Preeclampsia and cardiovascular disease in a large UK pregnancy cohort of linked electronic health records: a CALIBER study

Leon, Preeclampsia and cardiovascular disease

Lydia J Leon[1,2], PhD[#]; Fergus P McCarthy[1,3], PhD[#]; Kenan Direk[2], PhD; Arturo Gonzalez-Izquierdo[2], PhD; David Prieto-Merino[2,4], PhD; Juan P Casas[5], PhD*; Lucy Chappell[1], PhD*

1. Department of Women and Children's Health, King's College London, London, UK.

2. Institute of Health Informatics, University College London, London, UK.

3. The Irish Centre for Fetal and Neonatal Translational Research, University College Cork, Cork University Maternity Hospital, Cork, Ireland.

4. Applied Statistical Methods in Medical Research Group, Universidad Catolica San Antonio de Murcia, Murcia, Spain

5. Massachusetts Veterans Epidemiology Research and Information Center (MAVERIC), VA Boston Healthcare, MA, USA.

[#] These authors contributed jointly and are joint first authors

*These authors contributed jointly and are joint senior authors

Corresponding author: Dr Fergus McCarthy, The Irish Centre for Fetal and Neonatal Translational Research, University College Cork, Cork University Maternity Hospital, Wilton, Cork, Ireland. Email; Fergus.mccarthy@ucc.ie Tel: +353 212920609 Fax: +353 212920609

ABSTRACT

Background

The associations between pregnancy hypertensive disorders and common cardiovascular disorders have not been investigated at scale in a contemporaneous population. We aimed to investigate the association between preeclampsia, hypertensive disorders of pregnancy and subsequent diagnosis of 12 different cardiovascular disorders.

Methods

We used linked electronic health records from 1997 to 2016 to recreate a UK population-based cohort of 1.3 million women, mean age at delivery 28 years, with nearly 1.9 million completed pregnancies. We used multivariable Cox models to determine the associations between hypertensive disorders of pregnancy (HDP), and preeclampsia alone (term and preterm), with 12 cardiovascular disorders in addition to chronic hypertension. We estimated the cumulative incidence of a composite endpoint of any cardiovascular disorder according to preeclampsia exposure.

Results

During the 20 year study period, 18,624 incident cardiovascular disorders were observed, 65% of which had occurred in women under 40 years. Compared to women without hypertension in pregnancy, women who had one or more pregnancies affected by preeclampsia had a hazard ratio (HR) of 1.9 (95%CI 1.53-2.35) for any stroke, 1.67 (1.54-1.81) for cardiac atherosclerotic events, 1.82 (1.34-2.46) for peripheral events, 2.13 (1.64-2.76) for heart failure, 1.73 (1.38-2.16) for atrial fibrillation, 2.12 (1.49-2.99) for cardiovascular deaths and 4.47 (4.32-4.62) for chronic hypertension. Differences in cumulative incidence curves,

according to preeclampsia status, were apparent within one year of the first index pregnancy. Similar patterns of association were observed for HDP, whilst preterm preeclampsia conferred slightly further elevated risks.

Conclusions

Hypertensive disorders of pregnancy, including preeclampsia, have a similar pattern of increased risk across all 12 cardiovascular disorders and chronic hypertension, and the impact was evident soon after pregnancy. HDP should be considered as a natural screening tool for cardiovascular events, enabling cardiovascular risk prevention through national initiatives.

Keywords

Preeclampsia, hypertensive disorders of pregnancy, pregnancy, cardiovascular disease, hypertension.

Non-standard Abbreviations and Acronyms

CALIBER; Cardiovascular research using Linked Bespoke studies and Electronic health Records

HDP; hypertensive disorders of pregnancy

EHR; electronic health records

CPRD; Clinical Practice Research Datalink

HES; Hospital Episodes Statistics

ONS; Office for National Statistics

NOS; not otherwise specified

HR; Hazard ratios

CLINICAL PERSPECTIVE

What is new?

- Compared to women without preeclampsia, women who had one or more pregnancies affected by preeclampsia had elevated hazard ratios for any stroke, for cardiac atherosclerotic events, for peripheral events, heart failure, atrial fibrillation and cardiovascular deaths.
- Differences in cumulative incidence curves, according to preeclampsia status, were apparent within one year of the first index pregnancy.
- Similar patterns of association were observed for hypertensive disorders of pregnancy, whilst preterm preeclampsia conferred further elevated risks.

What are clinical implications?

- Women with any hypertensive disorders of pregnancy, in particular preterm preeclampsia, are at increased risk of all cardiovascular disorders.
- The age for cardiovascular screening may need to be reduced for women with a history of hypertensive disorders of pregnancy.

- Hypertensive disorders of pregnancy should be considered as a natural screening tool for premature cardiovascular events, enabling cardiovascular risk prevention through national initiatives.

INTRODUCTION

Preeclampsia affects 2-8% of pregnancies worldwide manifesting as hypertension and proteinuria in the second half of pregnancy.[1] Globally, preeclampsia is responsible for around 14% of maternal deaths[2] and is a major cause of perinatal morbidity and mortality.

Two decades of research have documented an association between preeclampsia, and major cardiovascular disorders in later life.[3-8] However, there are several limitations with the current evidence that prevent its translation into clinical care. Firstly, most research to date has focused on the use of composite end-points such as ischemic heart disease and cerebrovascular disorders, which include a heterogenous group of disorders with diverse aetiologies and clinical management. Secondly, over the past two decades the pattern of initial presentation of cardiovascular disorders has changed substantially in high-income countries such as the UK, with most of the events being neither myocardial infarction nor ischemic stroke.[9] Thirdly, only a minority of studies[4, 10, 11] have been able to adjust for post-pregnancy cardiovascular risk factors, such as hypertension, limiting their ability to examine potential mediating factors underlying this association.

To resolve these uncertainties, we used linked electronic health records to create a large-scale, contemporaneous pregnancy cohort of around 1.3 million participants over a twenty year period. We used this population-based cohort to investigate the association of preeclampsia, preterm preeclampsia, and other hypertensive disorders of pregnancy (HDP) with the 12 most common initial diagnoses of cardiovascular disorders in the UK, in addition to chronic hypertension. We also carried out exploratory analyses to investigate the extent to

which the preeclampsia/cardiovascular disease association was explained by post-pregnancy hypertension.

METHODS

Study design and data sources

We recreated a longitudinal population-based cohort study using linked routine electronic health records (EHR). Women were selected from the Cardiovascular research using Linked Bespoke studies and Electronic health Records (CALIBER)[12] resource (https://www.caliberresearch.org/portal) which contains verified and reproducible health phenotypes for hundreds of variables, including the 12 cardiovascular phenotypes investigated as outcomes in this analysis.[12] CALIBER and its associated phenotype algorithms combine longitudinal data from the Clinical Practice Research Datalink (CPRD),[13] Hospital Episodes Statistics (HES),[14] and Office for National Statistics (ONS) cause-specific mortality records.[15] See supplementary methods for further details. The analytic methods are available to other researchers for purposes of reproducing the results or replicating the procedure (in the online-only Data Supplement). We are not authorized to share the data

Participants

Participants were eligible for inclusion in the study if they were female, registered at any General Practitioner participating in the CPRD with 'up-to-standard' (UTS) data, and they had consented to linkage with HES and ONS.

Preeclampsia, the exposure of interest, is a syndrome of the latter half of pregnancy. Therefore, to avoid measurement error in the exposure, only non-preeclamptic pregnancies with a minimum length of 20 weeks' gestation were considered as the non-exposed group. The CPRD Pregnancy Register and HES Maternity File were used to identify women who had a completed pregnancy record.

Identifying a completed pregnancy

For our study, records from the CPRD pregnancy register were retained if they were considered 'complete' (>20 weeks' gestation), occurred between 1 January 1997 and 31 December 2016 (study period), and were between 11 to 49 years old (inclusive) at each estimated pregnancy end date (Figure 1 for details). Linked records from the HES maternity file within the study period and age eligibility range were retained if they related to a completed pregnancy event in which the recorded gestational age at delivery, or the inferred gestational age was at least 20 weeks. See supplementary methods for details.

Records of completed pregnancies from the CPRD pregnancy register and HES maternity file were then merged into a final 'dataset of completed pregnancies'. Records from the same participant that overlapped and had a combined length of less than 385 days (55 weeks) were analyzed as the same pregnancy. Otherwise, they were treated as separate pregnancies. For HES records, pregnancy end dates were estimated using the episode start date and a variable containing the number of days between episode start and delivery. In instances without a delivery record, the latest episode end date for that pregnancy was considered the pregnancy end date. Where multiple records for the same pregnancy were identified, the duplicate with the latest end date was retained; where both CPRD and HES data from the same pregnancy were retained, HES end dates were considered more reliable. **Figure 1** outlines the data linkage process between primary and secondary care pregnancy datasets. Supplementary methods and Figure S1 outlines further handling of duplicate entries.

Exposure

Following identification of eligible completed pregnancy records, each pregnancy was then defined as either preeclamptic or non-preeclamptic using CPRD and HES records. A record of preeclampsia was defined according to the presence of a pre-specified list of Read or ICD10 Codes relating to preeclampsia diagnosis (see Table S1). If a preeclampsia code was recorded within 20 weeks either side of a pregnancy end date, this pregnancy was labelled as preeclamptic. Preterm preeclampsia was a preeclamptic pregnancy in combination with a pregnancy record that ended prior to 37 weeks' gestation and women with term preeclampsia were excluded from this analysis. Preterm preeclampsia with delivery before 37 weeks' gestation commonly represents a more severe and complicated form of preeclampsia than preeclampsia occurring at term. Furthermore, preterm preeclampsia has been shown to have a more significant adverse effect on cardiovascular function six months postpartum.[16] As a result both term and preterm preeclampsia were considered in the analysis.

In analyses with hypertensive disorders of pregnancy (HDP) as the exposure, the same approach of matching to pregnancy records was taken. The International Society for the Study of Hypertension in Pregnancy (ISSHP) have recently redefined preeclampsia, and proteinuria is now not mandatory for a diagnosis of preeclampsia. Rather, this is diagnosed by the presence of de novo hypertension after 20 weeks' gestation accompanied by proteinuria and/or evidence of maternal acute kidney injury, liver dysfunction, neurological features, hemolysis or thrombocytopenia, and/or fetal growth restriction.[1] In this study, a record of HDP was based on any one of the Read and/or ICD10 codes listed in Table S1 relating to a diagnosis in primary and/or secondary care of preeclampsia, gestational hypertension, superimposed preeclampsia, or pre-existing hypertension during pregnancy. When there was more than one pregnancy the woman remained unexposed unless she had a preeclamptic

event, at which point she became exposed and remained exposed for the rest of follow up (even if in subsequent pregnancies she did not suffer from preeclampsia). Time to event restarted after each successive pregnancy.

Endpoints

The 12 cardiovascular disorders selected as outcomes were as follows: ischaemic stroke, intracerebral haemorrhage, subarachnoid haemorrhage, stroke not otherwise specified (NOS), myocardial infarction, stable angina, unstable angina, coronary heart disease NOS, peripheral arterial disease, abdominal aortic aneurysm, atrial fibrillation, and heart failure. These phenotypes were defined using all inferred or diagnosed cases of previously validated and replicable CALIBER EHR-algorithms (for details see https://www.caliberresearch.org/portal/).[9, 12, 17-19] In addition, for this analysis we created the end-point of cardiovascular death using ONS underlying cause of death data as recorded on death certificates. Any record of cause coded within the ICD10 'I' branch or the ICD9 39 to 45 branches was considered a cardiovascular death. Composite outcomes were created according to the following groupings: all stroke (ischaemic stroke, intracerebral haemorrhage, subarachnoid haemorrhage, stroke NOS), all peripheral disease (peripheral arterial disease, abdominal aortic aneurysm), cardiac atherosclerotic (myocardial infarction, stable angina, coronary heart disease NOS), other cardiovascular (heart failure, atrial fibrillation), and all cardiovascular (all 12 cardiovascular outcomes plus any record of cardiovascular mortality). Following peer review, chronic hypertension was added as an additional outcome. If a woman's first incident event in any of the single or composite endpoint analyses occurred within six weeks of a pregnancy end date this was excluded from the analysis. We judged these to be acute cardiovascular events related to pregnancy and of different aetiology to the

ones of interest in this analysis. In each individual analysis, pregnancies occurring after any cardiovascular event of the relevant type were excluded.

Linked participant records were followed from the start of their first completed pregnancy record until the study end date (31 December 2016), ONS registered death, or initial presentation of the cardiovascular disorder under analysis, whichever was earliest.

Statistical methods

Associations with each cardiovascular end-point were modelled independently and women could contribute to more than one incident event. Hazard ratios (HRs) were calculated using disease specific cox-proportional regressions with preeclampsia, HDP, or preterm preeclampsia as a time-varying exposure and time since end of each pregnancy to event/censoring as the timescale. The survival 2.41-3 package in R was used to calculate hazard ratios. The proportional hazards assumption was verified using statistical (global and per variable) and graphical diagnostics based on Schoenfeld residuals. The results are presented as HR with 95% confidence intervals (CIs). The main adjusted model controlled for maternal ethnicity,[20] maternal pre-pregnancy hypertension (prior to first pregnancy), maternal pre-pregnancy diabetes (prior to first pregnancy), index of multiple deprivation (IMD),[21] pregnancy number, and maternal age. The last two variables were included as time-varying confounders. To address the issue of correlation between pregnancies within a single woman we calculated robust standard errors for the adjusted analyses using a cluster term with patient ID within the Cox model.

In the Cox models, the exposures (preeclampsia, preterm preeclampsia, and HDP) were treated as time-varying, in which women who had more than one pregnancy during follow-up could contribute time to both unexposed and exposed groups, prior to the occurrence of a cardiovascular event or censoring (see Figure S2 for schematic example of time-varying models). Once a woman had a record of preeclampsia or HDP she remained in that group for the rest of her follow-up, such that only two categories existed: no preeclampsia/HDP, or one or more preeclampsia/HDP. For descriptive analyses, phenotype validation, and cumulative incidence analyses only the woman's first recorded pregnancy and relevant preeclampsia exposure were considered, and any later pregnancy events were ignored.

The preeclampsia exposure phenotype was validated by comparison to a published systematic review of risk factors for preeclampsia.[18] Study design attributes of the present study and the comparator study can be found in Table S2 and Figure S3.

Baseline variables and covariates

All baseline variables were compiled using the most recent measurement closest to 16 weeks' gestation of a woman's first recorded completed pregnancy. For body mass index (BMI), records prior to five years before first pregnancy start were excluded and for pre-pregnancy blood pressure readings used in phenotype validation, those prior to one year before first recorded pregnancy were also discarded. For full details on construction of all baseline variables see supplementary methods. Maternal age and pregnancy number (1 as first recorded completed pregnancy within study period) were included in models as time-varying covariates that were re-assessed with each pregnancy event.

An additional nested model was run to investigate the potential mediating associations of post-pregnancy hypertension on cardiovascular risk. Women with pre-pregnancy (inferred or diagnosed) hypertension (prior to first recorded pregnancy) were excluded from this model and an additional time-varying covariate for post-pregnancy hypertension diagnosis was added to the main adjusted cox-model. We quantified the percentage of excess risk (on an additive scale) explained by post-pregnancy hypertension by comparing this nested model to our main model using the formula ((HRModel-1-HRModel2)/(HRModel-1))x100. To evaluate the impact of pre-pregnancy BMI we ran a third nested model in a sub-sample with available data.

To increase comparability of our preeclampsia phenotype against the report of Bartsch et al.,[22] we calculated the unadjusted relative risks and 95% CIs using the R package epiR v0.9-79. Cumulative incidence comparisons were calculated using the etmCIF function in the R package etm v0.6-2, which uses a survival model in which time from end of first pregnancy to first appearance of any of our 12 cardiovascular events or cardiovascular death was used as the time-scale. All statistical and graphical analyses were carried out in R version 3.4.3.

Data access and analysis

The study was approved by the Independent Scientific Advisory Committee for the Medicines and Healthcare Products Regulatory Agency (protocol number 16_280R). The funders had no role in the study design, data collection, data analysis, data interpretation, or writing of the report. The first author had full access to all the data in the study and all authors had final responsibility for the decision to submit the manuscript for publication.

RESULTS

Participants

The study cohort included 1,899,150 unique pregnancies from 1,303,365 women (Figure 1). A total of 434,955 (33.37%) women had more than one pregnancy during the follow-up period. A total of 31,478 (2.42%) women had 33,344 preeclamptic pregnancies, of which 25,554 (76.64%) occurred in the first pregnancy, 5,811 (17.43%) in the second, and 1,979 (5.93%) in the third or later pregnancy.

Our EHR-derived preeclampsia phenotype reproduced a very similar pattern of association between 10 pre-pregnancy risk factors and preeclampsia, to those reported in a recently published systematic review (Figure S3).[23] We also reproduced the well-known association between pre-pregnancy blood pressure levels and preeclampsia risk (Figure 2), extending this to show a log-linear relationship across the full spectrum of measurements, without evidence of a threshold effect.

Distribution of maternal characteristics by preeclampsia status for each participant's first recorded completed pregnancy are shown in Table 1. As expected, women who had preeclampsia were more likely to be nulliparous, diabetic, hypertensive, and overweight or obese, and less likely to be smokers (all P-values <0.001). Pregnancies affected by preeclampsia were more likely to be delivered preterm (compared to those without) and had a lower mean infant birthweight (both P-values $<2.2 \times 10^{-16}$).

Outcome data

After a median follow-up of 9.25 (interquartile range=5.53-13.78) years, 18,624 first incident, and 21,798 total cardiovascular events of any type were recorded. In our cohort 12,129 (65.12%) of these 18,624 first events occurred in women below 40 (age at event density distribution is reported for any cardiovascular outcome in Figure S4)). Although in absolute numbers women with preeclampsia had fewer cardiovascular events than those without preeclampsia (861 v 17,763), in relative terms, the proportion in the preeclamptic group was approximately double that in the non-preeclamptic group (2.77% v 1.40%).

Hazard ratios from multivariable time-dependent Cox models describing the association between preeclampsia and each cardiovascular disorder are shown in the upper panel of Figure 3. Women exposed to preeclampsia had a higher hazard ratio than those without a preeclamptic pregnancy, for all 12 cardiovascular events, cardiovascular mortality and chronic hypertension, except for intracerebral haemorrhage and abdominal aortic aneurysm. Overall, having at least one preeclamptic event increased a woman's hazard of first incidence of any subsequent cardiovascular event by 1.69 (95% CI=1.57-1.81), and this strength of association was broadly consistent across major categories and single end-points. When the same analyses were repeated with HDP as the exposure of interest, a very similar pattern of association was observed, whilst with preterm preeclampsia the point estimates were generally higher but showed greater uncertainty (mid and bottom panel, Figure 3). Across all categories of hypertensive disorders of pregnancy, we observed an elevated adjusted hazard for chronic hypertension, the highest being in those pregnancies affected by preterm preeclampsia (HR 5.65 (95% CI 5.10-6.26)). Adjusted and unadjusted results (HR, 95% CIs and P values) are shown in Table S5, S6 and S7.

Adjustment for diagnosis of hypertension following end of pregnancy and preceding a first cardiovascular event, attenuated the associations between preeclampsia and cardiovascular events. Adjusted HR for all stroke was 1.68 (1.30-2.18), cardiac atherosclerotic was 1.45 (1.32-1.59), other cardiovascular disease was 1.43 (1.15-1.76), all peripheral disease was 1.60 (1.12-2.29), and for all cardiovascular events combined was 1.45 (1.34-1.57). Similar results were observed for HDP, and again tended to be higher but also attenuated with preterm preeclampsia (see Table S3). After we added post-pregnancy hypertension into the originally adjusted model to estimate its potentially mediating effect, we observed a 34.78% reduction in the point estimate of the HR for any CVD event. Of the 861 women who had a preeclampsia event prior to their first incident cardiovascular event, only 184 (21.37%) also had a pre-existing diagnosis of hypertension.

In a sub-sample of participants (N=375,009) with available data on pre-pregnancy BMI in addition to all previously used confounders, HRs showed the same increase in risk but were attenuated between 15.94% and 66.67% for all composite outcomes and exposures. Given the significant reduction in number of events there was also a substantial loss in precision of these estimates (see Table S4 for comparisons).

Figure 4 shows the cumulative incidence of any cardiovascular event according to preeclampsia status, using time since first pregnancy as the timescale. Difference in cumulative incidence was evident as early as one year after the index pregnancy. By two years after first pregnancy, the cumulative incidence of any cardiovascular event in women with preeclampsia was double that of women without preeclampsia.

DISCUSSION

This is the largest contemporaneous population-based cohort study that systematically considers the association between hypertensive disorders of pregnancy and the incidence of 12 common cardiovascular phenotypes. We have shown that HDP, preeclampsia and preterm preeclampsia have a largely similar pattern of increased risk for all cardiovascular disorders. The impact is evident as early as one year after the index pregnancy.

In line with previous studies we found that preeclampsia and HDP almost double the risk of a subsequent cardiovascular event and preterm preeclampsia leads to an even larger inflated risk.[4-8, 10, 11, 23-28] We report for the first time the association between preeclampsia, HDP, and preterm preeclampsia with sub-types of stroke, peripheral vascular disease, and atrial fibrillation, and we improve the resolution for others (heart failure, stable angina), in which prior evidence was limited.[4-8, 10, 11, 23-28] Secondly, the magnitude of the associations of preeclampsia with specific cardiovascular disorders was homogenous, with a similar pattern observed for HDP and a slightly further raised risk for preterm preeclampsia. Thirdly, the differences in cumulative incidence for cardiovascular events appear as early as one year post-partum. These novel findings suggest that a diagnosis of HDP (which is more common than preeclampsia; 7.41% of women had at least one diagnosis of HDP in our study versus 2.42% for preeclampsia only) can be used as an "opportunistic screening tool" to identify women at higher risk of cardiovascular events.

The fact that the magnitude of the association of HDP with all 12 cardiovascular disorders was homogenous favours the use of HDP as a screening tool for total cardiovascular risk (i.e. all 12 cardiovascular disorders). This is an expansion of the available recommendations by clinical

guidelines, such as those from the American Heart Association (AHA)[29] and the European Society of Cardiology[30] that concentrate on preeclampsia as the screening tool, and coronary heart disease and stroke as the outcomes to prevent. Clinical guidelines, such as those produced by the AHA, have made excellent steps towards early post-pregnancy detection and control of cardiovascular risk factors such as hypertension.[31, 32] In our UK cohort, one in five of the 861 preeclamptic women who suffered a cardiovascular event had been diagnosed with hypertension, and this, together with early separation in the survival curves (1-year post-partum), highlights the need for active screening and control of cardiovascular risk factors in this selected group.

Our study has multiple strengths that secure the validity of our findings. Firstly, the incidence of preeclampsia in our cohort was 2.42%, not dissimilar to previous estimates.[33] Secondly, unlike most previous studies,[3, 6-8, 10, 25-28] we conducted an extensive validation of our EHR preeclampsia phenotype beyond simple comparisons of baseline risk-factors. Thirdly, the 12 cardiovascular phenotypes used as outcomes have all been previously validated and shown to be highly reproducible.[9, 12, 17-19]

The quality and comprehensiveness of the EHR resources used for our study allowed us to include time-dependent exposures and covariates in our Cox models, which is preferable to baseline only models. However, it is important to highlight some limitations. Firstly, the substantial level of missingness in the EHR records of pre-pregnancy BMI and smoking (>60%) limited the validity of using multiple imputation and therefore precluded their inclusion in the final models. Taken together with the observational nature of our analysis this introduces the potential for residual confounding to be a partial, or total, explanation to our results.

However, adjustment for pre-pregnancy BMI (in a sub-sample) did not substantially change our conclusions. Other limitations include potential selection bias and inaccuracy of data collection. Although it is theoretically possible to have some degree of misclassification in the exposure (i.e. preeclampsia), we judged that inaccuracies in coding are unlikely to be related to hypertensive disorders. As a consequence the impact on exposure misclassification is just random error which is largely overturned by large dataset. It is possible that women who had a previous pregnancy prior to the start of the study period (either hypertensive or normotensive) might be mislabelled if their first pregnancy is not captured. However, we consider that it is likely that this has happened on a small scale and to a similar degree for previous hypertensive and normotensive pregnancies and therefore is not likely to have a substantial impact on our findings. The potential impact of non-CVD mortality as a competing risk on CVD events was assessed but had no impact on our results (Table S8). In the present study we did not evaluate whether addition of preeclampsia or HDP history has an incremental benefit over established risk-algorithms for cardiovascular disease in middle age (e.g. QRISK, Framingham).[34]

The precise mechanisms that explain the association between preeclampsia or HDP and cardiovascular disorders remain a matter of debate. A potential explanation is that shared risk factors between preeclampsia and cardiovascular disorders, such as hypertension and obesity, are at least partially responsible. Our findings are in favour of this hypothesis. Firstly, adjustment for hypertension that occurs post-pregnancy but prior to the cardiovascular event reduced the hazard ratios. This is in partial agreement with the findings by Canoy et al[10] who also identified hypertension between pregnancy and coronary heart disease or stroke as an important modifying factor in risk. Secondly, high blood pressure is a common risk factor for

HDP and preeclampsia and has homogenous associations with the 12 cardiovascular disorders we evaluated,[9] which could explain the similar pattern of homogenous associations that preeclampsia and HDP had on cardiovascular disorders evaluated. Thirdly, the almost immediate, post-pregnancy separation in the cumulative incidence on cardiovascular events by preeclampsia status (Figure 4) suggests it is more likely to be pre-pregnancy risk factors such as hypertension rather than only a de-novo insult created by preeclampsia that explains these observations.

Further work is required to estimate how many cardiovascular events could be avoided by implementing interventions such as active high blood pressure detection and control in women that had experienced HDP. The large-scale EHR resource we have created also serves as a unique opportunity to systematically test the theory that pregnancy can be conceived of as a "stress-test" for future development of chronic disease.[5]

In conclusion, we have showed that preeclampsia, preterm preeclampsia, and HDP have a similar pattern of increased risk for the 12 most common cardiovascular disorders observed in women in Britain, highlighting the opportunity to use HDP as a naturally occurring screening tool to detect women at high risk of cardiovascular events.

ACKNOWLEDGEMENTS

All authors have made substantial contributions to the research as follows: study design: LL, FMC, AG, JPC, LCC; study conduct, analyses: all authors; first draft of the manuscript: LL; manuscript revision and approval: all authors. The authors would like thank Dr Rhian Daniel for her invaluable help on study design and analysis.

SOURCES OF FUNDING

This report is independent research supported by the National Institute for Health Research Professorship, RP-2014-05-019. The views expressed in this publication are those of the authors and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health. FP McCarthy was funded by an NIHR Clinical Academic Fellowship. This project was supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre.

DISCLOSURES

The authors declare no conflict of interest.

REFERENCES

1. Brown MA, Magee LA, Kenny LC, Karumanchi SA, McCarthy FP, Saito S, Hall DR, Warren CE, Adoyi G, Ishaku S, et al. The hypertensive disorders of pregnancy: ISSHP classification, diagnosis & management recommendations for international practice. *Pregnancy hypertension*. 2018;13:291-310.

2. Say L, Chou D, Gemmill A, Tunçalp Ö, Moller A-B, Daniels J, Gülmezoglu AM, Temmerman M and Alkema L. Global causes of maternal death: a WHO systematic analysis. *The Lancet Global Health*. 2014;2:e323-e333.

3. Jarvie JL, Metz TD, Davis MB, Ehrig JC and Kao DP. Short-term risk of cardiovascular readmission following a hypertensive disorder of pregnancy. *Heart*. 2018;104:1187-1194.

4. Lykke JA, Langhoff-Roos J, Sibai BM, Funai EF, Triche EW and Paidas MJ. Hypertensive pregnancy disorders and subsequent cardiovascular morbidity and type 2 diabetes mellitus in the mother. *Hypertension (Dallas, Tex : 1979)*. 2009;53:944-951.

5. Bellamy L, Casas JP, Hingorani AD and Williams DJ. Pre-eclampsia and risk of cardiovascular disease and cancer in later life: systematic review and meta-analysis. *Bmj*. 2007;335:974.

6. Behrens I, Basit S, Melbye M, Lykke JA, Wohlfahrt J, Bundgaard H, Thilaganathan B and Boyd HA. Risk of post-pregnancy hypertension in women with a history of hypertensive disorders of pregnancy: nationwide cohort study. *Bmj*. 2017;358:j3078.

7. Ray JG, Vermeulen MJ, Schull MJ and Redelmeier DA. Cardiovascular health after maternal placental syndromes (CHAMPS): population-based retrospective cohort study. *Lancet*. 2005;366:1797-1803.

8. Smith GC, Pell JP and Walsh D. Pregnancy complications and maternal risk of ischaemic heart disease: a retrospective cohort study of 129,290 births. *Lancet*. 2001;357:2002-2006.

9. Rapsomaniki E, Timmis A, George J, Pujades-Rodriguez M, Shah AD, Denaxas S, White IR, Caulfield MJ, Deanfield JE, Smeeth L, et al. Blood pressure and incidence of twelve cardiovascular diseases: lifetime risks, healthy life-years lost, and age-specific associations in 1.25 million people. *Lancet*. 2014;383:1899-1911.

10. Canoy D, Cairns BJ, Balkwill A, Wright FL, Khalil A, Beral V, Green J and Reeves G. Hypertension in pregnancy and risk of coronary heart disease and stroke: A prospective study in a large UK cohort. *Int J Cardiol*. 2016;222:1012-1018.

11. Behrens I, Basit S, Lykke JA, Ranthe MF, Wohlfahrt J, Bundgaard H, Melbye M and Boyd HA. Association Between Hypertensive Disorders of Pregnancy and Later Risk of Cardiomyopathy. *Jama*. 2016;315:1026-1033.

12. Denaxas SC, George J, Herrett E, Shah AD, Kalra D, Hingorani AD, Kivimaki M, Timmis AD, Smeeth L and Hemingway H. Data resource profile: cardiovascular disease research using linked bespoke studies and electronic health records (CALIBER). *Int J Epidemiol*. 2012;41:1625-1638.

13. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T and Smeeth L. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *International Journal of Epidemiology*. 2015;44:827-836.

14. Herbert A, Wijlaars L, Zylbersztejn A, Cromwell D and Hardelid P. Data Resource Profile: Hospital Episode Statistics Admitted Patient Care (HES APC). *Int J Epidemiol*. 2017;46:1093-1093i.

15. Office for National Statistics. Mortality Statistics: Metadata 2015. Available at https://www.ons.gov.uk › file › mortalitymetadata2014_tcm77-241077 (last accessed 30.08.2019) .

16. Melchiorre K, Sutherland GR, Liberati M and Thilaganathan B. Preeclampsia is associated with persistent postpartum cardiovascular impairment. *Hypertension*. 2011;58:709-715.

17. Herrett E, Shah AD, Boggon R, Denaxas S, Smeeth L, van Staa T, Timmis A and Hemingway H. Completeness and diagnostic validity of recording acute myocardial infarction events in primary care, hospital care, disease registry, and national mortality records: cohort study. *BMJ. 2013 May 20;346:f2350. doi: 10.1136/bmj.f2350*.

18. Morley KI, Wallace J, Denaxas SC, Hunter RJ, Patel RS, Perel P, Shah AD, Timmis AD, Schilling RJ and Hemingway H. Defining Disease Phenotypes Using National Linked Electronic Health Records: A Case Study of Atrial Fibrillation. *PloS one*. 2014;9:e110900.

19. Koudstaal S, Pujades-Rodriguez M, Denaxas S, Gho J, Shah AD, Yu N, Patel RS, Gale CP, Hoes AW, Cleland JG, et al. Prognostic burden of heart failure recorded in primary care, acute hospital admissions, or both: a population-based linked electronic health record cohort study in 2.1 million people. *European journal of heart failure*. 2017;19:1119-1127.

20. George J, Mathur R, Shah AD, Pujades-Rodriguez M, Denaxas S, Smeeth L, Timmis A and Hemingway H. Ethnicity and the first diagnosis of a wide range of cardiovascular diseases: Associations in a linked electronic health record cohort of 1 million patients. *PloS one*. 2017;12:e0178945.

21. Pujades-Rodriguez M, Timmis A, Stogiannis D, Rapsomaniki E, Denaxas S, Shah A, Feder G, Kivimaki M and Hemingway H. Socioeconomic deprivation and the incidence of 12

cardiovascular diseases in 1.9 million women and men: implications for risk prediction and prevention. *PloS one*. 2014;9:e104671.

22. Bartsch E, Medcalf KE, Park AL and Ray JG. Clinical risk factors for pre-eclampsia determined in early pregnancy: systematic review and meta-analysis of large cohort studies. *Bmj*. 2016;353:i1753.

23. Riise HK, Sulo G, Tell GS, Igland J, Nygard O, Vollset SE, Iversen AC, Austgulen R and Daltveit AK. Incident Coronary Heart Disease After Preeclampsia: Role of Reduced Fetal Growth, Preterm Delivery, and Parity. *J Am Heart Assoc*. 2017;6(3)

24. Wikstrom AK, Haglund B, Olovsson M and Lindeberg SN. The risk of maternal ischaemic heart disease after gestational hypertensive disease. *BJOG : an international journal of obstetrics and gynaecology*. 2005;112:1486-1491.

25. Tooher J, Thornton C, Makris A, Ogle R, Korda A and Hennessy A. All Hypertensive Disorders of Pregnancy Increase the Risk of Future Cardiovascular Disease. *Hypertension (Dallas, Tex : 1979)*. 2017;70:798-803.

26. Mannisto T, Mendola P, Vaarasmaki M, Jarvelin MR, Hartikainen AL, Pouta A and Suvanto E. Elevated blood pressure in pregnancy and subsequent chronic disease risk. *Circulation*. 2013;127:681-690.

27. Funai EF, Friedlander Y, Paltiel O, Tiram E, Xue X, Deutsch L and Harlap S. Long-term mortality after preeclampsia. *Epidemiology*. 2005;16:206-215.

28. Kestenbaum B, Seliger SL, Easterling TR, Gillen DL, Critchlow CW, Stehman-Breen CO and Schwartz SM. Cardiovascular and thromboembolic events following hypertensive pregnancy. *American journal of kidney diseases : the official journal of the National Kidney Foundation*. 2003;42:982-989.

29. Mosca L, Benjamin EJ, Berra K, Bezanson JL, Dolor RJ, Lloyd-Jones DM, Newby LK, Pina IL, Roger VL, Shaw LJ, et al. Effectiveness-based guidelines for the prevention of cardiovascular disease in women--2011 update: a guideline from the american heart association. *Circulation*. 2011;123:1243-1262.

30. Piepoli MF, Hoes AW, Agewall S, Albus C, Brotons C, Catapano AL, Cooney MT, Corra U, Cosyns B, Deaton C, et al. 2016 European Guidelines on cardiovascular disease prevention in clinical practice: The Sixth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of 10 societies and by invited experts)Developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR). *European heart journal*. 2016;37:2315-2381.

31. Bushnell C, McCullough LD, Awad IA, Chireau MV, Fedder WN, Furie KL, Howard VJ, Lichtman JH, Lisabeth LD, Pina IL, et al. Guidelines for the prevention of stroke in women: a statement for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke*. 2014;45:1545-1588.

32. Goff DC, Jr., Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB, Gibbons R, Greenland P, Lackland DT, Levy D, O'Donnell CJ, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2014;129:S49-73.

33. Brown MA, Magee LA, Kenny LC, Karumanchi SA, McCarthy FP, Saito S, Hall DR, Warren CE, Adoyi G, Ishaku S, et al. Hypertensive Disorders of Pregnancy: ISSHP Classification, Diagnosis, and Management Recommendations for International Practice. *Hypertension*. 2018;72:24-43.

34. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, May M and Brindle P. Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. *Bmj*. 2007;335:136.

Figure 1: Flow diagram outlining construction of final pregnancy and preeclampsia cohort using CPRD and HES datasets. *strategy for dealing with duplicate entries is outlined in supplementary methods

Figure 2: Odds ratios (OR) for risk of preeclampsia from a subset of first recorded pregnancies with available blood pressure (BP) readings. BP taken up to 1 year prior to pregnancy start and within 16 weeks of gestation. A) Systolic blood pressure (SBP) and odds of preeclampsia in each consecutive group compared to lowest group (<95), B) Diastolic blood pressure (DBP) and odds of preeclampsia in each consecutive group compared to lowest group (DBP<55). Red dotted line indicates a weighted regression cubic spline with 95% CIs.

Figure 3: Forest plot of adjusted hazard ratios (HRs) with 95% confidence intervals (CIs) for 12 cardiovascular outcomes, relevant composites, and chronic hypertension, given exposure to A) Preeclampsia; B) Hypertensive disorders of pregnancy; C) Preeclampsia with preterm birth. All HRs were computed using a Cox-proportional hazards model with time-dependent exposure and adjusted for associations of maternal ethnicity, maternal age, pre-pregnancy diabetes, pre-pregnancy hypertension, index of multiple deprivation, and a cluster term to account for correlation within patients. Events are numbers of events in the exposure group only.

Figure 4: Cumulative incidence (with 95% confidence interval) of first cardiovascular event by preeclampsia status by years since end of first recorded pregnancy Numbers at risk in each group for each x-axis tick mark are listed by colour of exposure. Graph was only plotted up until the point that the numbers at risk in the control group were at least 10% of the maximum control group size.

Table 1. Summary of maternal demographic, first recorded pre-pregnancy, and delivery characteristics. Missingness recorded where present.

Maternal characteristic	**All (1,303,365)**	**No PE (N=1,277,811)**	**PE (N=25,554)**
Maternal age at delivery (years), mean (SD)	28.48 (6.15)	28.47 (6.15)	28.61 (6.29)
Ethnicity, n (%)			
Asian	69073 (5.30)	67849 (5.31)	1224 (4.79)
Black	48427 (3.72)	47124 (3.69)	1303 (5.10)
Other	63283 (4.86)	62338 (4.88)	945 (3.70)
Unknown	80233 (6.16)	79233 (6.20)	1000 (3.91)
White	1042349 (79.97)	1021267 (79.92)	21082 (82.50)
Practice IMD 2015, n (%)			
Quintile 1 (least deprived)	196375 (15.07)	192832 (15.09)	3543 (13.86)
Quintile 5 (most deprived)	320220 (24.57)	313612 (24.54)	6608 (25.86)
At first recorded pregnancy booking			
Diabetes, n (%)	3378 (0.26)	3020 (0.24)	358 (1.40)

Hypertension, n (%)	29144 (2.24)	27027 (2.12)	2117 (8.28)
Nulliparity, n (% of known)	294996 (64.66)	285189 (64.43)	9807 (72.06)
Missing (%)	847146 (65.00)	835202 (65.34)	11944 (46.74)
Pre-pregnancy BMI (kg/m2) (% of known)			
Underweight (<18.5)	18045 (4.77)	17720 (4.84)	325 (2.80)
Healthy (18.5-24.9)	226246 (59.85)	220853 (60.28)	5393 (46.46)
Overweight (25-29.9)	83692 (22.14)	80536 (21.98)	3156 (27.19)
Obese (30-39.9)	44466 (11.76)	42107 (11.49)	2359 (20.32)
Severely obese (>40)	5553 (1.47)	5177 (1.41)	376 (3.24)
Missing (%)	925363 (71.00)	911418 (71.00)	13945 (54.57)
Ever smoker, n (% of known)	199254 (39.18)	193893 (39.31)	5361 (34.87)
Missing (%)	794754 (60.98)]	784575 (61.40)	10179 (39.83)
Multi-fetal pregnancy, n (% of known)	30471 (2.86)	29261 (2.80)	1210 (6.00)
Missing (%)	227809 (17.48)	222490 (17.41)	5319 (20.81)
Gestational diabetes, n (%)	17199 (1.32)	16344 (1.28)	855 (3.35)

At delivery			
Gestational age at birth (weeks), mean (SD)	39.15 (2.53)	39.18 (2.5)	37.81 (3.14)
Preterm birth (<37 weeks), n (%)	53652 (4.12)	48500 (3.80)	5152 (20.16)
Infant birthweight* (grams), mean (SD)	3339 (590)	3348 (580)	2899.44 (845)
Missing (%)	311624 (23.91)	305661(23.92)	5963 (23.33)

*For multi-fetal pregnancies this refers to the last delivered infant