

A Model for Prediction of Spontaneous Preterm Birth in Asymptomatic Women

Kyung A. Lee, M.D., Ph.D.,¹ Moon Hee Chang,² Mi-Hye Park, M.D., Ph.D.,¹ Hyesook Park, M.D., Ph.D.,^{2,3}
Eun Hee Ha, M.D., Ph.D.,^{2,3} Eun Ae Park, M.D., Ph.D.,^{3,4} and Young Ju Kim, M.D., Ph.D.,^{1,3}

Abstract

Background: Preterm birth is a complex health problem with social, environmental, behavioral, and genetic determinants of an individual's risk and remains a major challenge in obstetrics. Recent research has caused improvements in predicting preterm birth; however, there is still controversy about the prediction of preterm birth in asymptomatic women. The purpose of this study was to determine if Bayesian filtering can be used in a clinical setting to predict spontaneous preterm birth in asymptomatic women.

Methods: A model of predicting spontaneous preterm birth using PopBayes based on a Bayesian filtering algorithm was developed using a previously collected dataset, then applied to a prospectively collected cohort of asymptomatic women who delivered singleton live newborns at or after 24 weeks of gestation. Cases complicated with major congenital malformations were excluded.

Results: The proportion of patients with spontaneous preterm birth was 18.4% (96 of 522) at <37 weeks gestation, 5.4% (28 of 522) at <34 weeks gestation, and 2.7% (14 of 522) at <32 weeks gestation. The match rates with the combination of demographic, clinical, and genetic factors using a Bayesian filtering method (PopBayes) were higher than the match rates using demographic and clinical factors only, including maternal age, maternal body mass index (BMI), prior preterm birth, education, occupation, income, and active and passive smoking. The match rates in preterm delivery before 32 weeks of gestation were higher than the match rates in preterm delivery before 37 and 34 weeks of gestation (94.3% vs. 84.7% and 82.0%, respectively). The negative predictive values for demographic, clinical, and genetic factors in predicting preterm delivery using PopBayes were consistently >90%.

Conclusions: We suggest that Bayesian filtering (PopBayes) is a customizable and useful tool in establishing a model for the prediction of preterm birth.

Introduction

PRETERM BIRTH REMAINS THE LEADING PROBLEM associated with perinatal morbidity and mortality worldwide.¹⁻⁵ To predict preterm birth is a key issue that needs to be addressed. There are numerous known factors related to preterm birth, including maternal age, body mass index (BMI), education, income, smoking, number of previous preterm births, cervical length (CL), fetal fibronectin (fFN), and genetic polymorphisms.^{1,2,4,6-17} However, every pregnant woman has a strong need for individualized risk assessment of preterm birth rather than a flood of information.

Recently, Bayesian filtering, a statistical technique, was introduced to sort out spam from e-mails. The basic principle of spam filtering is that Bayesian classifiers work by corre-

lating the use of words, with spam and nonspam e-mails, then using Bayesian statistics to calculate the probability that an e-mail is or is not spam. These processes, based on repetitive training or learning steps by using the existing dataset, make it possible to immediately determine whether or not a newly arrived e-mail is spam or not without iterating arduous statistical analyses. Smith et al.¹⁸ introduced a Bayesian modeling to develop a simple and robust method for predicting risk of cesarean section. With the use of Bayesian modeling, Bhattacharya et al.¹⁹ tried to identify women at a high recurrence risk of stillbirth in a second pregnancy.

The purpose of this study was to determine if Bayesian filtering is useful for prediction of spontaneous preterm birth in asymptomatic women and for counseling pregnant women for their individualized probability of preterm birth.

¹Department of Obstetrics and Gynecology, ²Department of Preventive Medicine, ³Medical Research Institute, and ⁴Department of Pediatrics, School of Medicine, Ewha Womans University, Ewha Womans University, Seoul, Korea.

Materials and Methods

Study population

The study population consisted of consecutive pregnant women who delivered singleton live newborns from 24 to 42 weeks of gestation at Ewha Womans University Hospital in Seoul, Korea, between 2003 and 2010. This study was approved by the Institutional Review Board of Ewha Womans University Hospital. We recruited patients who were registered for prenatal care at our hospital before 20 weeks of gestation. Participants who signed a consent form before enrollment were interviewed by trained interviewers, who recorded general epidemiologic and clinical data. Subject weights and heights were recorded, and blood samples were obtained according to standard protocols. Gestational age was determined in accordance with the date of onset of the last menstrual period (LMP) in most cases. When gestational age determined by the LMP date was not in agreement with fetal size verified by ultrasonography, however, we verified gestational age in agreement with dates at the first ultrasonographic estimation.

Women who had multiple gestations, intrauterine fetal death and fetuses complicated with major congenital malformations were excluded. To assess the clinical value of screening tests for women without symptoms at 20–28 weeks of gestation, we limited this study to nulliparas and multiparas who did not have spontaneous preterm labor or preterm premature rupture of membranes (PROM) at the time of enrollment. The principal outcome was spontaneous preterm birth before 37 weeks of gestation, which included births that followed spontaneous preterm labor, excluding medically indicated preterm birth with PROM or hypertensive disorders during pregnancy. Five hundred twenty-two women were enrolled, as follows: 96 women who underwent spontaneous preterm birth (<37 weeks gestation) with intact membranes and 426 women who delivered at term (\geq 37 weeks gestation).

Cervical length and fetal fibronectin

All CL measurements and fFN testing were done in an outpatient setting on asymptomatic patients at 20–28 weeks of gestation. All tests that were performed on labor and delivery were excluded because the tests were performed on symptomatic patients as part of a preterm labor evaluation. All CL measurements were measured using 4–8 MHz transvaginal probes (Accuvix XQ, Medison, Seoul, Korea) with an empty bladder. The optimal image was defined according to the criteria reported by Iams et al.⁶ The shortest functional CL was used because this has been found to be the most reproducible measurement. A short CL was defined as a CL <2.5 cm.

fFN testing was performed with a Dacron swab without the use of a speculum, which is according to an established protocol that has been validated previously at 20–28 weeks of gestation.²⁰ Vaginal swabs were sent for quantitative determination of the fFN concentration using an enzyme-linked immunosorbent assay (ELISA) (Fetal Fibronectin Immunoassay; Hologic Inc., Bedford, MA). An fFN concentration \geq 50 ng/mL was considered to be positive.

Genotyping analysis

Genotypes of paraoxonase (PON) 1 were analyzed according to an established protocol that has been validated

previously by our group and others.^{17,21} We extracted genomic DNA from maternal whole blood, which had been collected and stored at -80°C , using a QIAmp blood kit (Qiagen, Hilden, Germany). PON1 polymorphisms were analyzed according to Humbert et al.²¹ The genotyping of the PON1 polymorphism (rs662) was screened with a single base primer extension assay using an ABI PRISM SNaPshot Multiplex kit (ABI, Foster City, CA) according to the manufacturer's recommendation. Briefly, the genomic DNA flanking the single nucleotide polymorphism (SNP) of interest was amplified using a PCR reaction with forward and reverse primer pairs and standard PCR reagents in a 10- μL reaction volume containing 10 ng genomic DNA, 0.5 pM each oligonucleotide primer, 1 μL 10X PCR buffer, 250 mM dNTP (2.5 mM each), and 0.25 units i-StarTaq DNA polymerase (5 units/ μL) (iNtRON Biotechnology, Sungnam, Gyeonggi-Do, Korea). The PCR reactions were carried out as follows: 10 min at 95°C for 1 cycle, 35 cycles at 95°C for 30 s, 60°C for 1 min, 72°C for 1 min, and 1 cycle at 72°C for 10 min. After amplification, the PCR products were treated with 1 unit each of shrimp alkaline phosphatase (SAP) (Roche, Basle, Switzerland) and exonuclease I (USB Corporation, Cleveland, OH) at 37°C for 75 min and 72°C for 15 min to purify the amplified products.

The purified amplification products (1 μL) were added to a SNaPshot Multiplex Ready reaction mixture containing 0.15 pmol genotyping primer for the primer extension reaction. The primer extension reaction was carried out for 25 cycles at 96°C for 10 s, 50°C for 5 s, and 60°C for 30 s. The reaction products were treated with 1 unit SAP at 37°C for 1 hour and 72°C for 15 min to remove excess fluorescent dye terminators. The final reaction samples containing the extension products (1 μL) were added to 9 μL Hi-Di formamide (ABI). The mixture was incubated at 95°C for 5 min, followed by 5 min on ice, then analyzed by electrophoresis in an ABI Prism 3730xl DNA analyzer. Analysis was carried out using Genemapper software version 4.0 (ABI). The sequences of the primers were as follows: forward primer, GAGCACTTTATGGCACAATGA; reverse primer, ATAGTAGACAA C ATACGACCACGCTA; genotyping primer, TTTTCTTGA CCCCTACTTAC.

Determination of candidate risk factors for preterm birth

Initially, we searched for associations between preterm delivery and covariates, including maternal age, BMI at the time of delivery, parity, history of prior preterm delivery, education, occupation, income, active and passive smoking, CL, and fFN, using a chi-square test and Student's *t* test where appropriate. Logistic regression analyses were performed to evaluate the impact of the genetic polymorphisms of PON1 and haplotypes on the risk of preterm birth. All statistical analyses were performed using SPSS version 12.0 (SPSS Inc., Chicago, IL). Statistical significance was assumed at $p < 0.05$ (2-tailed).

Bayesian filtering (PopBayes)

After determination of candidate risk factors, we calculated conditional probabilities of preterm and term deliveries using PopBayes (Seoul, Korea), a statistical program based on Bayesian filtering, as follows:

TABLE 1. MATCH AND MISMATCH RATES OF RESULTS OF POPBAYES

	Preterm delivery	
	Yes	No
Positive test	a	b
Negative test	c	d

Match (%) = (a + d) / (a + b + c + d) × 100.
 Mismatch (%) = (b + c) / (a + b + c + d) × 100.

$$\Sigma \Pr(\text{Factor}|\text{Preterm}) = \Pr(\text{Factor1}|\text{Preterm}) \times \Pr(\text{Factor2}|\text{Preterm}) \times \Pr(\text{Factor3}|\text{Preterm}) \times \dots$$

$$\Sigma \Pr(\text{Factor}|\text{Term}) = \Pr(\text{Factor1}|\text{Preterm}) \times \Pr(\text{Factor2}|\text{Preterm}) \times \Pr(\text{Factor3}|\text{Preterm}) \times \dots$$

where Pr = probability.

The probability of preterm delivery was calculated as the total sum of conditional probabilities of demographic, clinical, and genetic factors, assuming that a gravida who delivered before 37 weeks of gestation and a gravida who delivered at term (≥ 37 weeks of gestation) were also calculated using the same methods. If the probability of preterm delivery was higher than the probability of term delivery, the test result was considered as preterm delivery, and *vice versa*.

The training dataset, a model development sample, was randomly selected from the original dataset of 522 patients and included two thirds of the study population, which consisted of 63 preterm deliveries before 37 weeks of gestation (positive outcome) and 286 term deliveries (negative outcome), thus containing a similar proportion of positive and negative outcomes to that of the test dataset, a model validation sample. The remaining one third of the study popu-

lation was added to the final test dataset, which included 96 preterm deliveries before 37 weeks of gestation and 426 term deliveries.

The final results are presented as match and mismatch rates (Table 1 and Fig. 1). If a test result using PopBayes was the same as the final outcome (a preterm or term delivery), the case was categorized as a match. If a test result was different from an outcome, the case was categorized as a mismatch. The match and mismatch rates indicate the numbers of match and mismatch cases among the total study population.

Results

The proportion of patients with spontaneous preterm births was 18.4% (96 of 522) at <37 weeks of gestation, 5.4% (28 of 522) at <34 weeks of gestation, and 2.7% (14 of 522) at <32 weeks of gestation. Basic demographic and clinical information are shown in Table 2. We enrolled multiparas and nulliparas in the study. There was no significant difference between the groups in terms of the rate of multiparity (*p* > 0.1). Cases in the preterm group had a significantly lower maternal age and BMI at the time of delivery and higher rates of prior preterm birth, maternal education ≤ 12 years, a short cervix, and a positive fFN than those in the term group (*p* < 0.05 for each). Prepregnancy BMI, rates of overweight and obesity, low education of parents, low income, and active and passive smoking were higher among the preterm group than the term group, but the differences were not statistically significant (*p* > 0.1).

Table 3 presents the frequencies of genotypes and alleles of the SNPs in the PON1 gene in the preterm and term groups. There was no significant difference between the two groups in terms of the frequency of genotypes in the PON1 gene. However, the preterm group had a significantly higher rate of allele A than the term group (*p* < 0.05).

The match rate, mismatch rate, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for demographic, clinical, and genetic factors in

Test results

P(test term)*P(term)	P(test preterm)*P(preterm)	Result	Category	Correct?
6.8633963559e-006	2.8901512097e-009	0	0	O
3.9796573747e-005	5.3013010108e-009	0	0	O
6.9906813370e-007	1.6829809019e-010	0	0	O
7.2899107693e-006	3.9759757581e-009	0	0	O
1.3907208454e-005	2.3326446493e-008	0	1	X
1.5403943026e-010	1.5586040659e-005	1	1	O
1.6188199671e-008	2.6586997036e-005	1	0	X
1.1535534208e-010	1.3082490605e-005	1	1	O
⋮	⋮	⋮	⋮	⋮
1.8112843602e-010	7.9364272908e-006	1	1	O
2.3322848424e-009	8.6308079068e-007	1	0	X
2.5732930093e-012	2.0957571735e-006	1	1	O
6.3122870480e-009	3.2524494469e-005	1	1	O
1.8890494300e-011	9.4644543197e-006	1	1	O
8.1104930091e-009	2.4638750051e-005	1	1	O
8.7774752915e-009	3.2732429230e-005	1	1	O

MATCH: 428/522 = 81.992337164751%, 428.0/522.0 = 81.992337164751%
 MISMATCH: 94/522 = 18.007662835249%
 UNKNOWN: 0/522 = 0.0%

FIG. 1. Test results presented as match and mismatch rates using PopBayes. The probability of preterm and term deliveries was calculated as the total sums of conditional probabilities of demographic, clinical, and genetic factors. If the probability of preterm delivery was higher than that of term delivery, the test result was considered as preterm delivery, and *vice versa*. Result, that is, a test result of an individual case, is shown as 0 (negative outcome, or term delivery) or 1 (positive outcome, or preterm delivery). Category represents an outcome that a woman delivered in preterm or term gestation. Correct? is shown as O (match) or X (mismatch).

TABLE 2. DEMOGRAPHIC AND CLINICAL CHARACTERISTICS OF STUDY POPULATION

Characteristic	Preterm (n=96)	Term (n=426)	p value
Maternal age, years ^a	30.6 ± 4.2	32.3 ± 4.1	< 0.001
20–24	8 (8.3%)	8 (1.9%)	< 0.01
25–29	29 (30.2%)	102 (23.9%)	
30–34	44 (45.8%)	186 (43.7%)	
35–39	13 (13.5%)	113 (26.5%)	
≥ 40	2 (2.1%)	17 (4.0%)	
Prepregnancy body weight (kg) ^a	54.5 ± 9.6	54.7 ± 8.2	NS
Maternal body weight at delivery (kg) ^a	64.1 ± 10.8	67.9 ± 8.6	< 0.01
Weight gain per week (kg) ^a	0.30 ± 0.1	0.34 ± 0.1	NS
Maternal height (cm) ^a	160.5 ± 4.5	161.5 ± 4.8	NS
Prepregnancy BMI ^{a,b}	21.1 ± 3.7	20.9 ± 3.1	NS
< 18.5	10 (22.7%)	76 (20.1%)	
18.5–22.9	24 (54.5%)	231 (60.9%)	
23.0–24.9	5 (11.4%)	34 (9.0%)	
≥ 25	5 (11.4%)	38 (10.0%)	
Maternal BMI at delivery ^{a,c}	24.9 ± 4.0	26.0 ± 3.0	< 0.01
< 18.5	0 (0.0%)	0 (0.0%)	< 0.001
18.5–22.9	33 (34.4%)	58 (13.8%)	
23.0–24.9	23 (24.0%)	111 (26.4%)	
≥ 25	40 (41.7%)	251 (59.8%)	
BMI increase per week ^a	0.12 ± 0.05	0.13 ± 0.04	NS
Gestational age at delivery (weeks) ^a	34.5 ± 2.5	39.6 ± 1.1	< 0.001
Multiparity ^d	57 (60.0%)	224 (52.7%)	NS
Prior preterm birth ^e	6 (11.3%)	10 (2.9%)	< 0.05
Maternal education ≤ 12 years ^f	37 (41.6%)	82 (19.9%)	< 0.001
Paternal education ≤ 12 years ^g	7 (23.3%)	52 (14.5%)	NS
With maternal occupation ^h	9 (28.1%)	130 (36.5%)	NS
Income < USD 2,500 ⁱ	15 (48.4%)	169 (46.9%)	NS
Smoking ^j	4 (12.5%)	37 (10.3%)	NS
Passive smoking ^k	16 (50.0%)	139 (38.7%)	NS
Cervical length (cm) ^a	2.83 ± 1.35	3.83 ± 0.66	< 0.001
Short cervix	32 (33.3%)	4 (0.9%)	< 0.001
Positive fetal fibronectin	20 (20.8%)	1 (0.2%)	< 0.001

^aValues are given as mean ± standard deviation.

^bDenominators are 44 for the preterm group and 379 for the term group.

^cA denominator is 420 for term group.

^dDenominators are 95 for the preterm group and 425 for the term group.

^eDenominators are 53 for the preterm group and 347 for the term group.

^fDenominators are 89 for the preterm group and 413 for the term group.

^gDenominators are 30 for the preterm group and 359 for the term group.

^hDenominators are 32 for the preterm group and 356 for the term group.

ⁱDenominators are 31 for the preterm group and 360 for the term group.

^jDenominators are 32 for the preterm group and 358 for the term group.

^kDenominators are 32 for the preterm group and 359 for the term group.

BMI, body mass index; NS, not significant; USD, US dollars.

predicting preterm delivery using PopBayes are shown in Table 4. The factors included in the final Bayesian model were maternal age, BMI, parity, history of prior preterm birth, maternal and paternal education, occupation, income, active and passive smoking, CL, fFN, and genetic analysis (PON1). The overall match rates, sensitivity, specificity, PPV, and NPV using demographic, clinical, and genetic factors were higher than those using demographic and clinical factors only, including maternal age, maternal BMI, prior preterm birth, education, occupation, income, and active and passive smoking. The match rates in the preterm delivery group before 32 weeks of gestation were higher than in the preterm delivery group before 37 and 34 weeks of gestation. The sensitivities for short cervix, fFN, and PON1 in the preterm delivery group before 37 weeks of gestation were 33.3%, 20.8%, and 43.5%, respectively, and those values were lower

than the values according to PopBayes, in which demographic, clinical, and genetic factors were considered as covariates.

Discussion

Principal findings

Our data showed that the match rates with the combination of demographic, clinical, and genetic factors using PopBayes were higher than the match rates using demographic and clinical factors only, including maternal age, maternal BMI, prior preterm birth, education, occupation, income, and active and passive smoking. Furthermore, our findings suggest that the match rates for preterm deliveries before 32 weeks of gestation were higher than preterm deliveries before 37 and 34 weeks of gestation (94.3% vs. 84.7% and 82.0%,

TABLE 3. FREQUENCIES OF GENOTYPES AND ALLELES OF SINGLE NUCLEOTIDE POLYMORPHISMS IN PON1 GENE IN PRETERM AND TERM GROUPS

Characteristic	Preterm n (%)	Term n (%)	p value	OR (95% CI)
Genotype ^a				
AA	13 (10.7)	33 (16.7)	NS	1.00
AG	38 (48.7)	137 (44.6)		0.70 (0.34-1.47)
GG	27 (34.6)	137 (44.6)		0.50 (0.23-1.07)
Allele ^b				
A	57 (43.5)	210 (32.9)	<0.05	1.00
G	74 (56.5)	429 (67.1)		0.64 (0.43-0.93)

^aDenominators are 78 for the preterm group and 307 for the term group.

^bDenominators are 131 for the preterm group and 639 for the term group.

A, adenine; CI, confidence interval; G, guanine; NS, not significant; OR, odds ratio.

respectively). The NPVs for demographic, clinical, and genetic factors in predicting preterm delivery using PopBayes were consistently >90%.

Maternal demographic and clinical factors

The frequencies of low maternal age (<25 years) and low maternal educational status (≤12 years) were higher in the preterm group than in the term group (8.3% vs. 1.9%, *p*<0.01, and 41.6% vs. 19.9%, *p*<0.001, respectively). These findings were consistent with previous studies. Low educational status and low maternal age are correlated with an increased risk of preterm birth.^{11,14}

Hendler et al.² showed that maternal thinness before pregnancy was associated with increased spontaneous preterm birth. In our study, the frequency of prepregnancy BMI <18.5 kg/m² was higher among the preterm group than the term group but failed to reach statistical significance (22.7% vs. 20.1%, *p*>0.1). Similarly, the preterm group had a lower BMI increase or weight gain per week during pregnancy than the term group, and these results reached statistical significance (0.12±0.05 vs. 0.13±0.04, *p*>0.1, and 0.30±0.1 vs. 0.34±0.1, *p*>0.1, respectively). According to a recent meta-

analysis, a maternal anthropometric measurement, such as BMI,²² determination of adequacy of weight gained, and height measurement, cannot be a single predictor for spontaneous preterm birth and should be taken into account with multiple other factors. Therefore, we considered prepregnancy BMI as a candidate factor for spontaneous preterm birth in combination with other factors by calculating the sum of conditional probabilities of each factor, using PopBayes based on Bayesian filtering.

One of the most important risk factors for preterm birth is prior preterm birth. Although most women who experience a preterm birth will deliver at term in subsequent pregnancies, the recurrence risk for preterm birth is ≥2-fold. In the current study, the frequency of women experiencing prior preterm birth was higher in the preterm group than in the term group (11.3% vs. 2.9%, *p*<0.05). This finding was consistent with previous studies.⁷ In a preterm prediction study, Mercer et al.⁷ demonstrated that the recurrent risk of preterm birth increases as the number of prior preterm births increases, with a 2-fold rise for each prior preterm birth.

CL and fFN

Cases in the preterm group had higher rates of a short cervix and positive fFN than those in the term group (*p*<0.05 for each). Of note, treatment of women with a short cervix with progesterone reduces the rate of spontaneous preterm delivery,¹² and, thus, transvaginal sonographic measurements of CL might be useful in both low-risk and high-risk populations. In the current study, the sensitivity for a short cervix in preterm delivery before 37 weeks of gestation was as low as 33.3%, which was consistent with previous studies.^{6,8,9} In the case of fFN, the sensitivity in predicting preterm birth was 20.8%. The low sensitivity of CL in predicting preterm birth in a low-risk population prompted us to consider CL or fFN not as a single marker but as one of the major risk factors for preterm birth.

Genetic factors (PON1)

Our group has demonstrated the relationship between various genetic polymorphisms and preterm birth in previous studies.¹⁵⁻¹⁷ In the current study, we showed that although there was no significant difference between the two groups in

TABLE 4. DIAGNOSTIC INDICES AND PREDICTIVE VALUES OF POPBAYES BASED ON BAYESIAN FILTERING FOR PREDICTION OF PRETERM DELIVERY

Preterm delivery ^a	<37 weeks		<34 weeks		<32 weeks	
	For patients ^b	For doctors ^c	For patients ^b	For doctors ^c	For patients ^b	For doctors ^c
Match	81.6 (426/522)	82.0 (428/522)	78.7 (411/522)	84.7 (442/522)	88.3 (461/522)	94.3 (492/522)
Mismatch	18.4 (96/522)	18.0 (94/522)	21.3 (111/522)	15.3 (80/522)	11.7 (61/522)	5.8 (30/522)
Sensitivity	66.7 (64/96)	68.8 (66/96)	57.1 (16/38)	57.1 (16/38)	50.0 (7/14)	64.3 (9/14)
Specificity	85.0 (362/426)	85.0 (362/426)	80.0 (395/494)	86.2 (426/494)	89.4 (454/508)	95.1 (483/508)
Positive predictive value	50.0 (64/128)	50.8 (66/130)	14.0 (16/115)	19.0 (16/84)	11.5 (7/61)	26.5 (9/34)
Negative predictive value	91.9 (32/394)	92.4 (30/392)	97.1 (395/407)	97.3 (426/438)	98.5 (454/461)	99.0 (483/488)

^aData are given as percentage (n/N).

^bSociodemographic data, including maternal age, BMI (body weight and height), parity, history of prior preterm birth, maternal and paternal education, occupation, income, and active and passive smoking, were included in analysis.

^cIn addition to sociodemographic data, cervical length, fetal fibronectin, and genetic analysis (PON1) were added to analysis for further counseling with a maternal-fetal medicine specialist.

terms of the frequency of genotypes in the PON1 gene, cases in the preterm group had a significantly higher rate of allele A than the term group ($p < 0.05$). Therefore, we included the PON1 gene in our final preterm prediction model.

Strengths and weaknesses

Our study differs from previous studies in that we introduced an analytic method based on Bayesian filtering into predicting preterm birth at <37 , <34 , and <32 weeks of gestation. A definite advantage of our method is to simplify analyses when new data emerge and to give each patient an individualized probability of preterm delivery, even though the probability is yes or no. Furthermore, it is more convenient to collect and analyze data on a large scale. Once the established data can be shared online, this will facilitate clinical counseling on preterm birth. Another advantage to an analytic method based on Bayesian filtering is that tailoring of candidate factors can be easily performed, as indicated in Table 4. Patients can self-check the probabilities of preterm birth by inputting their own sociodemographic and clinical data. Physicians can provide patients their personalized probabilities of preterm birth by collecting demographic and clinical data and additionally obtaining data from ultrasound examinations of CL, fFN, and genetic polymorphisms.

A limitation of our study findings was the lack of data in relation to such factors as sociodemographic factors (education, occupation, income, smoking, and genetic factors [PON1]). Although two major risk factors for spontaneous preterm birth (CL and fFN) were completely recorded, we admit that the problem of missing data was inevitable because the dataset for data training in this study was collected retrospectively. Despite this limitation, we introduced a Bayesian filter for predicting spontaneous preterm birth. A Bayesian filter is constantly self-adapting by learning from new cases of preterm births in a prospective study on a large scale.

Conclusions

We suggest that Bayesian filtering (PopBayes) is a useful tool in establishing a model of preterm birth prediction. In the future, the current study results can facilitate a prospective multicenter study involving the prediction of spontaneous preterm births.

Acknowledgments

This research was supported by Seoul Research and Business Development Grant No. 10621, Republic of Korea.

Disclosure Statement

The authors have no conflicts of interest to report.

References

- Honest H, Bachmann LM, Gupta JK, Kleijnen J, Khan KS. Accuracy of cervicovaginal fetal fibronectin test in predicting risk of spontaneous preterm birth: Systematic review. *BMJ* 2002;325:301.
- Hendler I, Goldenberg RL, Mercer BM, et al. The Preterm Prediction Study: Association between maternal body mass index and spontaneous and indicated preterm birth. *Am J Obstet Gynecol* 2005;192:882–886.
- Hill JL, Campbell MK, Zou GY, et al. Prediction of preterm birth in symptomatic women using decision tree modeling for biomarkers. *Am J Obstet Gynecol* 2008;198: e461–467.
- Fox NS, Saltzman DH, Klauser CK, Peress D, Gutierrez CV, Rebarber A: Prediction of spontaneous preterm birth in asymptomatic twin pregnancies with the use of combined fetal fibronectin and cervical length. *Am J Obstet Gynecol* 2009;201:311–315.
- Kramer MR, Cooper HL, Drews-Botsch CD, Waller LA, Hogue CR: Metropolitan isolation segregation and black-white disparities in very preterm birth: A test of mediating pathways and variance explained. *Soc Sci Med* 2010;71: 2108–2116.
- Iams JD, Goldenberg RL, Meis PJ, et al. The length of the cervix and the risk of spontaneous premature delivery. National Institute of Child Health and Human Development Maternal Fetal Medicine Unit Network. *N Engl J Med* 1996;334:567–572.
- Mercer BM, Goldenberg RL, Das A, et al. The preterm prediction study: A clinical risk assessment system. *Am J Obstet Gynecol* 1996;174:1885–1893.
- Iams JD, Goldenberg RL, Mercer BM, et al. The Preterm Prediction Study: Can low-risk women destined for spontaneous preterm birth be identified? *Am J Obstet Gynecol* 2001;184:652–655.
- Berghella V, Bega G, Tolosa JE, Berghella M. Ultrasound assessment of the cervix. *Clin Obstet Gynecol* 2003;46:947–962.
- Cnattingius S. The epidemiology of smoking during pregnancy: Smoking prevalence, maternal characteristics, and pregnancy outcomes. *Nicotine Tobacco Res* 2004;6 (Suppl 2):S125–140.
- Thompson JM, Irgens LM, Rasmussen S, Daltveit AK. Secular trends in socioeconomic status and the implications for preterm birth. *Paediatr Perinat Epidemiol* 2006; 20:182–187.
- Fonseca EB, Celik E, Parra M, Singh M, Nicolaides KH. Progesterone and the risk of preterm birth among women with a short cervix. *N Engl J Med* 2007;357:462–469.
- Nabet C, Lelong N, Ancel PY, Saurel-Cubizolles MJ, Kaminski M. Smoking during pregnancy according to obstetric complications and parity: Results of the EUROPOP study. *Eur J Epidemiol* 2007;22:715–721.
- Smith LK, Draper ES, Manktelow BN, Dorling JS, Field DJ. Socioeconomic inequalities in very preterm birth rates. *Arch Dis Child Fetal Neonatal Ed* 2007;92:F11–14.
- Suh YJ, Ha EH, Park H, Kim YJ, Kim H, Hong YC. GSTM1 polymorphism along with PM10 exposure contributes to the risk of preterm delivery. *Mutat Res* 2008;656:62–67.
- Suh YJ, Kim YJ, Park H, Park EA, Ha EH. Oxidative stress-related gene interactions with preterm delivery in Korean women. *Am J Obstet Gynecol* 2008;198:541–547.
- Lee BE, Park H, Park EA, et al. Paraoxonase 1 gene and glutathione S-transferase mu 1 gene interaction with preterm delivery in Korean women. *Am J Obstet Gynecol* 2010; 203:569.
- Smith GC, Dellens M, White IR, Pell JP. Combined logistic and Bayesian modeling of cesarean section risk. *Am J Obstet Gynecol* 2004;191:2029–2034.
- Bhattacharya S, Prescott GJ, Black M, Shetty A. Recurrence risk of stillbirth in a second pregnancy. *Br J Obstet Gynaecol* 2010;117:1243–1247.
- Goldenberg RL, Iams JD, Das A, et al. The Preterm Prediction Study: Sequential cervical length and fetal fibronectin

- testing for the prediction of spontaneous preterm birth. National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network. *Am J Obstet Gynecol* 2000;182:636–643.
21. Humbert R, Adler DA, Distechi CM, Hassett C, Omiecinski CJ, Furlong CE. The molecular basis of the human serum paraoxonase activity polymorphism. *Nat Genet* 1993;3:73–76.
22. Honest H, Bachmann LM, Ngai C, Gupta JK, Kleijnen J, Khan KS. The accuracy of maternal anthropometry measurements as predictor for spontaneous preterm birth—A systematic review. *Eur J Obstet Gynecol Reprod Biol* 2005;119:11–20.

Address correspondence to:

Young Ju Kim, M.D., Ph.D.

Department of Obstetrics and Gynecology

Ewha Womans University School of Medicine

Seoul, 158-056

Korea

E-mail: kkyj@ewha.ac.kr

