

Original Article
Medical Informatics



Association of Gastroesophageal Reflux Disease with Preterm Birth: Machine Learning Analysis

Kwang-Sig Lee ,^{1*} Eun Sun Kim ,^{2*} Do-young Kim,^{1,3} In-Seok Song ,⁴ and Ki Hoon Ahn ³

¹AI Center, Korea University Anam Hospital, Seoul, Korea

²Department of Gastroenterology, Korea University Anam Hospital, Seoul, Korea

³Department of Obstetrics and Gynecology, Korea University Anam Hospital, Seoul, Korea

⁴Department of Oral and Maxillofacial Surgery, Korea University Anam Hospital, Seoul, Korea



Received: Apr 22, 2021

Accepted: Sep 16, 2021

Address for Correspondence:

Ki Hoon Ahn, MD, PhD

Department of Obstetrics and Gynecology,
Korea University Anam Hospital, 73
Goryeodae-ro, Seongbuk-gu, Seoul 02841,
Korea.

E-mail: akh1220@hanmail.net

*Kwang-Sig Lee and Eun Sun Kim contributed equally to this work.

© 2021 The Korean Academy of Medical Sciences.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ORCID iDs

Kwang-Sig Lee
<https://orcid.org/0000-0002-0576-0098>

Eun Sun Kim
<https://orcid.org/0000-0003-1820-459X>

In-Seok Song
<https://orcid.org/0000-0002-0763-8838>

Ki Hoon Ahn
<https://orcid.org/0000-0002-6314-4621>

Funding

This work was supported by the Korea University Medical Center (No. K1925051) and the Ministry of Science and ICT of South Korea under the Information Technology

ABSTRACT

Background: This study used machine learning and population data for testing the associations of preterm birth with gastroesophageal reflux disease (GERD) and periodontitis.

Methods: Retrospective cohort data came from Korea National Health Insurance Service claims data for all women who aged 25–40 years and gave births for the first time as singleton pregnancy during 2015–2017 (405,586 women). The dependent variable was preterm birth during 2015–2017 and the independent variables were GERD (coded as no vs. yes) for each of the years 2002–2014, periodontitis (coded as no vs. yes) for each of the years 2002–2014, age (year) in 2014, socioeconomic status in 2014 measured by an insurance fee, and region (city) (coded as no vs. yes) in 2014. Random forest variable importance was adopted for finding main predictors of preterm birth and testing its associations with GERD and periodontitis.

Results: Based on random forest variable importance, main predictors of preterm birth during 2015–2017 were socioeconomic status in 2014, age in 2014, GERD for the years 2012, 2014, 2010, 2013, 2007 and 2009, region (city) in 2014 and GERD for the year 2006. The importance rankings of periodontitis were relatively low.

Conclusion: Preterm birth has a stronger association with GERD than with periodontitis. For the prevention of preterm birth, preventive measures for GERD would be essential together with the improvement of socioeconomic status for pregnant women. Especially, it would be vital to promote active counseling for general GERD symptoms (neglected by pregnant women).

Keywords: Premature Birth; Gastroesophageal Reflux; Periodontitis; Socioeconomic Status; Age

INTRODUCTION

Preterm birth, defined as “birth between 20 and 37 weeks of gestation”, is a major contributor to disease burden for infants and children in the world.¹⁻⁴ One of every 10 newborns were preterm babies in the United States during 2003–2012, i.e., 5,042,982 (12.2%) of 41,206,315 births.¹ The share of preterm births increased rapidly from 4.3% to 6.0% in Korea during 2001–2010.² On a global level, 15 million newborns are preterm babies every year and this is responsible for 965,000 neonatal deaths and 125,000 deaths among children aged 1–5 years.

Research Center support program supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation) (No. IITP-2018-0-01405). The funder had no role in the design of the study, the collection, analysis and interpretation of the data and the writing of the manuscript.

Disclosure

The authors have no potential conflicts of interest to disclose.

Data Statement

The data presented in this study are not publicly available. But the data are available from the corresponding author upon reasonable request and under the permission of Korea National Health Insurance Service.

Author Contributions

Conceptualization: Lee KS, Kim ES, Song IS, Ahn KH. Data curation: Lee KS, Kim ES, Kim DY, Ahn KH. Formal analysis: Lee KS, Kim ES, Kim DY, Ahn KH. Investigation: Lee KS, Kim ES, Kim DY, Ahn KH. Methodology: Lee KS, Kim ES, Kim DY, Ahn KH. Software: Lee KS, Kim ES, Kim DY, Ahn KH. Validation: Lee KS, Kim ES, Kim DY, Ahn KH. Writing - original draft: Lee KS, Kim ES, Song IS, Ahn KH. Writing - review & editing: Lee KS, Kim ES, Song IS, Ahn KH.

Three quarters of this mortality is estimated to be avoidable with cost-effective interventions.^{3,4} Independent suggestions have been made on a positive association between gastroesophageal reflux disease (GERD) and periodontitis and on a positive linkage between periodontitis and preterm birth (PTB).⁵⁻¹¹ Likewise, it would be natural to expect a positive relationship between GERD and preterm birth. But little endeavor has been made in this direction.

One previous study adopted machine learning and general hospital data for testing an association among spontaneous preterm labor and birth (SPLB), GERD and periodontitis.¹² Data for this study came from 731 obstetric patients in Anam Hospital (Seoul, Korea) during January 5, 1995–August 28, 2018. Random forest variable importance, the contribution of a certain variable for the performance of the random forest, was adopted for finding main predictors of SPLB and testing its associations with GERD and periodontitis. Based on random forest variable importance, main predictors of SPLB were age, education, GERD, *Helicobacter pylori*, upper gastrointestinal tract symptoms, various types of body mass index (BMI), diabetes and hypertension, parity, pelvic inflammatory disease, and medication such as progesterone, calcium channel blocker. Periodontitis ranked twenty second. It was concluded that SPLB had a stronger association with GERD than with periodontitis and, for preventing SPLB, preventive measures for hypertensive disorder, GERD and diabetes mellitus would be very important together with appropriate BMI management and effective progesterone and calcium-channel-blocker medications.

However, the data in this previous study was relatively small (731 participants). Machine learning analysis with population data is expected to bring more conclusive evidence on the association among the three diseases above. In this context, this study uses machine learning and population data for testing the following hypotheses:

Hypothesis 1: Preterm birth associates positively with GERD.

Hypothesis 2: Preterm birth associates positively with periodontitis.

METHODS

Participants

Retrospective cohort data came from Korea National Health Insurance Service claims data for all women in Korea who aged 25–40 years and gave births for the first time as singleton pregnancy during 2015–2017 (405,586 women). The dependent variable was preterm labor and birth during 2015–2017 (birth between 20 weeks and 0 day and 36 weeks and 6 days of gestation). Four types of preterm labor and birth were defined based on International Classification of Diseases, 10th Revision (ICD-10) code: 1) PTB 1: preterm birth with premature rupture of membranes (PROM) only; 2) PTB 2: preterm labor and birth without PROM; 3) PTB 3: PTB 1 or PTB 2; 4) PTB 4: PTB 3 or other indicated preterm birth (**Supplementary Table 1**). This variable was coded as “no” vs. “yes”. And the following 29 independent variables were included in this study: 1) GERD (coded as no vs. yes) for each of the years 2002–2014, i.e., GERD_2002, ..., GERD_2014; 2) periodontitis (coded as no vs. yes) for each of the years 2002–2014, i.e., Periodontitis_2002, ..., Periodontitis_2014; 3) age (year) in 2014; 4) socioeconomic status in 2014 measured by an insurance fee with the range of 1 (the highest group) to 20 (the lowest group); and 5) region (city) (coded as no vs. yes) in 2014. GERD (GERD_2002, ..., GERD_2014) and periodontitis (Periodontitis_2002, ..., Periodontitis_2014) were screened from ICD-10 codes (**Supplementary Table 1**).

Analysis

Logistic regression, the random forest and the artificial neural network were adopted and compared for predicting preterm birth.^{12,13} Data on 402,092 observations with full information were split into training and validation sets with a 70:30 ratio (281,464 vs. 120,628 observations). Accuracy, a ratio of correct predictions among 120,628 observations, was used as a standard for validating the models. Random forest variable importance, the contribution of a certain variable for the performance of the random forest, was used for finding main predictors of preterm birth and testing its associations with GERD and periodontitis. R-Studio 1.3.959 (R-Studio Inc., Boston, MA, USA) was employed for the analysis during September 2019–May 2020.

Ethics statement

This retrospective study complied with the tenets of the Helsinki Declaration and was approved by the Institutional Review Board (IRB) of Korea University Anam Hospital on November 5, 2018 (2018AN0365). Informed consent was waived by the IRB.

RESULTS

Descriptive statistics are shown for participants' preterm birth and its categorical predictors in **Table 1**. Among 405,586 participants, 21,961 (5.41%), 9,034 (2.23%), 28,054 (6.92%) and 29,167 (7.19%) belonged to PTB 1, 2, 3 and 4, respectively. Among the participants, 376,384 (92.8%) resided in urban areas. Indeed, the average age and socioeconomic status of the participants were 31 and 11, respectively. During 2002–2014, the prevalence of GERD was consistently higher for those with preterm birth than for those without the outcome: 10.54%, 10.33%, 10.50%, 10.47% and 7.90% for PTB 1, PTB 2, PTB 3, PTB 4 and the normal in 2014, respectively; and 0.79%, 0.80%, 0.78%, 0.79% and 0.62% for PTB 1, PTB 2, PTB 3, PTB 4 and the normal in 2002, respectively (**Table 2**). A similar pattern was found for the prevalence of periodontitis: 0.09%, 0.09%, 0.08%, 0.08% and 0.05% for PTB 1, PTB 2, PTB 3, PTB 4 and the normal in 2014, respectively; and 0.07%, 0.07%, 0.07%, 0.07% and 0.04% for PTB 1, PTB 2, PTB 3, PTB 4 and the normal in 2002, respectively.

The accuracy of the random forest was similar with those of logistic regression and the artificial neural network (94.56%, 97.79%, 93.07% and 92.80% for PTB 1, PTB 2, PTB 3 and PTB 4 in **Table 3**, respectively). Based on random forest variable importance in **Fig. 1**, main predictors of preterm birth during 2015–2017 were socioeconomic status in 2014 (240.28), age in 2014 (221.13), GERD for the years 2012 (42.24), 2014 (38.86), 2010 (37.76), 2013 (36.64), 2007 (35.01) and 2009 (34.39), region in 2014 (34.36) and GERD for the year 2006 (31.98). These values were the average for PTB 1, PTB 2, PTB 3 and PTB 4 (**Supplementary Fig. 1** for each of PTB 1, PTB 2, PTB 3 and PTB 4). Top-10 predictors for PTB 2 were socioeconomic status in 2014 (121.02), age in 2014 (112.27), GERD for the years 2012 (21.31), 2010 (19.61), 2011 (19.53), 2014 (19.26) and 2009 (19.22), region in 2014 (16.58) and GERD for the years 2013 (16.35) and 2006 (15.80). Likewise, top-10 predictors for PTB 4 were socioeconomic status in 2014 (306.92), age in 2014 (294.18), GERD for the years 2012 (56.84), 2014 (49.77), 2010 (48.33) and 2013 (47.89), region in 2014 (46.74) and GERD for the years 2007 (46.47), 2008 (42.78) and 2009 (41.78). The importance rankings of periodontitis were relatively low across board.

The results of logistic regression are shown in **Tables 4** and **5**. The odds of PTB 2 (or PTB 4) will increase by 6.6% (or 3.1%) if age increases by 1 year. The odds of PTB 4 will go up by

Table 1. Descriptive statistics on PTB and its categorical predictors

Variables	No	Yes	Yes (%)
PTB 1 ^a	383,625	21,961	5.41
PTB 2	396,552	9,034	2.23
PTB 3	377,532	28,054	6.92
PTB 4	376,419	29,167	7.19
GERD_2002	402,549	3,037	0.75
GERD_2003	401,602	3,984	0.98
GERD_2004	400,395	5,191	1.28
GERD_2005	398,792	6,794	1.68
GERD_2006	396,672	8,914	2.20
GERD_2007	392,581	13,005	3.21
GERD_2008	389,565	16,021	3.95
GERD_2009	383,767	21,819	5.38
GERD_2010	379,903	25,683	6.33
GERD_2011	376,073	29,513	7.28
GERD_2012	372,081	33,505	8.26
GERD_2013	370,426	35,160	8.67
GERD_2014	366,591	38,995	9.61
Periodontitis_2002	405,388	198	0.05
Periodontitis_2003	405,321	265	0.07
Periodontitis_2004	405,181	405	0.10
Periodontitis_2005	405,155	431	0.11
Periodontitis_2006	404,994	592	0.15
Periodontitis_2007	405,272	314	0.08
Periodontitis_2008	405,282	304	0.07
Periodontitis_2009	405,233	353	0.09
Periodontitis_2010	405,243	343	0.08
Periodontitis_2011	405,287	299	0.07
Periodontitis_2012	405,327	259	0.06
Periodontitis_2013	405,312	274	0.07
Periodontitis_2014	405,342	244	0.06
Region (city) in 2014	29,202	376,384	92.80

PTB = preterm birth, GERD = gastroesophageal reflux disease.

^aPTB during 2015–2017.

3.2% if socioeconomic status go down by 10 units, e.g., from 1 (the highest group) to 10 (the middle group). The odds of PTB 4 is greater by 8.0% for those with GERD in 2011 than those without it. Likewise, the odds of PTB 4 is higher by 7.9% for those with GERD in 2009 than those without it. Logistic regression requires an unrealistic assumption of *ceteris paribus*, i.e., “all the other variables staying constant”. In this context, the results of logistic regression would serve as supplementary information to random forest variable importance. According to the findings of this study above, 1) the hypothesis cannot be rejected on a positive association between preterm birth and GERD (hypothesis 1) while 2) the hypothesis can be rejected on a positive association between preterm birth and periodontitis (hypothesis 2).

DISCUSSION

In summary, based on random forest variable importance, main predictors of preterm birth during 2015–2017 were socioeconomic status in 2014, age in 2014, GERD for the years 2012, 2014, 2010, 2013, 2007 and 2009, region (city) in 2014 and GERD for the year 2006. The importance rankings of periodontitis were relatively low.

According to the findings above, this study draws the following clinical implications. Firstly, the results of this study promote active counseling on GERD for effective prenatal

Table 2. Percentages of GERD and periodontitis during 2002–2014

Determinants	PTB 1 ^a	PTB 2	PTB 3	PTB 4	Normal
GERD_2002	0.79	0.80	0.78	0.79	0.62
GERD_2003	1.20	1.30	1.23	1.22	0.77
GERD_2004	1.49	1.59	1.54	1.55	1.02
GERD_2005	1.98	3.42	2.04	2.07	1.30
GERD_2006	2.49	2.55	2.50	2.48	1.78
GERD_2007	3.61	3.56	3.63	3.64	2.61
GERD_2008	4.45	4.35	4.45	4.43	3.22
GERD_2009	6.07	5.73	6.00	6.00	4.39
GERD_2010	6.94	6.64	6.90	6.91	5.21
GERD_2011	8.24	7.72	8.17	8.18	5.93
GERD_2012	8.88	8.27	8.73	8.71	6.86
GERD_2013	9.32	9.41	9.37	9.37	7.15
GERD_2014	10.54	10.33	10.50	10.47	7.90
Periodontitis_2002	0.07	0.07	0.07	0.07	0.04
Periodontitis_2003	0.07	0.08	0.08	0.08	0.05
Periodontitis_2004	0.07	0.09	0.08	0.08	0.09
Periodontitis_2005	0.15	0.07	0.13	0.12	0.09
Periodontitis_2006	0.18	0.29	0.20	0.20	0.11
Periodontitis_2007	0.10	0.08	0.09	0.09	0.06
Periodontitis_2008	0.07	0.07	0.08	0.08	0.06
Periodontitis_2009	0.11	0.10	0.10	0.10	0.07
Periodontitis_2010	0.09	0.07	0.09	0.08	0.07
Periodontitis_2011	0.07	0.09	0.08	0.08	0.06
Periodontitis_2012	0.06	0.06	0.06	0.06	0.06
Periodontitis_2013	0.08	0.08	0.09	0.09	0.05
Periodontitis_2014	0.09	0.09	0.08	0.08	0.05

PTB = preterm birth, GERD = gastroesophageal reflux disease.

^aPTB during 2015–2017.

care. Independent suggestions have been made on a positive association between GERD and periodontitis^{14,15} and on a positive linkage between periodontitis and preterm birth.⁶ Likewise, it would be natural to expect a positive relationship between GERD and preterm birth. This study supports this expectation given the high importance rankings of GERD from the random forest in this study: GERD for the years 2012, 2014, 2010, 2013, 2007, 2009 and 2006 ranked within the top ten. GERD is reported to cause cytokine-triggered inflammation,¹⁶⁻¹⁸ which is expected to cause uterine contraction and preterm birth.¹⁹⁻²² As a matter of fact, GERD symptoms such as nausea and vomiting are quite general but usually ignored among pregnant women. Based on the relatively small data (731 participants), the previous study drew a clinical implication that the promotion of active counseling on GERD is vital for effective prenatal care.¹² This study confirms this clinical implication, using the population data (405,586 women). Secondly, the findings of this study highlight the importance of socioeconomic status as a main predictor of preterm birth. A few studies indicate that low socioeconomic status is a risk factor for preterm birth: education below high school graduation for 731 Korean women in a single-center retrospective study (odds ratio, 1.660),¹² income below \$1,000 for 2,645 Korean women in a multi-center prospective

Table 3. Model performance

Models	Accuracy				AUC			
	PTB 1 ^a	PTB 2	PTB 3	PTB 4	PTB 1	PTB 2	PTB 3	PTB 4
Logistic regression	0.9456	0.9779	0.9307	0.9280	0.5300	0.5700	0.5400	0.5400
Artificial neural network	0.9456	0.9779	0.9307	0.9280	0.5000	0.5000	0.5000	0.5000
Random forest	0.9456	0.9779	0.9307	0.9280	0.5100	0.5000	0.5100	0.5100

PTB = preterm birth, AUC = area under the receiver-operating-characteristic curve.

^aPTB during 2015–2017.

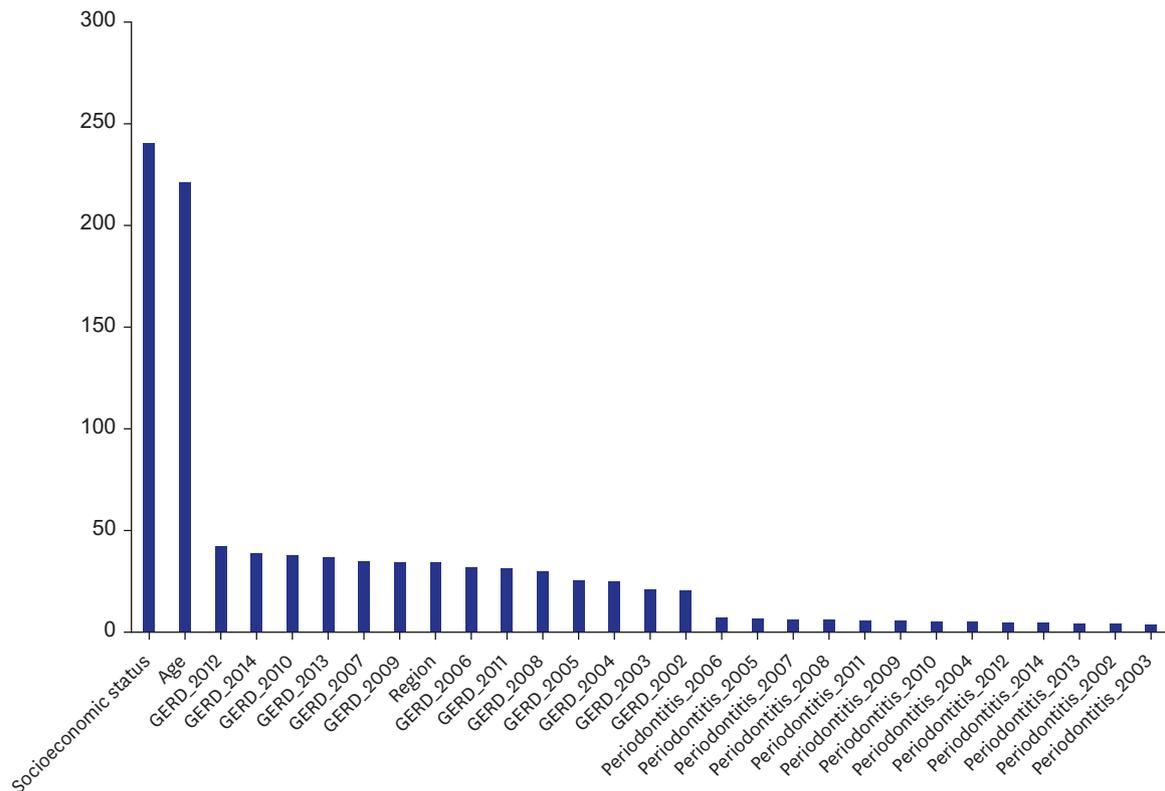


Fig. 1. Variable importance from the random forest: average for PTB 1, PTB 2, PTB 3 and PTB 4. PTB = preterm birth, GERD = gastroesophageal reflux disease.

study (odds ratio, 5.1)²³ and employment in heavy work for 7,634 Italian women in a multi-center retrospective study (odds ratio, 1.9).²⁴ Socioeconomic status was measured by an insurance fee with the range of 1 (the highest group) to 20 (the lowest group) in this study and lower socioeconomic status was a risk factor for preterm birth (odds ratio, 1.0032: The odds of PTB 4 will go up by 3.2% if socioeconomic status go down by 10 units, e.g., from 1 [the highest group] to 10 [the middle group]). This result is consistent with the existing literature and, it would be needed for preventing preterm birth to promote socioeconomic status for pregnant women.

To our best knowledge, this study is the first attempt to use machine learning and population data for finding main predictors of preterm birth and evaluating its association with GERD and periodontitis. The terms “artificial intelligence” and “machine learning” have garnered great attention on a global scale in the past decades and a vast improvement of computing power has been a main factor for the unprecedented performance and popularity of these cutting-edge approaches. This study will be a good starting point in this direction to find main predictors of preterm birth and draw effective implications for its prevention and management. However, this study still had some limitations. Firstly, this study did not examine possible mediating effects among variables. Secondly, this study adopted the binary category of preterm birth as no vs. yes. But preterm birth can have multiple categories and it will be a good topic for future study to compare different predictors for various categories of preterm birth. Thirdly, it was not the scope of this study to explore and evaluate various pathways between GERD and preterm birth. Little literature is available and more effort

Table 4. Coefficients of predictors from logistic regression for each type of PTB

Determinants	PTB 1 ^a	PTB 2	PTB 3	PTB 4
GERD_2002	0.9544	1.0025	0.9592	0.9708
GERD_2003	1.0716	1.2522*	1.1327*	1.1078
GERD_2004	1.0650	1.0778	1.0824	1.0915
GERD_2005	1.0950	1.0906	1.1172*	1.1459*
GERD_2006	1.0723	1.1095	1.0648	1.0511
GERD_2007	1.0202	1.0735	1.0343	1.0413
GERD_2008	1.0462	1.0971	1.0730*	1.0609
GERD_2009	1.0768*	1.0381	1.0766*	1.0792*
GERD_2010	1.0424	0.9652	1.0349	1.0400
GERD_2011	1.0862*	1.0414	1.0834*	1.0800*
GERD_2012	1.0420	0.9931	1.0244	1.0230
GERD_2013	1.0099	1.0615	1.0278	1.0317
GERD_2014	1.0368	1.0526	1.0393	1.0360
Periodontitis_2002	1.3104	1.0205	1.2382	1.1803
Periodontitis_2003	0.9851	1.1937	1.1632	1.1980
Periodontitis_2004	0.6649	0.7627	0.7044	0.7245
Periodontitis_2005	1.4547*	0.2311*	1.1135	1.0654
Periodontitis_2006	1.2991	2.1960*	1.5374*	1.4755*
Periodontitis_2007	1.2063	0.5800	1.1270	1.0833
Periodontitis_2008	0.8798	1.0610	1.0316	1.0571
Periodontitis_2009	1.5617*	1.0737	1.2373	1.2520
Periodontitis_2010	0.8903	1.0944	0.9954	0.9444
Periodontitis_2011	0.7997	1.0785	0.8899	0.9991
Periodontitis_2012	0.7696	0.9035	0.8063	0.8393
Periodontitis_2013	1.1700	1.1032	1.2903	1.3994
Periodontitis_2014	1.4784	1.6822	1.4200	1.3511
Age in 2014	1.0174*	1.0664*	1.0279*	1.0305*
Socioeconomic status in 2014	1.0033*	1.0021	1.0031*	1.0032*
Region in 2014	1.0803*	0.9004*	1.0088	1.0244

PTB = preterm birth, GERD = gastroesophageal reflux disease.

^aPTB during 2015–2017.

* $P < 0.10$.

is needed in this direction. Fourthly, GERD was screened from ICD-10 codes in this study. Additional validation based on endoscopic finding and motility testing are expected to improve external validity for this line of research. Fifthly, this study excluded education, employment, upper gastrointestinal tract symptoms, *Helicobacter pylori*, pelvic inflammatory disease, various types of BMI, diabetes and hypertension, parity and medication such as progesterone, calcium channel blocker. Sixthly, the area under the receiver-operating-characteristic curves of this study belonged to the range of 0.51–0.57 and they were not high enough to serve as diagnostic tests. Further improvement is needed and one solution would be to extend this study to consider various information on health status and medication history. Finally, further investigations of single vs. multiple gestation, first vs. recurrent preterm birth, would deliver more insights and more detailed clinical implications.^{25,26}

In conclusion, preterm birth has a stronger association with GERD than with periodontitis. For the prevention of preterm birth, preventive measures for GERD would be essential together with the improvement of socioeconomic status for pregnant women. Especially, it would be vital to promote active counseling for general GERD symptoms (neglected by pregnant women).

Table 5. P values: coefficients of predictors from logistic regression for each type of PTB

Determinants	PTB 1 ^a	PTB 2	PTB 3	PTB 4
GERD_2002	0.630	0.986	0.626	0.723
GERD_2003	0.395	0.046*	0.077*	0.142
GERD_2004	0.377	0.474	0.207	0.154
GERD_2005	0.144	0.347	0.043*	0.011*
GERD_2006	0.207	0.202	0.201	0.303
GERD_2007	0.670	0.305	0.417	0.320
GERD_2008	0.290	0.143	0.061*	0.111
GERD_2009	0.044*	0.506	0.024*	0.017*
GERD_2010	0.227	0.510	0.264	0.193
GERD_2011	0.010*	0.415	0.005*	0.006*
GERD_2012	0.181	0.885	0.381	0.402
GERD_2013	0.746	0.193	0.310	0.239
GERD_2014	0.206	0.238	0.130	0.157
Periodontitis_2002	0.436	0.972	0.500	0.601
Periodontitis_2003	0.963	0.697	0.577	0.492
Periodontitis_2004	0.187	0.550	0.190	0.214
Periodontitis_2005	0.075*	0.040*	0.604	0.760
Periodontitis_2006	0.187	0.001*	0.010*	0.020*
Periodontitis_2007	0.500	0.351	0.638	0.753
Periodontitis_2008	0.681	0.897	0.905	0.827
Periodontitis_2009	0.058*	0.864	0.355	0.319
Periodontitis_2010	0.697	0.829	0.985	0.822
Periodontitis_2011	0.515	0.869	0.687	0.997
Periodontitis_2012	0.472	0.842	0.492	0.561
Periodontitis_2013	0.602	0.830	0.320	0.169
Periodontitis_2014	0.178	0.214	0.184	0.255
Age in 2014	0.000**	0.000**	0.000**	0.000**
Socioeconomic status in 2014	0.069*	0.449	0.051*	0.040*
Region in 2014	0.022*	0.029*	0.764	0.403

PTB = preterm birth, GERD = gastroesophageal reflux disease.

^aPTB during 2015–2017.

* $P < 0.100$; ** $P < 0.001$.

SUPPLEMENTARY MATERIALS

Supplementary Table 1

ICD-10 code for PTB, GERD and periodontitis

[Click here to view](#)

Supplementary Fig. 1

Variable importance from the random forest: separate for PTB 1, PTB 2, PTB 3 and PTB 4.

[Click here to view](#)

REFERENCES

1. Magro Malosso ER, Saccone G, Simonetti B, Squillante M, Berghella V. US trends in abortion and preterm birth. *J Matern Fetal Neonatal Med* 2018;31(18):2463-7.
[PUBMED](#) | [CROSSREF](#)
2. Lee NH. International trends and implications for preterm birth. *Health Welf Policy Forum* 2013;(200):116-27.
3. World Health Organization. News: Preterm birth. <http://www.who.int/news-room/fact-sheets/detail/preterm-birth>. Updated 2018. Accessed June 1, 2019.

4. Harrison MS, Goldenberg RL. Global burden of prematurity. *Semin Fetal Neonatal Med* 2016;21(2):74-9.
[PUBMED](#) | [CROSSREF](#)
5. Eke PI, Dye BA, Wei L, Slade GD, Thornton-Evans GO, Borgnakke WS, et al. Update on prevalence of periodontitis in adults in the United States: NHANES 2009 to 2012. *J Periodontol* 2015;86(5):611-22.
[PUBMED](#) | [CROSSREF](#)
6. Puertas A, Magan-Fernandez A, Blanc V, Revelles L, O'Valle F, Pozo E, et al. Association of periodontitis with preterm birth and low birth weight: a comprehensive review. *J Matern Fetal Neonatal Med* 2018;31(5):597-602.
[PUBMED](#) | [CROSSREF](#)
7. Vakil N, van Zanten SV, Kahrilas P, Dent J, Jones RGlobal Consensus Group. The Montreal definition and classification of gastroesophageal reflux disease: a global evidence-based consensus. *Am J Gastroenterol* 2006;101(8):1900-20.
[PUBMED](#) | [CROSSREF](#)
8. Patrick L. Gastroesophageal reflux disease (GERD): a review of conventional and alternative treatments. *Altern Med Rev* 2011;16(2):116-33.
[PUBMED](#)
9. Vinesh E, Masthan K, Kumar MS, Jeyapriya SM, Babu A, Thinakaran M. A clinicopathologic study of oral changes in gastroesophageal reflux disease, gastritis, and ulcerative colitis. *J Contemp Dent Pract* 2016;17(11):943-7.
[PUBMED](#) | [CROSSREF](#)
10. Deppe H, Mücke T, Wagenpfeil S, Kesting M, Rozej A, Bajbouj M, et al. Erosive esophageal reflux vs. non erosive esophageal reflux: oral findings in 71 patients. *BMC Oral Health* 2015;15:84.
[PUBMED](#) | [CROSSREF](#)
11. Ali RA, Egan LJ. Gastroesophageal reflux disease in pregnancy. *Best Pract Res Clin Gastroenterol* 2007;21(5):793-806.
[PUBMED](#) | [CROSSREF](#)
12. Lee KS, Song IS, Kim ES, Ahn KH. Determinants of spontaneous preterm labor and birth including gastroesophageal reflux disease and periodontitis. *J Korean Med Sci* 2020;35(14):e105.
[PUBMED](#) | [CROSSREF](#)
13. Han J, Micheline K. *Data Mining: Concepts and Techniques*. 2nd ed. San Francisco, CA, USA: Elsevier; 2006.
14. Song JY, Kim HH, Cho EJ, Kim TY. The relationship between gastroesophageal reflux disease and chronic periodontitis. *Gut Liver* 2014;8(1):35-40.
[PUBMED](#) | [CROSSREF](#)
15. Adachi K, Mishiro T, Tanaka S, Yoshikawa H, Kinoshita Y. A study on the relationship between reflux esophagitis and periodontitis. *Intern Med* 2016;55(18):2523-8.
[PUBMED](#) | [CROSSREF](#)
16. Souza RF, Huo X, Mittal V, Schuler CM, Carmack SW, Zhang HY, et al. Gastroesophageal reflux might cause esophagitis through a cytokine-mediated mechanism rather than caustic acid injury. *Gastroenterology* 2009;137(5):1776-84.
[PUBMED](#) | [CROSSREF](#)
17. Kahrilas PJ. Turning the pathogenesis of acute peptic esophagitis inside out. *JAMA* 2016;315(19):2077-8.
[PUBMED](#) | [CROSSREF](#)
18. Dunbar KB, Agoston AT, Odze RD, Huo X, Pham TH, Cipher DJ, et al. Association of acute gastroesophageal reflux disease with esophageal histologic changes. *JAMA* 2016;315(19):2104-12.
[PUBMED](#) | [CROSSREF](#)
19. Hagberg H, Mallard C, Jacobsson B. Role of cytokines in preterm labour and brain injury. *BJOG* 2005;112 Suppl 1:16-8.
[PUBMED](#) | [CROSSREF](#)
20. Cappelletti M, Della Bella S, Ferrazzi E, Mavilio D, Divanovic S. Inflammation and preterm birth. *J Leukoc Biol* 2016;99(1):67-78.
[PUBMED](#) | [CROSSREF](#)
21. Sivarajasingam SP, Imami N, Johnson MR. Myometrial cytokines and their role in the onset of labour. *J Endocrinol* 2016;231(3):R101-19.
[PUBMED](#) | [CROSSREF](#)
22. Pandey M, Chauhan M, Awasthi S. Interplay of cytokines in preterm birth. *Indian J Med Res* 2017;146(3):316-27.
[PUBMED](#)
23. Kim YJ, Lee BE, Park HS, Kang JG, Kim JO, Ha EH. Risk factors for preterm birth in Korea: a multicenter prospective study. *Gynecol Obstet Invest* 2005;60(4):206-12.
[PUBMED](#) | [CROSSREF](#)

24. Di Renzo GC, Giardina I, Rosati A, Clerici G, Torricelli M, Petraglia F, et al. Maternal risk factors for preterm birth: a country-based population analysis. *Eur J Obstet Gynecol Reprod Biol* 2011;159(2):342-6.
[PUBMED](#) | [CROSSREF](#)
25. Ouh YT, Park JH, Ahn KH, Hong SC, Oh MJ, Kim HJ, et al. Recurrent risk of preterm birth in the third pregnancy in Korea. *J Korean Med Sci* 2018;33(24):e170.
[PUBMED](#) | [CROSSREF](#)
26. Kim S, Park HS, Kwon H, Seol HJ, Bae JG, Ahn KH, et al. Effect of cervical cerclage on the risk of recurrent preterm birth after a twin spontaneous preterm birth. *J Korean Med Sci* 2020;35(11):e66.
[PUBMED](#) | [CROSSREF](#)