| All Breast Cancer (n = 3133) | 0.97 (0.87, 1.09) | | 0.62 |
| HR+ Breast Cancer (n = 2149) | 0.88 (0.77, 1.01) | | 0.06 |
| Non-breast Cancer (n = 3459) | 0.91 (0.83, 1.01) | | 0.08 |

[a]Outcome in participants with history of preeclampsia, non-breast cancers include all cancers except breast and non-melanoma skin cancers
[b]HR = Hazard Ratio; CI = Confidence Interval
[c]All breast and HR+ (hormone receptor-positive) breast cancer adjusted for age, ethnicity, BMI (Body Mass Index), parity, age at first birth, family history of breast cancer, smoking history, diet, and exercise; non-breast cancers adjusted for age, ethnicity, BMI, smoking history, and diet

## Association of IGF1R rs2016347 genotype with study outcomes

The association of rs2016347 genotypes with the outcomes of breast, HR+ breast, and non-breast cancers did not reveal any significant association in participants without preeclampsia, as noted in Table 2. Similarly, risk for the dominant genetic model revealed no significant associations (see Fig. 3).

## Impact of rs2016347 genotype on risk of cancer by preeclampsia status

There was no statistically significant change in risk for cancer in women with a history of preeclampsia when not considering *IGF1R* SNP rs20162347 genotype, although the impact on HR+ breast cancer was suggestive. Analyses stratified by preeclampsia history do demonstrate that genotype may in fact further influence risk. When looking at the risk for all breast cancers, carrying at least one T allele resulted in an HR 0.87; 95% CI, 0.62–1.24, $P = 0.45$. However, when looking at HR+ breast cancers only, carrying a T allele was protective, HR 0.67, 95% CI, 0.46–0.97, $P = 0.04$, with an interaction term for Preeclampsia*T allele of HR 0.68, 95% CI, 0.45–1.02, $P = 0.06$. Efforts to quantify the relative excess risk due to interaction (RERI) required removal of the additional targeted genotyping (31 of the 126 cases of HR+ breast cancer), and was suggestive of substantial interaction with an additive RERI of −0.46; 95% CI, −1.13–0.22, and multiplicative RERI of 0.69; 95% CI, 0.43–1.10. There were too few adequately documented HR− genotyped participants for meaningful analysis.

When looking at the risk of developing a non-breast cancer in women with preeclampsia, carrying at least one T allele resulted in an HR 0.76, 95% CI, 0.53–1.08, $P = 0.12$, with an interaction term for Preeclampsia*T allele HR 0.77,

**Table 2** Association of rs2016347 genotype with outcomes among participants without preeclampsia

| *IGF1R* SNP rs2016347 genotype[a] | | | | |
|---|---|---|---|---|
| Outcome[b] | GG | GT | TT | *P*-value[c] |
| Breast cancer (*N* = 1033) | 24.5% | 49.0% | 26.5% | 0.69 |
| HR + Breast cancer (*N* = 793) | 23.7% | 49.3% | 27.0% | 0.92 |
| Non-breast cancer (*N* = 271) | 23.6% | 46.9% | 29.5% | 0.78 |
| No cancer (*N* = 4368) | 23.5% | 48.8% | 27.7% | ----- |

[a]Values represent percent of each genotype with specified outcome

[b]N is the total number of participants with genotyping for each outcome

[c]P-values are for $\chi^2$ goodness-of-fit test using distribution of genotypes for no cancer as reference group

95% CI, 0.49–1.21, $P = 0.26$. Results of the impact of rs2016347 for the dominant model are provided in Fig. 3.

Risks for all breast cancer, HR+ breast cancer only, and non-breast cancer demonstrated no impact from carrying a T allele in women without a history of preeclampsia, with HR 0.95, 95% CI, 0.82–1.10, $P = 0.48$, HR 0.99, 95% CI, 0.84–1.17, $P = 0.92$, and HR 1.03, 95% CI, 0.78–1.36, $P = 0.86$, respectively, as noted in Fig. 3. Of note, participants with genotyping did not differ in their risk of breast and non-breast cancer from the entire study population.

## Discussion

Results of this NHS2 study of predominantly White (> 90%) women demonstrate a 33% lower risk of developing HR+ breast cancer in those with a self-reported history of preeclampsia who also inherited the protective *IGF1R* T allele of rs2016347. Without considering inheritance of this functional *IGF1R* SNP, these women exposed to preeclampsia, much like those in the CTS initially evaluated without *IGF1R* genotyping, would otherwise appear to show no significant reduction in their later life breast cancer incidence (Fig. 2). Likewise, reviews of multi-national cohort studies performed before awareness of this *IGF1R* SNP contribution concluded that the general breast cancer risk-reducing impact of preeclampsia is modest (10–20%), if present at all [10]. Even today, for understudied populations like Black women whose age-adjusted prevalence of preeclampsia can exceed 12%, the question of preeclampsia's potential protective effect on later life breast cancer incidence remains incomplete, as these studies have not accounted for *IGF1R* rs2016347 genotype differences [1].

Our current NHS2 findings, as well as those from the prior CTS study, now strongly suggest that the degree of breast cancer protection from preeclampsia depends on a mechanistic interaction between two population exposures: preeclampsia and inheritance of the functionally blunted *IGF1R* T allele of rs2016347 that causes a significant reduction in IGF1R mRNA expression observable in multiple normal organ and host tissues [38]. Moreover, both NHS2 and CTS studies also demonstrate that this protective mechanistic interaction appears strongest against HR+ breast cancer, producing a 30–40% reduction in this most common subtype of breast cancer. Furthermore, these concordant NHS2 and CTS findings are consistent with numerous preclinical and clinical studies examining the role of the IGF-1 axis in driving HR+ breast cancers [27–30]. At a mechanistic level, it is now well-established that tumorigenic crosstalk occurs between the estrogen receptor (ER) axis and epithelial signal transduction downstream of activated IGF1R, and that this crosstalk plays a role in both HR+ breast cancer development and progression [52–55]. Our findings

**Fig. 3** Outcome risk for NHS2 participants carrying at least one T allele of *IGF1R* SNP rs2016347 by preeclampsia history



| Outcome[a] | HR (95% CI)[b] | | P value[c] | P value for interaction[d] |
|---|---|---|---|---|
| **All Breast Cancer** | | | | 0.67 |
| Preeclampsia + (N = 163) | 0.87 (0.62, 1.24) | | 0.45 | |
| Preeclampsia − (N = 1033) | 0.95 (0.82, 1.10) | | 0.48 | |
| **HR+ Breast Cancer only** | | | | 0.06 |
| Preeclampsia + (N= 126) | 0.67 (0.46, 0.97) | | 0.04 | |
| Preeclampsia - (N = 793) | 0.99 (0.84, 1.17) | | 0.92 | |
| **Non-breast Cancer** | | | | 0.26 |
| Preeclampsia + (N = 154) | 0.76 (0.53, 1.08) | | 0.12 | |
| Preeclampsia - (N = 271) | 1.03 (0.78, 1.36) | | 0.86 | |

[a]Outcome by preeclampsia history, HR+: hormone receptor-positive, non-breast cancer includes all cancers except breast and non-melanoma skin cancers
[b]HR: Hazard Ratio; CI: Confidence Interval
[c]P value is for dominant genetic model, TT and GT genotype compared to GG reference group. All breast and HR+ (hormone receptor-positive) breast cancer adjusted for age, ethnicity, parity, age at first birth, BMI (Body Mass Index), family history of breast cancer, smoking history, diet, and exercise; non-breast cancers adjusted for age, ethnicity, BMI, smoking history, and diet
[d]P value for interaction for preeclampsia*TT/GT genotype

are also consistent with those of Winder who found better survival outcomes for women with HR+ breast cancers treated with the anti-estrogen tamoxifen who carry the T allele of rs2016347 [40], as well as with those of Bhargava who found that while IGF1R overexpression is most common in HR+ relative to HR− breast tumors, patients with HR+ breast cancers expressing lower levels of IGF1R have significantly better survival outcomes following ER-targeted endocrine therapy [56].

Many population studies are unable to discern HR+ from HR− breast cancer outcomes, and this likely explains why a recent UK Biobank cohort analysis which lacked data on HR status was unable to show a significant lowering of overall breast cancer risk in women with HDP carrying the rs2016347 T allele [38]. However, that same UK Biobank study did detect a statistically significant 41% reduction in the development of non-breast cancers as a pooled group with > 40% comprising either colorectal, uterine or ovarian cancers; this observation strongly suggests that the cancer protecting mechanistic interaction between preeclampsia and T allele inheritance may extend to endocrine-independent tumor types beyond HR+ breast cancer. Our current NHS2 study suggests a 24% reduction in development of another pooled group of non-breast

cancers (Fig. 3), in which 35% consisted of colorectal, uterine, or ovarian cancers. In summary, beyond our conclusive findings about protection against HR+ breast cancers, our observations about non-breast cancer outcomes suggest that future studies be performed to quantitate the non-breast cancer protective interactions between preeclampsia and *IGF1R* rs2016347 T allele inheritance in populations statistically powered to assess ovarian, colorectal and uterine cancer development.

A limitation of this study may have been that preeclampsia exposure was determined by self-report, yet studies of maternal recall of preeclampsia have found that misclassification is minimal; and a validation study of preeclampsia in the NHS2 demonstrated a positive predictive value of 89% [57–59]. As well, any misclassification resulting from recall bias would more likely be non-discriminatory due to the lack of awareness of any causal association between preeclampsia and cancer, and thus would be inclined to bias the results toward the null. Another limitation was that data linking preeclampsia to a specific pregnancy was sometimes lacking, and as a result, date of first pregnancy was used in these cases as preeclampsia is known to be much more common in the first pregnancy [60]. Low numbers did not permit analysis of HR− breast cancers as a group, and data also did not allow analysis by subtypes of preeclampsia such as early/late, mild/severe, or preeclampsia superimposed on chronic hypertension. Genotyping for this study was not random, and although there was no selection based on preeclampsia and all breast cancer cases were genotyped, this could have introduced ascertainment bias which may have impacted the results. However, the genotype has not previously been associated with the outcome of any of the case control studies.

When comparing genotype-preeclampsia interactions on breast cancer risk between NHS2, UK Biobank and CTS cohort studies, it is important to note that rs2016347 allele frequencies are very similar in each of the 3 cohorts. However, differences in their population risk factors may play a role in their varied outcome risks. All three cohorts are largely White and match in most characteristics except age, with participants in the NHS2 being 20 years younger at study entry and thus potentially more likely to have developed premenopausal breast cancer [3, 18]. Likewise, this age difference accounts for differences in the number and type of tumors comprising our non-breast cancer group. In particular, the younger age of the NHS2 cohort resulted in a larger percentage of thyroid cancer and melanoma cases, and a lower percentage of colorectal cancer cases. In addition, the targeted genotyping of a small number of participants with ovarian cancer may also have contributed to the different HRs noted between the non-breast cancer groups for NHS2 (HR 0.76) and UK Biobank (HR 0.59) cohorts.

## Conclusion

This retrospective NHS2 cohort study in largely White non-Hispanic women validates earlier CTS findings and confirms that women with preeclampsia who have inherited the common T allele of *IGF1R* rs2016347 have a 30–40% lower risk of developing HR+ breast cancer, pointing to the need to consider incorporating this exposure information (preeclampsia history and *IGF1R* rs2016347 genotyping) into risk assessment models and personalized breast cancer screening recommendations. Importantly, this breast cancer protective interaction has yet to be validated in non-White women who have the highest exposure rates to preeclampsia and may be more likely to develop premenopausal HR− breast cancer. Of interest, this NHS2 study also observed a nominal reduction in non-breast cancer incidence, supporting the statistically significant UK Biobank cohort findings of reduced non-breast cancer development in women with both exposures. Taken together, these cohort studies strongly implicate a mechanistic role of the IGF-1 axis in cancer development, supporting targeted prevention strategies designed to blunt IGF-1 axis expression.

**Data availability** The datasets utilized for this study and the code used for the analyses are available from the corresponding author upon reasonable request.

## Declarations

**Conflict of interest** The authors declare that they have no known potential financial conflicts of interest or personal relationships that could have appeared to influence the research or publication of this article.

**Informed consent** The NHS2 study protocol was approved by the Institutional Review Board (IRB) of the Brigham and Women's Hospital, and the IRB allowed participants' completion of questionnaires as implied consent.

# References

1. Boakye E, Kwapong YA, Obisesan O et al (2021) Nativity-related disparities in preeclampsia and cardiovascular disease risk among a racially diverse cohort of US women. JAMA Netw Open 4(12):e2139564. https://doi.org/10.1001/jamanetworkopen.2021.39564

2. Bigham Z, Robles Y, Freund KM, Palmer JR, Bertrand KA (2022) Hypertensive diseases of pregnancy and risk of breast cancer in the Black women's health study. Breast Cancer Res Treat 194(1):127–135. https://doi.org/10.1007/s10549-022-06606-3

3. Powell MJ, Von Behren J, Neuhausen S, Reynolds P, Benz CC (2017) Functional IGF1R variant predicts breast cancer risk in women with preeclampsia in California teachers study. Cancer Causes Control 28(10):1027–1032. https://doi.org/10.1007/s10552-017-0942-7

4. Powell M, Benz C, Ereman R, Prebil LA, Jamshidian F, Rose S. (2013) Abstract P3–07–03: insulin-like growth factor-1 receptor variant associated with decreased breast cancer risk in women with pregnancy-induced hypertension. Abstracts: thirty-sixth annual CTRC-AACR san antonio breast cancer symposium - Dec 10- 14, 2013; San Antonio, TX. Cancer Res 73(24suppl).

5. Rana S, Lemoine E, Granger JP, Karumanchi SA (2019) Preeclampsia: pathophysiology, challenges, and perspectives [published correction appears in circ res. 2020 Jan 3,126(1):e8]. Circ Res 124(7):1094–1112. https://doi.org/10.1161/CIRCRESAHA.118.313276

6. Abalos E, Cuesta C, Grosso AL, Chou D, Say L (2013) Global and regional estimates of preeclampsia and eclampsia: a systematic review. Eur J Obstet Gynecol Reprod Biol 170(1):1–7. https://doi.org/10.1016/j.ejogrb.2013.05.005

7. Wu P, Haththotuwa R, Kwok CS et al (2017) Preeclampsia and future cardiovascular health: a systematic review and meta-analysis. Circ Cardiovasc Qual Outcomes 10(2):e003497. https://doi.org/10.1161/CIRCOUTCOMES.116.003497

8. Benschop L, Duvekot JJ, Roeters van Lennep JE (2019) Future risk of cardiovascular disease risk factors and events in women after a hypertensive disorder of pregnancy. Heart 105(16):1273–1278. https://doi.org/10.1136/heartjnl-2018-313453

9. Vatten LJ, Romundstad PR, Trichopoulos D, Skjaerven R (2002) Preeclampsia in pregnancy and subsequent risk for breast cancer. Br J Cancer 87(9):971–973. https://doi.org/10.1038/sj.bjc.6600581

10. Kim JS, Kang EJ, Woo OH et al (2013) The relationship between preeclampsia, pregnancy-induced hypertension and maternal risk of breast cancer: a meta-analysis. Acta Oncol 52(8):1643–1648. https://doi.org/10.3109/0284186X.2012.750033

11. Opdahl S, Romundstad PR, Alsaker MD, Vatten LJ (2012) Hypertensive diseases in pregnancy and breast cancer risk. Br J Cancer 107(1):176–182. https://doi.org/10.1038/bjc.2012.195

12. Terry MB, Perrin M, Salafia CM et al (2007) Preeclampsia, pregnancy-related hypertension, and breast cancer risk. Am J Epidemiol 165(9):1007–1014. https://doi.org/10.1093/aje/kwk105

13. Sun M, Fan Y, Hou Y, Fan Y (2018) Preeclampsia and maternal risk of breast cancer: a meta-analysis of cohort studies. J Matern Fetal Neonatal Med 31(18):2484–2491. https://doi.org/10.1080/14767058.2017.1342806

14. Bhattacharya S, Prescott GJ, Iversen L, Campbell DM, Smith WC, Hannaford PC (2012) Hypertensive disorders of pregnancy and future health and mortality: a record linkage study. Pregnancy Hypertens 2(1):1–7. https://doi.org/10.1016/j.preghy.2011.08.116

15. Wang F, Zhang W, Cheng W, Huo N, Zhang S (2021) Preeclampsia and cancer risk in women in later life: a systematic review and meta-analysis of cohort studies. Menopause 28(9):1070–1078. https://doi.org/10.1097/GME.0000000000001806

16. Serrand C, Mura T, Fabbro-Peray P et al (2021) Assessment of all-cause cancer incidence among individuals with preeclampsia or eclampsia during first pregnancy. JAMA Netw Open 4(6):e2114486. https://doi.org/10.1001/jamanetworkopen.2021.14486

17. Behrens I, Basit S, Jensen A et al (2016) Hypertensive disorders of pregnancy and subsequent risk of solid cancer–a nationwide cohort study. Int J Cancer 139(1):58–64. https://doi.org/10.1002/ijc.30065

18. Powell MJ, Dufault SM, Gunderson EP, Benz CC (2020) Cancer and cardiovascular risk in women with hypertensive disorders of pregnancy carrying a common IGF1R variant. Mayo Clin Proc 95(12):2684–2696. https://doi.org/10.1016/j.mayocp.2020.03.037

19. Kappou D, Vrachnis N, Sifakis S (2012) Recent insights into the role of the insulin-like growth factor axis in preeclampsia. In: Sifakis S (ed) From Preconception to Postpartum. InTech, pp 147–160 (**ISBN: 978- 953-51-0353-0**)

20. Peng HY, Xue M, Xia AB (2011) Study on changes of IGF-I and leptin levels in serum and placental tissue of preeclampsia patients and their associativity. Xi Bao Yu Fen Zi Mian Yi Xue Za Zhi 27(2):192–194

21. Olmos A, Díaz L, Avila E et al (2013) Associations between insulin-like growth factor I, vascular endothelial growth factor and its soluble receptor 1 in umbilical serum and endothelial cells obtained from normotensive and preeclamptic pregnancies. Growth Factors 31(4):123–129. https://doi.org/10.3109/08977194.2013.802692

22. Dubova EA, Pavlov KA, Lyapin VM, Kulikova GV, Shchyogolev AI, Sukhikh GT (2014) Expression of insulin-like growth factors in the placenta in preeclampsia. Bull Exp Biol Med 157(1):103–107. https://doi.org/10.1007/s10517-014-2502-4

23. Kharb S, Nanda S (2017) Patterns of biomarkers in cord blood during pregnancy and preeclampsia. Curr Hypertens Rev 13(1):57–64. https://doi.org/10.2174/1573402113666170126101914

24. Altinkaynak K, Hü AH, Bakan E, Kumtepe Y (2003) Serum IGF-I and IGFBP-3 in healthy pregnancies and patients with preeclampsia. Clin Biochem 36(3):221–223. https://doi.org/10.1016/s0009-9120(03)00008-0

25. Liao S, Vickers MH, Taylor RS et al (2017) Maternal serum IGF-1, IGFBP-1 and 3, and placental growth hormone at 20weeks' gestation in pregnancies complicated by preeclampsia. Pregnancy Hypertens 10:149–154. https://doi.org/10.1016/j.preghy.2017.07.148

26. Wu X, Zhang X, Hao Y, Li J (2021) Obesity-related protein biomarkers for predicting breast cancer risk: an overview of systematic reviews. Breast Cancer 28(1):25–39. https://doi.org/10.1007/s12282-020-01182-0

27. Christopoulos PF, Msaouel P, Koutsilieris M (2015) The role of the insulin-like growth factor-1 system in breast cancer. Mol Cancer 14:43. https://doi.org/10.1186/s12943-015-0291-7

28. Key TJ, Appleby PN, Reeves GK, Roddam AW, Endogenous Hormones and Breast Cancer Collaborative Group (2010) Insulin-like growth factor 1 (IGF1), IGF binding protein 3 (IGFBP3), and breast cancer risk: pooled individual data analysis of 17 prospective studies. Lancet Oncol 11(6):530–542. https://doi.org/10.1016/S1470-2045(10)70095-4

29. Tamimi RM, Colditz GA, Wang Y et al (2011) Expression of IGF1R in normal breast tissue and subsequent risk of breast cancer. Breast Cancer Res Treat 128(1):243–250. https://doi.org/10.1007/s10549-010-1313-1

30. Monson KR, Goldberg M, Wu HC, Santella RM, Chung WK, Terry MB (2020) Circulating growth factor concentrations and breast cancer risk: a nested case-control study of IGF-1, IGFBP-3, and breast cancer in a family-based cohort. Breast Cancer Res 22(1):109. https://doi.org/10.1186/s13058-020-01352-0

31. Friedrich RE, Hagel C, Bartel-Friedrich S (2010) Insulin-like growth factor-1 receptor (IGF-1R) in primary and metastatic undifferentiated carcinoma of the head and neck: a possible target of immunotherapy. Anticancer Res 30(5):1641–1643

32. Fu S, Tang H, Liao Y et al (2016) Expression and clinical significance of insulin-like growth factor 1 in lung cancer tissues and perioperative circulation from patients with non-small-cell lung cancer. Curr Oncol 23(1):12–19. https://doi.org/10.3747/co.23.2669

33. Shiratsuchi I, Akagi Y, Kawahara A et al (2011) Expression of IGF-1 and IGF-1R and their relation to clinicopathological factors in colorectal cancer. Anticancer Res 31(7):2541–2545

34. Nurwidya F, Andarini S, Takahashi F, Syahruddin E, Takahashi K (2016) Implications of insulin-like growth factor 1 receptor activation in lung cancer. Malays J Med Sci 23(3):9–21

35. Ahearn TU, Peisch S, Pettersson A et al (2018) Expression of IGF/insulin receptor in prostate cancer tissue and progression to lethal disease. Carcinogenesis 39(12):1431–1437. https://doi.org/10.1093/carcin/bgy112

36. Powell MJ, Dufault SM, Henry JE, Allison AC, Cora R, Benz CC (2019) Pregnancy hypertension and a commonly Inherited IGF1R variant (rs2016347) reduce breast cancer risk by enhancing mammary gland involution. J Oncol 2019:6018432. https://doi.org/10.1155/2019/6018432

37. Prebil LA, Ereman RR, Powell MJ et al (2014) First pregnancy events and future breast density: modification by age at first pregnancy and specific VEGF and IGF1R gene variants. Cancer Causes Control 25(7):859–868. https://doi.org/10.1007/s10552-014-0386-2

38. GTEx portal. Variant page. (2022) The Broad Institute of MIT and Harvard. https://www.gtexportal.org/home/snp/rs2016347. Last accessed April 10, 2022

39. de Groot S, Charehbili A, van Laarhoven HW et al (2016) Insulin-like growth factor 1 receptor expression and IGF1R 3129G > T polymorphism are associated with response to neoadjuvant chemotherapy in breast cancer patients: results from the NEO-ZOTAC trial (BOOG 2010–01). Breast Cancer Res 18(1):3. https://doi.org/10.1186/s13058-015-0663-3

40. Winder T, Giamas G, Wilson PM et al (2014) Insulin-like growth factor receptor polymorphism defines clinical outcome in estrogen receptor-positive breast cancer patients treated with tamoxifen. Pharmacogenomics J 14(1):28–34. https://doi.org/10.1038/tpj.2013.8

41. Bao Y, Bertoia ML, Lenart EB et al (2016) Origin, methods, and evolution of the three nurses' health studies. Am J Pub Health 106(9):1573–1581. https://doi.org/10.2105/AJPH.2016.303338

42. Colditz GA, Philpott SE, Hankinson SE (2016) The impact of the nurses' health study on population health: prevention, translation, and control. Am J Pub Health 106(9):1540–1545. https://doi.org/10.2105/AJPH.2016.303343

43. Chavarro JE, Rich-Edwards JW, Gaskins AJ et al (2016) Contributions of the nurses' health studies to reproductive health research. Am J Public Health 106(9):1669–1676. https://doi.org/10.2105/AJPH.2016.303350

44. Lindström S, Loomis S, Turman C et al (2017) A comprehensive survey of genetic variation in 20,691 subjects from four large cohorts. PLoS ONE 12(3):e0173997. https://doi.org/10.1371/journal.pone.0173997

45. Powell MJ, Fuller S, Gunderson EP, Benz CC (2022) Reduced cardiovascular risks in women with endometriosis or polycystic ovary syndrome carrying a common functional IGF1R variant. Hum Reprod 37(5):1083–1094. https://doi.org/10.1093/humrep/deac059

46. Austin PC (2018) Statistical power to detect violation of the proportional hazards assumption when using the cox regression model. J Stat Comput Simul 88(3):533–552. https://doi.org/10.1080/00949655.2017.1397151

47. R Core Team (2022). R: A language and environment for statistical computing. Foundation for Statistical Computing, Vienna, Austria. URL: http://www.r-project.org/index.html

48. R Project for Statistical Computing. (2020) The R foundation. A Package for Survival Analysis in S. Version 2.38. https://www.R-project.org/. Accessed June 6, 2020

49. Therneau T, Grambsch P (2000) Modeling survival data: extending the cox model. Springer-Verlag, New York, New York, NY

50. Wickham H (2016) Ggplot2: elegant graphics for data analysis. Springer International Publishing, New York, NY

51. Sjoberg D, Whiting K, Curry M, Lavery J, Larmarange J (2021) Reproducible summary tables with the gtsummary package. The R J 13:570–580. https://doi.org/10.32614/RJ-2021-053

52. Lanzino M, Morelli C, Garofalo C et al (2008) Interaction between estrogen receptor alpha and insulin/IGF signaling in breast cancer. Curr Cancer Drug Targ 8(7):597–610. https://doi.org/10.2174/156800908786241104

53. Hawsawi Y, El-Gendy R, Twelves C, Speirs V, Beattie J (1836) (2013) Insulin-like growth factor—oestradiol crosstalk and mammary gland tumourigenesis. Biochim Biophys Acta 2:345–353. https://doi.org/10.1016/j.bbcan.2013.10.005

54. Iida M, Tsuboi K, Niwa T, Ishida T, Hayashi SI (2019) Compensatory role of insulin-like growth factor 1 receptor in estrogen receptor signaling pathway and possible therapeutic target for hormone therapy-resistant breast cancer. Breast Cancer 26(3):272–281. https://doi.org/10.1007/s12282-018-0922-0

55. Engels CC, de Glas NA, Sajet A et al (2016) The influence of insulin-like growth factor-1-receptor expression and endocrine treatment on clinical outcome of postmenopausal hormone receptor positive breast cancer patients: a Dutch team substudy analysis. Mol Oncol 10(4):509–516. https://doi.org/10.1016/j.molonc.2015.10.010

56. Bhargava R, Beriwal S, McManus K, Dabbs DJ (2011) Insulin-like growth factor receptor-1 (IGF-1R) expression in normal breast, proliferative breast lesions, and breast carcinoma. Appl Immunohistochem Mol Morphol 19(3):218–225. https://doi.org/10.1097/PAI.0b013e3181ffc58c

57. Stuart JJ, Tanz LJ, Missmer SA et al (2018) Hypertensive disorders of pregnancy and maternal cardiovascular disease risk factor development: an observational cohort study. Ann Intern Med 169(4):224–232. https://doi.org/10.7326/M17-2740

58. Stuart JJ, Bairey Merz CN, Berga SL et al (2013) Maternal recall of hypertensive disorders in pregnancy: a systematic review. J Womens Health (Larchmt) 22(1):37–47. https://doi.org/10.1089/jwh.2012.3740

59. Carter EB, Stuart JJ, Farland LV et al (2015) Pregnancy complications as markers for subsequent maternal cardiovascular disease: validation of a maternal recall questionnaire. J Womens Health (Larchmt) 24(9):702–712. https://doi.org/10.1089/jwh.2014.4953

60. Burton GJ, Redman CW, Roberts JM, Moffett A (2019) Pre-eclampsia: pathophysiology and clinical implications. BMJ 366:l2381. https://doi.org/10.1136/bmj.l2381