Predicting complications in hypertensive disorders of pregnancy: external validation of a prognostic model for adverse perinatal outcomes

Klaartje M. Olde Loohuis, MD; Kim Luijken, MSc, PhD; Hannah Brown Amoakoh, MD, PhD; Kwame Adu-Bonsaffoh, MD, PhD; Diederick E. Grobbee, MD, PhD, FESC; Kerstin Klipstein-Grobusch, MSc PhD; Emmanuel Srofenyoh, MD; Mary Amoakoh-Coleman, MD, PhD; Joyce L. Browne, MD, PhD

BACKGROUND: Prediction models can be used as simple evidence-based tools to identify fetuses at risk of perinatal death. Payne et al developed a prognostic model for perinatal death in women with hypertensive disorders of pregnancy, a leading cause of maternal/fetal morbidity and mortality.

OBJECTIVE: This study aimed to externally validate the predictive performance of this model in pregnant women with hypertensive disorders of pregnancy admitted between 26 and 34 weeks of gestation in Ghana.

STUDY DESIGN: The perinatal model was applied in the SPOT (Severe Pre-eclampsia adverse Outcome Triage) study, a cohort of women with hypertensive disorders of pregnancy admitted between 26 and 34 weeks of gestation to referral facilities in Ghana. Predictive performance was assessed by calibration (calibration-in-the-large coefficient and calibration slope) and discrimination (based on the c-statistic).

RESULTS: Of the 543 women included in the validation analysis, 87 (16%) experienced perinatal death from delivery until hospital discharge. Predictive performance of the model was poor. The calibration-in-the-large coefficient was 1.12 (95% confidence interval, 0.87−1.36, 0 for good calibration), calibration slope was 0.08 (95% confidence interval, −0.21 to 0.36, 1 for good calibration), and c-statistic was 0.52 (95% confidence interval, 0.44−0.59).

CONCLUSION: This perinatal prediction model performed poorly in this cohort in Ghana. Possible reasons include differences in case mix, clinical management strategies, or data collection procedures between development and validation settings; suboptimal modeling strategies at development; or omission of important predictors. Given the burden of perinatal mortality and importance of risk stratification, new prediction model development and validation is recommended.

Key words: external validation, hypertensive disorders of pregnancy, perinatal death, preeclampsia, pregnancy, prognosis, risk prediction

From the Department of Global Public Health and Bioethics, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands (Olde Loohuis, Brown Amoakoh, Adu-Bonsaffoh, Grobbee, Klipstein-Grobusch, Amoakoh-Coleman, and Browne); Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands (Luijken); Department of Epidemiology, Noguchi Memorial Institute for Medical Research, University of Ghana, Accra, Ghana (Brown Amoakoh and Amoakoh-Coleman); Department of Obstetrics and Gynaecology, University of Ghana Medical School, Accra, Ghana (Adu-Bonsaffoh); Department of Obstetrics and Gynaecology, Korle-Bu Teaching Hospital, Accra, Ghana (Adu-Bonsaffoh); Division of Epidemiology and Biostatistics, School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa (Klipstein-Grobusch); Department of Obstetrics and Gynaecology, Greater Accra Regional Hospital, Ghana Health Service, Accra, Ghana (Srofenyoh)

The authors report no conflict of interest.

This study was supported by the Gates Foundation (grant number OPP1197342) and the Dutch Research Council (grant number NWA.1160.18.159).

All participants provided written informed consent before their inclusion in the study. For participants aged <18 years, parental or guardian consent was obtained in accordance with ethical guidelines.

Cite this article as: Olde Loohuis KM, Luijken K, Brown Amoakoh H, et al. Predicting complications in hypertensive disorders of pregnancy: external validation of a prognostic model for adverse perinatal outcomes. Am J Obstet Gynecol Glob Rep 2025;XX:x.ex−x.ex.

Corresponding author: Klaartje M. Olde Loohuis, MD. k.m.oldeloohuis@umcutrecht.nl

2666-5778/$36.00

http://dx.doi.org/10.1016/j.xagr.2025.100455

Introduction

The global burden of stillbirth and neonatal mortality remains high, with an estimated yearly 2.6 million stillbirths[1] and 2.5 million neonatal deaths.[2] The vast majority of these deaths occur in South Asia and sub-Saharan Africa.[2,3] Between 6% and 20% of all perinatal deaths are associated with hypertensive disorders of pregnancy (HDPs).[4] In Ghana, HDPs are one of the leading causes of maternal and perinatal deaths.[5,6] Estimates suggest that more than half of all perinatal deaths are preventable and that major progress could be achieved by improving both the coverage and quality of care at birth and for newborns.[7]

HDPs consist of a spectrum of diseases including chronic hypertension, gestational hypertension, preeclampsia, superimposed preeclampsia, HELLP (hemolysis, elevated liver enzymes, and low platelet count) syndrome, and

AJOG Global Reports at a Glance

Why was this study conducted?

This study aimed to externally validate Payne et al's prognostic model for predicting perinatal death in women with hypertensive disorders of pregnancy using data from a low- and middle-income country (LMIC) setting.

Key findings

The model had poor predictive performance in the validation cohort from Accra, Ghana. Potential explanations for this include differences in case mix, clinical management, or data collection procedures, as well as model overfitting and the omission of important predictors.

What does this add to what is known?

This study highlights the need for the external validation of prediction models in diverse LMIC settings, and emphasizes the importance of context-specific model recalibration and standardized data collection practices to improve maternal and perinatal outcomes.

eclampsia.[8] Pregnancies complicated by severe and early-onset HDPs (ie, before 34 weeks of pregnancy), are at the highest risk of (serious) adverse maternal and perinatal complications, and delivery is the only ultimate cure.[9–11] However, for clinically stable women with severe early-onset HDPs, there is currently no consensus on whether management should be expectant or focused on inducing the delivery.[12–14] Furthermore, these different management strategies might involve conflicting maternal and fetal risks and benefits.[14] Therefore, risk prediction models could assist health care providers and pregnant women in making the right decision about their management strategy.

Clinical decision support could improve the quality of health care, especially in resource-constrained settings, by guiding the provision of health care and delivering effective triage. In these settings, few risk prediction models have been developed so far. Ngwenya et al[15] (2020) developed a model to predict both maternal and perinatal outcomes in a population exclusively comprising women with severe preeclampsia. Another prognostic model by Payne et al[16] for predicting perinatal mortality in women with HDPs was developed in a multicenter prospective cohort of women from Fiji, Uganda, South Africa, Brazil, and Pakistan. Predictors were maternal age, a count of symptoms, and dipstick proteinuria. With an area under the receiver operating characteristic curve of 0.75 (95% confidence interval [CI], 0.71–0.80), the model performed moderately well with good internal calibration. A potential limitation of this model could be that it did not comprise all known risk factors for perinatal outcomes because it was developed by performing a secondary analysis of a cohort established to predict adverse maternal outcomes. The predictive performance of this model in a new data set remains unknown given that, to our knowledge, no external validation studies have yet been performed.

Thus, in the current study, we externally validated the prognostic model by Payne et al[16] for predicting perinatal mortality in women with HDPs remote from term in a low-resource (peri)urban setting in Ghana.

Methods

Study design, setting, and population

This research is embedded in the ongoing SPOT (Severe Pre-eclampsia adverse Outcome Triage) studies, which include the SPOT cohort. The SPOT cohort is a multicenter prospective cohort that is being conducted in different referral hospitals in the Greater Accra Region and the Eastern Region of Ghana. The SPOT cohort commenced in November 2017, and data were collected until November 2023.

The participating referral hospitals are the Greater Accra Regional Hospital, Tema General Hospital, La General Hospital, Lekma General Hospital, Eastern Regional Hospital, and Korle-Bu Teaching Hospital. Four of these hospitals are within the Greater Accra Region, and the Eastern Regional Hospital is in the Eastern Region of Ghana. These hospitals were selected because of their referral function, the availability of a neonatal intensive care unit, their large volume of patients/workload, and sufficient infrastructure to conduct this study. Furthermore, these settings are relevant because of the availability of diagnostic and prognostic tests for prognostic modeling. Together, the hospitals have an estimated annual number of deliveries of >30,000. The estimated incidence of HDPs within these facilities is between 8% and 39%.[17,18] Women aged ≥16 years admitted to a hospital with HDPs (classification included in Table 1[19–21]) between the gestational ages of 26 and 34 weeks were eligible for inclusion. Exclusion criteria were stillbirth before hospital admission, spontaneous active labor during admission, and/or the presence of any of the severe maternal outcomes before inclusion or before collecting the predictor variables because the prognostic model would not be applied in such cases. The maternal outcomes were maternal mortality, serious morbidity of the central nervous system, cardiorespiratory and renal morbidity, hepatic morbidity, hematologic morbidity, or other serious morbidities (further details are included in Supplemental File S1: maternal adverse outcomes). The only maternal outcome we chose not to exclude was maternal eclampsia during admission. This decision was based on the expectation that eclampsia would neither interfere with the model's predictive accuracy for perinatal risk assessment nor diminish the clinical relevance of including women with eclampsia at admission. During admission, all women received standard care according to their hospital's guidelines.

TABLE 1
Classification of hypertensive disorders of pregnancy applied in the SPOT study[a]

Disorder	Definition
Preeclampsia	Blood pressure ≥140/90 mm Hg (at least 1 component, twice, ≥4 h apart, ≥20+0 wk) and either: • Proteinuria (of ≥2+ by dipstick, ≥0.3 g per day by 24-h collection, or ≥30 g/mol by urinary protein-to-creatinine ratio), or • Hyperuricemia (greater than upper limit of nonpregnancy normal range)
Superimposed preeclampsia	Rapidly increasing requirements for antihypertensive drugs, systolic blood pressure >170 mm Hg or diastolic blood pressure >120 mm Hg, new proteinuria, or new hyperuricemia
Severe preeclampsia	• Elevated blood pressure, systolic ≥160 mm Hg or diastolic ≥110 mm Hg, at least 1 component, twice, ≥4 h apart, ≥20+0 wk; and proteinuria (of ≥2+ by dipstick, ≥0.3 g per day by 24-h collection, or ≥30 g/mol by urinary protein-to-creatinine ratio) • HELLP syndrome: hemolysis, elevated liver enzymes, and low platelets; even in the absence of hypertension or proteinuria[19]; preeclampsia with severe symptoms (headache, blurred vision, right upper quadrant pain, etc.)
Gestational hypertension	Blood pressure ≥140/90 mm Hg (at least 1 component, twice, ≥4 h apart, ≥20+0 wk) without significant proteinuria
Chronic hypertension	Blood pressure ≥140/90 mm Hg before 20+0 wk of gestation
Partial HELLP	Hemolysis and low platelets OR low platelets and elevated liver enzymes[20]
Eclampsia	The presence of preeclampsia and seizures

SPOT, Severe Pre-eclampsia adverse Outcome Triage.

[a] For these definitions, we used the ISSHP (International Society for the Study of Hypertension in Pregnancy) classification[19] together with definitions used in the miniPIERS/fullPIERS studies.[20,21]

Olde Loohuis. Predicting complications in hypertensive disorders of pregnancy. Am J Obstet Gynecol Glob Rep 2025.

Some women were admitted to hospital multiple times during pregnancy (n=126). For these women, only predictor measurements at the first hospital admission were used to predict adverse perinatal outcomes occurring at any time until final hospital discharge of the mother.

Outcomes

The primary perinatal outcome in this study was a composite outcome of stillbirth and early neonatal death (up to discharge of the mother from the hospital) from any cause. Stillbirth was defined using the existing World Health Organization definition as the death of an infant after the gestational age of 28 weeks but before or during birth.[4] Early neonatal death was defined as "neonatal death up to one week after birth."[22] In this analysis, the same definition as that used in the model development study —"neonatal death up to hospital discharge"—was applied to facilitate more accurate validation. To identify outcomes, clinical files (during admission) and hospital discharge data were used.

Predictors

This external validation study used the predictor variables from Payne's perinatal death model: maternal age, count of symptoms, and dipstick proteinuria (Table 2). In line with the methodology used in the model development study, data on predictors were gathered from the time of admission up to 24 hours after admission. Symptoms included in the symptom count were headache, visual disturbances, chest pain, dyspnea, abdominal pain, and nausea and vomiting. Proteinuria was assessed using dipstick analysis of maternal urine.

Study procedures and data quality

In this study, data were collected from both electronic and paper-based medical records, and supplemented with face-to-face interviews of women with HDPs using standardized data collection forms administered by trained research assistants. Regular data quality checks were performed by local researchers. To optimize data quality, the paper questionnaires were subsequently double-entered into a REDCap (Vanderbilt University, Nashville, TN) database.[23] Further details of the study procedures are provided on the SPOT study website.[24]

Ethical approval

The SPOT study protocol was approved by the Ghana Health Service Ethics

TABLE 2
Predictors included in the prognostic model predicting adverse perinatal outcomes

Predictors	Data format
Maternal age	Years (Continuous)
Count of symptoms	0, 1, 2 (Categorical)
Dipstick proteinuria	Negative, 1+, 2+, 3+, 4+ (Categorical)

Olde Loohuis. Predicting complications in hypertensive disorders of pregnancy. Am J Obstet Gynecol Glob Rep 2025.

Review Committee (protocol ID GHSERC-GHSERC015/09/17) and the Ethical and Protocol Review Committee of the College of Health Sciences, University of Ghana (protocol ID GHSERC-CHS-EtM.4-P1.2/2017-2018). All participants provided written informed consent before study enrollment. Where applicable, consent of a parent or guardian was also obtained for women aged <18 years.

We followed the TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis) guidelines in our reporting.[25] Details are presented in Supplemental Table S9.

Statistical analysis

Missing data. Missing data were observed across variables, generally ranging from 0% to 4.8%, except for body mass index (BMI) at enrollment (37.2%) (Supplemental Table S3 includes detailed information). Missing predictor values were imputed because we assumed that when the prognostic model is used in clinical practice, all necessary data for the model will be collected. We assumed that missing values were missing at random and used multiple imputation by chained equations (MICE). The number of imputed data sets was 50. We found this reasonable given that the highest proportion of missing data was 37%.[26] Missing values were predicted on the basis of all other predictors, the outcome, and additional patient characteristics (Supplemental File S3 includes details on missing values and imputation techniques). Patients with a missing outcome were excluded after imputation (n=53).[27]

Description of baseline characteristics. Baseline characteristics of participants were summarized using descriptive statistics for the entire validation cohort and stratified by perinatal death status.

Model performance. The equation of the perinatal death model[16] is as follows:

$$\begin{aligned}\text{Logit (perinatal death)} = {} & -4.75 + 0.024 \\ & \times \text{maternal age} + 0.389 \times \text{presence of} \\ & 1\ \text{symptom} + 1.338 \times \text{presence of} \\ & \geq 2\ \text{symptoms} + 1.119 \\ & \times \text{dipstick proteinuria } 2+ \text{ or } 3+ \\ & + 1.457 \times \text{dipstick proteinuria } 4+\end{aligned}$$

To evaluate predictive performance, we assessed calibration, discrimination, and overall accuracy. Calibration refers to the alignment between the predicted and observed outcomes, and was evaluated on the basis of the calibration-in-the-large coefficient and calibration slope, which were pooled across imputed data sets on a logit scale using Rubin's rules.[28] In addition, we presented a smooth calibration curve, created using locally estimated scatterplot smoothing (LOESS)-smoothed plots on stacked imputed data sets to visualize estimated risks vs observed proportions.[29]

To assess discriminative performance, we calculated the c-statistic to determine how well the estimated risks differentiate between women who experienced perinatal death and those who did not. The c-statistic is equal to the area under the receiver operating characteristic curve in a binary logistic model.[30] Standard errors were estimated using the DeLong method, and estimates were pooled across imputed data sets on a logit scale using Rubin's rules.

Model updating. Subsequently, we updated the model by 2 additional steps.[31] First, we adjusted the intercept, which might improve model performance in a validation setting with a different average outcome risk. We expected a higher risk of perinatal death due to the early gestational age at which women presented with symptoms of HDPs relative to the development cohort. The intercept was updated by adding the calibration-in-the-large coefficient to the linear predictor of the model. Second, we adjusted all predictor regression coefficients by a single overall adjustment factor based on the calibration slope, after which the intercept was updated accordingly.

All statistical analyses were performed in R, version 4.0.3 (R Foundation for Statistical Computing, Vienna, Austria).

Sample size calculation. Guidance for sample sizes in external validation studies was not yet developed at the time when our study was planned.[32] Therefore, we based our planning on the recommendations by Vergouwe et al[33] and Steyerberg,[34] which indicate that at least 100 (but preferably ≥250) events are required for valid external validation. We expected a stillbirth rate of approximately 14%, based on a previous study reporting on maternal near-miss and perinatal outcomes in the SPOT cohort.[35] This is consistent with the perinatal mortality rate reported in an Ethiopian study.[36] Therefore, we expected that a sample size of 600 to 800 women with HDPs in our cohort would be sufficient for externally validating the model.

Additional analyses

In addition, because of the differences in gestational age between the validation cohort and the development cohort, we validated the model in a restricted cohort of similar gestational age (from 32+1 weeks). Furthermore, we stratified women according to their HDP diagnoses on admission (preeclampsia vs other diagnoses) to assess differences in adverse outcome risks and their effect on the model validation.

Results

Descriptive statistics

Eligibility criteria were met by 596 women, of whom 53 were excluded from analysis after missing data imputation because of a missing outcome (Table 3 and Supplemental Table S3 present cohort descriptive characteristics before imputation). The mean age of the 543 included women was 32 years (SD, 5.6 years). Mean gestational age on admission was 214 days (SD, 15.4), corresponding to 31 weeks of pregnancy. Most (74%) of the participating women were multiparous, and 49% of women

TABLE 3
Baseline characteristics measured within 24 hours of hospital admission in the eligible cohort from the ongoing SPOT study

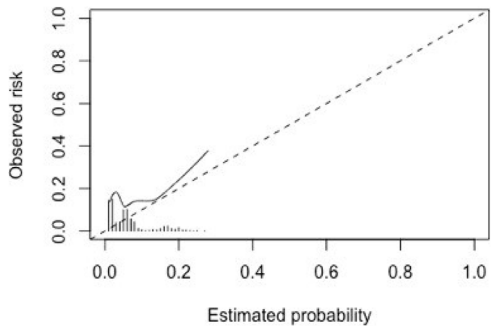
Variable	Overall cohort (N=543)	Women who experienced perinatal death (n=87)	Women who did not experience perinatal death (n=456)
Maternal age (y), mean (SD)	32.0 (5.7)	33.0 (6.1)	31.9 (5.6)
≤18, n (%)	6 (1.1)	0 (0)	6 (1.3)
18–24, n (%)	44 (8.1)	8 (9.2)	36 (7.9)
25–29, n (%)	127 (23.4)	16 (18.4)	111 (24.3)
30–34, n (%)	178 (32.8)	23 (26.4)	155 (34)
≥35, n (%)	188 (34.6)	40 (46.0)	149 (32.7)
Gestational age (d), mean (SD)	214.0 (15.4)	215.3 (15.3)	213.8 (15.4)
Multiple parity, n (%)	401 (73.9)	68 (78.2)	333 (73.1)
Gravidity, n (%)			
1	64 (11.8)	7 (8.0)	57 (12.5)
2	75 (13.8)	15 (17.2)	60 (13.2)
3	134 (24.7)	12 (13.8)	121 (26.5)
4	128 (23.6)	24 (27.6)	103 (22.6)
5	61 (11.2)	13 (14.9)	48 (10.5)
6	45 (8.3)	5 (5.7)	40 (8.8)
7	22 (4.1)	8 (9.2)	14 (3.1)
8+	15 (2.8)	2 (2.3)	13 (2.9)
Singleton, n (%)	508 (93.6)	82 (94.3)	427 (93.6)
BMI (kg/m^2), mean (SD)	33.1 (7.1)	32.1 (7.6)	33.2 (7)
Smoking, n (%)	1 (0.2)	0 (0)	1 (0.2)
Count symptoms, n (%)			
1	164 (30.2)	24 (27.6)	139 (30.5)
2	73 (13.4)	15 (17.2)	58 (12.7)
3	29 (5.3)	5 (5.7)	24 (5.3)
4	10 (1.8)	1 (1.1)	9 (2)
5	3 (0.6)	0 (0)	3 (0.7)
Highest systolic blood pressure within 24 h after admission (mm Hg), mean (SD)	162 (22.5)	161 (23.2)	162 (22.4)
Highest diastolic blood pressure within 24 h after admission (mm Hg), mean (SD)	102 (17.1)	102 (16.9)	103 (17.2)
Diagnosis at hospital admission	??	??	??
Chronic hypertension or gestational hypertension, n (%)	138 (25.4)	23 (26.4)	115 (25.2)
Mild or severe preeclampsia, n (%)	361 (66.5)	58 (66.7)	303 (66.4)
Eclampsia, n (%)	44 (8.1)	6 (6.9)	38 (8.3)

Information is based on imputed data sets generated using Rubin's rules (Supplemental File includes descriptive statistics of the original data).

BMI, body mass index; *SPOT*, Severe Pre-eclampsia adverse Outcome Triage.

Olde Loohuis. Predicting complications in hypertensive disorders of pregnancy. Am J Obstet Gynecol Glob Rep 2025.

FIGURE
Smooth calibration curve of Payne's model in SPOT validation cohort



The curve was obtained by predictions over stacked imputed data sets.

SPOT, Severe Pre-eclampsia adverse Outcome Triage.

Olde Loohuis. Predicting complications in hypertensive disorders of pregnancy. Am J Obstet Gynecol Glob Rep 2025.

experienced no symptoms within the first 24 hours of admission. The mean highest systolic blood pressure within the first 24 hours of admission was 162 mm Hg (SD, 22.4). The mean highest diastolic blood pressure within the first 24 hours of admission was 102 mm Hg (SD, 17.1). A total of 87 women (16%) experienced a perinatal death outcome. Table 3 presents descriptive statistics comparing women with and without an observed outcome.

The descriptive statistics of the validation data set and the development data set were broadly similar (Supplemental File S5). The perinatal death rate in our data set was higher than that in the development data set (15% in SPOT vs 7% in development cohort).

Distribution of estimated risks

The estimated risks of adverse perinatal outcomes over all 50 imputed data sets ranged from 0.01 to 0.28, with an average of 0.06 (Figure). Supplemental Figure S6 presents the distribution of the linear predictor, which can inform sample size calculations for future external validations.

Performance of the model

The predictive performance of the model was poor (Table 3). The calibration-in-the-large coefficient was 1.12 (95% CI, 0.87−1.36), indicating that the average risk of perinatal death was underestimated. The calibration slope of 0.08 (95% CI, −0.21 to 0.36) suggests that the estimated risks were excessively divergent from the observed risk, resembling a scenario where a model is overfitted to the development data. The smooth calibration curve is shown in the Figure. The c-statistic of the model was 0.52 (95% CI, 0.44−0.59), indicating poor discrimination between women who experienced a perinatal death and those who did not.

Model updating

Subsequently, the model was updated as planned, but predictive performance was not restored, as indicated by updated smooth calibration curves (Supplemental Figure S7). The performance of the model after updating the intercept and coefficients with a single factor is presented in Table 4.

Additional analysis

We assessed model performance in stratified groups based on gestational age (>32 or <32 weeks) and HDP diagnosis (preeclampsia vs other diagnoses of HDP). Performance of the model remained poor (Table 5; Supplemental Figure S8).

Discussion

This external validation study assessed the performance of Payne's prognostic model for adverse perinatal outcomes[16] in the SPOT cohort of pregnant women with HDPs admitted to hospital at 26 to 34 weeks of gestation. Despite a c-statistic of 0.75 in the development cohort, model performance at external validation was suboptimal with both poor calibration and discrimination. Even after updating the intercept and coefficients, performance remained suboptimal. Furthermore, model performance was not improved by stratifying for different gestational age and HDP diagnoses.

In the evaluation of a prediction model and before its implementation in clinical practice, external validation is

TABLE 4
Performance of the original prognostic model for predicting adverse perinatal outcomes in the basic model and after updating

Performance measure	Model		
	Basic model	Model after updating intercept	Model after updating intercept and coefficient
Calibration-in-the-large coefficient (95% CI)	1.12 (0.87−1.36)	0 (−0.24 to 0.24)	0 (−0.22 to 0.22)
Calibration slope (95% CI)	0.08 (−0.21 to 0.36)	0.08 (−0.21 to 0.36)	1 (−2.8 to 4.8)
C-statistic (95% CI)	0.52 (0.44−0.59)	0.52 (0.44−0.59)	0.52 (0.44−0.59)

CI, confidence interval.

Olde Loohuis. Predicting complications in hypertensive disorders of pregnancy. Am J Obstet Gynecol Glob Rep 2025.

TABLE 5
Performance of the prognostic model predicting adverse perinatal outcomes in subgroups

Performance measure	Gestational age ≤32 wk (n=371, with n=57 perinatal deaths)	Gestational age >32 wk (n=172, with n=30 perinatal deaths)	Preeclampsia (n=405, with n=64 perinatal deaths)	Other diagnoses of HDP (n=138, with n=23 perinatal deaths)
Calibration-in-the-large coefficient (95% CI)	1.05 (0.76−1.35)	1.25 (0.83−1.66)	0.92 (0.64−1.2)	1.85 (1.38−2.32)
Calibration slope (95% CI)	0.12 (−0.24 to 0.48)	−0.01 (−0.52 to 0.50)	0.13 (−0.23 to 0.49)	0.04 (−0.68 to 0.75)
C-statistic (95% CI)	0.53 (0.44−0.62)	0.49 (0.36−0.62)	0.53 (0.44−0.61)	0.57 (0.44−0.68)

CI, confidence interval; *HDP*, hypertensive disorders of pregnancy.

Olde Loohuis. Predicting complications in hypertensive disorders of pregnancy. Am J Obstet Gynecol Glob Rep 2025.

an essential step. In general, predictive performance tends to decrease in new data sets because of overfitting during model development.[31,34] Findings from 2 recent reviews show that in low- and middle-income country (LMIC) settings, limited prediction studies have been performed within maternal health care.[37,38] Only the study by Ukah et al[39] used external validation to assess performance of a model predicting maternal outcomes in women with HDP in different settings.[21]

There are several possible explanations for the suboptimal external performance of the prognostic model predicting adverse perinatal outcomes.

First, gestational age at the first presentation of HDP is an important prognostic factor for adverse perinatal outcomes, but it was not included in the model.[40−42] Incorporating gestational age as a predictor in a prognostic model may enhance its overall performance, potentially maintaining effectiveness across diverse populations with varying gestational ages. This is particularly relevant given the gestational age differences in the development and validation cohort, with the SPOT study recruiting women at admission to hospital between 26 and 34 weeks rather than at any gestational age.

Second, the modeling strategy in the development study might have resulted in model overfitting.[31] The selection of candidate predictors was based on statistical criteria rather than background clinical knowledge, and it was not performed in a multivariable manner, such as through backward elimination or lasso regression.[43] This approach could have resulted in omitting important predictors and their interactions. No shrinkage was applied to adjust coefficients of the prognostic model. Furthermore, as previously mentioned, the model was developed from a cohort established to predict adverse *maternal* outcomes, which could potentially not have included important predictors for *perinatal* outcomes. Finally, the prediction horizon was not explicitly defined in terms of time (limited to the admission period, which can vary across settings), potentially affecting the model's accuracy.

Third, women in the SPOT data set might differ from those in the development study, who received antenatal care in Fiji, Uganda, South Africa, Brazil, and Pakistan. Although both studies were conducted in LMIC settings, the relationships among maternal age, presenting symptoms, and proteinuria may vary, leading to suboptimal external predictive performance. Women in the SPOT data set were also of younger gestational age at first presentation with symptoms of HDP, suggesting more severe illness. Furthermore, we chose to include women with eclampsia upon admission, which could have resulted in a higher caseload of severe preeclampsia in our cohort. Correspondingly, the average outcome proportion was higher than in the development cohort, which might explain the observed miscalibration. However, our additional analysis performed in a subgroup of similar gestational age did not show improved calibration.

Fourth, differences in data collection procedures between the development and validation study might explain the suboptimal model performance. Regarding data collection, our data showed that 52% of the SPOT women did not experience any symptoms within the first 24 hours of admission. This finding seems inconsistent with the expectation that these women who were remote from term with HDP and admitted to a hospital are severely ill. Although it is technically feasible for women with only high blood pressure and proteinuria (without evident symptoms) to be admitted, this apparent inconsistency could be attributed to potential misclassifications in (electronic) health record data registration of symptoms.

Finally, the model assumes that standard care is provided to all women with HDP when predicting perinatal adverse outcomes in practice. In our context, the critical question is how realistic this assumption is, given that we collected data from various referral hospitals across diverse settings in Ghana. These settings include both urban and periurban areas, with women referred from both private and public facilities, and involving populations with varying income levels in Ghana. In addition, variations in treatment effects might influence the model's performance in our setting. Factors such as financial support available to women could also influence care-seeking and treatment choices, which may differ from those in the development settings. Including treatment options in the model building stage is a known method to improve model performance.[44]

Our study had important strengths. First, the prospective SPOT cohort and

the number of adverse events are relatively large. This contributes to improving prediction research in LMIC settings, where prognostic modeling studies, and especially external validation studies, are scarce. This is especially important in settings where the global burden of HDP-related deaths is highest.[39] Second, by externally validating an existing model, we contributed to evaluating the quality, outcomes, and applicability of previously performed prediction research in women with HDPs. Finally, the validated model predicts an outcome with clear definition that can be assessed consistently across settings. The uniformity in outcome definition between the development and validation cohorts suggests a low risk of discrepancies in predictive performance due to outcome assessment.

Some limitations also need to be considered in the interpretation of these findings. First, and most importantly, there might be data quality–related issues in our data set. Unlike the development cohort, which had no missing data for candidate predictors, SPOT data had missing values more often (generally <4.8%, except for BMI at enrollment). Despite applying data quality measures, missingness could not be completely prevented. The impact of missing values in our validation study was minimized by applying multiple imputation, but the actual predictor distribution might not have been reconstructed perfectly. Although we assumed that missing values were missing at random, it is possible that some of the missingness might have been informative because of the data being collected from electronic health records.[45] Second, the additional analyses were performed with only small groups of patients, making it difficult to draw conclusions from these results.

The next step following this prognostic study is to evaluate whether different models, such as those recently published by Ngwenya et al,[15] could better predict adverse perinatal outcomes. Future studies should ensure that there are sufficient data for meaningful validation, address power-related constraints, and avoid prematurely assessing models in settings with potential case mix and data collection differences that could impact predictive performance. Next, if necessary, a new model for predicting perinatal death could be developed that stratifies by gestational age or includes gestational age as a predictor tailored to local settings. Ideally, such risk prediction models should be adapted to inform and thus support clinical decisions about interventions (or whether or not to intervene) in the clinical course of pregnant women with HDPs.[46] Finally, incorporating the timing of clinical interventions into (dynamic) models is of great importance for guiding clinicians, pregnant women, and families in determining whether and when to intervene to optimize outcomes for both mothers and children. This also highlights the importance of standardized, comprehensive data collection in clinical settings, for example through the expanding use of electronic health record systems. Such standardized data collection can facilitate these efforts and support postimplementation evaluation.

Conclusion

The predictive performance of Payne's perinatal death model in selected hospitals in Ghana was poor. Given the importance of clinical decision-making support for settings where the perinatal morbidity and mortality burden associated with HDPs is highest, continued efforts to validate existing models, develop new models, and evaluate their impact after implementation are recommended.

Data sharing statement

Will individual participant data be available (including data dictionaries)?	Yes
What data in particular will be shared?	Individual participant data that underlie the results reported in this article, after deidentification (text, tables, figures, and appendices)
What other documents will be available?	Study protocol, analytic code on special request
When will data be available (start and end dates)?	Beginning 3 months and ending 5 years following article publication
With whom?	Researchers who provide a methodologically sound proposal
For what types of analyses?	To achieve aims in the approved proposal
By what mechanism will data be made available?	Proposals should be directed to hamoakoh@noguchi.ug.edu.gh; to gain access, data requestors will need to sign a data access agreement

CRediT authorship contribution statement

Klaartje M. Olde Loohuis: Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Kim Luijken:** Writing – review & editing, Visualization, Supervision, Methodology, Formal analysis, Data curation, Conceptualization. **Hannah Brown Amoakoh:** Writing – review & editing, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Kwame Adu-Bonsaffoh:** Writing – review & editing, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Diederick E. Grobbee:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Kerstin Klipstein-Grobusch:** Writing – review & editing, Supervision, Methodology, Formal analysis. **Emmanuel Srofenyoh:** Writing – review & editing, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Mary Amoakoh-Coleman:** Writing – review & editing, Supervision, Funding acquisition, Formal analysis, Conceptualization. **Joyce L. Browne:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work, the authors used ChatGPT/4o to improve readability of the text. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Supplementary materials

Supplementary material associated with this article can be found in the online

version at doi:10.1016/j.xagr.2025.100455.

REFERENCES

1. De Bernis L, Kinney MV, Stones W, et al. Stillbirths: ending preventable deaths by 2030. Lancet 2016;387:703–16.
2. Hug L, Alexander M, You D, Alkema L. United Nations Inter-Agency Group for Child Mortality Estimation. National, regional, and global levels and trends in neonatal mortality between 1990 and 2017, with scenario-based projections to 2030: a systematic analysis. Lancet Glob Health 2019;7:e710–20.
3. Cousens S, Blencowe H, Stanton C, et al. National, regional, and worldwide estimates of stillbirth rates in 2009 with trends since 1995: a systematic analysis. Lancet 2011;377:1319–30. https://doi.org/10.1016/S0140-6736(10)62310-0.
4. Lawn JE, Blencowe H, Pattinson R, et al. Stillbirths: where? When? Why? How to make the data count? Lancet 2011;377:1448–63. https://doi.org/10.1016/S0140-6736(10)62187-3.
5. Angell JN, Abdul-Mumin AS, Gold KJ. Determining the cause of stillbirth in Kumasi, Ghana. Int J Gynaecol Obstet 2019;147:173–8.
6. Adu-Bonsaffoh K, Ntumy MY, Obed SA, Seffah JD. Perinatal outcomes of hypertensive disorders in pregnancy at a tertiary hospital in Ghana. BMC Pregnancy Childbirth 2017;17:388.
7. Bhutta ZA, Das JK, Bahl R, et al. Can available interventions end preventable deaths in mothers, newborn babies, and stillbirths, and at what cost? Lancet 2014;384:347–70.
8. Hutcheon JA, Lisonkova S, Joseph KS. Epidemiology of pre-eclampsia and the other hypertensive disorders of pregnancy. Best Pract Res Clin Obstet Gynaecol 2011;25:391–403. https://doi.org/10.1016/j.bpobgyn.2011.01.006.
9. Lisonkova S, Joseph KS. Incidence of preeclampsia: risk factors and outcomes associated with early- versus late-onset disease. Am J Obstet Gynecol 2013;209:544.e1–544.e12. https://doi.org/10.1016/j.ajog.2013.08.019.
10. Lisonkova S, Sabr Y, Mayer C, Young C, Skoll A, Joseph KS. Maternal morbidity associated with early-onset and late-onset Preeclampsia. Obstet Gynecol 2014;124:771–81.
11. Sibai BM, Barton JR. Expectant management of severe preeclampsia remote from term: patient selection, treatment, and delivery indications. Am J Obstet Gynecol 2007;196:514.e1–9.
12. The American College of Obstetricians and Gynecologists. Gestational hypertension and preeclampsia. ACOG Pract Bull 2018;222:e237–60.
13. Pels A, Von Dadelszen P, Engelbrecht S, et al. Timing and mode of delivery. In: Magee LA, Dadelszen PV, Stones W, Mathai M, eds. FIGO textbook pregnancy hypertension, Carlisle, England: The Alliance for Global Women's Medicine; 2018:305–11.
14. Steegers EAP, Von Dadelszen P, Duvekot JJ, Pijnenborg R. Pre-eclampsia. Lancet 2010;376:631–44.
15. Ngwenya S, Jones B, Mwembe D, Nare H, Heazell AEP. Development and validation of risk prediction models for adverse maternal and neonatal outcomes in severe preeclampsia in a low-resource setting, Mpilo Central Hospital, Bulawayo, Zimbabwe. Pregnancy Hypertens 2020;23:18–26.
16. Payne BA, Groen H, Ukah UV, et al. Development and internal validation of a multivariable model to predict perinatal death in pregnancy hypertension. Pregnancy Hypertens 2015;5:315–21.
17. Awuah SP, Okai I, Ntim EA. Bedu-Addo K. Prevalence, placenta development, and perinatal outcomes of women with hypertensive disorders of pregnancy at Komfo Anokye Teaching Hospital. PLoS One 2020;15:e0233817.
18. Adu-Bonsaffoh K, Oppong SA, Binlinla G, Obed SA. Maternal deaths attributable to hypertensive disorders in a tertiary hospital in Ghana. Int J Gynaecol Obstet 2013;123:110–3.
19. Tranquilli AL, Brown MA, Zeeman GG, Dekker G, Sibai BM. The definition of severe and early-onset preeclampsia. Statements from the International Society for the Study of Hypertension in Pregnancy (ISSHP). Pregnancy Hypertens 2013;3:44–7.
20. Payne BA, Hutcheon JA, Ansermino JM, et al. A risk prediction model for the assessment and triage of women with hypertensive disorders of pregnancy in low-resourced settings: the miniPIERS (Pre-eclampsia Integrated Estimate of RiSk) multi-country prospective cohort study. PLoS Med 2014;11:e1001589.
21. Von Dadelszen P, Payne B, Li J, et al. Prediction of adverse maternal outcomes in pre-eclampsia: development and validation of the fullPIERS model. Lancet 2011;377:219–27. https://doi.org/10.1016/S0140-6736(10)61351-7.
22. Lehtonen L, Gimeno A, Parra-Llorca A, Vento M. Early neonatal death: a challenge worldwide. Semin Fetal Neonatal Med 2017;22:153–60.
23. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research Electronic Data Capture (REDCap)-a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42:377–81.
24. SPOT study in Ghana; 2024. Available at: https://spotstudyghana.com/en. Accessed 1st December 2024.
25. Moons KGM, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med 2015;162:W1–73.
26. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. Stat Med 2011;30:377–99.
27. von Hippel PT. Regression with missing Ys: an improved strategy for analyzing multiply imputed data. Sociol Methodol 2007;37:83–117.
28. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010;21:128–38.
29. Van Calster B, McLernon DJ, Van Smeden M, Wynants L, Steyerberg EW. Topic Group 'Evaluating Diagnostic Tests and Prediction Models' of the STRATOS Initiative. Calibration: the Achilles heel of predictive analytics. BMC Med 2019;17:230.
30. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. Eur Heart J 2014;35:1925–31.
31. Moons KGM, Kengne AP, Grobbee DE, et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart 2012;98:691–8.
32. Riley RD, Debray TPA, Collins GS, et al. Minimum sample size for external validation of a clinical prediction model with a binary outcome. Stat Med 2021;40:4230–51.
33. Vergouwe Y, Steyerberg EW, Eijkemans MJC, Habbema JDF. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. J Clin Epidemiol 2005;58:475–83.
34. Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating. New York, NY: Springer; 2019.
35. Drechsel KCE, Adu-Bonsaffoh K, Olde Loohuis KM, et al. Maternal near-miss and mortality associated with hypertensive disorders of pregnancy remote from term: a multicenter observational study in Ghana. AJOG Glob Rep 2022;2:100045.
36. Asseffa NA, Demissie BW. Perinatal outcomes of hypertensive disorders in pregnancy at a referral hospital, Southern Ethiopia. PLoS One 2019;14:e0213240.
37. Heestermans T, Payne B, Kayode GA, et al. Prognostic models for adverse pregnancy outcomes in low-income and middle-income countries: a systematic review. BMJ Glob Health 2019;4:e001759.
38. Antwi E, Amoakoh-Coleman M, Vieira DL, et al. Systematic review of prediction models for gestational hypertension and preeclampsia. PLoS One 2020;15:e0230955.
39. Ukah UV, Payne B, Lee T, Magee LA, Von Dadelszen P. fullPIERS and miniPIERS Working Groups. External validation of the fullPIERS model for predicting adverse maternal outcomes in pregnancy hypertension in low- and middle-income countries. Hypertension 2017;69:705–11.
40. Nakamura N, Ushida T, Nakatochi M, et al. Mortality and neurological outcomes in extremely and very preterm infants born to

mothers with hypertensive disorders of pregnancy. Sci Rep 2021;11:1729.
41. Jaleta DD, Gebremedhin T, Jebena MG. Perinatal outcomes of women with hypertensive disorders of pregnancy in Jimma Medical Center, southwest Ethiopia: retrospective cohort study. PLoS One 2021;16:e0256520.
42. Lugobe HM, Muhindo R, Kayondo M, et al. Risks of adverse perinatal and maternal outcomes among women with hypertensive disorders of pregnancy in southwestern Uganda. PLoS One 2020;15:e0241207.
43. Heinze G, Wallisch C, Dunkler D. Variable selection – a review and recommendations for the practicing statistician. Biom J 2018;60:431–49.
44. Groenwold RHH, Moons KGM, Pajouheshnia R, et al. Explicit inclusion of treatment in prognostic modeling was recommended in observational and randomized settings. J Clin Epidemiol 2016;78:90–100.
45. Groenwold RHH. Informative missingness in electronic health record systems: the curse of knowing. Diagn Progn Res 2020;4:8.
46. van Geloven N, Swanson SA, Ramspek CL, et al. Prediction meets causal inference: the role of treatment in clinical prediction models. Eur J Epidemiol 2020;35:619–30.