Simple Model to Predict Gestational Diabetes in Nulliparous Women in Early Pregnancy

Juan Jesus FERNANDEZ ALBA (✉ jjesus.fernandez@uca.es)

Hospital Universitario Puerto Real. Servicio de Obstetricia y Ginecología

Maria CASTILLO LARA

Hospital Universitario Puerto Real. Servicio de Obstetricia y Ginecología

Jose Manuel JIMENEZ HERAS

Institute of Research and Innovation in Biomedical Sciences of the Province of Cadiz (INiBICA). Hospital Universitario Puerta del Mar. Avda. Ana de Viya

Rocio MORENO CORTES

Hospital Universitario Puerto Real. Servicio de Obstetricia y Ginecología

CARMEN GONZALEZ MACIAS

Hospital Universitario Puerto Real. Servicio de Obstetricia y Ginecología

Angel VILAR SANCHEZ

Hospital Universitario Puerto Real. Servicio de Obstetricia y Ginecología

Florentino CARRAL SAN LAUREANO

Hospital Universitario Puerto Real

Luis MORENO CORRAL

University of Cádiz (UCA)

Research Article

Keywords: Gestational diabetes mellitus, clinical risk score, nulliparous, chronic hypertension, predictive model, Thyroid Stimulating Hormone

Posted Date: November 10th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1063893/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background

The aim of this study was to develop a simple tool using anthropometric, clinical, and analytical variables to predict the risk of gestational diabetes mellitus in the first trimester of pregnancy.

Methods

A historical cohort study was conducted with 1,946 Caucasian nulliparous pregnant women at the Department of Obstetrics and Gynecology of the University Hospital of Puerto Real (Cádiz/Spain). The predictive model used was a multivariate logistic regression evaluated by 10-fold cross-validation with five iterations. Receiver-operating characteristic and prediction recall curves were plotted with the predictions of the model. Optimal cut-off points of the receiver-operating characteristic curve were estimated using the Youden Index and minimum distance to point 1,1.

Results

The final logistic regression model included both maternal weight and height, thyroid-stimulating hormone, family history of diabetes mellitus or hypertension, and the presence of chronic hypertension as predictive variables. The model showed an accuracy of 0.93 and a receiver-operating characteristic area under the curve of 0.791.

Conclusions

This report provides a tool to predict the onset of gestational diabetes mellitus in nulliparous women during the first trimester of pregnancy. This tool is more accurate than the currently available tools and can facilitate the implementation of preventative interventions.

1. Background

Gestational diabetes mellitus (GDM) is defined as glucose intolerance that is diagnosed for the first time during pregnancy, but usually resolves soon after delivery.¹ Currently, there is an increase in the incidence of obesity and type II diabetes mellitus, and this has led to an increase in the incidence of GDM.^{2,3} GDM is associated with adverse perinatal outcomes, including macrosomia, shoulder dystocia, cesarean delivery, preeclampsia, and neonatal hypoglycemia.^{4,5} Moreover, GDM is a marker of susceptibility to health problems later in life for both the mother and her child, including obesity, type 2 diabetes, metabolic syndrome, and cardiovascular disease.⁶

Nowadays, GDM is mostly diagnosed at the third trimester of pregnancy when the fetal and maternal deleterious effects have already started. This is because, in most cases, the diagnosis is established in women who return a positive screening test result (O'Sullivan's test), and this test is usually performed between 24 and 28 weeks of gestation.⁷

There are several models for predicting the risk of GDM. The performance of these models varies widely with receiver-operating characteristic area under the curve (ROC-AUC)⁸ results ranging from 0.7 to 0.9.⁹⁻¹⁴ Most of these models are produced using multivariate logistic regression,¹⁵ but some are developed using learning machine techniques.^{16,17} No one technique has been shown to have clear superiority over the others for general diabetes or in GDM.¹⁸

Consequently, there is no common model used in clinical practice. This may be due to limitations inherent in the models, including the inclusion of too many variables to be practical for clinical practice. In addition, other models include variables that are not measured in the context of the management of a normal pregnancy.^{19,20} Moreover, some models perform considerably worse in certain groups, such as primiparous women¹⁰ or women with a negative gestational diabetes history.¹⁹ These groups are the ones with the most uncertainty, and with the greatest need for determination of the risk. Therefore, a simple model with variables regularly assessed in clinical practice and high predictive capability for most groups is needed. If we were able to determine which pregnant women are at risk of developing GDM in the first trimester, we could implement interventions aimed at minimizing the maternal and fetal impacts of diabetes before the disease is established, as suggested by some authors.²¹⁻²³

We hypothesized that it is possible to predict the risk of GDM in the first trimester of gestation using anthropometric, clinical, and analytic variables. The aim of this study was to develop a predictive model using these variables.

2. Methods

A historical cohort study was conducted at the Department of Obstetrics and Gynecology of the University Hospital of Puerto Real (Cádiz/Spain). Medical records of all consecutive singleton births that occurred between January 2014 and December 2019 were retrieved from our clinical information system. This study was approved by the Ethic Committee of Biomedical Research of Cádiz.

Pregnant women identified as having preexisting diabetes mellitus were excluded. Additionally, pregnancies with deliveries before gestational week 32 or after gestational week 42 were excluded. Only nulliparous Caucasian women were included in the study.

Categorical data were summarized as counts and percentages. The distributions of continuous data were assessed using the Shapiro-Wilk test. Continuous data with a normal distribution were summarized as mean and standard deviation. In contrast, when the data showed a non-normal distribution, the median and the interquartile range were used.

The predictive model involved multivariate logistic regression evaluated with 10-fold cross-validation and five iterations trained towards improvement of the accuracy using R software version 4.0.5.²⁴ The normality and homogeneity of the model residuals were tested using the Shapiro-Francia and Levene tests, respectively.

Variable selection was performed using a backward elimination procedure. The inclusion criteria for the variables was 95% statistical significance for both dichotomous and quantitative variables. To maintain

model consistency, when variables were polytomous, only one of the subcategories was required for it to be significant and subsequently included.

Quantitative variables were transformed into categorical variables and the predictive impact in the model was tested using an F test and Chi-square test with an H^0 that considered no predictive changes in the model.

We modeled all the interactions between the two variables. If the new variable originating from the interaction was statistically significant, we evaluated the predictive impact of the interaction. This evaluation consisted of a MANOVA by Roy's largest root and Pillai's Trace. The values of ROC-AUC and prediction recall (PR)-AUC of the inclusion of the interaction were analyzed.

ROC and PR curves were plotted with the predictions of the model. Optimal cut-off points of the ROC curve were estimated by the Youden index and minimum distance to point 1,1 (sensitivity and specificity equal to 1). Additionally, we estimated the cut-off point needed to obtain a sensitivity equal to 80% as a rule-in alternative to diagnose GDM.

An online calculator was developed using the Shiny package²⁵ for the R programming language. This tool is accessible in the link <https://obgynreference.shinyapps.io/calcdiabgest/>.

3. Results

The initial study population involved 7,505 pregnancies. Following application of the exclusion criteria (associated pathologies [n=4,034], input errors [n=168], delivery before 32 weeks or after 42 weeks of gestation [n=25], previous pregnancy [n=1,332]), 1946 nulliparous pregnant women of Caucasian race were included (Figure 1).

Table 1 shows the epidemiological characteristics of the study population. A family history of diabetes was found in 20% of pregnant women. Body mass index (BMI) percentages were adjusted to the expected values for a Spanish female population.²⁶

Based on the multivariate logistic regression model, GDM was predicted in nulliparous women with an accuracy of 0.93 and a ROC-AUC of 0.791. These predictive values were achieved using independent variables that are commonly assessed in clinical practice, such as weight and height, thyroid-stimulating hormone (TSH) levels, family history of diabetes mellitus or hypertension, and the presence of chronic hypertension. (Table 2).

There are some points worth noting regarding the features of the independent variables of the model. The most impactful variable was expected to be either a family history of diabetes or obesity. However chronic hypertension and twin pregnancy increased the risk of GDM considerably more.

There was a statistically significant interaction between age and chronic hypertension. This interaction shows that young women with chronic hypertension are more prone to develop GDM. Our population showed more incidence of hypertension in older population. (Figure 2).

Although age does not have remarkable importance when comparing two pregnant women with a small age difference, it can become the most impactful factor when comparing women with an age difference of 19 years, and this surpasses the impact of twin pregnancy. For a pregnant woman aged 40 years, the impact of chronic hypertension was ameliorated to an extent to be lower than the impact of family history of diabetes, and this was remarkable considering that a model without this variable interaction still showed chronic hypertension as the most impactful variable.

The model reached an accuracy of 0.93. The Shapiro-Francia test indicated that the residual variables did not show a normal distribution; thus, the Levene test was performed with some different splits of the data, and showed homogeneity.

The prediction capability of the model was tested with an ROC curve (Figure 3). The ROC-AUC of the model was 0.79 (95% confidence interval [CI]: 0.75–0.82). The optimal cut-off point obtained using the Youden method provided a sensitivity of 0.65 (95%CI: 0.56–0.73) and a specificity of 0.80 (95%CI: 0.78–0.82) for a threshold of 7.8%. We did consider estimating a cut-off point with higher sensitivity for use as a rule-in tool. For a sensitivity of 0.80 (95%CI: 0.72–0.87), we obtained a specificity of 0.59 (0.57–0.61) with a threshold of 4.3%.

Since the model had a considerable ROC-AUC, we calculated a PR curve for this model to properly compare it with other results. The PR curve can be seen in Figure 4. For the final model, the value of the PR curve was 0.30.

4. Discussion

We found an increase in the risk of GDM development due to age, BMI, twin pregnancy, hypertension, diabetes, hypertension history, and TSH level, and there was an interaction between age and hypertension. The relation between these features and GDM is already known. We only identified one previous model that considered chronic hypertension¹¹ as a risk factor for GDM and no models had considered TSH levels or family history of hypertension.

Based on our findings, we developed a model to predict GDM in nulliparous women with a ROC-AUC of 0.79, which provided an 80% sensitivity and 59% specificity for a 4.3% threshold.

Untreated GDM carries significant risk of perinatal morbidity at all severity levels. Prompt treatment can enhance outcomes in newborns. Our model provides a tool to assess the risk of GDM in the first trimester of pregnancy, allowing early intervention and minimization of the maternal and fetal impact of diabetes before the disease manifests.

Although there are many resources available for prediction of GDM, they all have limitations. The GDM models that we are aware of are shown in Table 3. However, direct comparison between some models might not be appropriate owing to the difference in GDM prevalence in different populations.

Savvidou et al.¹⁹ reported a 0.86 ROC-AUC, but this dropped when used to predict the GDM risk in women with a negative GDM history. They also used variables that are not measured in common clinical practice,

making its clinical application more difficult. Models, such as that of Van Leeuwen et al.²³ or Zheng et al.¹³ are simple to administer, but our model has a better AUC.

Sweeting et al.⁹ developed a model which performs better with a 0.91 ROC-AUC, although their population size was small. Moreover, when this model was used in nulliparous women, the AUC dropped to 0.76 when using biomarkers, and to 0.70 without biomarkers, which is considerably less than the 0.79 ROC-AUC found for our model. Moreover, the model itself uses a high number of variables, and many of them are not currently measured in common clinical practice, especially lipocalin-2.

Artzi et al.¹⁶ compared many different variables and reported a GDM prevalence similar to ours. Although they obtained an ROC-AUC of 0.86, it was achieved using a model with 255 variables, which is of no use for current practice. Moreover, for nulliparous women, this model drops from a PR-AUC of 0.32 to a PR-AUC of 0.22, while our model has a PR-AUC of 0.30 for nulliparous women. The same authors, using the criteria established by the National Institute of Health (NIH),²⁷ obtained an ROC-AUC of 0.68. Our model provided an ROC-AUC of 0.79 (95%CI 0.75–0.82), thus surpassing the predictive capability of the NIH criteria. It worth mentioning that this difference might well be wider, since one of these rules considers previous pregnancies.

Other models, such as that of Syngelaki et al.¹¹ or Nanda et al.,¹² have good predictive capability, the former having an ROC-AUC of 0.84 without including previous history of GDM. Based on the variables used in these models, the ROC-AUC drop may be less severe. However, these models did not incorporate data on nulliparous women, or data from another cohort.

One of the strengths of our model is its high predictive capability. Another strength is its simplicity. As such, the model can be used by most current clinics since the variables it contains are commonly determined in regular clinical practice. Moreover, since those variables are evaluated in the first trimester, the model allows prediction of the development of GDM before it appears. Another point is that our study covers a group of great interest—nulliparous women—for whom clinicians do not have information about previous pregnancies.

Our model does have limitations. First, some potentially impactful variables were not studied owing to the limitations of the database or current practice, such as glucose levels. Second, this study was performed mainly on Caucasian women; thus, we should be careful about generalizing the results to women of other ethnicities. Third, since the model was designed for nulliparous women, it was expected to underperform in women with previous pregnancies.

Although the results obtained by cross-validation are good, to use our model in another population, an external validation study would need to be carried out. It is also our desire to improve the model in the future using some variables that were not studied in this retrospective study, such is glucose levels and ethnicity. Furthermore, a future study in multiparous pregnant women, including those with a history of GDM as a variable, would be desirable.

5. Conclusions

This study provides a tool to predict the onset of GDM in nulliparous women during the first trimester. It is more accurate than the any of the other currently available tools and allows earlier treatment intervention.

6. Abbreviations

GDM: gestational diabetes mellitus, ROC-AUC: receiver-operating characteristic area under the curve

Declarations

1. Ethics approval and consent to participate

This study was approved by the Ethic Committee of Biomedical Research of Cádiz at 13th July of 2020. In the present study due to the retrospective nature of the study, no informed consent was required.

2. Consent for publication

Not applicable

3. Availability of data and material

The data that support the findings of this study are available from the corresponding author (Fernandez Alba JJ), upon reasonable request.

4. Competing interests

The authors declare that they have no competing interests.

5. Funding:

Not applicable, no funding was received for the development of this study.

6. Author contributions

Fernández Alba JJ: Conceptualization, Methodology, Software, Reviewing and Editing

Castillo Lara M: Data curation, Writing, Reviewing and Editing

Jiménez Heras JM: Methodology, Data curation, Software, Reviewing and Editing

Moreno Cortés R: Investigation, Data curation

González Macías C: Investigation

Vilar Sánchez A: Investigation

Carral San Laureano F: Investigation

Moreno Corral L: Supervision

7. Acknowledgements

We would like to express our gratitude to Professor Emilio Sanchez Cantalejo, from Andalusian Public Health School, for his teaching in the applicability of predictive models in medicine.

References

- 1 Reece EA, Leguizamon G, Wiznitzer A. Gestational diabetes: the need for a common ground. *Lancet* 2009; 373:1789–97. [https://doi.org/10.1016/S0140-6736\(09\)60515-8](https://doi.org/10.1016/S0140-6736(09)60515-8).
- 2 Bardenheier BH, Elixhauser, A, Imperatore G, et al. Variation in prevalence of gestational diabetes mellitus among hospital discharges for obstetric delivery across 23 states in the United States. *Diabetes Care* 2013; 36:1209–14. <https://doi.org/10.2337/dc12-0901>.
- 3 Kim SY, Saraiva C, Curtis CM, Wilson HG, Troyan J, Sharma AJ. Fraction of gestational diabetes mellitus attributable to overweight and obesity by race/ethnicity, California, 2007–2009. *Am J Public Health* 2013; 10:65–72. <https://doi.org/10.2105/AJPH.2013.301469>.
- 4 Hillier TA, Pedula KL, Vesco KK, et al. Excess gestational weight gain. *Obstet Gynecol* 2008; 112:1007–14. <https://doi.org/10.1097/AOG.0b013e31818a9779>.
- 5 Metzger BE, Coustan DR, Trimble ER. Hyperglycemia and adverse pregnancy outcomes. *Clin Chem* 2019; 65:937–8. <https://doi.org/10.1373/clinchem.2019.303990>.
- 6 Farahvar S, Walfisch A, Sheiner E. Gestational diabetes risk factors and long-term consequences for both mother and offspring: a literature review. *Expert Rev Endocrinol Metab* 2018; 14:63–74. <https://doi.org/10.1080/17446651.2018.1476135>.
- 7 American Diabetes Association. Classification and diagnosis of diabetes: standards of medical care in diabetes-2020. *Diabetes Care* 2020;43:S14–S31. <https://doi.org/10.2337/dc20-S002>.
- 8 Bowers AJ, Zhou X. Receiver operating characteristic (ROC) area under the curve (AUC): a diagnostic measure for evaluating the accuracy of predictors of education outcomes. *J Educ Stud Placed Risk*. 2019; 24:20–46. <https://doi.org/10.1080/10824669.2018.1523734>.
- 9 Sweeting AN, Wong J, Appelblom H, et al. A novel early pregnancy risk prediction model for gestational diabetes mellitus. *Fetal Diagn Ther* 2019; 45:76–84. <https://doi.org/10.1159/000486853>.
- 10 Lovati E, Beneventi F, Simonetta M, et al. Gestational diabetes mellitus: including serum pregnancy-associated plasma protein-A testing in the clinical management of primiparous women? A case-control study. *Diabetes Res Clin Pract* 2013; 100:340–7. <https://doi.org/10.1016/j.diabres.2013.04.002>.
- 11 Syngelaki A, Kotecha, R, Pastides A, Wright A, Nicolaides WA. First-trimester biochemical markers of placentation in screening for gestational diabetes mellitus. *Metabolism* 2015; 64:1485–9. <https://doi.org/10.1016/j.metabol.2015.07.015>.

- 12 Nanda S, Savvidou M, Syngelaki A, Akolekar R, Nicolaides KH. Prediction of gestational diabetes mellitus by maternal factors and biomarkers at 11 to 13 weeks. *Prenat Diagn* 2011; 31:135–41. <https://doi.org/10.1002/pd.2636>.
- 13 Zheng T, Ye W, Wang X, et al. A simple model to predict risk of gestational diabetes mellitus from 8 to 20 weeks of gestation in Chinese women. *BMC Pregnancy Childbirth* 2019; 19:252–62. <https://doi.org/10.1186/s12884-019-2374-8>.
- 14 Maged, A.M., Moety, G.A., Mostafa, W.A., Hamed, D.A. Comparative study between different biomarkers for early prediction of gestational diabetes mellitus. *J Matern Fetal Neonatal Med* 2014; 27:1108–12. <https://doi.org/10.3109/14767058.2013.850489>.
- 15 Ahmad M, Becerra B, Hernandez D, Okpala P, Olney A, Becerra M. “You do it without their knowledge.” Assessing knowledge and perception of stealthing among college students. *Int J Environ Res Public Health* 2020; 17:3527. <https://doi.org/10.3390/ijerph17103527>.
- 16 Artzi NS, Shilo S, Hadar E, et al. Prediction of gestational diabetes based on nationwide electronic health records. *Nat Med* 2020; 26:71–6. <https://doi.org/10.1038/s41591-019-0724-8>.
- 17 Pintaudi B, Fresa R, Dalfra M, et al. “The risk stratification of adverse neonatal outcomes in women with gestational diabetes (STRONG) study. *Acta Diabetol* 2018; 55:1261–73. <https://doi.org/10.1007/s00592-018-1208-x>.
- 18 Evangelia C, Jie MA, Collins GJ, Steyerberg EW, Verbakel JY, Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol* 2019; 10:12–22. <https://doi.org/10.1016/j.jclinepi.2019.02.004>.
- 19 Savvidou M, Nelson SM, Makgoba M, Messow CM, Sattar N, Nicolaides K. First-trimester prediction of gestational diabetes mellitus: examining the potential of combining maternal characteristics and laboratory measures. *Diabetes* 2010; 59:3017–22. <https://doi.org/10.2337/db10-0688>.
- 20 Thériault S, Giguère Y, Massé J, Girouard J, Forest JC. Early prediction of gestational diabetes: a practical model combining clinical and biochemical markers. *Clin Chem Lab Med* 2016; 54:509–18. <https://doi.org/10.1515/cclm-2015-0537>.
- 21 Chang W, Cheng J, Allaire JJ, et al. (2021). shiny: Web Application Framework for R. R package version 1.6.0. <https://CRAN.R-project.org/package=shiny>
- 22 Landon MB, Spong CY, Thom E, et al. A multicenter, randomized trial of treatment for mild gestational diabetes. *N Engl J Med* 2009; 14:1339–48. <https://doi.org/10.1056/NEJMoa0902430>.
- 23 Crowther CA, Hiller JE, Moss JR, McPhee AJ, Jeffries WS, Robinson JS. Effect of treatment of gestational diabetes mellitus on pregnancy outcomes. *N Engl J Med* 2005; 24:2477–86. <https://doi.org/10.1056/NEJMoa042973>.

- 24 van Leeuwen M, Opmeer BC, Zweers EJ, et al. Estimating the risk of gestational diabetes mellitus: a clinical prediction model based on patient characteristics and medical history. BJOG 2010; 117:69–75. <https://doi.org/10.1111/j.1471-0528.2009.02425.x>.
- 25 R Core Team.2020. R: A language and environment for statistical computing. [online] R-project.org. Available at: <https://www.r-project.org/> [Accessed 17 September 2021].
- 26 Spain's National Health System Information System. 2017. Gobierno de España, Ministerio de Sanidad. [online] Available at: <https://www.mscbs.gob.es/estadEstudios/sanidadDatos/home.htm> [Accessed 5 May 2021].
- 27 US Department of Health and Human Services, National Institutes of Health & Eunice Kennedy Shriver National Institute of Child Health and Human Development. Am I at Risk for Gestational Diabetes? 2021. Available at: https://www.nichd.nih.gov/sites/default/files/publications/pubs/Documents/gestational_diabetes_2012.pdf [Accessed 2 May 2021].

Tables

Table 1. Characteristics of the population (n=1,946)

Gestational diabetes, n (%)	127 (6.5%)
Age, median (IQR)	31.79 (7.92)
Diabetes family history, n (%)	304 (15.6%)
Miscarriage, n (%)	
0	1598 (82.11%)
1	275 (14.13%)
2+	73 (3.6%)
BMI, n (%)	
Underweight (<19.5 kg/m ²)	56 (2.9%)
Normal (19.5 kg/m ² – 25 kg/m ²)	1121 (57.6%)
Overweight (25 kg/m ² – 30 kg/m ²)	499 (25.6%)
Obesity (>30 kg/m ²)	270 (13.9%)
Hypertension family history, n (%)	1622 (83.3%)
Twin pregnancy, n (%)	65 (3.3%)
Chronic hypertension, n (%)	16 (0.8%)
TSH first trimester, median (IQR)	1.89 (1.35)

IQR, interquartile range; BMI, body mass index; TSH, thyroid-stimulating hormone

Table 2. Variables of the model

Variables	Class	Coefficient	Std. Error	Z value	Pr (> z)
Intercept	Quantitative	-8.11	0.79	-10.27	<2e ⁻¹⁶
Age	Quantitative	0.09	0.02	4.69	2.74e ⁻⁰⁶
Underweight	Binomial	0.23	0.64	0.63	0.71
Overweight	Binomial	0.45	0.23	1.92	0.54
Obesity	Binomial	1.00	0.26	3.85	1.2e ⁻⁰⁴
Hypertension family history	Binomial	0.81	0.30	2.73	6e ⁻⁰³
Diabetes family history	Binomial	1.35	0.21	6.21	5.15e ⁻¹⁰
Twin pregnancy	Binomial	1.43	0.36	4.02	5.81e ⁻⁰⁵
Chronic hypertension	Binomial	8.47	3.38	2.51	1.22e ⁻⁰²
TSH first trimester	Quantitative	0.43	0.08	5.44	5.23e ⁻⁰⁸
Int Age x chr_Hta	Quantitative	-0.2	0.10	-2.07	0.038

Std. Error = standard error of the coefficient. Z value = 95% significance. Pr (<|z|) = level of significance (Z value) superior to 0.05 (with the exception BMI). The variable intercept is the intercept of the model itself. Age = modeled as a continuous variable in years. BMI = the gap is made with BMI values of 19.5, 25, 30 kg/m² for underweight, overweight, and obesity, respectively. TSH first trimester = level of TSH on the first trimester on mIU/L. Int Age x chr_Hta = interaction between age and chronic hypertension. Abbreviations: BMI, body mass index; TSH, thyroid-stimulating hormone

Table 3. Known GDM prediction models and their characteristics

Main author	Population size	GDM prevalence	Number of variables	Previous GDM	Type of model	ROC-AUC
Sweeting et al. ⁹	980	25.3%	10	Yes	Logistic reg.	0.901
Artzi et al. ¹⁶	588622	3.9%	255	Yes	Linear Machine Tree	0.85
Syngelaki et al. ¹¹	31225	2.5%	9	No	Logistic reg.	0.84
Savvidou et al. ¹⁹	372	33.3%	11	Yes	Logistic reg.	0.86
Nanda et al. ¹²	11464	2.6%	9	No	Logistic reg.	0.84
Thériault et al. ²⁰	7929	4.9%	6	Yes	Logistic reg.	0.79
Zheng et al. ¹³	4671	13.48%	8	No	Logistic reg.	0.766
Maged et al. ¹⁴	269	10%	14	Yes	Logistic reg.	0.74
Leeuwen et al. ²⁴	995	3.2%	8	No	Logistic reg.	0.77
Lobati et al. ¹⁰	673	45.6%	5	No	Logistic reg.	0.7

Figures

Figure 1

Population study algorithm selection

Figure 2

Boxplot showing the correlation between age (in years) and chronic hypertension. Women with chronic hypertension are significantly older than normal women.

Figure 3

ROC curve of the model The cut-off point represents the point closest to the coordinates 1,1 (sensitivity and specificity equal to 1). Additional cut-off points of interest: Youden with 0.65 sensitivity and 0.80 specificity, a high sensitivity point at 0.80 sensitivity and 0.59 specificity. Abbreviations: ROC, receiver-operating characteristic; AUC, area under the curve

Figure 4

PR curve of the model Abbreviations: PR, prediction recall; AUC, area under the curve