Journal of the American Heart Association

ORIGINAL RESEARCH

Development and Internal Validation of a Model Predicting Premature Cardiovascular Disease Among Women With Hypertensive Disorders of Pregnancy: A Population-Based Study in Quebec, Canada

U. Vivian Ukah, MPH, PhD; Natalie Dayan, MD, MSc; Nathalie Auger, MD, MSc; Siyi He, MSc; Robert W. Platt, PhD

BACKGROUND: Hypertensive disorders of pregnancy (HDP) are associated with an increased risk of premature cardiovascular disease (CVD), but existing cardiovascular prediction models do not adequately capture risks in young women. We developed a model to predict the 10-year risk of premature CVD and mortality among women who have HDP.

METHODS AND RESULTS: Using a population-based cohort of women with HDP who delivered between April 1989 and March 2017 in Quebec, Canada, we developed a 10-year CVD risk model using Cox proportional hazards regression. Women aged 18 to 45 years were followed from their first HDP-complicated delivery until March 2018. We assessed performance of the model based on discrimination, calibration, and risk stratification ability. Internal validity was assessed using the bootstrap method. The cohort included 95 537 women who contributed 1 401 084 person-years follow-up. In total, 4024 (4.2%) of women were hospitalized for CVD, of which 1585 events (1.6%) occurred within 10 years of follow-up. The final model had modest discriminatory performance (area under the receiver operating characteristic curve, 0.66; 95% CI, 0.65–0.67) and good calibration with slope of 0.95 and intercept of −0.19. There was moderate classification accuracy (likelihood ratio+: 5.90; 95% CI, 5.01–6.95) in the highest-risk group upon risk stratification.

CONCLUSIONS: Overall, our model had modest performance in predicting the 10-year risk of premature CVD for women with HDP. We recommend the addition of clinical variables, and external validation, before consideration for clinical use.

Key Words: cardiovascular ■ prediction modeling ■ preeclampsia ■ pregnancy hypertension ■ prognosis

Cardiovascular disease (CVD) is a leading cause of morbidity and mortality among women worldwide,[1,2] and more women die of CVD than men at all ages.[3] Improving CVD prediction is a priority since the majority of CVD is preventable. Several clinical prediction models are routinely used for predicting CVD risk in the general population, such as the Framingham risk score model,[4] QRISK score,[5] and the Adult Treatment Panel model.[6] However, most of these models underestimate the risk of CVD in young women.[7,8] Professional bodies such as the American Heart Association and Society of Obstetricians and Gynaecologists of Canada[9] have highlighted the need to identify female-specific factors that can aid in predicting CVD for women.

Pregnancy involves substantial cardiovascular changes, and a woman's response to pregnancy can be indicative of future cardiovascular health.[10] Numerous

Correspondence to: U. Vivian Ukah, MPH, PhD, Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Quebec, Canada H3A 1A2. E-mail: vivian.ukah@mail.mcgill.ca

Supplementary Materials for this article are available at https://www.ahajournals.org/doi/suppl/10.1161/JAHA.120.017328

For Sources of Funding and Disclosures, see page 9.

JAHA is available at: www.ahajournals.org/journal/jaha

CLINICAL PERSPECTIVE

What Is New?

- We developed a 10-year cardiovascular risk prediction model for women with hypertension in pregnancy.
- Our model, developed using large administrative data, performed moderately (likelihood ratio+ >5 in the highest-risk group).

What Are the Clinical Implications?

- Our model has the potential to identify young women at the highest risk of cardiovascular disease.
- However, the model must be improved and externally validated before clinical use.
- Early identification of high-risk women may aid in cardiovascular disease prevention through lifestyle and pharmaceutical interventions.

Nonstandard Abbreviation and Acronym

HDP	hypertensive disorders of pregnancy

studies have consistently reported a 2-fold higher risk of developing premature CVD among women who experience hypertensive disorders of pregnancy (HDP).[11–14] The postpartum period is now widely regarded as a possible window of opportunity for CVD prevention in women.[15–17] However, there is a lack of guidance on risk factor management in this population.[18]

There have been attempts to improve the prediction of CVD in women by adding pregnancy complications, especially HDP, as potential variables to existing predictive models.[7,19–25] However, these studies failed to improve the performance of models.[20] For example, a study by Markovitz et al[21] reported an improvement of only 0.004 in the c-index upon the addition of pregnancy complications to NORRISK 2, an existing prediction model.[26] These studies used a population of older women (>40 years of age), possibly diluting the importance of pregnancy factors and limiting generalizability to younger women.[27] Since the absolute risk of CVD is low in young women, merely adding pregnancy complications to an existing model is likely insufficient to permit discrimination between women with and without CVD.[20]

In response to these gaps in clinical prediction, we developed and internally validated a 10-year model to predict premature CVD and all-cause mortality before 60 years of age among a population of young women with HDP, a target population with an established higher baseline risk of CVD. This time frame was chosen based on previous recommendation as a clinically important period for CVD prevention.[20] We hypothesized that easily measured clinical and pregnancy/reproductive factors available in administrative data would be sufficient to predict CVD in this high-risk population.

METHODS

Data and Study Population

Requests to access the data set from qualified researchers trained in human subject confidentiality protocols may be sent to Institut de la statistique du Québec (https://www.stat.gouv.qc.ca/research/#/accueil). We used a longitudinal pregnancy cohort, constructed from the Maintenance and Use of Data for the Study of Hospital Clientele registry, consisting of all hospital-based deliveries in Quebec, Canada from April 1989 until March 2016.[12] This cohort includes >99% of all deliveries in Quebec, a province that comprises more than one quarter of the Canadian population.

We selected women aged 18 to 45 years who had HDP in at least 1 of their pregnancies during the study period. HDP was defined as preexisting hypertension in pregnancy, gestational hypertension, preeclampsia/eclampsia/hemolysis, elevated liver enzymes, and low platelet count syndrome, superimposed pre-eclampsia, or unspecified hypertension during pregnancy.[9,28] HDP was identified using diagnostic codes from the *International Classification of Diseases, Ninth and Tenth Revisions* (*ICD-9, ICD-10*).[12] A validation study reported high sensitivity and specificity (>80% each) for identifying HDP using *ICD* codes.[29]

We excluded women with preexisting CVD before study entry (n=659), and those who died in their first affected pregnancy (n=46) (Figure 1).

Outcomes

The primary outcome was a composite of premature (before the age of 60 years) CVD events and/or premature all-cause in-hospital mortality. CVD events included hospitalization for heart disease (heart failure, ischemic heart disease, myocardial infarction, angina, cardiac arrest, inflammatory heart disease, conduction defects, valve disorders, cardiomyopathy, pulmonary heart disease), cerebrovascular disease (ischemic, hemorrhagic), atherosclerotic disease, aortic aneurysm or dissection, other aneurysm, arterial embolism, cardiovascular interventions (coronary angioplasty, coronary artery bypass graft, pacemaker, valve surgery, cardiac transplant), and admission to a coronary care unit. CVD diagnoses and procedures were identified using *ICD*-9 and *-10*, Canadian Classification of Diagnostic, Therapeutic, and Surgical Procedures, and Canadian Classification of Health Intervention codes.[12]

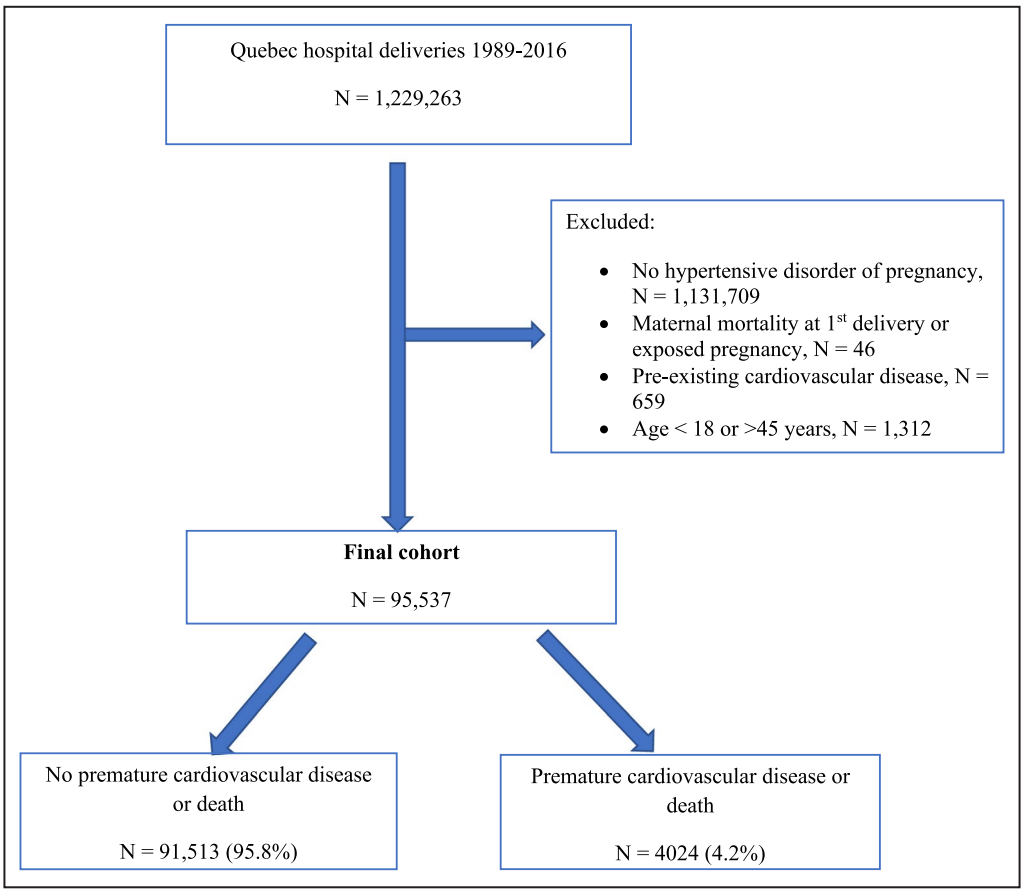

Figure 1. Development of study cohort.

These codes have been previously used for the identification of CVD outcomes.[12]

Follow-up

Women were followed from the delivery hospitalization for their first HDP-complicated pregnancy until 10 years thereafter, until they reached 60 years of age, or the end of the study period (March 31, 2018),whichever came first, to identify subsequent hospitalization for CVD or death.

Sample Size

To avoid model overfitting, we required at least 20 events per degree of freedom.[30] Based on this and the total of 4024 events (1585 within 10 years of follow-up), we had sufficient sample size to consider up to 75 degrees of freedom for candidate predictors.

Candidate Predictors, Variable Selection, and Coding

Demographic, clinical, and reproductive factors known to be associated with an increased risk of CVD were considered candidate variables. These included maternal age, smoking, obesity, socioeconomic deprivation, and preexisting medical conditions such as diabetes mellitus; pregnancy factors such as parity, multifetal pregnancy; and complications during pregnancy such as type of HDP, gestational diabetes mellitus, severe maternal morbidity, stillbirth, preterm delivery, low birth weight, and admission into neonatal intensive care unit and adult intensive care unit. Socioeconomic deprivation was measured using a composite score of neighborhood income, education, and employment.[12] Severe maternal morbidity was defined according to the published Canadian Perinatal System Surveillance criteria.[31] All candidate predictors were measured at the time of the index delivery.

Variable selection was done in different stages to include only candidate predictors that are routinely and reliably collected.[32] Variables that were clinically related but with very low incidence were combined. For example, tobacco, alcohol, and illicit drug use were combined as substance use, while any previous pregnancy complications (previous stillbirth, preterm delivery, low birthweight, admission to neonatal

intensive care unit, neonatal death, or severe maternal morbidity) were combined as a composite variable for previous history of obstetric complications. Variables were assessed for collinearity, and where collinear (r>0.5), the most clinically relevant or commonly reported variable was selected. Previous history of obstetric complications was combined with parity coded in 3 levels as these variables were collinear; categories were previous obstetric complication (multiparous), no previous obstetric complication (multiparous), and no previous obstetric complication (primiparous).

The final candidate predictors were the following: prenatal substance use, obesity, socioeconomic deprivation, diabetes mellitus (preexisting diabetes mellitus/gestational diabetes mellitus/no diabetes mellitus), HDP subtype (preexisting hypertension, gestational hypertension, preeclampsia/hemolysis, elevated liver enzymes, and low platelet count syndrome/eclampsia, superimposed preeclampsia, or unspecified hypertension), cesarean delivery, severe maternal morbidity, stillbirth, neonatal death, admission to neonatal intensive care unit; and previous pregnancy complications (yes/multiparous, no/multiparous, primiparous). Continuous candidate predictor variables included maternal age, gestational age at delivery, and length of stay in intensive care unit, modeled using restricted cubic splines with 3 clinically relevant knot locations.[30]

Possible interactions between variables were assessed and included if statistically significant at α=0.10.[33] The final variable selection was done using Least Absolute Selection and Shrinkage Operator regression.[32,33] A Cox proportional hazards regression model was fit using the final variables to predict the onset of premature cardiovascular outcomes and all-cause mortality.

Statistical Analysis

Cox proportional hazards regression models were developed according to the steps outlined by Steyerberg and Vergouwe,[32] and Harrell,[30] to predict outcomes. We reported the model development process based on the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis guidelines (Table S1).[34]

Model Performance and Internal Validation

The 10-year risk of developing the outcome was estimated using the final selected predictors, and the prediction performance of the model was assessed based on discriminatory, calibration, and risk stratification abilities.[33] Discrimination measures the ability of the model to accurately distinguish between women who developed an adverse outcome and those who did not. This was measured by the c-statistic, which is analogous to the area under the receiver operating characteristic curve for a binary outcome but taking censoring into account.[30] A c-statistic of 0.5 was considered not useful, 0.5 to 0.6 poor, 0.6 to <0.7 modest, and ≥0.7 good.[35] Calibration measures the agreement between the observed and predicted risks. Calibration performance was examined by plotting the mean observed events versus the mean predicted risks by decile of the predicted risk. Calibration slopes were interpreted as noninformative (slope ≤0.5), poor (0.5 < slope ≤0.7), or good (slope >0.7).

The model's ability to stratify the population into low- and high-risk categories was examined using a risk classification table. We divided the population into the top 2% of the population with the highest calculated risk; the next >2% to 5%, >5% to 10%, and the rest of the population were grouped according to their calculated risk of disease occurrence. These cut-offs corresponded to the distribution of events in the population.[36] Likelihood ratios (LR) were computed for each group using Deeks and Altman`s method for multiple categories, to assess classification accuracy within each group.[37] For clinical use, positive LRs of >5 or >10 were respectively interpreted as moderate or good "rule-in" tests (a value reflecting a high probability of disease occurrence). For negative LRs, <0.2 was considered a moderate rule-out test (ie, high probability that the disease will not occur), and <0.1 was considered a good rule-out test.

The model was assessed for internal validity using the Efron bootstrap method with 200 iterations and the over-optimism (ie, the degree to which a model is overfit to the sample data) was reported.[33]

Secondary Analyses

For secondary analyses, we predicted the 15-year risk of developing the outcome using the same final selected predictors. Similar to the primary analyses, women were followed from their first HDP-complicated pregnancy until 15 years thereafter, 60 years of age, or the study period end, whichever came first, to identify study outcomes. The model was re-assessed based on discriminatory performance, as described above.

Sensitivity Analyses

Recognizing that pregnancy complications could occur up to 42 days after delivery and an outcome in this time window could be related to the pregnancy itself, we ran the models after excluding women with outcomes within this time period. We also ran the primary model

excluding women with preexisting hypertension and diabetes mellitus.

All analyses were conducted using R version 3.5.1 (The R Project for Statistical Computing).[38]

Ethics review for this study was waived by the institutional review board of the University of Montreal Hospital Centre, because the data were de-identified and informed consent was not required.

RESULTS

Participants

Of the 1 229 263 women in the original cohort, HDP occurred in 7.9% of the population, with the majority (81.6%) occurring in the first pregnancy. Steps in the selection of the final cohort are shown in Figure 1. The final cohort consisted of 95 537 women between the ages of 18 and 45 years who contributed a total length of follow-up of 1 401 084 person-years. In total, 4024 (4.2%) women experienced premature CVD or death, of which 1585 (1.6%) and 2435 (2.6%) occurred within 10 and 15 years of follow-up, respectively.

Table 1 (and Table S2) show the characteristics of women in the total population according to the occurrence of CVD/mortality. Women who developed these outcomes were more likely to be older, socioeconomically deprived, use substances (alcohol/tobacco/illicit drugs), have preexisting diabetes mellitus or hypertension, and have preterm delivery, at the time of their complicated pregnancy compared with women who did not develop an outcome.

The types and incidence rates of CVD/mortality are presented in Table 2. The incidence of the outcome was 28.72 per 10 000 person-years and the most frequent outcomes were ischemic heart disease (9.49/10 000 person-years), conduction disorders/arrhythmia (8.84/10 000 person-years), and atherosclerotic disease (7.59/10 000 person-years).

Model Development and Performance

The final prediction model contained 10 variables with 15 degrees of freedom and the model equation is presented in Table 3. The model included maternal age, socioeconomic deprivation, substance use, gestational age at delivery, obesity, diabetes mellitus, cesarean delivery, previous complications, type of HDP, and number of days in the intensive care unit. The model had moderate discriminatory performance with a *c*-statistic of 0.66 (95% CI, 0.65–0.67), albeit minimal overfitting (optimism of 0.008) upon internal validation (Figure 2).

The calibration slope was 0.95 and the intercept was −0.19, indicating a good agreement between

Table 1. Characteristics of Women With Hypertensive Disorders of Pregnancy According to Outcome (Premature Cardiovascular Events or Death), Quebec, 1989 to 2018, N=95 537

	No. Women (%)	
	No Outcome (N=4024)	Outcome (N=91 513)
Maternal age, y		
<20	96 (2.4)	2890 (3.2)
20–24	590 (14.6)	17 546 (19.2)
25–29	1222 (30.4)	31 438 (34.4)
30–34	1165 (29.0)	24 695 (27.0)
35–39	767 (19.0)	11 916 (13.0)
≥40	184 (4.6)	3028 (3.2)
Multiple pregnancy		
Yes	104 (2.6)	3228 (3.5)
No	3920 (97.4)	88 285 (96.5)
Time period at delivery		
1989–1995	2093 (52.0)	22 448 (24.5)
1996–2002	1097 (27.3)	20 219 (22.1)
2003–2009	596 (14.8)	23 610 (25.8)
2010–2016	237 (5.9)	25 230 (27.6)
Socioeconomic disadvantage		
Yes	964 (24.0)	18 694 (20.4)
No	2722 (67.6)	67 731 (74.0)
Unknown	338 (8.4)	5088 (5.6)
Substance use*		
Yes	116 (2.9)	1397 (1.5)
No	3908 (97.1)	90 116 (98.5)
Morbid obesity		
Yes	198 (4.9)	4047 (4.4)
No	3826 (95.1)	87 466 (95.6)
Preexisting diabetes mellitus		
Yes	375 (9.3)	2569 (2.8)
No	3649 (90.7)	88 944 (97.2)
Primiparity at complicated pregnancy		
Yes	3310 (82.3)	74 688 (81.6)
No	714 (17.7)	16 825 (18.4)
Hypertensive disorder of pregnancy subgroups		
Preexisting hypertension	688 (17.1)	8583 (9.4)
Gestational hypertension	715 (17.7)	29 300 (32.0)
Preeclampsia/HELLP syndrome	1766 (43.9)	38 208 (41.8)
Superimposed preeclampsia	143 (3.6)	1813 (2.0)
Unspecified hypertension	712 (17.7)	13 609 (14.8)
Gestational diabetes mellitus		
Yes	420 (10.4)	8387 (9.2)
No	3604 (89.6)	83 126 (90.8)
Cesarean section		
Yes	1541 (38.3)	30 104 (32.9)

(*Continued*)

Table 1. Continued

	No. Women (%)	
	No Outcome (N=4024)	Outcome (N=91 513)
No	2483 (61.7)	61 409 (67.1)
Severe maternal morbidity		
Yes	799 (19.9)	18 638 (20.4)
No	3225 (80.1)	72 875 (79.6)
Preterm delivery (GA <37 wk)		
Yes	952 (23.7)	15 733 (17.2)
No	3072 (76.3)	75 780 (82.8)
Stillbirth		
Yes	50 (1.2)	532 (0.6)
No	3974 (98.8)	90 981 (99.4)
Low birth weight (<2500 g)		
Yes	910 (22.6)	15 301 (16.7)
No	3112 (77.4)	76 204 (83.3)
Admission to NICU		
Yes	260 (6.5)	7266 (8.0)
No	3728 (93.5)	83 853 (92.0)
Neonatal death		
Yes	31 (0.8)	290 (0.3)
No	3957 (99.2)	90 829 (99.7)
Previous pregnancy complication*,† (among multiparous women)		
Yes	163 (4.1)	3298 (3.6)
No	551 (13.7)	16 825 (14.8)

GA indicates gestational age; HELLP, hemolysis, elevated liver enzymes, and low platelet count; and NICU, neonatal intensive care unit.

*Substance use includes tobacco, alcohol, and illicit drug use.

†Previous complications include any previous stillbirth, preterm delivery, low birth weight, admission to NICU, neonatal death, or severe maternal morbidity.

the predicted and observed rates of outcomes within each decile of calculated predicted probabilities (Figure 3).

Risk stratification and classification accuracy are presented in Table 4. In the 2% of the population with the highest risk, the calculated risk probability was ≥0.07%. In this highest-risk category consisting of 1601 (1.7%) women, the model was able to moderately rule-in the likelihood of outcome occurrence with LR+ of 5.90 (5.01–6.95), showing good risk classification accuracy. However, among the women in this high-risk group, only 9.1% had the outcome, indicating poor risk stratification.

Secondary Analyses

The 15-year prediction model showed similar discriminatory performance with a *c*-statistic of 0.65 (95% CI, 0.63–0.66), although the calibration performance decreased with a slope of 0.89 and intercept of −0.38 for calibration.

Table 2. Incidence Rates of Hospitalization for Different Cardiovascular Diseases for Women With Hypertensive Disorders of Pregnancy, Quebec, 1989 to 2018

	No. Cardiovascular Hospitalizations	Incidence Rate Per 10 000 Person-Years (95% CI)
Heart failure	471	3.36 (3.07–3.68)
Ischemic heart disease	1329	9.49 (8.99–10.01)
Myocardial infarction	754	5.38 (5.01–5.78)
Angina	417	2.98 (2.70–3.28)
Cardiac arrest	105	0.75 (0.62–0.81)
Inflammatory heart disease	105	0.75 (0.62–0.81)
Conduction defects	1239	8.84 (8.36–9.35)
Valve disorders	386	2.76 (2.49–3.04)
Cardiomyopathy	260	1.86 (1.64–2.10)
Pulmonary heart disease	165	1.18 (1.01–1.37)
Ischemic cerebrovascular disease	567	4.05 (3.72–4.39)
Cerebrovascular hemorrhage	266	1.90 (1.68–2.14)
Atherosclerotic disease	1063	7.59 (7.14–8.06)
Aortic aneurysm or dissection	44	0.31 (0.23–0.42)
Other aneurysm	158	1.13 (0.97–1.32)
Coronary angioplasty	532	3.80 (3.49–4.13)
Coronary artery bypass graft	149	1.06 (0.91–1.25)
Pacemaker	142	1.01 (0.86–1.20)
Valve surgery	118	0.84 (0.70–1.01)
Cardiac transplant	<5	0.03 (0.01–0.08)
Coronary care unit admission	621	4.43 (4.10–4.80)
Arterial embolism	124	0.89 (0.74–1.06)
Death	626	4.47 (4.13–4.83)
Outcome (cardiovascular disease or death)	4024	28.72 (27.85–29.62)

N=95 537.

Sensitivity Analysis

After excluding women with outcomes occurring within 42 days of delivery (n=108), the total number of women in the analysis was 95 429, with 1477 women having premature CVD/mortality within 10 years of follow-up. There was a slight decrease in the discriminatory performance of the model: area under the receiver operating characteristic curve of 0.64 (95% CI, 0.63–0.66) for 10-year risk prediction.

Exclusion of women with preexisting hypertension and diabetes mellitus (n=13 322) resulted in a population of 82 215 women, of which 1133 had an outcome within 10 years of follow-up. This resulted in decreased discriminatory performance of the model with area under the receiver operating characteristic curve of 0.62 (95% CI, 0.61–0.64) for 10-year risk prediction.

Table 3. Final Model With Coefficients for Prediction of Cardiovascular Risk Among Women With Hypertensive Disorders of Pregnancy, Quebec, 1989 to 2018

Variables	Coefficients
Maternal age*	
Socioeconomically deprived	
No	Reference
Yes	0.252
Substance use	
No	Reference
Yes	0.694
Gestational age at delivery*	
Obesity	
No	Reference
Yes	0.260
Diabetes mellitus	
None	Reference
GDM	0.027
Preexisting diabetes mellitus	0.590
Cesarean delivery	
No	Reference
Yes	0.031
Previous complications	
No, multiparous	Reference
Yes, multiparous	−0.047
Primiparous	−0.206
Type of HDP	
Preexisting hypertension	Reference
Gestational hypertension	−0.538
Preeclampsia/HELLP	−0.567
Preeclampsia superimposed on preexisting hypertension	0.022
Unspecified hypertension	−0.424
NICU admission days*	
GDM*cesarean section	−0.004
Preexisting diabetes mellitus*cesarean section	0.446

GDM indicates gestational diabetes mellitus; HELLP, hemolysis, elevated liver enzymes, and low platelet count; HDP, hypertensive disorders of pregnancy; and NICU, neonatal intensive care unit.

*Variables fit as splines, therefore no coefficients.

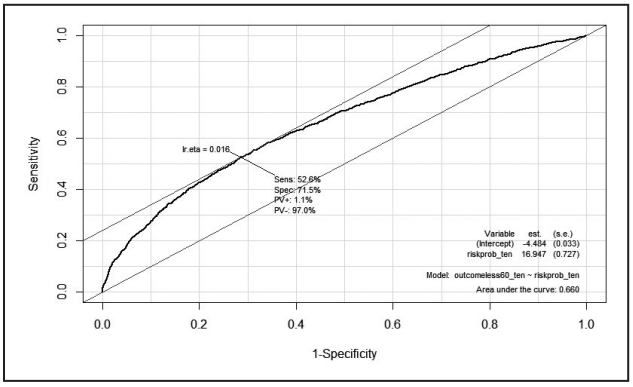

Figure 2. Area under the receiver operating characteristic curve (AUROC) showing model discrimination (0.66 95% CI: 0.65-0.66) for 10-year prediction of cardiovascular disease or death.

DISCUSSION

Summary

In this study, we assessed whether the risk of premature CVD and mortality for women with a history of HDP can be predicted using easily measured factors at the time of delivery. Overall, our model had modest discrimination and calibration, and marginal overfitting, for the 10-year prediction of study outcomes among high-risk women with HDP. The resulting LR was sufficient to accurately rule-in adverse outcomes for the highest-risk category. However, only 3.6% of the total outcomes were captured in this group. Although our model showed strong potential for predicting CVD among women with HDP, its performance, particularly discriminatory and risk stratification abilities, needs to be improved before adoption in clinical practice.

Comparison With Existing Literature

To our knowledge, this is the first study to develop a prediction model using a population-based cohort of only high-risk young women with HDP. In studies that added HDP to preexisting models, the change in discriminatory performance ranged from no increase[25] to only 0.003.[21] The population in these studies differed from ours. In the study by Markovitz et al,[21] women below the age of 40 years at the time of predictor measurements were excluded, while the studies by Timpka et al[25] and Stuart et al[24] assessed models in a combined cohort of women with and without pregnancy complications. To be clinically applicable, a model has to be developed and tested in a sample of women with characteristics similar to those of the population of interest. Therefore, our model development approach better addresses CVD prediction among young women and enhances the understanding of risk assessment and management in the postpartum period for women at high risk of CVD.[27]

Strengths

Our study has important methodological strengths that can help guide the further refinement of sex-specific CVD prediction efforts. First, we developed the model using a large population-based cohort of women representative of Canadians. This allowed for sufficient sample size and power for model development. Quebec is made up of diverse ethnic populations, thus increasing

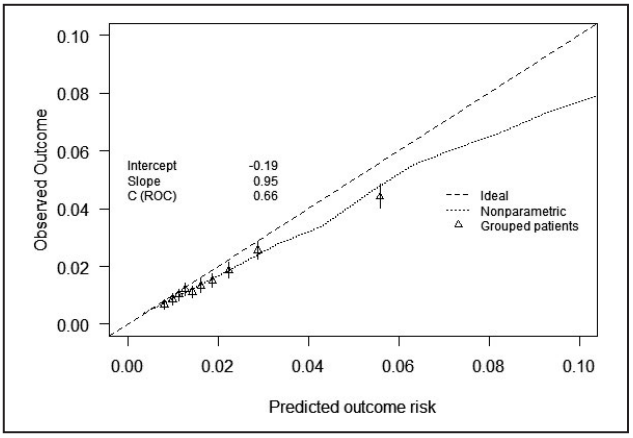

Figure 3. Calibration plot of observed vs predicted 10-year risk of outcome using deciles of predicted probability.
ROC indicates receiver operating characteristic.

the generalizability of our findings. Secondly, we preselected variables based on clinical knowledge that could be associated with CVD and final variable selections were based on techniques recommended by methodological experts (data reduction and Least Absolute Selection and Shrinkage Operator), rather than the use of forward and backward selection procedures known to produce biased estimates.[33] We assessed interactions between variables based on clinical importance such as age and type of pregnancy complications. Thirdly, our data had adequate follow-up, which enabled us to identify women who developed the outcome—an important feature because of the young age at cohort entry. We chose 10 and 15 years for prediction periods because these have been suggested as important time points for interventions to prevent CVD.[20,39] Finally, we followed the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis statement to ensure transparency in modeling and the reporting process.[34]

Limitations

A limitation in our study is the reliance on *ICD* codes for defining predictors, which may not be sensitive for some medical conditions. We tried to minimize this influence by dropping variables known to be inconsistently recorded with *ICD* codes. However, we cannot ignore the possibility of coding errors and potential effect on model performance.

We were also limited by the number of clinically relevant continuous variables available in the data set; for example, we did not have blood pressure or lipid levels or highly sensitive C-reactive protein, which are commonly used in cardiovascular clinical prediction models. The inclusion of continuous data would increase heterogeneity between women, allowing for better model discrimination.[20] In addition, we lacked information on ethnicity/race and family history of CVD, which may also add incremental value.

Implications

The resulting LR+ in the highest-risk group (5.90) suggests that the model would have good clinical utility if risk stratification is improved. If the prediction accuracy of our model is substantially improved (for example, area under the receiver operating characteristic curve ≥0.7 and >50% of total outcome within the highest-risk group), then the model may be used immediately after external validation,[32] that is, after testing in a different data set. Ideally, the model would be applied immediately after delivery, before hospital discharge, to identify which women are at the highest risk of CVD and who might benefit from targeted screening and risk reduction. The ability to capture these women at delivery hospitalization provides a unique opportunity for positive health interventions, including those focused on physical activity, diet, and smoking cessation, as well as pharmacological approaches, following recommended guidelines.[40]

Although the use of administrative data offers many advantages, it also poses several limitations as discussed above. We acknowledge the possibility that measurements available in administrative health data may not be sufficient to predict CVD in this population of women with HDP. Further studies are required to test whether the addition of paraclinical variables, such as blood pressure, would increase model performance. If such clinical variables would add incremental value to our model, we believe that a useful clinical prediction tool based on easily established variables is a real possibility in postpartum women.

Table 4. Risk Classification and Stratification Table (N=95 537)

Calculated Risk Probability	No. Women in Risk Group (%)	No. Women With Outcomes (%)	No. Women Without Outcomes (%)	Likelihood Ratio (95% CI)
<0.035	86 546 (90.6)	1176 (1.4)	85 370 (98.6)	0.82 (0.73–0.92)
≥0.035 to <0.045	4027 (4.2)	126 (3.1)	3901 (96.9)	1.91 (1.61–2.27)
≥0.045 to <0.07	3363 (3.5)	138 (4.1)	3225 (95.9)	2.54 (2.15–3.00)
≥0.07	1601 (1.7)	145 (9.1)	1456 (90.9)	5.90 (5.01–6.95)
Total	95 537	4024	91 513	

Of note, there was a slight decrease in model discriminatory performance when women with preexisting hypertension and diabetes mellitus were excluded. This suggests that women with preexisting diseases are easier to distinguish from women without preexisting conditions, and supports the notion that disease risk is higher in women with premorbid conditions. Therefore, preexisting conditions are important to include in CVD prediction models.

CONCLUSIONS

Our clinical prediction model developed using administrative health data alone can reasonably predict the presence of premature CVD or mortality among women with a history of HDP. However, further model refinement is needed to reliably stratify women into low and high risk of developing CVD. Upon model improvement and external validation, the goal is to ultimately develop a risk calculator for use in clinical practice to guide cardiovascular risk reduction in women with complicated pregnancies.

ARTICLE INFORMATION

Received May 7, 2020; accepted August 27, 2020.

Affiliations

From the Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Quebec, Canada (U.V.U., N.D., N.A., R.W.P.); Institut national de santé publique du Québec, Montreal, Quebec, Canada (U.V.U., N.A., S.H.); Research Institute - McGill University Health Centre, Montreal, Quebec, Canada (N.D., R.W.P.); University of Montreal Hospital Research Centre, Montreal, Quebec, Canada (N.A., S.H.); Department of Social and Preventive Medicine, School of Public Health, University of Montreal, Quebec, Canada (N.A.); Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, Quebec, Canada (R.W.P.); and Department of Pediatrics, McGill University, Montreal, Quebec, Canada (R.W.P.).

Sources of Funding

This study was supported by the Canadian Institutes of Health Research Foundation grant (FDN-143297) and the Heart & Stroke Foundation of Canada (G-18-0021776). U. Vivian Ukah is supported by the Fonds de recherche du Québec-Santé (FRQS) postdoctoral award. Natalie Dayan (Junior Track 1) and Nathalie Auger (Junior Track 2) acknowledge Clinical Research Scholar career awards from FRQS.

Disclosures

None.

Supplementary Materials

Tables S1–S2

REFERENCES

1. Roth GA, Johnson C, Abajobir A, Abd-Allah F, Abera SF, Abyu G, Ahmed M, Aksut B, Alam T, Alam K, et al. Global, regional, and national burden of cardiovascular diseases for 10 causes, 1990 to 2015. *J Am Coll Cardiol*. 2017;70:1–25.
2. Shaw LJ, Bairey Merz CN, Pepine CJ, Reis SE, Bittner V, Kelsey SF, Olson M, Johnson BD, Mankad S, Sharaf BL, et al. Insights from the NHLBI-sponsored Women's Ischemia Syndrome Evaluation (WISE) study: part I: gender differences in traditional and novel risk factors, symptom evaluation, and gender-optimized diagnostic strategies. *J Am Coll Cardiol*. 2006;47:S4–S20.
3. Appelman Y, van Rijn BB, Ten Haaf ME, Boersma E, Peters SA. Sex differences in cardiovascular risk factors and disease prevention. *Atherosclerosis*. 2015;241:211–218.
4. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation*. 1998;97:1837–1847.
5. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Brindle P. Performance of the QRISK cardiovascular risk prediction algorithm in an independent UK sample of patients from general practice: a validation study. *Heart*. 2008;94:34–39.
6. Conroy RM, Pyörälä K, Fitzgerald AP, Sans S, Menotti A, De Backer G, De Bacquer D, Ducimetière P, Jousilahti P, Keil U, et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J*. 2003;24:987–1003.
7. Baart SJ, Dam V, Scheres LJJ, Damen JAAG, Spijker R, Schuit E, Debray TPA, Fauser BCJM, Boersma RE, Moons KGM, et al. Cardiovascular risk prediction models for women in the general population: a systematic review. *PLoS One*. 2019;14:e0210329.
8. Damen JAAG, Hooft L, Schuit E, Debray TPA, Collins GS, Tzoulaki I, Lassale CM, Siontis GCM, Chiocchia V, Roberts C, et al. Prediction models for cardiovascular disease risk in the general population: systematic review. *BMJ*. 2016;353:i2416.
9. Magee LA, Pels A, Helewa M, Rey E, von Dadelszen P; Canadian Hypertensive Disorders of Pregnancy (HDP) Working Group. Diagnosis, evaluation, and management of the hypertensive disorders of pregnancy. *Pregnancy Hypertens*. 2014;4:105–145.
10. Williams D. Pregnancy: a stress test for life. *Curr Opin Obstet Gynecol*. 2003;15:465–471.
11. Bellamy L, Casas J, Hingorani AD, Williams DJ. Pre-eclampsia and risk of cardiovascular disease and cancer in later life: systematic review and meta-analysis. *BMJ*. 2007;335:974–977.
12. Auger N, Fraser WD, Schnitzer M, Leduc L, Healy-Profitós J, Paradis G. Recurrent pre-eclampsia and subsequent cardiovascular risk. *Heart*. 2017;103:235–243.
13. Grandi SM, Filion KB, Yoon S, Ayele HT, Doyle CM, Hutcheon JA, Smith GN, Gore GC, Ray JG, Nerenberg K, et al. Cardiovascular disease-related morbidity and mortality in women with a history of pregnancy complications: systematic review and meta-analysis. *Circulation*. 2019;139:1069–1079.
14. Brown MC, Best KE, Pearce MS, Waugh J, Robson SC, Bell R. Cardiovascular disease risk in women with pre-eclampsia: systematic review and meta-analysis. *Eur J Epidemiol*. 2013;28:1–19.
15. Haas DM, Ehrenthal DB, Koch MA, Catov JM, Barnes SE, Facco F, Parker CB, Mercer BM, Bairey-Merz CN, Silver RM, et al. Pregnancy as a window to future cardiovascular health: design and implementation of the nuMoM2b heart health study. *Am J Epidemiol*. 2016;183:519–530.
16. Rich-Edwards JW, Fraser A, Lawlor DA, Catov JM. Pregnancy characteristics and women's future cardiovascular health: an underused opportunity to improve women's health? *Epidemiol Rev*. 2014;36:57–70.
17. Zhou X, Niu J, Ji W, Zhang Z, Wang PP, Ling XB, Li Y. Precision test for precision medicine: opportunities, challenges and perspectives regarding pre-eclampsia as an intervention window for future cardiovascular disease. *Am J Transl Res*. 2016;8:1920.
18. Dayan N, Udell JA. Moving toward sex-specific cardiovascular risk estimation. *Can J Cardiol*. 2020;36:13–15.
19. Freibert SM, Mannino DM, Bush H, Crofford LJ. The association of adverse pregnancy events and cardiovascular disease in women 50 years of age and older. *J Womens Health*. 2011;20:287–293.
20. Gunnarsson OS, Timpka S. Pregnancy complication history in 10-year cardiovascular disease risk prediction: a review of recent evidence. *Curr Epidemiol Rep*. 2019;6:321–328.
21. Markovitz AR, Stuart JJ, Horn J, Williams PL, Rimm EB, Missmer SA, Tanz LJ, Haug EB, Fraser A, Timpka S, et al. Does pregnancy complication history improve cardiovascular disease risk prediction? Findings from the HUNT study in Norway. *Eur Heart J*. 2019;40:1113–1120.
22. Callaway LK, Mamun A, McIntyre HD, Williams GM, Najman JM, Nitert MD, Lawlor DA. Does a history of hypertensive disorders of pregnancy help predict future essential hypertension? Findings

from a prospective pregnancy cohort study. *J Hum Hypertens*. 2013;27:309–314.
23. Gaziano TA, Young CR, Fitzmaurice G, Atwood S, Gaziano JM. Laboratory-based versus non-laboratory-based method for assessment of cardiovascular disease risk: the NHANES I follow-up study cohort. *Lancet*. 2008;371:923–931.
24. Stuart JJ, Tanz LJ, Cook NR, Spiegelman D, Missmer SA, Rimm EB, Rexrode KM, Mukamal KJ, Rich-Edwards JW. Hypertensive disorders of pregnancy and 10-year cardiovascular risk prediction. *J Am Coll Cardiol*. 2018;72:1252–1263.
25. Timpka S, Fraser A, Schyman T, Stuart JJ, Åsvold BO, Mogren I, Franks PW, Rich-Edwards JW. The value of pregnancy complication history for 10-year cardiovascular disease risk prediction in middle-aged women. *Eur J Epidemiol*. 2018;33:1003.
26. Selmer R, Igland J, Ariansen I, Tverdal A, Njølstad I, Furu K, Tell GS, Klemsdal TO. NORRISK 2: a Norwegian risk model for acute cerebral stroke and myocardial infarction. *Eur J Prev Cardiol*. 2017;24:773–782.
27. Grandi SM, Smith GN, Platt RW. The relative contribution of pregnancy complications to cardiovascular risk prediction: are we getting it wrong? *Circulation*. 2019;140:1965–1967.
28. Tranquilli AL, Dekker G, Magee L, Roberts J, Sibai BM, Steyn W, Zeeman GG, Brown MA. The classification, diagnosis and management of the hypertensive disorders of pregnancy: a revised statement from the ISSHP. *Pregnancy Hypertens*. 2014;4:97–104.
29. Joseph KS, Fahey J; Canadian Perinatal Surveillance System. Validation of perinatal data in the discharge abstract database of the Canadian institute for health information. *Chronic Dis Can*. 2009;29:96–100.
30. Harrell FE Jr. *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*. SpringerLink ebooks—Mathematics and Statistics. 2nd ed. New YorkSpringer; 2015. DOI: 10.1007/978-3-319-19425-7.
31. Dzakpasu S, Deb-Rinker P, Arbour L, Darling EK, Kramer MS, Liu S, Luo W, Murphy PA, Nelson C, Ray JG, et al. Severe maternal morbidity in Canada: temporal trends and regional variations, 2003–2016. *J Obstet Gynaecol Can*. 2019;41:1589–1598.e16.
32. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J*. 2014;35:1925–1931.
33. Steyerberg EW. *Clinical Prediction Models: a Practical Approach to Development, Validation, and Updating*. SpringerLink ebooks—Mathematics and Statistics. New York, NYSpringer; 2009.
34. Moons KGM, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, Vickers AJ, Ransohoff DF, Collins GS. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*. 2015;162:W1–W73.
35. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;143:29–36.
36. Janes H, Pepe MS, Gu W. Assessing the value of risk predictions by using risk stratification tables. *Ann Intern Med*. 2008;149:751–760.
37. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. *BMJ*. 2004;329:168–169.
38. R Core Team. *R: a Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2013. Available at: https://www.r-project.org/. Accessed January 6, 2020.
39. Lloyd-Jones DM. Cardiovascular risk prediction: basic concepts, current status, and future directions. *Circulation*. 2010;121:1768–1777.
40. Heida KY, Bots ML, De Groot CJM, van Dunné FM, Hammoud NM, Hoek A, Laven JS, Maas AH, van Lennep JER, Velthuis BK, et al. Cardiovascular risk management after reproductive and pregnancy-related disorders: a Dutch multidisciplinary evidence-based guideline. *Eur J Prev Cardiol*. 2016;23:1863–1879.

SUPPLEMENTAL MATERIAL

Table S1. TRIPOD Checklist: Prediction Model Development and Validation.

Section/ Topic		Checklist Item		Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	Title page
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	3
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	4-5
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	5
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	5
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	5-6
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	5
	5b	D;V	Describe eligibility criteria for participants.	5-6
	5c	D;V	Give details of treatments received, if relevant.	NA
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	6
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	NA
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	7
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	NA
Sample size	8	D;V	Explain how the study size was arrived at.	7-8,
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	NA
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	7-8
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	NA
	10c	V	For validation, describe how the predictions were calculated.	NA

	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	9
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	NA
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	9
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	NA
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	10, Figure 1
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	11, Table 1
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	NA
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	11
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	NA
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	Table 3
	15b	D	Explain how to the use the prediction model.	NA
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	11-12
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	NA
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	14-15
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	NA
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	15
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	15-16
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	S

Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	16

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.

Table S2. Characteristics of women at their first HDP-complicated pregnancy by outcome (cardiovascular disease or death), Quebec, 1989 to 2018 (for women with at least 10 years follow up), N = 58,932.

	No. women (%)	
	No Outcome (N= 3572)	**Outcome (N= 55,360)**
Maternal age, years		
<20	86 (2.4)	2004 (3.6)
20-24	544 (15.2)	11,837 (21.4)
25-29	1107 (31.0)	19,688 (35.6)
30-34	1033 (28.9)	14,125 (27.0)
35-39	653 (18.3)	6307 (11.4)
≥40	149 (4.2)	1399 (2.5)
Multiple pregnancy		
Yes	104 (2.6)	3228 (3.5)
No	3920 (97.4)	88,285 (96.5)
Time period at delivery		
1989–1995	2093 (58.6)	22,448 (40.6)
1996–2002	1097 (30.7)	20,219 (36.5)
2003–2006	381 (10.7)	12,687 (22.9)
Socioeconomically deprived		
Yes	854 (23.9)	10,999 (19.9)
No	2402 (67.3)	40,860 (73.8)
Unknown	316 (8.9)	3501 (6.3)
Substance use*		

Yes	116 (2.9)	1397 (1.5)
No	3908 (97.1)	90,116 (98.5)
Morbid obesity		
Yes	152 (4.3)	1136 (2.1)
No	3420 (95.7)	54,224 (97.9)
Preexisting diabetes		
Yes	329 (9.2)	1760 (3.2)
No	3243 (90.8)	53,600 (96.8)
Primiparity at complicated pregnancy		
Yes	3014 (84.4)	46,560 (84.1)
No	558 (15.6)	8800 (15.9)
Hypertensive disorder of pregnancy sub groups		
Pre-existing hypertension	623 (17.4)	5840 (10.6)
Gestational hypertension	502 (14.1)	11,032 (19.9)
Pre-eclampsia/HELLP syndrome	1639 (45.9)	25,677 (46.4)
Superimposed pre-eclampsia	121 (3.4)	986 (1.8)
Unspecified hypertension	687 (19.2)	11,825 (21.4)
Gestational diabetes		
Yes	367 (10.3)	4025 (7.3)
No	3205 (89.7)	51,335 (92.7)
Cesearean section		
Yes	1348 (37.7)	17,297 (31.2)
No	2224 (62.3)	38,063 (68.8)

Severe maternal morbidity			
Yes	681 (19.1)	10,075 (18.2)	
No	2891 (81.9)	45,285 (81.8)	
Preterm delivery (GA <37 weeks)			
Yes	833 (23.3)	9305 (16.8)	
No	2739 (76.7)	46,055 (83.2)	** *
Stillbirth			
Yes	44 (1.2)	338 (0.6)	
No	3528 (98.8)	55,022 (99.4)	
Low birth weight (<2500g)			
Yes	808 (22.6)	9050 (16.4)	
No	2763 (77.4)	46,305 (83.6)	
Admission into NICU			
Yes	176 (5.0)	2855 (5.2)	
No	3360 (95.0)	52,145 (94.8)	
Neonatal death			
Yes	27 (0.8)	167 (0.3)	
No	3509 (99.2)	54,833 (99.7)	
Previous pregnancy complication[±] **(among multiparous women)**			
Yes	128 (3.6)	1574 (2.8)	
No	430 (12.0)	7226 (13.1)	

Abbreviations: GA, gestational age, NICU, neonatal intensive care unit

*Substance use incudes tobacco, alcohol, and illicit drug use

± Previous complications include any previous stillbirth, preterm delivery, low birthweight, admission to NICU, neonatal death, or severe maternal morbidity