BMJ Open

Development and validation of a prediction model for pregnancy induced hypertension in a Ghanaian cohort

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-012670
Article Type:	Research
Date Submitted by the Author:	19-May-2016
Complete List of Authors:	ANTWI, EDWARD; Julius Center for Health Sciences and Primary Care, Julius Global Health; Ghana Health Service, Accra Groenwold, Rolf; UMC Utrecht, Julius Center Browne, Joyce; University Medical Center Utrecht, Julius Global Health, Julius Centrum for Health Sciences and Primary Care Franx, Arie; University Medical Center Utrecht, Obstetrics and gynecology Agyepong, Irene; School of Public Health, College of Health Sciences, University of Ghana, Legon, Accra, Ghana , Health Policy Planning and Management Koram, Kwadwo; Noguchi Memorial Institute for Medical Research, College of Health Sciences, University of Ghana, Legon, , Epidemiology Klipstein-Grobusch, Kerstin; University Medical Center, Julius Global Health, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht; School of Public Health, University of the Witwatersrand, Johannesburg, South Africa , Division of Epidemiology & Biostatistics, Grobbee, Diederick; UNIVERSITY MEDICAL CENTER, Julius Global Health, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Obstetrics and gynaecology
Keywords:	predictors, pregnancy induced hypertension, prediction model, hypertensive disorders of pregnancy, risk scores

Development and validation of a prediction model for pregnancy induced hypertension in a Ghanaian cohort

Edward Antwi[a,b], MD, MSc, Rolf H.H. Groenwold[a], MD, PhD, Joyce L.Browne[a], MD, MSc, Arie Franx[c], MD, PhD,Irene Agyepong[d], MD, DrPH, Kwadwo A. Koram[e], MD, PhD, Kerstin Klipstein-Grobusch[a,f], MSc, PhD, Diederick E Grobbee[a], MD, PhD.

Affiliations:

a. Julius Global Health, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, The Netherlands

b. Ghana Health Service, Accra, Ghana

c. Department of Obstetrics and Gynecology, University Medical Center Utrecht, The Netherlands

d. School of Public Health, College of Health Sciences, University of Ghana, Legon, Accra, Ghana

e. Noguchi Memorial Institute for Medical Research, College of Health Sciences, University of Ghana, Legon, Accra, Ghana

f. Division of Epidemiology & Biostatistics, School of Public Health, University of the Witwatersrand, Johannesburg, South Africa

Name and address for correspondence :

Dr. Edward Antwi

Email: ed_antwi@yahoo.com

Telephone: +233-244561328

Abstract

Objective: To develop and validate a prediction model for identifying women at increased risk of developing pregnancy induced hypertension (PIH) in Ghana.

Design: A prospective study. We used frequencies for descriptive analysis, Chi square test for associations and logistic regression to derive the prediction model. Discrimination was estimated by the c-statistic. Calibration was assessed by calibration plot of actual versus predicted probability.

Setting: Primary care antenatal clinics in Ghana.

Participants: Two thousand five hundred and twenty nine pregnant women in the development cohort and 647 pregnant women in the validation cohort. Inclusion criterion was women without chronic hypertension.

Primary outcome: Pregnancy induced hypertension.

Results: Predictors of PIH were diastolic blood pressure, family history of hypertension in parents, history of PIH in a previous pregnancy, parity, height and weight.

The c-statistic of the original model was 0.71 (95% C.I: 0.64-0.78) and 0.69 (95% CI: 0.60-0.78) in the validation cohort. Calibration was good in both cohorts. The negative predictive value (NPV) of women in the development cohort at high risk of PIH was 95.1% compared to 92.0% in the validation cohort.

Conclusion

The prediction model showed adequate performance after validation in an independent cohort and can be used to classify women into high, moderate or low risk of developing PIH. It

contributes to efforts to provide clinical decision-making support to improve maternal health and birth outcomes.

Key words

Predictors, pregnancy induced hypertension, prediction model, hypertensive disorders of pregnancy, risk scores.

Article summary

1. Use of prospectively collected data from antenatal period through to delivery.
2. Data was collected in primary care setting and reflected practice.
3. The prediction model validated in a different cohort of pregnant women.
4. Limitation of using only maternal clinical characteristics to predict PIH.
5. The study had PIH as only outcome and not pre-eclampsia or eclampsia.

Introduction

Hypertensive disorders of pregnancy (HDP), which include pregnancy induced hypertension (PIH), pre-eclampsia, eclampsia and the haemolysis, elevated liver enzymes and low platelets (HELLP) syndrome are the third leading cause of maternal deaths globally(1), with most of these deaths occurring in low- and middle-income countries (LMICs). HDPs are the leading cause of maternal death in Latin America and the Caribbean accounting for 25.7% of mortality; in Africa they rank third (9.1%) (2). In Ghana, 14% of all female deaths are pregnancy related with HDPs being the third leading cause of maternal deaths (9%) after haemorrhage (22%) and induced abortion (11%) (3).

The underlying causes of HDPs are not fully known (4),however accurate prediction of women at increased risk of HDP could lead to better antenatal care (ANC) and a reduction of complications from the condition.

Clinical prediction models estimate the probability of individuals having certain health conditions or obtaining defined health outcomes (5-8). They combine two or more items of patient data to predict clinical outcome and prior to application in clinical practice should be externally validated (5-11). The main approaches to predicting the occurrence of PIH include the use of maternal clinical characteristics, Uterine Artery Doppler and biomarkers (12-14). Although a number of prediction models for HDP, mainly pre-eclampsia and eclampsia have been developed in high-income countries, they may not be suitable for low- and middle-income countries because of differences in the availability and the cost of diagnostic tools (15).

The aim of this study was to develop and externally validate a contextual appropriate and low cost clinical prediction model for PIH based on maternal characteristics obtained at the first antenatal care visit for use in primary care settings in Ghana and potentially other LMIC.

Methods

Study design and population

(i) Development cohort

The prediction model was developed in a prospective cohort of 2,529 pregnant women attending antenatal care in primary care setting in six hospitals in the Greater Accra region of Ghana between February and May 2010. The eligibility criterion was pregnant women without chronic hypertension. The exclusion criteria were a history of hypertension or having hypertension before 20 weeks gestation as per blood pressure (BP) measurements. After potential participants had given written informed consent, they were enrolled and followed up at ANC visits until they delivered. Ethical approval for the study was granted by the Ethical Review Committee of the Ghana Health Service (Ethical Clearance ID No: GHS-ERC 02/1/10).

Sample size estimation was based on the incidence of HDPs in the Ghanaian population and on the principle of ten outcome events per variable (16). The Ghana Maternal Health Survey of 2007(3) had estimated that 9% of all maternal deaths were due to HDP. Using an estimated incidence of PIH of 10% in the study population and for 10 predictors, we aimed to enrol 2500 women but actually enrolled 2,529.

Data was obtained from the women's medical records as measured by the midwives during routine antenatal care. The midwives had been given standardized training in data collection. Candidate predictors were selected based on a review of the literature on variables known to be associated with PIH. Information on the following predictors: maternal age, diabetes mellitus (confirmed diagnosis of diabetes mellitus), family history of hypertension (confirmed diagnosis of hypertension in parents or siblings), family history of diabetes (confirmed diagnosis of

diabetes in parents or siblings), and family history of multiple pregnancies were obtained during the first antenatal clinic visit. Blood pressure (measured with a mercury sphygmomanometer), height (measured in centimetres with a stadiometer), weight (measured in kilogrammes with a bath room scale) and urine protein (defined as 2+ or more on urine dipstick) were also obtained during the first and subsequent antenatal clinic visits. Pregnancy outcomes were obtained from the hospital maternity register. Data was checked for accuracy and entered into SPSS software (version 20.0, IBM SPSS Statistics Inc., Chicago, Illinois, USA) and R statistical software (version 3.1.0 (2014-04-10)) for analysis.

(ii) Validation cohort

For external validation of the derived prediction model, data from 647 adult pregnant women recruited as part of a prospective cohort study conducted between July 2012 and March 2014 at Ridge Regional Hospital and Maamobi General Hospital in Accra were utilized. These hospitals provide primary antenatal care similar to that received by the women in the derivation study. Ethical approval for the validation study was granted by the Ethical Review Committee of the Ghana Health Service (GHS-ERC 07/09/11). The inclusion criteria were women less than 17 weeks pregnant and 18 years or older with no pre-existing hypertension. Pregnant women were included in the study after they had given written informed consent and were interviewed by trained research assistants using a structured questionnaire for socio-demographic characteristics and obstetric history. Weight, height, blood pressure, urine protein at the initial and subsequent ANC visits was obtained from the maternal health record books. Pregnancy outcomes were obtained from the hospital maternity register. Data was entered by trained data clerks using EpiDataEntry (EpiData Association, Odense, Denmark, 2010) and validated by double entry, cleaned and checked for missing data.

Outcome

The outcome, PIH, was defined as a systolic BP of 140mmHg or more and or a diastolic BP of 90mmHg or more on at least two separate occasions, and present for the first time after 20 weeks of pregnancy(17). Blood pressure was measured with a mercury sphygmomanometer with the woman in the sitting position, in line with standard antenatal clinic practice in both the development and validation cohort.

Data analysis

The mean and standard deviation of continuous predictors were calculated for women who developed PIH and those who did not. Means were compared using the independent T-test; percentages for categorical data were assessed by Chi-square test. Missing data were imputed by multiple imputation using "Multivariate Imputation by Chained Equations (MICE)" function in R(18). Missing values were imputed 10 times and Rubin's rule (19), was applied to pool results over the ten imputed datasets. Predictors that were related to PIH by a pre-determined p-value of 0.20 or less were selected and used in a multivariable logistic regression model. Stepwise backward selection using $p<0.20$ was used to derive the model which was internally validated using the bootstrapping technique. The resulting shrinkage factor after bootstrapping was used to adjust the regression coefficients, thus correcting for model overfitting.

The performance of the models in the development and validation cohort was assessed by discrimination and calibration. Discrimination is the ability of the model to distinguish between women who develop PIH and those who do not and was assessed using the c-statistic. The c-statistic or area under the receiver operating characteristic curve (AUC) ranges from 0.5 (no discrimination) to 1.0 (perfect discrimination)(11) . Calibration of the model was assessed by the calibration plot of actual probability versus the predicted probability.

For application of the model, a score chart was derived using the regression coefficients of the predictors. The total score of each woman was related to her risk of developing PIH. Cut-off points based on a total score of less than one, between two and six and equal or greater than seven were used to classify women into low, moderate and high risk of PIH respectively. The sensitivity, specificity, negative and positive predictive values of the cut off points were calculated.

Reporting and analysis of study results was conducted according to the TRIPOD checklist (15;20) statistical data analysis was done by use of SPSS software (version 20.0, IBM SPSS Statistics Inc., Chicago, Illinois, USA) and R statistical software (version 3.1.0 (2014-04-10)).

Results

Table 1 describes the baseline characteristics of the development and validation cohorts at the first ANC visit.

Table 1: **Characteristics of the development and validation cohort at first antenatal visit stratified by PIH**

	Development Cohort				Validation Cohort			
Variable ***Mean (SD), resp. N (%)***	**PIH (Yes) N=261**	**PIH (No) N=2268**	**O.R (95% C.I)**	**P-value**	**PIH (Yes) N= 42**	**PIH (No) N= 605**	**O.R (95% C.I)**	**P-value**
Age (years)	28.9 (5.9)	28.0 (5.8)	1.03(1.01-1.05)	0.013	29.8(5.6)	28.2(5.0)	1.06 (0.99-1.13)	0.053
Height (cm)	159.9 (6.7)	160.6 (7.4)	0.98 (0.97-1.01)	0.19	161.4(9.5)	161.1(7.5)	1.01(0.97-1.05)	0.757
Weight (kg)	73.3 (19.0)	66.2 (13.2)	1.03 (1.02-1.04)	<0.001	74.0(14.8)	65.9(7.5)	1.05(1.02-1.07)	<0.001
Systolic Blood Pressure (mmHg)	116.0 (15.2)	108.7(10.8)	1.05(1.04-0.06)	<0.001	115.6(14.5)	111.6(12.2)	1.02(1.00-1.046)	0.046
Diastolic Blood Pressure (mmHg)	71.9 (11.6)	66.2(9.1)	1.06 (1.05-1.08)	<0.001	75.2(12.6)	69.1(10.5)	1.05 (1.02-1.08)	<0.001
Gestational age (weeks)	21.9 (6.1)	20.5(6.9)	1.03 (1.01-1.05)	0.003	10.9(2.9)	11.4(2.9)	0.95(0.85-1.05)	0.298
Employed	243 (93.1%)	2092 (92.2%)	1.14 (0.69-1.88)	0.62	42(100%)	604(100%)	11.2 x10^7	0.62
Married	194 (74.3%)	1775 (78.3%)	0.80 (0.60-1.08)	0.15	38(90.5%)	501(82.8%)	1.97(0.69-5.65)	0.21
Educational level								
None	30(11.8%)	230(10.4%)	Referent		4(9.5%)	60(9.9%)	ReferentReferent	
Primary	55 (21.7%)	424 (19.1%)	0.84(0.47-1.47)	0.53	4(9.5%)	75(12.4%)	1.89(0.41-8.75)	0.42
Junior High School	101(39.9%)	999 (44.9%)	0.83 (0.50-1.38)	0.47	20(47.6%)	260(43.0%)	1.51(0.33-6.97)	0.59
Senior High School	42 (16.6%)	410 (18.4%)	0.65 (0.41-1.03)	0.07	11(26.2%)	125(20.7%)	2.18(0.63-7.52)	0.217
Tertiary	25 (9.9%)	160 (7.2%)	0.66(0.39- 1.11)	0.12	3(7.1%)	85(14.0%)	2.49(0.68-9.20)	0.17
Family history of hypertension (Parents)	70 (26.8 %)	392 (17.2%)	1.75 (1.29-2.34)	0.001	5(29.2%)	22(3.6%)	3.45(1.24-9.62)	0.018
Previous history of PIH	40 (15.3%)	23 (1.0%)	17.8 (10.4-30.2)	<0.001	1(2.4%)	20(3.3%)	0.72(0.09-5.49)	0.749

Development Cohort

Women with and without PIH differed with respect to age (28.9 (SD 5.9) years vs. 28.0 (SD: 5.8) years, p=0.01). There was no difference in mean height between women who developed PIH and those without PIH (159.9 cm (SD 6.7) vs. 160.6 cm (SD 7.4), p=0.19). The mean weight differed between women with and without PIH (73.3 kg (SD 19.0) vs. 66.2 kg (SD 13.2), p<0.001). The mean diastolic blood pressure also differed between women who developed PIH and those who did not (71.9mmHg (SD 11.6) vs. 66.2mmHg (SD 9.1), p<0.001).

About 27% of women with PIH had a parent with hypertension compared to 17.2% of women without PIH (p <0.001). Furthermore 15.3% of women with PIH had a history of PIH in a previous pregnancy compared to 1.0% of women without PIH (p <0.001).

Validation cohort

Mean age of women who developed PIH (29.8(SD 5.6) years) was higher than in those who did not. (28.2(SD 5.0) years, p=0.053). There was no difference in mean height between women with and without PIH (161.4cm (SD 9.5) vs. 161.1cm (SD 7.5), p=0.75). However there was a difference in the mean weight of women with and without PIH (74.0 kg (SD 14.8) vs. 65.9kg (SD 7.5), p<0.001). The mean diastolic blood pressure differed between women who developed PIH and those who did not (75.2mmHg (SD 12.6) vs. 69.1mmHg (SD 10.5), p<0.001), as did mean systolic blood pressure (115.6mmHg (SD 14.5) vs. 111.6 mmHg (SD 12.2), p=0.046).

Of the women who developed PIH, 29.2% reported a family history of hypertension in parents compared to 3.6 % of those who did not (p=0.02). Percentage of women with previous history of PIH did not materially differ between those who developed PIH and those who did not.

Table 2 shows the adjusted Odds ratios of predictors of PIH in the development cohort.

Table 2. Adjusted Odds ratio of predictors of PIH at the first antenatal care visit in a cohort of 2,529 pregnant women.

	Adjusted O.R (95% C.I)	P-value
PIH in a previous pregnancy	12.34 (7.02-21.68)	<0.001
Hypertension in parents	1.53 (1.11-2.12)	0.027
Diastolic BP (mmHg)	1.05 (1.03-1.06)	<0.001
Height (cm)	0.97 (0.95-0.99)	0.001
Weight (kg)	1.03 (1.03-1.043)	<0.001
Parity	1.02 (0.92-1.15)	0.50
Intercept	1.18 (0.01- 4.23)	

These are maternal height, weight, diastolic blood pressure, a history of hypertension in the parents, a previous history of PIH in the mother and parity. The c-statistic of the model was 0.71 (95% CI 0.64 - 0.78).

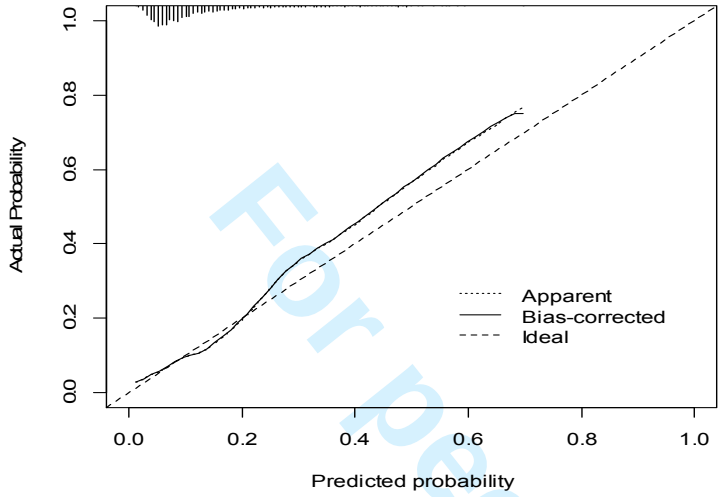
The final prediction model was:

Logit (PIH) = -1.48 -0.034*Height+0.42*Hypertension in parents+2.46*Previous PIH + 0.025*Weight + 0.044*Diastolic BP + 0.027*Parity.

The c-statistic after external validation was 0.69 (95% CI 0.60-0.78).

Figure 1 shows the calibration plot for the development cohort.

Figure 1. Calibration plot for the development cohort.



The dotted 45° line denotes the perfect agreement between predicted risk (x-axis) and observed risk (y-axis).The smoothed line approximates the agreement between predicted and observed risks across subgroups of pregnant women ranked by increasing predicted risks.

The calibration plot shows a good fit for probabilities between 0.05 and 2.0 where most of the events occur. Figure 2 shows the calibration plot in the validation cohort. Again the plot shows a good fit for probabilities between 0.05 and 2.0, where most of the events occur.

Figure 2. Calibration plot for the validation cohort.

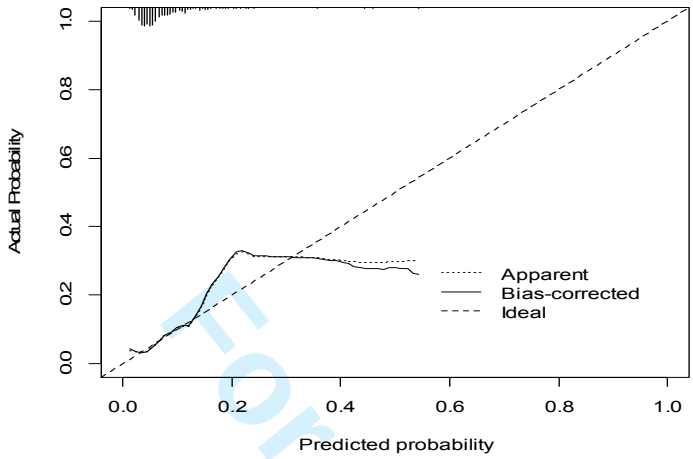


Table 3 presents the score chart for obtaining the total risk score of each woman.

Table 3. Score chart for the risk of developing pregnancy induced hypertension in a cohort of pregnant women from Ghana.

Predictor	Score
History of hypertension in parents	No=0 Yes=4
PIH‡ in a previous pregnancy	No=0 Yes=24
Diastolic blood pressure (mmHg)	< 60=0 61-70 = 1 71-80 = 2 81-90 = 3 >90 = 4
Height(cm)	≥ 161=0 56-160=1 151-155=2 0-150=3
Weight (kg)	≤ 70=0

	71-80=1 81-90=2 ≥91=3
Parity	0=0 ≥1=1

Table 4 shows the categorization of the development cohort into low, moderate and high risk. Three hundred and twenty one women were classified as being at high risk of developing PIH and 80 of them eventually developed PIH giving a positive predictive value (PPV) of 24.9% and a negative predictive value of 92.0%.

Table 4. Categorization of development cohort into low, moderate and high risk.

	PIH‡ (Yes)	PIH‡ (No)	Sensitivity	Specificity	NPV§	PPV**
Low risk (N=402) (Score ≤ 1)	15 (3.7%)	387 (96.3%)	93.6%	19.0%	96.3%	11.8%
Moderate risk (N=1,546) (Score 2-6)	14 (9.1 %)	1405(90.9%)	33.9%	88.2%	92.0 %	24.9 %
High risk (N=321) (Score≥ 7)	80 (24.9 %)	241 (75.1%)				

PIH‡, pregnancy induced hypertension; NPV§, Negative predictive value; PPV**, Positive predictive value.

Table 5. Categorization of the validation cohort into low, moderate and high risk.

	PIH‡ (Yes)	PIH‡ (No)	Sensitivity	Specificity	NPV§	PPV**
Low risk (N=323)	11 (3.4%)	312 (96.6%)	73.8%	51.6 %	96.6 %	9.6 %
Moderate risk (N=229)	16 (7.0 %)	213 (93.0 %)	35.7%	86.8 %	95.1 %	15.8 %
High risk (N=95)	15 (15.8 %)	80 (84.2 %)				

Table 5 presents information on the categorization of the validation cohort into low, moderate and high risk of PIH. Ninety-five women were classified as high risk and 15 of them eventually developed PIH, giving a PPV of 15.8% and a negative predictive value of 95.1%.

Discussion

We developed and externally validated a simple prediction model for PIH in two different cohorts of pregnant women attending ANC clinics in similar settings in line with the general recommendation that before being applied in clinical practice, prediction models should be externally validated (5-11) The c-statistic of the model in the original cohort (0.71(95% CI: 0.64-0.78)) was only slightly reduced (0.69(95% CI: 0.60-0.78)) after external validation, consistent with findings from other studies (21-24). Nijdam et al(25) in the Netherlands derived a prediction model for identifying nulliparous women who developed hypertension before 36

weeks of gestation using systolic blood pressure, diastolic blood pressure and weight. The AUC of the original model of 0.78 (95% CI 0.75-0.82) reduced to 0.75 (95% CI 0.68-0.81) after external validation. The small decrease in c-statistic in our study implies that the model predicts well based on data routinely collected as part of antenatal care and can be applied to the pregnant women in the study setting.

Most prediction models for HDPs, such as the miniPIERS model (27), have focussed on pre-eclampsia and eclampsia which are severer forms of the disorder. However milder forms such as PIH are also associated with less favourable pregnancy outcomes. Given that PIH can be managed to prevent progression to severer forms, a model that identifies women at risk is useful.

A limitation of our study was the application of clinical characteristics only, excluding biomarkers and Uterine Artery Doppler in our prediction model. This is because of the non-routine use of these parameters in ANC in the Ghanaian setting. Both approaches are expensive and the equipment for analysing these biomarkers is generally not available in many low resource settings. However, future research could assess the added value of these biomarkers as recent systematic review for first trimester prediction of preeclampsia showed that a combination of Uterine Artery Doppler, maternal characteristics and two or more biomarkers yielded detection rates of 38% to 100% (13). The best rates were reported for the combination of Inhibin A, PLGF, PAPP-A, Uterine Artery Doppler and maternal characteristics (13). The difficulty of predicting PIH using only maternal clinical characteristics has been pointed out (26), however, the feasibility of applying these models in low resource settings currently remains limited due to constraints in the availability of diagnostic equipment and the high cost of the tests which are beyond the means of most people who require them. Thus despite the increased predictive value of adding biomarkers to the predictive model; the need to derive reasonably accurate prediction

models that use variables, which are routinely easy to obtain for low resource settings is important.

In the development cohort, 321(12.7%) women were classified as being at high risk of developing PIH. Eighty of them eventually developed PIH giving a PPV of 24.9% and NPV of 92%. In the validation cohort, 95(14.7%) women were classified as being at high risk of PIH and 15 of them developed the condition. The PPV was 15.8% and the NPV 95.1%. Classifying women into different risk categories allows for closer monitoring of pregnant women at high risk. This will include more frequent ANC visits or referral for specialist care.

Given that the addition of biomarkers in the screening of women could enhance the identification of those at high risk of PIH, future research should explore the added value of biomarkers in the early identification of pregnant women at increased risk of HDPs in LMICs. Such studies should be accompanied by comparative cost effectiveness of the routine data only predictive models and the models that combine routine data and biomarkers to provide essential health technology assessment information for future decision making. In the interim however, despite the fact that the modest PPV in both the development and validation cohorts show the limitation and difficulty of predicting PIH using only demographic and clinical characteristics the model has the potential of identifying pregnant women at increased risk of PIH for subsequent care and monitoring(27). Its further validation and use is worth serious consideration in low resource settings.

Conclusion

We developed and validated a prediction model for PIH at the first ANC visit using maternal data prospectively collected in a LMIC setting. Our results are easily converted into a simple user friendly clinical decision making support tool for use in antenatal clinics in low resource

settings that enables frontline providers of maternal health services to use a score chart to quickly categorize women into different risk levels. The strength of this model is the use of a few maternal clinical variables already routinely obtained by care-givers during routine ANC. Such a simple predictive model to aid frontline providers of maternal care to estimate the probability of PIH later on in the pregnancy and take relevant precautions is potentially life saving. Obtaining the information does not involve expensive procedures such as Uterine Artery Doppler (28). The application of the model at the ANC should aid in the early detection of women at risk of PIH and contribute to efforts to provide clinical decision-making support to improve maternal health outcomes. We would recommend its validation in other low-income settings as well as implementation research to inform implementation, monitoring and evaluation at scale in Ghana.

Acknowledgement

We acknowledge Ms. Helenmary Bainson, Mrs. Cecelia Opong-Peprah and Mrs. Emma Antwi who supervised the data collection. We thank the midwives who collected the data and the data entry staff. I gratefully acknowledge the DWS scholarship from Utrecht University which enabled me to conduct this study and finalize this manuscript. Finally I acknowledge funding from the Ghana Health Service for the initial data collection.

Contributors

EA designed the study, collected data, carried out data analysis and wrote the initial draft of the manuscript. RHHG assisted with data analysis. DEG, RHHG, IA, KAK,KK-G,JLB and AF provided scientific guidance and were also actively involved in the preparation and review of the manuscript and approved it.

Declaration of conflicting interests: None declared.

Funding: This research received funding from the DWS Scholarship of Utrecht University.

Data Sharing: No additional data available.

Reference List

(1) Kassebaum NJ, Bertozzi-Villa A, Coggeshall MS, Shackelford KA, Steiner C, Heuton KR, et al. Global, regional, and national levels and causes of maternal mortality during 1990 to 2013: a systematic analysis for the Global Burden of Disease Study 2013. The Lancet384(9947):980-1004.

(2) Khan KS, Wojdyla D, Say L, Gulmezoglu AM, Van Look PF. WHO analysis of causes of maternal death: a systematic review. The Lancet367(9516):1066-74.

(3) Ghana Statistical Service (GSS) GHSGaMII. Ghana Maternal Health Survey 2007. Calverton, Maryland, USA; 2009 May 5.

(4) Conde-Agudelo A, Belizan JM. Risk factors for pre-eclampsia in a large cohort of Latin American and Caribbean women. BJOG: An International Journal of Obstetrics & Gynaecology 2000 Jan 1;107(1):75-83.

(5) Altman DG, Vergouwe Y, Royston P, Moons KGM. Prognosis and prognostic research: validating a prognostic model. BMJ 2009 May 28;338.

(6) Moons KGM, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how?
3544. BMJ 2009 Feb 23;338.

(7) Moons KGM, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. BMJ 2009 Jun 4;338.

(8) Royston P, Moons KGM, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. BMJ 2009 Mar 31;338.

(9) Moons KGM, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. External validation, model updating, and impact assessment
3562. Heart 2012 Mar 7.

(10) Reilly BM, Evans AT. Translating Clinical Research into Clinical Practice: Impact of Using Prediction Rules To Make Decisions
3563. Ann Intern Med 2006 Feb 7;144(3):201-9.

(11) Steyerberg E. Clinical prediction models. A practical approach to development, validation and updating. ; 2009. Springer Science + Media; 2009.

(12) Akolekar R, Syngelaki A, Sarquis R, Zvanca M, Nicolaides KH. Prediction of early, intermediate and late pre-eclampsia from maternal factors, biophysical and biochemical markers at 11 to 13 weeks. Prenat Diagn 2011 Jan 1;31(1):66-74.

(13) Kuc S, Wortelboer EJ FAU - van Rijn B, van Rijn BB FAU - Franx A, Franx AF, Visser GH FAU - Schielen P, Schielen PC. Evaluation of 7 serum biomarkers and uterine artery Doppler ultrasound for first-trimester prediction of preeclampsia: a systematic review.(1533-9866 (Electronic)).

(14) Poon LC, Nicolaides KH. First-trimester maternal factors and biomarker screening for preeclampsia. Prenat Diagn 2014 Jul 1;34(7):618-27.

(15) Ukah UV. Preliminary external validation of the fullPIERS risk prediction model for women with pre-eclampsia using the miniPIERS dataset. Pregnancy Hypertens 2015 Jan 3;5(1):124-5.

(16) Harrell Jr FE. Regression modelling strategies with application to linear models, logistic regression, and survival analysis. 1 ed. New York: Springer-Verlag New York; 2001.

(17) Report of the National High Blood Pressure Education Program Working Group on High Blood Pressure in Pregnancy. American Journal of Obstetrics & Gynecology183(1):s1-s22.

(18) Buuren SV, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software 2011;45(3):-67.

(19) Rubin DB. Multiple Imputation for Nonresponse in Surveys. http://sites stat psu edu/ 2015 August 15 [cited 2015 Aug 16];Available from: URL: http://sites.stat.psu.edu/~jls/mifaq.html#howto

(20) Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): the TRIPOD Statement
74. Br J Surg 2015 Feb 1;102(3):148-58.

(21) Hukkelhoven CW, Steyerberg EW, Habbema JD. Predicting outcome after traumatic brain injury: development and validation of a prognostic score based on admission characteristics. J Neurotrauma 2005;22(10):1025-39.

(22) Signorini DF, Andrews PJ, Jones PA. Predicting survival using simple clinical variables: a case study in traumatic brain injury
3588. J Neurol Neurosurg Psychiatry 1999;66(1):20-5.

(23) Signorini DF, Andrews PJ, Jones PA. Adding insult to injury: the prognostic value of early secondary insults for survival after traumatic brain injury
3589. J Neurol Neurosurg Psychiatry 1999;66(1):26-31.

(24) Moons KGM, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart 2012 Mar 7.

(25) Nijdam ME, Janssen KJ, Moons KG, Grobbee DE, van der Post JA, Bots ML, et al. Prediction model for hypertension in pregnancy in nulliparous women using information obtained at the first antenatal visit. Journal of Hypertension 2010;28(1).

(26) Angeli F, Angeli E, Reboldi G, Verdecchia P. Hypertensive disorders during pregnancy: clinical applicability of risk prediction models. Journal of Hypertension 2011;29(12).

(27) Payne B, Hodgson S, Hutcheon JA, Joseph KS, Li J, Lee T, et al. Performance of the fullPIERS model in predicting adverse maternal outcomes in pre-eclampsia using patient data from the PIERS (Pre-eclampsia Integrated Estimate of RiSk) cohort, collected on admission
3560. BJOG: An International Journal of Obstetrics & Gynaecology 2013 Jan 1;120(1):113-8.

(28) Onwudiwe N, Yu CKH, Poon LCY, Spiliopoulos I, Nicolaides KH. Prediction of pre-eclampsia by a combination of maternal history, uterine artery Doppler and mean arterial pressure. Ultrasound Obstet Gynecol 2008 Dec 1;32(7):877-83.

TRIPOD Checklist: Prediction Model Development and Validation

Section/Topic	Item		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.	,1
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	2
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
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	4
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.	5,6
	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.	-
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	7
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	-
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	5,6
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	-
Sample size	8	D;V	Explain how the study size was arrived at.	5
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.	7
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	7
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	7
	10c	V	For validation, describe how the predictions were calculated.	-
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	7,8
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	-
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	8
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	5-7
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.	
	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.	9
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	9
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	9
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	
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).	11
	15b	D	Explain how to the use the prediction model.	8,12, 13
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	11,12, 13
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	-
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	16,17
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	15
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	15,16, 17
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	17,18
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	-
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	19

*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.

BMJ Open

Development and validation of a prediction model for gestational hypertension in a Ghanaian cohort

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-012670.R1
Article Type:	Research
Date Submitted by the Author:	09-Sep-2016
Complete List of Authors:	ANTWI, EDWARD; Julius Center for Health Sciences and Primary Care, Julius Global Health; Ghana Health Service, Accra Groenwold, Rolf; UMC Utrecht, Julius Center Browne, Joyce; University Medical Center Utrecht, Julius Global Health, Julius Centrum for Health Sciences and Primary Care Franx, Arie; University Medical Center Utrecht, Obstetrics and gynecology Agyepong, Irene; School of Public Health, College of Health Sciences, University of Ghana, Legon, Accra, Ghana , Health Policy Planning and Management Koram, Kwadwo; Noguchi Memorial Institute for Medical Research, College of Health Sciences, University of Ghana, Legon, , Epidemiology Klipstein-Grobusch, Kerstin; University Medical Center, Julius Global Health, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht; School of Public Health, University of the Witwatersrand, Johannesburg, South Africa , Division of Epidemiology & Biostatistics, Grobbee, Diederick; UNIVERSITY MEDICAL CENTER, Julius Global Health, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Obstetrics and gynaecology
Keywords:	predictors, prediction model, hypertensive disorders of pregnancy, risk scores, gestational hypertension

Development and validation of a prediction model for gestational hypertension in a Ghanaian cohort

Edward Antwi[a,b], MD, MSc, Rolf H.H. Groenwold[a], MD, PhD, Joyce L.Browne[a], MD, MSc, Arie Franx[c], MD, PhD,Irene Agyepong[d], MD, DrPH, Kwadwo A. Koram[e], MD, PhD, Kerstin Klipstein-Grobusch[a,f], MSc, PhD, Diederick E Grobbee[a], MD, PhD.

Affiliations:

a. Julius Global Health, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, The Netherlands

b. Ghana Health Service, Accra, Ghana

c. Department of Obstetrics and Gynecology, University Medical Center Utrecht, The Netherlands

d. School of Public Health, College of Health Sciences, University of Ghana, Legon, Accra, Ghana

e. Noguchi Memorial Institute for Medical Research, College of Health Sciences, University of Ghana, Legon, Accra, Ghana

f. Division of Epidemiology & Biostatistics, School of Public Health, University of the Witwatersrand, Johannesburg, South Africa

Name and address for correspondence :

Dr. Edward Antwi

Email: ed_antwi@yahoo.com

Telephone: +233-244561328

Abstract

Objective: To develop and validate a prediction model for identifying women at increased risk of developing gestational hypertension (GH) in Ghana.

Design: A prospective study. We used frequencies for descriptive analysis, Chi square test for associations and logistic regression to derive the prediction model. Discrimination was estimated by the c-statistic. Calibration was assessed by calibration plot of actual versus predicted probability.

Setting: Primary care antenatal clinics in Ghana.

Participants: Two thousand five hundred and twenty nine pregnant women in the development cohort and 647 pregnant women in the validation cohort. Inclusion criterion was women without chronic hypertension.

Primary outcome: Gestational hypertension.

Results: Predictors of GH were diastolic blood pressure, family history of hypertension in parents, history of GH in a previous pregnancy, parity, height and weight.

The c-statistic of the original model was 0.71 (95% C.I: 0.64-0.78) and 0.69 (95% CI: 0.60-0.78) in the validation cohort. Calibration was good in both cohorts. The negative predictive value (NPV) of women in the development cohort at high risk of GH was 95.1% compared to 92.0% in the validation cohort.

Conclusion

The prediction model showed adequate performance after validation in an independent cohort and can be used to classify women into high, moderate or low risk of developing GH. It

contributes to efforts to provide clinical decision-making support to improve maternal health and birth outcomes.

Key words

Predictors, gestational hypertension, prediction model, hypertensive disorders of pregnancy, risk scores.

Article summary

1. Use of prospectively collected data from antenatal period through to delivery.
2. Data was collected in primary care setting and reflected practice.
3. The prediction model validated in a different cohort of pregnant women.
4. Limitation of using only maternal clinical characteristics to predict GH.
5. The study had GH as only outcome and not pre-eclampsia or eclampsia.

Introduction

Hypertensive disorders of pregnancy (HDP), which include gestational hypertension (GH), pre-eclampsia, eclampsia and the haemolysis, elevated liver enzymes and low platelets (HELLP) syndrome are the third leading cause of maternal deaths globally(1), with most of these deaths occurring in low- and middle-income countries (LMICs). The International Society for the Study of Hypertension in Pregnancy (ISSHP) classifies HDPs as chronic hypertension, gestational hypertension, pre-eclampsia-de novo or superimposed on chronic hypertension and white coat hypertension (2). HDPs are the leading cause of maternal death in Latin America and the Caribbean accounting for 25.7% of mortality; in Africa they rank third (9.1%) (3). In Ghana, 14% of all female deaths are pregnancy related with HDPs being the third leading cause of maternal deaths (9%) after haemorrhage (22%) and induced abortion (11%) (4).

The underlying causes of HDPs are not fully known (5),however accurate prediction of women at increased risk of HDP could lead to better antenatal care (ANC) and a reduction of complications from the condition.

Clinical prediction models estimate the probability of individuals having certain health conditions or obtaining defined health outcomes (6-9). They combine two or more items of patient data to predict clinical outcome and prior to application in clinical practice should be externally validated (6-12). The main approaches to predicting the occurrence of GH include the use of maternal clinical characteristics, Uterine Artery Doppler and biomarkers (13-15). Although a number of prediction models for HDP, mainly pre-eclampsia and eclampsia have been developed in high-income countries, they may not be suitable for low- and middle-income countries because of differences in the availability and the cost of diagnostic tools (16).

The aim of this study was to develop and externally validate a contextual appropriate and low cost clinical prediction model for GH based on maternal characteristics obtained at the first antenatal care visit for use in primary care settings in Ghana and potentially other LMIC.

Methods

Study design and population

(i) Development cohort

The prediction model was developed in a prospective cohort of 2,529 pregnant women attending antenatal care in primary care setting in six hospitals in the Greater Accra region of Ghana between February and May 2010. The eligibility criterion was pregnant women without chronic hypertension. The exclusion criteria were a history of hypertension or having hypertension before 20 weeks gestation as per blood pressure (BP) measurements. After potential participants had given written informed consent, they were enrolled and followed up at ANC visits until they delivered. Ethical approval for the study was granted by the Ethical Review Committee of the Ghana Health Service (Ethical Clearance ID No: GHS-ERC 02/1/10).

Sample size estimation was based on the incidence of HDPs in the Ghanaian population and on the principle of ten outcome events per variable (17). The Ghana Maternal Health Survey of 2007(3) had estimated that 9% of all maternal deaths were due to HDP. Using an estimated incidence of GH of 10% in the study population and for 10 predictors, we aimed to enrol 2500 women but actually enrolled 2,529.

Data was obtained from the women's medical records as measured by the midwives during routine antenatal care. The midwives had been given standardized training in data collection. Candidate predictors were selected based on a review of the literature on variables known to be

associated with GH (18-22). Information on the following predictors: maternal age, diabetes mellitus (confirmed diagnosis of diabetes mellitus), family history of hypertension (confirmed diagnosis of hypertension in parents or siblings), family history of diabetes (confirmed diagnosis of diabetes in parents or siblings), and family history of multiple pregnancies were obtained during the first antenatal clinic visit. Blood pressure (measured with a mercury sphygmomanometer), height (measured in centimetres with a stadiometer), weight (measured in kilogrammes with a bath room scale) and urine protein (defined as 2+ or more on urine dipstick) were also obtained during the first and subsequent antenatal clinic visits. Pregnancy outcomes were obtained from the hospital maternity register.

(ii) Validation cohort

For external validation of the derived prediction model, data from 647 adult pregnant women recruited as part of a prospective cohort study conducted between July 2012 and March 2014 at Ridge Regional Hospital and Maamobi General Hospital in Accra were utilized. These hospitals provide primary antenatal care similar to that received by the women in the derivation study. Ethical approval for the validation study was granted by the Ethical Review Committee of the Ghana Health Service (GHS-ERC 07/09/11). The inclusion criteria were women less than 17 weeks pregnant and 18 years or older with no pre-existing hypertension. Pregnant women were included in the study after they had given written informed consent and were interviewed by trained research assistants using a structured questionnaire for socio-demographic characteristics and obstetric history. Weight, height, blood pressure, urine protein at the initial and subsequent ANC visits was obtained from the maternal health record books. Pregnancy outcomes were obtained from the hospital maternity register. Data was entered by trained data clerks using

EpiDataEntry (EpiData Association, Odense, Denmark, 2010) and validated by double entry, cleaned and checked for missing data.

Outcome

The outcome, GH, was defined as a systolic BP of 140mmHg or more and or a diastolic BP of 90mmHg or more on at least two separate occasions, and present for the first time after 20 weeks of pregnancy(23). In both cohorts blood pressure measurements were taken using a mercury sphygmomanometer by trained midwives. The appropriate adult sized cuff was placed on the bare left upper arm with the woman comfortably seated and her back supported and the legs uncrossed. The arm was at the level of the heart and neither the patient nor the observer talked during the measurement. Korotkoff phase V sounds were used (24). Two readings were taken at interval of five minutes and the average used to represent the woman's BP.The sphygmomanometers at the clinics are calibrated periodically to ensure accurate readings

The gestational age at which GH was diagnosed is available for both cohorts.

Data analysis

The mean and standard deviation of continuous predictors were calculated for women who developed GH and those who did not. Means were compared using the independent T-test; percentages for categorical data were assessed by Chi-square test. Missing data were imputed by multiple imputation using "Multivariate Imputation by Chained Equations (MICE)" function in R(25). Missing values were imputed 10 times and Rubin's rule (26), was applied to pool results over the ten imputed datasets. Predictors that were related to GH by a pre-determined p-value of 0.20 or less were selected and used in a multivariable logistic regression model. Stepwise backward selection using $p<0.20$ was used to derive the model which was internally validated

using the bootstrapping technique. The resulting shrinkage factor after bootstrapping was used to adjust the regression coefficients, thus correcting for model overfitting.

The performance of the models in the development and validation cohort was assessed by discrimination and calibration. Discrimination is the ability of the model to distinguish between women who develop GH and those who do not and was assessed using the c-statistic. The c-statistic or area under the receiver operating characteristic curve (AUC) ranges from 0.5 (no discrimination) to 1.0 (perfect discrimination)(12). Calibration of the model was assessed by the calibration plot of actual probabilities versus predicted probabilities.

For application of the model, a score chart was derived using the regression coefficients of the predictors. The total score of each woman was related to her risk of developing GH. Cut-off points based on a total score of less than one, between two and six and equal or greater than seven were used to classify women into low, moderate and high risk of GH respectively. The sensitivity, specificity, negative and positive predictive values of the cut off points were calculated.

Reporting and analysis of study results was conducted according to the TRIPOD checklist (27) statistical data analysis was done by use of SPSS software (version 20.0, IBM SPSS Statistics Inc., Chicago, Illinois, USA) and R statistical software (version 3.1.0 (2014-04-10)).

Results

Table 1 describes the baseline characteristics of the development and validation cohorts at the first ANC visit.

Table 1: **Characteristics of the development and validation cohort at first antenatal visit stratified by GH**

	Development Cohort				Validation Cohort			
Variable ***Mean (SD), resp. N (%)***	**GH (Yes) N=261**	**GH (No) N=2268**	**O.R (95% C.I)**	**P-value**	**GH (Yes) N= 42**	**GH (No) N= 605**	**O.R (95% C.I)**	**P-value**
Age (years)	28.9 (5.9)	28.0 (5.8)	1.03(1.01-1.05)	0.013	29.8(5.6)	28.2(5.0)	1.06 (0.99-1.13)	0.053
Height (cm)	159.9 (6.7)	160.6 (7.4)	0.98 (0.97-1.01)	0.19	161.4(9.5)	161.1(7.5)	1.01(0.97-1.05)	0.757
Weight (kg)	73.3 (19.0)	66.2 (13.2)	1.03 (1.02-1.04)	<0.001	74.0(14.8)	65.9(7.5)	1.05(1.02-1.07)	<0.001
Systolic Blood Pressure (mmHg)	116.0 (15.2)	108.7(10.8)	1.05(1.04-0.06)	<0.001	115.6(14.5)	111.6(12.2)	1.02(1.00-1.046)	0.046
Diastolic Blood Pressure (mmHg)	71.9 (11.6)	66.2(9.1)	1.06 (1.05-1.08)	<0.001	75.2(12.6)	69.1(10.5)	1.05 (1.02-1.08)	<0.001
Gestational age (weeks)	21.9 (6.1)	20.5(6.9)	1.03 (1.01-1.05)	0.003	10.9(2.9)	11.4(2.9)	0.95(0.85-1.05)	0.298
Employed	243 (93.1%)	2092 (92.2%)	1.14 (0.69-1.88)	0.62	37 (88.1%)	523 (86.4%)	0.86 (0.33-2.26)	0.76
Married	194 (74.3%)	1775 (78.3%)	0.80 (0.60-1.08)	0.15	38(90.5%)	501(82.8%)	1.97(0.69-5.65)	0.21
Educational level								
None	30(11.8%)	230(10.4%)	Referent		4(9.5%)	60(9.9%)	ReferentReferent	
Primary	55 (21.7%)	424 (19.1%)	0.84(0.47-1.47)	0.53	4(9.5%)	75(12.4%)	1.89(0.41-8.75)	0.42
Junior High School	101(39.9%)	999 (44.9%)	0.83 (0.50-1.38)	0.47	20(47.6%)	260(43.0%)	1.51(0.33-6.97)	0.59
Senior High School	42 (16.6%)	410 (18.4%)	0.65 (0.41-1.03)	0.07	11(26.2%)	125(20.7%)	2.18(0.63-7.52)	0.217
Tertiary	25 (9.9%)	160 (7.2%)	0.66(0.39- 1.11)	0.12	3(7.1%)	85(14.0%)	2.49(0.68-9.20)	0.17
Family history of hypertension (Parents)	70 (26.8 %)	392 (17.2%)	1.75 (1.29-2.34)	0.001	5(29.2%)	22(3.6%)	3.45(1.24-9.62)	0.018
Previous history of GH	40 (15.3%)	23 (1.0%)	17.8 (10.4-30.2)	<0.001	1(2.4%)	20(3.3%)	0.72(0.09-5.49)	0.749

Development Cohort

Women with and without GH differed with respect to age (28.9 (SD 5.9) years vs. 28.0 (SD: 5.8) years, p=0.01). There was no difference in mean height between women who developed GH and those without GH (159.9 cm (SD 6.7) vs. 160.6 cm (SD 7.4), p=0.19). The mean weight differed between women with and without GH (73.3 kg (SD 19.0) vs. 66.2 kg (SD 13.2), p<0.001). The mean diastolic blood pressure also differed between women who developed GH and those who did not (71.9mmHg (SD 11.6) vs. 66.2mmHg (SD 9.1), p<0.001).

About 27% of women with GH had a parent with hypertension compared to 17.2% of women without GH (p <0.001). Furthermore 15.3% of women with GH had a history of GH in a previous pregnancy compared to 1.0% of women without GH (p <0.001).

Validation cohort

Mean age of women who developed GH (29.8(SD 5.6) years) was higher than in those who did not. (28.2(SD 5.0) years, p=0.053). There was no difference in mean height between women with and without GH (161.4cm (SD 9.5) vs. 161.1cm (SD 7.5), p=0.75). However there was a difference in the mean weight of women with and without GH (74.0 kg (SD 14.8) vs. 65.9kg (SD 7.5), p<0.001). The mean diastolic blood pressure differed between women who developed GH and those who did not (75.2mmHg (SD 12.6) vs. 69.1mmHg (SD 10.5), p<0.001), as did mean systolic blood pressure (115.6mmHg (SD 14.5) vs. 111.6 mmHg (SD 12.2), p=0.046).

Of the women who developed GH, 29.2% reported a family history of hypertension in parents compared to 3.6 % of those who did not (p=0.02). Percentage of women with previous history of GH did not materially differ between those who developed GH and those who did not.

Table 2 shows the adjusted Odds ratios of predictors of GH in the development cohort.

Table 2. Adjusted Odds ratio of predictors of GH at the first antenatal care visit in a cohort of 2,529 pregnant women.

	Adjusted O.R (95% C.I)	P-value
GH in a previous pregnancy	12.34 (7.02-21.68)	<0.001
Hypertension in parents	1.53 (1.11-2.12)	0.027
Diastolic BP (mmHg)	1.05 (1.03-1.06)	<0.001
Height (cm)	0.97 (0.95-0.99)	0.001
Weight (kg)	1.03 (1.03-1.043)	<0.001
Parity	1.02 (0.92-1.15)	0.50
Intercept	1.18 (0.01- 4.23)	

These are maternal height, weight, diastolic blood pressure, a history of hypertension in the parents, a previous history of GH in the mother and parity. The c-statistic of the model was 0.71 (95% CI 0.64 - 0.78).

The final prediction model was:

Logit (GH) = -1.48 -0.034*Height+0.42*Hypertension in parents+2.46*Previous GH + 0.025*Weight + 0.044*Diastolic BP + 0.027*Parity.

The c-statistic after external validation was 0.69 (95% CI 0.60-0.78).

Figure 1 shows the calibration plot for the development cohort.

The dotted 45° line denotes the perfect agreement between predicted risk (x-axis) and observed risk (y-axis).The smoothed line approximates the agreement between predicted and observed risks across subgroups of pregnant women ranked by increasing predicted risks.

The calibration plot shows a good fit for probabilities between 0.05 and 2.0 where most of the events occur. Figure 2 shows the calibration plot in the validation cohort. Again the plot shows a good fit for probabilities between 0.05 and 2.0, where most of the events occur.

Table 3 presents the score chart for obtaining the total risk score of each woman.

Table 3. Score chart for the risk of developing gestational hypertension in a cohort of pregnant women from Ghana.

Predictor	Score
History of hypertension in parents	No=0 Yes=4
GH‡ in a previous pregnancy	No=0 Yes=24
Diastolic blood pressure (mmHg)	< 60=0 61-70 = 1 71-80 = 2 81-90 = 3 >90 = 4
Height(cm)	≥ 161=0 56-160=1 151-155=2 0-150=3
Weight (kg)	≤ 70=0 71-80=1 81-90=2 ≥91=3
Parity	0=0 ≥1=1

Table 4 shows the categorization of the development cohort into low, moderate and high risk. Three hundred and twenty one women were classified as being at high risk of developing GH and 80 of them eventually developed GH giving a positive predictive value (PPV) of 24.9% and a negative predictive value of 92.0%. The likelihood ratio positive was 1.16 for low risk and 2.87 for moderate risk while the likelihood ratio negative was 0.34 for low risk and 0.75 for moderate risk.

Table 4. Categorization of development cohort into low, moderate and high risk.

	GH (Yes)	GH (No)	Sensitivity	Specificity	NPV	PPV	LR+	LR-
Low risk (N=402) (Score ≤ 1)	15 (3.7%)	387 (96.3%)	93.6%	19.0%	96.3%	11.8%	1.16	0.34
Moderate risk (N=1,546) Score (2-6)	14 (9.1%)	1,405 (90.9%)	33.9%	88.2%	92.0%	24.9%	2.87	0.75
High risk (N=321) (Score ≥ 7)	80 (24.9%)	241 (75.1%)						

GH, gestational hypertension; NPV, Negative predictive value; PPV, Positive predictive value; LR+, Likelihood ratio positive; LR-, Likelihood ratio negative.

Table 5. Categorization of the validation cohort into low, moderate and high risk.

	GH (Yes)	GH (No)	Sensitivity	Specificity	NPV	PPV	LR+	LR-
Low risk (N=323)	11 (3.4%)	312 (96.6%)	73.8%	51.6%	96.6%	9.6%	1.53	0.51
Moderate risk (N=229)	16 (7.0%)	213 (93.0%)	35.7%	86.8%	95.1%	15.8%	2.59	0.74
High risk (N=95)	15 (15.8%)	80 (84.2%)						

GH, gestational hypertension; NPV, Negative predictive value; PPV, Positive predictive value; LR+, Likelihood ratio positive; LR-, Likelihood ratio negative.

Table 5 presents information on the categorization of the validation cohort into low, moderate and high risk of GH. Ninety-five women were classified as high risk and 15 of them eventually developed GH, giving a PPV of 15.8% and a negative predictive value of 95.1%. The likelihood ratio positive was 1.53 for low risk and 2.59 for moderate risk while the likelihood ratio negative was 0.51 for low risk and 0.74 for moderate risk. Table 6 shows the number of observations and missing values (with percentage missing) for the development and validation cohorts. Table 7 compares characteristics of women in the development and validation cohorts before and after imputation.

Discussion

We developed and externally validated a simple prediction model for GH in two different cohorts of pregnant women attending ANC clinics in similar settings in line with the general recommendation that before being applied in clinical practice, prediction models should be externally validated (6-12) The c-statistic of the model in the original cohort (0.71(95% CI: 0.64-0.78)) was only slightly reduced (0.69(95% CI: 0.60-0.78)) after external validation, consistent with findings from other studies (28-31). Nijdam et al(32) in the Netherlands derived a prediction model for identifying nulliparous women who developed hypertension before 36 weeks of gestation using systolic blood pressure, diastolic blood pressure and weight. The AUC of the original model of 0.78 (95% CI 0.75-0.82) reduced to 0.75 (95% CI 0.68-0.81) after external validation. The small decrease in c-statistic in our study implies that the model predicts well based on data routinely collected as part of antenatal care and can be applied to the pregnant women in the study setting.

Most prediction models for HDPs, such as the SCOPE model (16), have focussed on pre-eclampsia and eclampsia which are severer forms of the disorder. However milder forms such as GH are also associated with less favourable pregnancy outcomes. Given that GH can be managed to prevent progression to severer forms, a model that identifies women at risk is useful.

A limitation of our study was the application of clinical characteristics only, excluding biomarkers and Uterine Artery Doppler in our prediction model. This is because of the non-routine use of these parameters in ANC in the Ghanaian setting. Both approaches are expensive and the equipment for analysing these biomarkers is generally not available in many low resource settings. However, future research could assess the added value of these biomarkers as recent systematic review for first trimester prediction of preeclampsia showed that a combination

of Uterine Artery Doppler, maternal characteristics and two or more biomarkers yielded detection rates of 38% to 100% (13). The best rates were reported for the combination of Inhibin A, PLGF, PAPP-A, Uterine Artery Doppler and maternal characteristics (13). The difficulty of predicting GH using only maternal clinical characteristics has been pointed out (33), however, the feasibility of applying these models in low resource settings currently remains limited due to constraints in the availability of diagnostic equipment and the high cost of the tests which are beyond the means of most people who require them. Thus despite the increased predictive value of adding biomarkers to the predictive model; the need to derive reasonably accurate prediction models that use variables, which are routinely easy to obtain for low resource settings is important.

In the development cohort, 321(12.7%) women were classified as being at high risk of developing GH. Eighty of them eventually developed GH giving a PPV of 24.9% and NPV of 92%. In the validation cohort, 95(14.7%) women were classified as being at high risk of GH and 15 of them developed the condition. The PPV was 15.8% and the NPV 95.1%. Classifying women into different risk categories allows for closer monitoring of pregnant women at high risk. This will include more frequent ANC visits or referral for specialist care.

Given that the addition of biomarkers in the screening of women could enhance the identification of those at high risk of GH, future research should explore the added value of biomarkers in the early identification of pregnant women at increased risk of HDPs in LMICs. Such studies should be accompanied by comparative cost effectiveness of the routine data only predictive models and the models that combine routine data and biomarkers to provide essential health technology assessment information for future decision making. In the interim however, despite the fact that the modest PPV in both the development and validation cohorts show the limitation and

difficulty of predicting GH using only demographic and clinical characteristics the model has the potential of identifying pregnant women at increased risk of GH for subsequent care and monitoring. Its further validation and use is worth serious consideration in low resource settings.

Conclusion

We developed and validated a prediction model for GH at the first ANC visit using maternal data prospectively collected in a LMIC setting. Our results are easily converted into a simple user friendly clinical decision making support tool for use in antenatal clinics in low resource settings that enables frontline providers of maternal health services to use a score chart to quickly categorize women into different risk levels. The strength of this model is the use of a few maternal clinical variables already routinely obtained by care-givers during routine ANC. Such a simple predictive model to aid frontline providers of maternal care to estimate the probability of GH later on in the pregnancy and take relevant precautions is potentially life saving. Obtaining the information does not involve expensive procedures such as Uterine Artery Doppler (34). The application of the model at the ANC should aid in the early detection of women at risk of GH and contribute to efforts to provide clinical decision-making support to improve maternal health outcomes. We would recommend its validation in other low-income settings as well as implementation research to inform implementation, monitoring and evaluation at scale in Ghana.

Acknowledgement

We acknowledge Ms. Helenmary Bainson, Mrs. Cecelia Opong-Peprah and Mrs. Emma Antwi who supervised the data collection. We thank the midwives who collected the data and the data entry staff. I gratefully acknowledge the DWS scholarship from Utrecht University which

enabled me to conduct this study and finalize this manuscript. Finally I acknowledge funding from the Ghana Health Service for the initial data collection.

Contributors

EA designed the study, collected data, carried out data analysis and wrote the initial draft of the manuscript. RHHG assisted with data analysis. DEG, RHHG, IA, KAK,KK-G,JLB and AF provided scientific guidance and were also actively involved in the preparation and review of the manuscript and approved it.

Declaration of conflicting interests: None declared

Funding: This research received funding from the DWS Scholarship of Utrecht University. The funders played no role in the study design, data collection, data analysis and interpretation as well as writing of the manuscript.

Data sharing statement: No additional data available.

Reference List

(1) Kassebaum NJ, Bertozzi-Villa A, Coggeshall MS, Shackelford KA, Steiner C, Heuton KR, et al. Global, regional, and national levels and causes of maternal mortality during 1990 to 2013: a systematic analysis for the Global Burden of Disease Study 2013. The Lancet384(9947):980-1004.

(2) Tranquilli AL, Dekker G, Magee L, Roberts J, Sibai BM, Steyn W, et al. The classification, diagnosis and management of the hypertensive disorders of pregnancy: A revised statement from the ISSHP. Pregnancy Hypertension: An International Journal of Women's Cardiovascular Health 4(2):97-104.

(3) Khan KS, Wojdyla D, Say L, Golmezoglu AM, Van Look PF. WHO analysis of causes of maternal death: a systematic review. The Lancet367(9516):1066-74.

(4) Ghana Statistical Service (GSS) GHSGaMII. Ghana Maternal Health Survey 2007. Calverton, Maryland, USA; 2009 May 5.

(5) Conde-Agudelo A, Belizan JM. Risk factors for pre-eclampsia in a large cohort of Latin American and Caribbean women. BJOG: An International Journal of Obstetrics & Gynaecology 2000 Jan 1;107(1):75-83.

(6) Altman DG, Vergouwe Y, Royston P, Moons KGM. Prognosis and prognostic research: validating a prognostic model. BMJ 2009 May 28;338.

(7) Moons KGM, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? BMJ 2009 Feb 23;338.

(8) Moons KGM, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. BMJ 2009 Jun 4;338.

(9) Royston P, Moons KGM, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. BMJ 2009 Mar 31;338.

(10) Moons KGM, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart 2012 Mar 7.

(11) Reilly BM, Evans AT. Translating Clinical Research into Clinical Practice: Impact of Using Prediction Rules To Make Decisions. Ann Intern Med 2006 Feb 7;144(3):201-9.

(12) Steyerberg E. Clinical prediction models. A practical approach to development, validation and updating. ; 2009. Springer Science + Media; 2009.

(13) Akolekar R, Syngelaki A, Sarquis R, Zvanca M, Nicolaides KH. Prediction of early, intermediate and late pre-eclampsia from maternal factors, biophysical and biochemical markers at 11ΓÇô13 weeks. Prenat Diagn 2011 Jan 1;31(1):66-74.

(14) Kuc S, Wortelboer EJ FAU - van Rijn B, van Rijn BB FAU - Franx A, Franx AF, Visser GH FAU - Schielen P, Schielen PC. Evaluation of 7 serum biomarkers and uterine artery Doppler ultrasound for first-trimester prediction of preeclampsia: a systematic review.(1533-9866 (Electronic)).

(15) Poon LC, Nicolaides KH. First-trimester maternal factors and biomarker screening for preeclampsia. Prenat Diagn 2014 Jul 1;34(7):618-27.

(16) Kenny LC, Black MA, Poston L, Taylor R, Myers JE, Baker PN, et al. Early Pregnancy Prediction of Preeclampsia in Nulliparous Women, Combining Clinical Risk and Biomarkers The Screening for Pregnancy Endpoints (SCOPE) International Cohort Study. Hypertension 2014;64(3):644-52.

(17) Harrell JrFE. Regression modelling strategies with application to linear models, logistic regression, and survival analysis. 1 ed. New York: Springer-Verlag New York; 2001.

(18) Dalmaz CA, Santos KGd, Botton MR, Roisenberg I. Risk factors for hypertensive disorders of pregnancy in southern Brazil. Revista da Assoc. Med. Brasileira 2011;57:692-6.

(19) Poon LCY, Kametas NA, Chelemen T, Leal A, Nicolaides KH. Maternal risk factors for hypertensive disorders in pregnancy: a multivariate approach. Journal of human hypertension 2010;24(2):104-10.

(20) Ros HS, Cnattingius S, Lipworth L. Comparison of Risk Factors for Preeclampsia and Gestational Hypertension in a Population-based Cohort Study. American Journal of Epidemiology 1998 Jun 1;147(11):1062-70.

(21) Sibai BM, Gordon T, Thom E, Caritis SN, Klebanoff M, McNellis D, et al. Risk factors for preeclampsia in healthy nulliparous women: A prospective multicenter study. American Journal of Obstetrics and Gynecology 1995;172(2):642-8.

(22) Tebeu PM, Foumane P, Mbu R, Fosso Gl, Biyaga PT, Fomulu JN. Risk factors for hypertensive disorders in pregnancy: A report from the maroua regional hospital, Cameroon. Journal of Reproduction & Infertility 2011;12(3):227.

(23) Report of the National High Blood Pressure Education Program Working Group on High Blood Pressure in Pregnancy. American Journal of Obstetrics & Gynecology183(1):s1-s22.

(24) Pickering TG, Hall JE, Appel LJ, Falkner BE, Graves J, Hill MN, et al. Recommendations for blood pressure measurement in humans and experimental animals part 1: blood pressure measurement in humans: a statement for professionals from the Subcommittee of Professional and Public Education of the American Heart Association Council on High Blood Pressure Research. Hypertension 2005;45(1):142-61.

(25) Buuren Sv, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software 2011;45(3):-67.

(26) Rubin DB. Multiple Imputation for Nonresponse in Surveys. http://sites stat psu edu/ 2015 August 15 [cited 2015 Aug 16];Available from: URL: http://sites.stat.psu.edu/~jls/mifaq.html#howto

(27) Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): the TRIPOD Statement. Br J Surg 2015 Feb 1;102(3):148-58.

(28) Hukkelhoven CW, Steyerberg EW, Habbema JD. Predicting outcome after traumatic brain injury: development and validation of a prognostic score based on admission characteristics. J Neurotrauma 2005;22(10):1025-39.

(29) Signorini DF, Andrews PJ, Jones PA. Predicting survival using simple clinical variables: a case study in traumatic brain injury. J Neurol Neurosurg Psychiatry 1999;66(1):20-5.

(30) Signorini DF, Andrews PJ, Jones PA. Adding insult to injury: the prognostic value of early secondary insults for survival after traumatic brain injury. J Neurol Neurosurg Psychiatry 1999;66(1):26-31.

(31) Moons KGM, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart 2012 Mar 7.

(32) Nijdam ME, Janssen KJ, Moons KG, Grobbee DE, van der Post JA, Bots ML, et al. Prediction model for hypertension in pregnancy in nulliparous women using information obtained at the first antenatal visit. Journal of Hypertension 2010;28(1).

(33) Angeli F, Angeli E, Reboldi G, Verdecchia P. Hypertensive disorders during pregnancy: clinical applicability of risk prediction models. Journal of Hypertension 2011;29(12).

(34) Onwudiwe N, Yu CKH, Poon LCY, Spiliopoulos I, Nicolaides KH. Prediction of pre-eclampsia by a combination of maternal history, uterine artery Doppler and mean arterial pressure. Ultrasound Obstet Gynecol 2008 Dec 1;32(7):877-83.

Appendix

Table 6. Number of observations and missing values (with percentage missing) for the development and validation cohorts.

	Development cohort		Validation cohort	
Variable	No. of observations	Missing (%)	No. of observations	Missing (%)
Age	2514	15 (0.6)	647	0 (0)
History of hypertension in parents	2498	31(1.2)	647	0 (0)
Height	2435	94 (3.7)	646	1(0.2)
Weight	2522	7 (0.3)	646	1(0.2)
Systolic Blood Pressure	2523	6 (0.23)	646	1(0.2)
Diastolic Blood Pressure	2522	7 (0.3)	646	1(0.2)
Parity	2527	2 (0.08)	647	0(0)
Previous history of gestational hypertension	2395	134 (5.3)	504	143(22.1)

Table 7. Comparison of characteristics of women in the development and validation cohorts before and after imputation.

Variable	Development cohort	Development cohort after imputation	Validation cohort	Validation cohort after imputation
Age (years)	28.1 (5.8)	28.1 (5.8)	28.3 (5.1)	28.3 (5.1)
Height (cm)	160.5 (7.4)	160.5 (7.4)	161.1 (7.6)	161.1 (7.6)
Weight (Kg)	66.9 (14.1)	66.9 (14.1)	66.4 (12.9)	66.4 (12.9)
Diastolic BP (mmHg)	66.8 (11.6)	66.8 (11.6)	69.5 (10.7)	69.5 (10.7)
Systolic BP (mmHg)	109.4 (11.6)	109.4 (11.6)	111.9 (12.4)	111.9 (12.4)

History of hypertension in parents	462 (18.5%)	470 (18.5%)	27 (4.2%)	27 (4.2%)

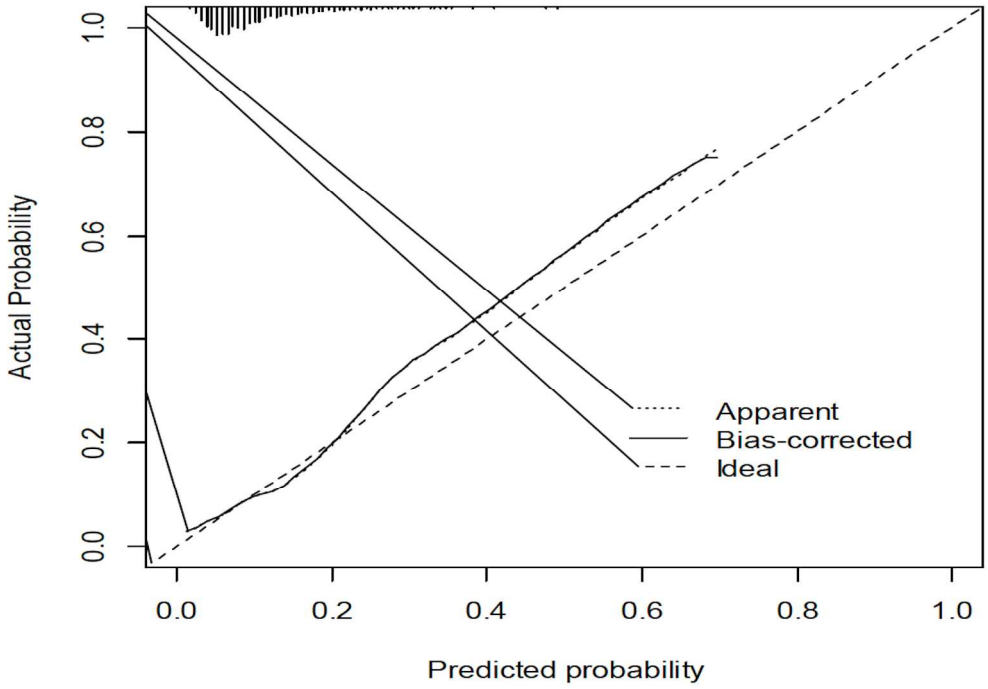


Figure 1. Calibration plot for the development cohort.
Page 12
388x268mm (300 x 300 DPI)

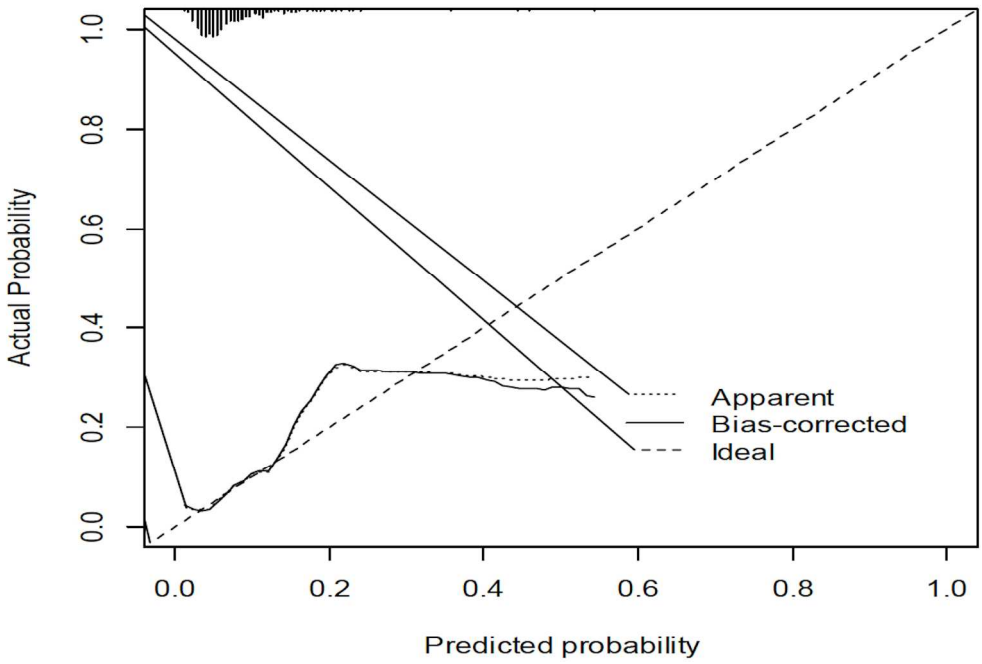


Figure 2. Calibration plot for the validation cohort.
Page 13
396x268mm (300 x 300 DPI)

Correction

Antwi E, Groenwold RHH, Browne JL, *et al*. Development and validation of a prediction model for gestational hypertension in a Ghanaian cohort. *BMJ Open* 2017;7: e012670.

The author affiliations are listed incorrectly. The correct author list and affiliations are as follows:

Edward Antwi,[1,2] Rolf H H Groenwold,[1] Joyce L Browne,[1] Arie Franx,[3] Irene A Agyepong,[2] Kwadwo A Koram,[4] Kerstin Klipstein-Grobusch,[1,5] Diederick E Grobbee,[1]

Author affiliations:

1. Julius Global Health, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands
2. Ghana Health Service, Accra, Ghana
3. Department of Obstetrics and Gynecology, University Medical Center Utrecht, Utrecht, The Netherlands
4. Noguchi Memorial Institute for Medical Research, College of Health Sciences, University of Ghana, Legon, Accra, Ghana
5. Division of Epidemiology & Biostatistics, School of Public Health, University of the Witwatersrand, Johannesburg, South Africa

BMJ Open 2017;**7**:e012670corr1. doi:10.1136/bmjopen-2016-012670corr1