




OPEN

## Risk scores for predicting small for gestational age infants in Japan: The TMM birthree cohort study

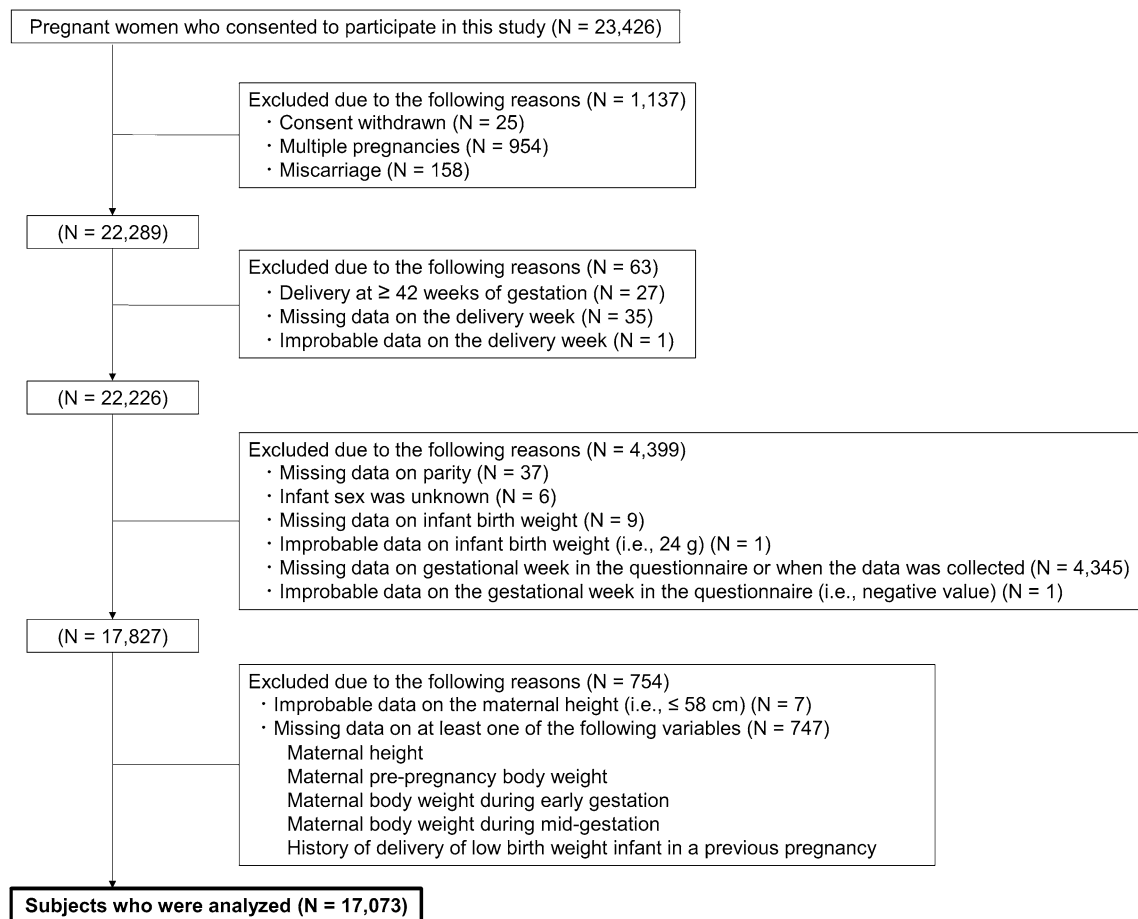
Noriyuki Iwama<sup>1,2</sup>, Taku Obara<sup>2,3,4</sup>, Mami Ishikuro<sup>2,3</sup>, Keiko Murakami<sup>2,3</sup>, Fumihiko Ueno<sup>2,3</sup>, Aoi Noda<sup>2,3,4</sup>, Tomomi Onuma<sup>2,3</sup>, Fumiko Matsuzaki<sup>2</sup>, Tetsuro Hoshiai<sup>1</sup>, Masatoshi Saito<sup>1,5</sup>, Hirohito Metoki<sup>6,7</sup>, Junichi Sugawara<sup>1,6,8</sup>, Nobuo Yaegashi<sup>1,6,8,9</sup> & Shinichi Kuriyama<sup>2,3,10</sup>

This study aimed to construct a prediction model for small-for-gestational-age (SGA) infants in Japan by creating a risk score during pregnancy. A total of 17,073 subjects were included in the Tohoku Medical Megabank Project Birth and Three-Generation Cohort Study, a prospective cohort study. A multiple logistic regression model was used to construct risk scores during early and mid-gestational periods (11–17 and 18–21 weeks of gestation, respectively). The risk score during early gestation comprised the maternal age, height, body mass index (BMI) during early gestation, parity, assisted reproductive technology (ART) with frozen-thawed embryo transfer (FET), smoking status, blood pressure (BP) during early gestation, and maternal birth weight. The risk score during mid-gestation also consisted of the maternal age, height, BMI during mid-gestation, weight gain, parity, ART with FET, smoking status, BP level during mid-gestation, maternal birth weight, and estimated fetal weight during mid-gestation. The C-statistics of the risk scores during early- and mid-gestation were 0.658 (95% confidence interval [CI]: 0.642–0.675) and 0.725 (95% CI: 0.710–0.740), respectively. In conclusion, the predictive ability of the risk scores during mid-gestation for SGA infants was acceptable and better than that of the risk score during early gestation.

Small for gestational age (SGA) infants, a common surrogate of fetal growth restriction (FGR), is a risk factor for adverse perinatal events<sup>1–4</sup>. Infants born SGA or with FGR are at risk for stillbirth, non-reassuring fetal status, perinatal asphyxia, neonatal death, and neurological developmental delay<sup>1–4</sup>. Infants who are diagnosed as SGA after birth have a higher risk of developing severe adverse complications than those who are diagnosed as SGA before birth<sup>5,6</sup>. By judgement of the delivery management method following fetal surveillance using ultrasonography and cardiotocography, the risk of fetal and/or adverse neonatal complications, including severe fetal distress and fetal/infant death, may decrease<sup>5,6</sup>. Therefore, detection of pregnant women at a high risk of delivering SGA infants is important for careful fetal surveillance<sup>7</sup>. Healthcare providers evaluate infant birth weight as the estimated fetal weight (EFW) during the fetal period using ultrasonography (a simple and accepted method) at the prenatal checkup; however, not all SGA infants are detected due to overestimation of the EFW<sup>8,9</sup>. Therefore, using a prediction model for the early identification of pregnant women at a high risk of delivering SGA infants, and subsequent careful prenatal checkups and fetal surveillance before delivery, may lead to preferable outcomes.

Thus far, previous studies in several countries have reported prediction models based on the clinical risk factors, ultrasound findings, and biomarkers of SGA infants<sup>10–14</sup>. Furthermore, several studies in Japan have also reported certain factors associated with infant birth weight; however, racial differences exist in the maternal and neonatal characteristics, including the maternal physique and infant birth weight<sup>15–22</sup>. Pre-pregnancy body

<sup>1</sup>Department of Obstetrics and Gynecology, Tohoku University Hospital, 1-1, Seiryomachi, Sendai, Miyagi 980-8574, Japan. <sup>2</sup>Division of Molecular Epidemiology, Department of Preventive Medicine and Epidemiology, Tohoku Medical Megabank Organization, Tohoku University, Sendai, Japan. <sup>3</sup>Division of Molecular Epidemiology, Graduate School of Medicine, Tohoku University, Sendai, Japan. <sup>4</sup>Department of Pharmaceutical Sciences, Tohoku University Hospital, Sendai, Miyagi, Japan. <sup>5</sup>Department of Maternal and Fetal Therapeutics, Tohoku University Graduate School of Medicine, Sendai, Miyagi, Japan. <sup>6</sup>Tohoku Medical Megabank Organization, Tohoku University, Sendai, Miyagi, Japan. <sup>7</sup>Division of Public Health, Hygiene and Epidemiology, Tohoku Medical Pharmaceutical University, Sendai, Miyagi, Japan. <sup>8</sup>Environment and Genome Research Center, Tohoku University Graduate School of Medicine, Sendai, Miyagi, Japan. <sup>9</sup>Department of Obstetrics and Gynecology, Tohoku University Graduate School of Medicine, Sendai, Miyagi, Japan. <sup>10</sup>International Research Institute of Disaster Science, Tohoku University, Sendai, Miyagi, Japan. ✉email: noriyuki.iwama@med.tohoku.ac.jp



**Figure 1.** Flow chart of the study participants.

mass index (BMI) and recommendations for weight gain during pregnancy are different across countries<sup>20,21</sup>. A possible reason for the difference in mean infant birth weight between Japanese and White people is differences in maternal physique and weight gain during pregnancy<sup>23</sup>.

Due to lack of previous studies on pregnant Japanese women reporting a prediction model for SGA infants, this study aimed to construct a prediction model for SGA infants in Japan by creating a risk score during pregnancy.

## Results

**Maternal and neonatal characteristics of the study subjects.** Figure 1 presents a flow chart of the study subjects. Of the 23,425 pregnant women who consented to participate in the Tohoku Medical Megabank Project (TMM) Birth and Three-Generation Cohort Study (TMM BirThree Cohort Study), participants were excluded due to the following reasons: withdrawn consent (N = 25); multiple pregnancies (N = 954); miscarriages (N = 158); delivery week at  $\geq 42$  weeks of gestation (N = 27); missing data (N = 35) and improbable data (N = 1) on the delivery week; missing data on parity (N = 37); unknown infant sex (N = 6); missing data on infant birth weight (N = 9); improbable data on infant birth weight (i.e., 24 g) (N = 1); missing data on the gestational week in the questionnaire or when the data was collected (N = 4,345); and improbable data on the gestational week in the questionnaire (N = 1). Furthermore, seven women with improbable data on the maternal height were excluded followed by 747 women who were excluded due to missing data on at least one of the following variables: maternal height, pre-pregnancy body weight (BW), BW during early gestation, BW during mid-gestation, and history of delivery of low birth weight (LBW) infants in a previous pregnancy. Finally, 17,073 women remained and were subsequently analyzed. Women with “No answer” on questions regarding the maternal birth weight and smoking status were not excluded considering the clinical use of risk scores to predict SGA infants.

Table 1 shows the maternal and neonatal characteristics of the study subjects. The number of SGA infants was 1,126 (6.6%).

**Risk scores during early and mid-gestation for predicting SGA infants.** Supplementary Table S1 and the supplementary information present the results of the univariate logistic regression model. Supplementary Tables S2, S3, and S4 show the adjusted odds ratios, regression coefficients, and integer scores during the early- and mid-gestation periods based on the multiple logistic regression model. The selected explanatory vari-

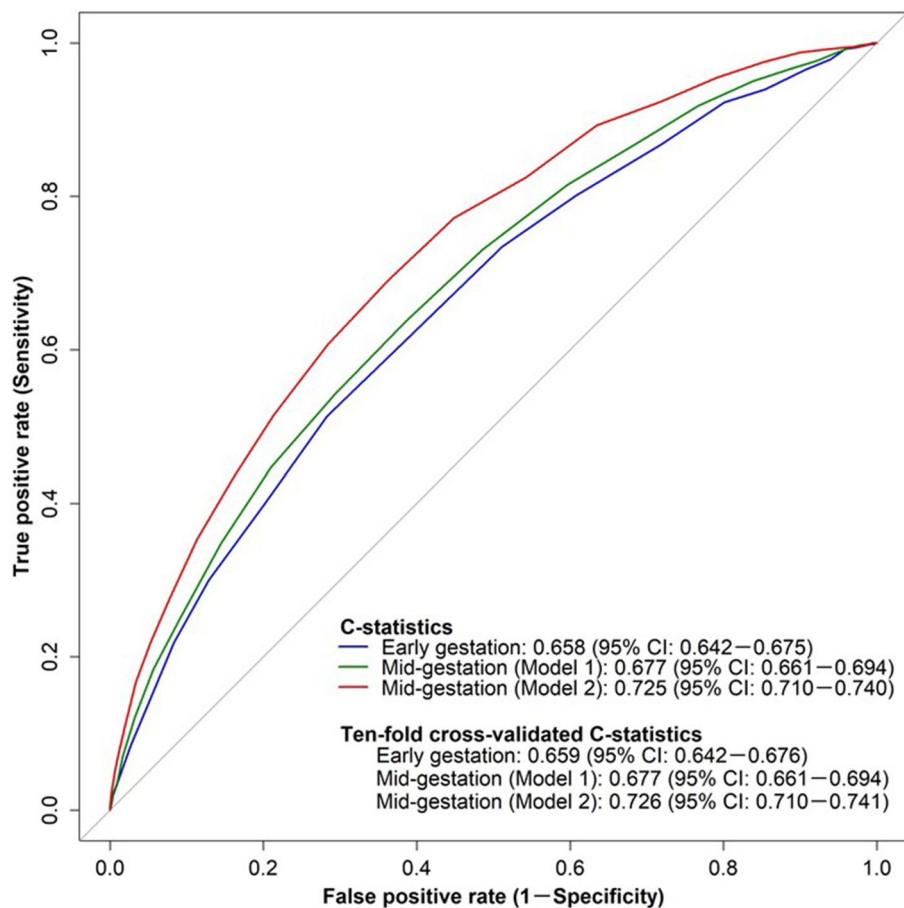
Variables	Values of the study subjects (N = 17,073)	The numbers and proportion of SGA infants in each category, cases/number (%)
<b>Maternal characteristics</b>		
<b>Gestational week at prenatal checkup during early gestation, median (IQR)</b>	12.4 (11.7–13.1)	–
<b>Gestational week at prenatal checkup during mid-gestation, median (IQR)</b>	20.1 (19.4–20.6)	–
<b>Age during early gestation, years</b>	31.8 ± 5.0	–
< 25 years, n (%)	1,496 (8.8)	92/1,496 (6.2)
25–29.9 years, n (%)	4,685 (27.4)	308/4,685 (6.6)
30–34.9 years, n (%)	6,271 (36.7)	403/6,271 (6.4)
≥ 35 years, n (%)	4,621 (27.1)	323/4,621 (7.0)
<b>Age during mid-gestation, years</b>	31.9 ± 5.0	–
< 25 years, n (%)	1,412 (8.3)	86/1,412 (6.1)
25–29.9 years, n (%)	4,609 (27.0)	302/4,609 (6.6)
30–34.9 years, n (%)	6,275 (36.8)	404/6,275 (6.4)
≥ 35 years, n (%)	4,777 (28.0)	334/4,777 (7.0)
<b>Height, cm</b>	158.3 ± 5.4	–
Quartile 1 (< 155 cm), n (%)	4,122 (24.1)	374/4,122 (9.1)
Quartile 2 (155–158.9 cm), n (%)	4,986 (29.2)	352/4,986 (7.1)
Quartile 3 (159–161.9 cm), n (%)	3,345 (19.6)	205/3,345 (6.1)
Quartile 4 (≥ 162 cm), n (%)	4,620 (27.1)	195/4,620 (4.2)
<b>Pre-pregnancy BW, kg</b>	54.3 ± 9.5	–
<b>Initial BW during early gestation, kg</b>	54.7 ± 9.4	–
<b>Initial BW during mid-gestation, kg</b>	57.2 ± 9.2	–
<b>Pre-pregnancy BMI, kg/m<sup>2</sup></b>	21.6 ± 3.5	–
Quartile 1 (< 19.3 kg/m <sup>2</sup> ), n (%)	4,237 (24.8)	379/4,237 (9.0)
Quartile 2 (≥ 19.3 and < 20.8 kg/m <sup>2</sup> ), n (%)	4,063 (23.8)	268/4,063 (6.6)
Quartile 3 (≥ 20.8 and < 23.0 kg/m <sup>2</sup> ), n (%)	4,462 (26.1)	276/4,462 (6.2)
Quartile 4 (≥ 23.0 kg/m <sup>2</sup> ), n (%)	4,311 (25.3)	203/4,311 (4.7)
<b>Initial BMI during early gestation, kg/m<sup>2</sup></b>	21.8 ± 3.5	–
Quartile 1 (< 19.5 kg/m <sup>2</sup> ), n (%)	4,263 (25.0)	381/4,263 (8.9)
Quartile 2 (≥ 19.5 and < 21.1 kg/m <sup>2</sup> ), n (%)	4,217 (24.7)	292/4,217 (6.9)
Quartile 3 (≥ 21.1 and < 23.2 kg/m <sup>2</sup> ), n (%)	4,225 (24.8)	242/4,225 (5.7)
Quartile 4 (≥ 23.2 kg/m <sup>2</sup> ), n (%)	4,368 (25.6)	211/4,368 (4.8)
<b>Initial BMI during mid-gestation, kg/m<sup>2</sup></b>	22.8 ± 3.4	–
Quartile 1 (< 20.5 kg/m <sup>2</sup> ), n (%)	4,101 (24.0)	387/4,101 (9.4)
Quartile 2 (≥ 20.5 and < 22.1 kg/m <sup>2</sup> ), n (%)	4,263 (25.0)	303/4,263 (7.1)
Quartile 3 (≥ 22.1 and < 24.2 kg/m <sup>2</sup> ), n (%)	4,377 (25.6)	246/4,377 (5.6)
Quartile 4 (≥ 24.2 kg/m <sup>2</sup> ), n (%)	4,332 (25.4)	190/4,332 (4.4)
<b>Weight gain between pre-pregnancy and early gestation (Initial BW during early gestation – pre-pregnancy BW), kg</b>	0.5 ± 2.5	–
Quartile 1 (< -0.6 kg), n (%)	4,370 (25.6)	288/4,370 (6.6)
Quartile 2 (≥ -0.6 and < 0.6 kg), n (%)	4,257 (24.9)	289/4,257 (6.8)
Quartile 3 (≥ 0.6 and < 1.7 kg), n (%)	4,086 (23.9)	268/4,086 (6.6)
Quartile 4 (≥ 1.7 kg), n (%)	4,360 (25.5)	281/4,360 (6.4)
<b>Weight gain between early and mid-gestation (Initial BW during mid-gestation – Initial BW during early gestation), kg</b>	2.4 ± 1.8	–
Quartile 1 (< 1.5 kg), n (%)	4,098 (24.0)	361/4,098 (8.8)
Quartile 2 (≥ 1.5 and < 2.4 kg), n (%)	4,527 (26.5)	346/4,527 (7.6)
Quartile 3 (≥ 2.4 and < 3.4 kg), n (%)	4,333 (25.4)	240/4,333 (5.5)
Quartile 4 (≥ 3.4 kg), n (%)	4,115 (24.1)	179/4,115 (4.4)
<b>Parity, n (%)</b>		
Primipara	8,073 (47.3)	529/8,073 (6.6)
Multipara without HDP or delivery of LBW infants in a previous pregnancy	7,676 (45.0)	427/7,676 (5.6)
Multipara with HDP and/or delivery of LBW infants in a previous pregnancy	1,324 (7.8)	170/1,324 (12.8)
<b>Conception method, n (%)</b>		
Continued		

Variables	Values of the study subjects (N = 17,073)	The numbers and proportion of SGA infants in each category, cases/number (%)
Natural pregnancy	15,883 (93.0)	1,059/15,883 (6.7)
Non-ART (ovulation induction or AIH)	388 (2.3)	27/388 (7.0)
ART (conventional IVF or ICSI) with fresh ET	80 (0.5)	5/80 (6.3)
ART (conventional IVF or ICSI) with FET	489 (2.9)	18/489 (3.7)
ART (conventional IVF or ICSI) without information on the method of ET	203 (1.2)	14/203 (6.9)
Others	30 (0.2)	3/30 (10.0)
<b>Maternal birth weight, n (%)</b>		
< 2,500 g	783 (4.6)	91/783 (11.6)
2,500–2,999 g	3,160 (18.5)	282/3,160 (8.9)
3,000–3,499 g	4,925 (28.9)	245/4,925 (5.0)
≥ 3,500 g	1,298 (7.6)	31/1,298 (2.4)
Unknown or No answer	6,907 (40.5)	477/6,907 (6.9)
<b>Medical history, n (%)</b>		
Diabetes mellitus		
No	17,010 (99.6)	1,122/17,010 (6.6)
Yes	63 (0.4)	4/63 (6.4)
SLE and/or APS		
No	17,042 (99.8)	1,122/17,042 (6.6)
Yes	31 (0.2)	4/31 (12.9)
Chronic kidney disease		
No	17,022 (99.7)	1,123/17,022 (6.6)
Yes	51 (0.3)	3/51 (5.9)
Hyperthyroidism		
No	16,683 (98.9)	1,115/16,683 (6.6)
Yes	190 (1.1)	11/190 (5.8)
Hypothyroidism		
No	16,836 (98.6)	1,108/16,836 (6.6)
Yes	237 (1.4)	18/237 (7.6)
<b>Smoking status, n (%)</b>		
Never	10,202 (59.8)	683/10,202 (6.7)
Quit smoking before conception	4,003 (23.5)	248/4,003 (6.2)
Quit smoking after conception	2,388 (14.0)	137/2,388 (5.7)
Continue smoking during pregnancy	412 (2.4)	49/412 (11.9)
No answer	68 (0.4)	9/68 (13.2)
<b>Alcohol drinking, n (%)</b>		
Constitutionally never drinker	990 (5.8)	69/990 (7.0)
Almost never drinking or Quit drinking	12,738 (74.6)	836/12,738 (6.6)
Continue drinking during pregnancy	3,290 (19.3)	218/3,290 (6.6)
No answer	55 (0.3)	3/55 (5.5)
<b>Initial clinic blood pressure level during early gestation</b>		
Normal blood pressure (SBP is < 120 mmHg and DBP is < 80 mmHg)	12,726 (74.5)	854/12,726 (6.7)
High normal blood pressure (SBP is 120–129 mmHg and DBP is < 80 mmHg)	2,520 (14.8)	145/2,520 (5.8)
Elevated blood pressure (SBP is 130–139 mmHg and/or DBP is 80–89 mmHg)	1,423 (8.3)	89/1,423 (6.3)
Grade 1 or higher hypertension (SBP is ≥ 140 mmHg and/or DBP is ≥ 90 mmHg)	404 (2.4)	38/404 (9.4)
<b>Initial clinic blood pressure level during mid-gestation</b>		
Normal blood pressure (SBP is < 120 mmHg and DBP is < 80 mmHg)	13,465 (78.9)	892/13,465 (6.6)
High normal blood pressure (SBP is 120–129 mmHg and DBP is < 80 mmHg)	2,436 (14.3)	147/2,436 (6.0)
Elevated blood pressure (SBP is 130–139 mmHg and/or DBP is 80–89 mmHg)	926 (5.4)	61/926 (6.6)
Grade 1 or higher hypertension (SBP is ≥ 140 mmHg and/or DBP is ≥ 90 mmHg)	246 (1.4)	26/246 (10.6)
<b>Initial SD value of the EFW during mid-gestation, SD</b>	0.3 ± 0.9	–
Continued		

Variables	Values of the study subjects (N = 17,073)	The numbers and proportion of SGA infants in each category, cases/number (%)
≥ +1.5 SD	1,441 (8.4)	35/1,441 (2.4)
≥ +0.5 SD and < +1.5 SD	5,414 (31.7)	208/5,414 (3.8)
> -0.5 SD and < +0.5 SD	7,821 (45.8)	524/7,821 (6.7)
> -1.5 SD and ≤ -0.5 SD	2,228 (13.1)	320/2,228 (14.4)
≤ -1.5 SD	169 (1.0)	39/169 (23.1)
<b>Obstetric complications, n (%)</b>		
HDP	846 (5.0)	124/846 (14.7)
Gestational diabetes mellitus	438 (2.6)	28/438 (6.4)
Placental abruption	34 (0.2)	5/34 (14.7)
Placenta previa	102 (0.6)	3/102 (2.9)
Low-lying placenta	67 (0.4)	1/67 (1.5)
Intrauterine fetal death	8 (0.1)	4/8 (50.0)
Stillbirth	12 (0.1)	6/12 (50.0)
<b>Neonatal characteristics</b>		
Infant sex (male/female), n (%) / n (%)	8,837 (51.8) / 8,236 (48.2)	583/8,837 (6.6) / 543/8,236 (6.6)
Major congenital anomalies	328 (1.9)	35/328 (10.7)
<b>Chromosomal abnormality</b>		
Trisomy 21	19 (0.1)	2/19 (10.5)
Trisomy 18	5 (0.03)	4/5 (80.0)
Trisomy 13	2 (0.01)	1/2 (50.0)
<b>Skeletal dysplasia</b>		
Thanatophoric dysplasia	3 (0.01)	0/3 (0.0)
Achondrogenesis	1 (0.01)	0/1 (0.0)
Achondroplasia	1 (0.01)	0/1 (0.0)
Osteogenesis imperfecta	0 (0.0)	–
<b>Other major congenital anomalies</b>		
302 (1.8)		29/302 (9.6)
<b>Gestational age at delivery, weeks</b>		
39.2 ± 1.7		38.7 ± 2.7
Preterm birth (Delivery at less than 37 weeks of gestation)	880 (5.2)	125/880 (14.2)
Preterm birth (Delivery at less than 34 weeks of gestation)	196 (1.2)	55/196 (28.1)
Preterm birth (Delivery at less than 32 weeks of gestation)	131 (0.8)	36/131 (27.5)
<b>Infant birth weight</b>		
Grams	3,031 ± 430	2,349 ± 431 in SGA infants
SD value	0.2 ± 1.0	-1.8 ± 0.6 in SGA infants
LBW infants (birth weight < 2,500 g), n (%)	1,395 (8.2)	638/1,395 (45.7)
SGA infants (birth weight < 10 <sup>th</sup> percentile), n (%)	1,126 (6.6)	–
Preterm SGA infants (Birth weight < 10th percentile and preterm birth from 22 to < 37 weeks of gestation)	125 (0.7)	–
Term SGA infants (Birth weight < 10th percentile and birth from 37 to < 42 weeks of gestation)	1,001 (5.9)	–

**Table 1.** Maternal and neonatal characteristics of the study subjects. Continuous variables and categorical variables are expressed as mean ± SD and number (%), respectively. Abbreviations: AIH, artificial insemination by the husband; APS, antiphospholipid syndrome; ART, assisted reproductive technology; BMI, body mass index; BW, body weight; DBP, diastolic blood pressure; EFW, estimated fetal weight; ET, embryo transfer; FET, frozen-thawed ET; HDP, hypertensive disorders of pregnancy; ICSI, intracytoplasmic sperm injection; IQR, interquartile range; IVF, in vitro fertilization; LBW, low birth weight; SBP, systolic blood pressure; SD, standard deviation; SGA, small for gestational age; SLE, systemic lupus erythematosus.

ables in all risk scores were maternal height, parity, assisted reproductive technology (ART; conventional in vitro fertilization or intracytoplasmic sperm injection) with frozen-thawed embryo transfer (FET), continued smoking during pregnancy, and maternal birth weight. The maternal age during early gestation, maternal BMI during early pregnancy, and hypertension (grade 1 or higher) during early gestation were also selected while creating the risk scores during early gestation. When the risk score during mid-gestation was constructed, the maternal age during mid-gestation, maternal BMI during mid-gestation, weight gain between the initial body weight (BW) during early gestation and the BW during mid-gestation, and hypertension (grade 1 or higher) during mid-gestation were selected for model 1. For model 2 during mid-gestation, the initial standard deviation (SD) value of the EFW during mid-gestation was also selected in addition to the parameters selected for model 1.



**Figure 2.** Receiver operating characteristics curve of each risk score for predicting SGA infants. Abbreviations: CI, confidence interval; EFW, estimated fetal weight; SGA, small for gestational age.

**Model performance and calibration plot of the risk scores for predicting SGA infants.** The discrimination performance of each risk score is shown in Fig. 2. The risk score during early gestation showed a poor discrimination performance. Models 1 and 2 (i.e., risk scores during mid-gestation) showed poor and acceptable discrimination performances, respectively. The C-statistics and ten-fold cross-validated C-statistics of model 2 during mid-gestation were 0.725 (95% confidence interval [CI]: 0.710–0.740) and 0.726 (95% CI: 0.710–0.741), respectively. The results of the sensitivity analysis performed using multiple imputation by a chained equation are described in the supplementary information.

Table 2 shows the observed proportion of the SGA infants and the predicted probability of the SGA infants according to the quintiles of each risk score. As shown in Fig. 3, the possibility of miscalibration was low, because the calibration curve in each risk score was close to the diagonal line (i.e., the line of perfect calibration). Figure 4 shows the result of the decision curve analysis. The risk score during early gestation, model 1, and model 2 during mid-gestation had a higher net benefit (NB) than that of either all or no subjects considered to be at a high risk of delivering SGA infants when the threshold probabilities were 0.029–0.183 (risk score = –3 to 10); 0.023–0.195 (risk score = –5 to 9), and 0.019–0.312 (risk score = –6 to 11), respectively.

**Comparison of model performance between different risk scores for predicting SGA infants.** Supplementary Table S5 shows the results of a comparison of the model performance between different risk scores for the prediction of SGA infants. In terms of discrimination and reclassification, model 2 showed a better performance for predicting SGA infants during mid-gestation as compared to the risk score during early gestation or model 1 during mid-gestation. As shown in Fig. 4, the NB in model 2 during mid-gestation was higher than that in the risk score during early gestation and model 1 during mid-gestation.

**Discrimination and clinical utility based on different cut-off values of the risk scores for prediction of SGA infants.** Table 3 shows the discrimination performance and the NB based on the lowest risk score of quintile 5 as the cut-off value in each risk score. In model 2 during mid-gestation, the lowest risk score of quintile 5 as the cut-off value (i.e., risk score = 4) was equivalent to a threshold probability of 0.110 and had an NB of 0.011. Supplementary Table S6 also shows the discrimination performance and NB when each risk score that was closest to several threshold probabilities was set as the cut-off value. Supplementary Table S7 illustrates

Total risk score	Predicted probability of SGA infants (95% CI), %	Observed proportion of SGA infants (Cases/total), %
<b>Early gestation</b>		
Quintile 1 ( $\leq -2$ )	2.6 (2.1–3.2)	2.7 (87/3,262)
Quintile 2 (-1 to 0)	4.3 (3.6–5.0)	4.3 (137/3,222)
Quintile 3 (1 to 2)	5.8 (4.9–6.4)	5.7 (208/3,679)
Quintile 4 (3 to 4)	7.5 (6.6–8.4)	7.5 (242/3,217)
Quintile 5 ( $\geq 5$ , High risk)	12.1 (11.2–13.3)	12.2 (452/3,693)
Predicted probability of SGA infants based on the risk score	$\exp(\text{logit})/(1 + \exp[\text{logit}])$ where $\text{logit} = -3.0394 + 0.1545 \times (\text{risk score})$	–
<b>Model 1 during mid-gestation</b>		
Quintile 1 ( $\leq -3$ )	2.4 (1.9–2.9)	2.4 (92/3,808)
Quintile 2 (-2 to -1)	4.1 (3.4–4.8)	4.1 (116/2,831)
Quintile 3 (0 to 1)	5.5 (4.9–6.4)	5.6 (200/3,564)
Quintile 4 (2 to 3)	7.6 (6.2–8.0)	7.1 (215/3,023)
Quintile 5 ( $\geq 4$ , High risk)	12.9 (12.0–14.1)	13.1 (503/3,847)
Predicted probability of SGA infants based on the risk score	$\exp(\text{logit})/(1 + \exp[\text{logit}])$ where $\text{logit} = -2.9199 + 0.1671 \times (\text{risk score})$	–
<b>Model 2 during mid-gestation (Model 1 + EFW)</b>		
Quintile 1 ( $\leq -5$ )	1.7 (1.1–1.9)	1.5 (51/3,199)
Quintile 2 (-4 to -2)	3.4 (3.0–4.1)	3.6 (146/4,099)
Quintile 3 (-1 to 0)	5.1 (4.2–5.8)	5.0 (151/3,017)
Quintile 4 (1 to 3)	7.8 (7.4–9.2)	8.3 (292/3,538)
Quintile 5 ( $\geq 4$ , High risk)	16.5 (14.7–17.4)	16.1 (486/3,028)
Predicted probability of SGA infants based on the risk score	$\exp(\text{logit})/(1 + \exp[\text{logit}])$ where $\text{logit} = -2.8282 + 0.1853 \times (\text{risk score})$	–

**Table 2.** Predicted probability and observed proportion of SGA infants according to the quintiles of each risk score. The risk score during early gestation comprised the maternal age, height, BMI during early gestation, parity, ART with FET, smoking status, BP during early gestation, and maternal birth weight. The risk score during mid-gestation (model 1) consisted of the maternal age, height, BMI during mid-gestation, weight gain, parity, ART with FET, smoking status, BP level during mid-gestation, and maternal birth weight. The risk score during mid-gestation (model 2) also consisted of the maternal age, height, BMI during mid-gestation, weight gain, parity, ART with FET, smoking status, BP level during mid-gestation, maternal birth weight, and estimated fetal weight during mid-gestation. Abbreviations: ART; assisted reproductive technology, BMI; body mass index; BP, blood pressure; CI, confidence interval; exp, exponential; EFW, estimated fetal weight; FET, frozen-thawed embryo transfer; SGA, small for gestational age.

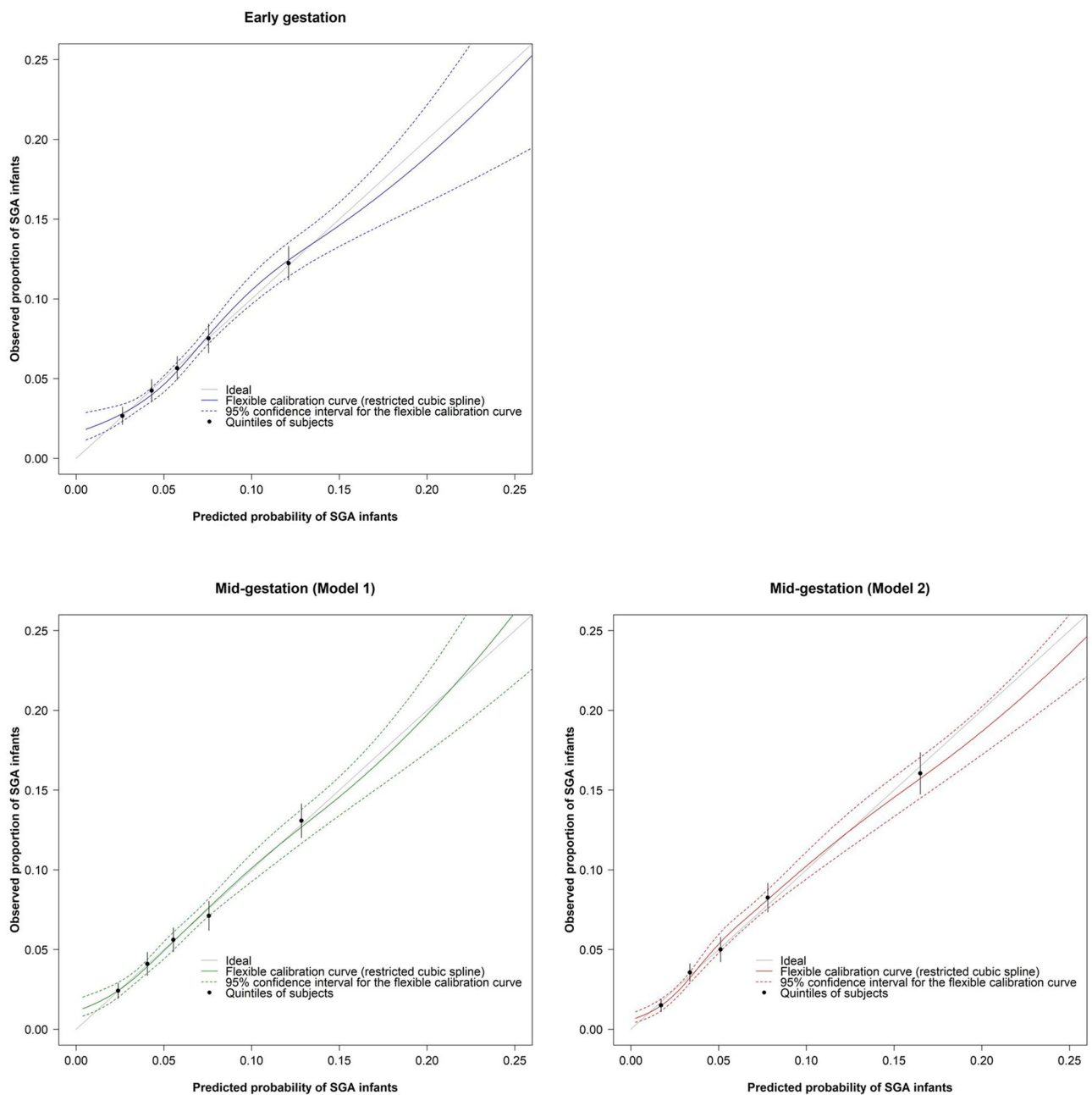
the discrimination performance and NB for each of the risk score that had the maximum Youden index that was set as the cut-off value.

**Model performance, calibration, and clinical utility of each risk score for predicting preterm and term SGA infants (Results of an additional analysis).** The results of the model performance, calibration, and clinical utility of each risk score for predicting preterm and term SGA infants, an additional analysis performed in the present study, are described in the Supplementary Information.

## Discussion

This is the first study in Japan to construct a prediction model for SGA infants based on risk scores. Since the predictive ability of the risk score including the EFW during mid-gestation was acceptable, it can be incorporated into prenatal checkup protocols for the early detection of pregnant women at a high risk of delivering SGA infants in Japan.

The predictive ability of the risk scores for SGA infants during early gestation was poor. However, by using a risk score, healthcare providers and pregnant women can collectively identify certain risk factors to improve outcomes, for instance, the smoking status during pregnancy, which can be modified through smoking cessation<sup>24,25</sup>. Maternal smoking during pregnancy is associated with higher resistance within the umbilical artery flow<sup>26</sup>. Decrease in endothelial nitric oxide synthase (eNOS) activity in fetal umbilical and chorionic vessels, caused by maternal smoking, may be a possible mechanism<sup>27</sup>. In addition, eNOS activity in pregnant women who quit smoking during pregnancy was higher than in those who smoked during pregnancy<sup>27</sup>. Therefore, smoking cessation may lead to a decrease in risk of SGA infants by increasing eNOS activity in fetal umbilical and chorionic vessels. Probably because model 1 during mid-gestation included the maternal weight gain and was closer to the time of delivery than the risk score during early gestation, it had a higher ability to predict SGA infants. By using a risk score during mid-gestation, healthcare providers can recognize maternal weight gain, which may

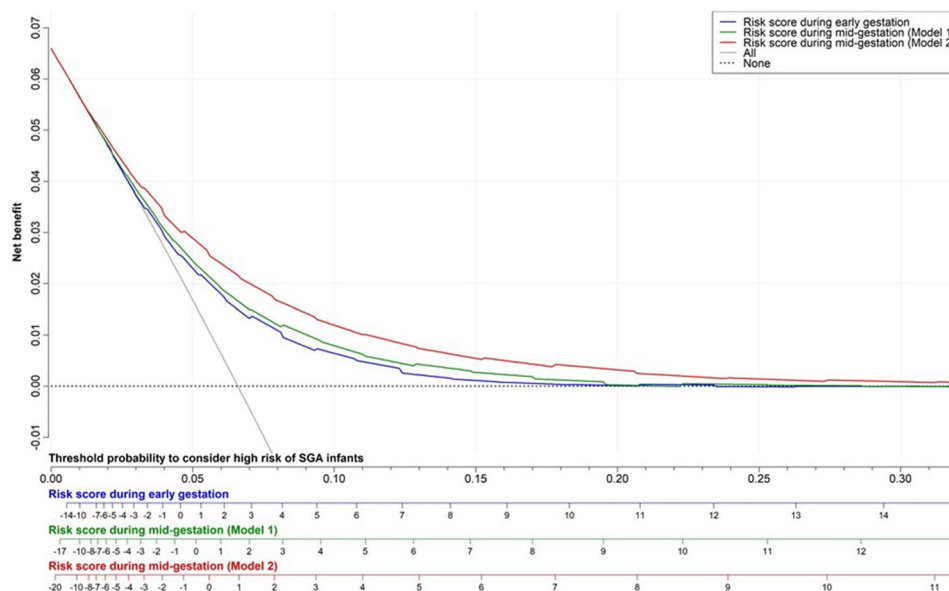


**Figure 3.** Calibration plot of each risk score for predicting SGA infants. “Ideal” indicates the line of perfect calibration. Each black circle and vertical line show the mean and 95% confidence interval of the proportion of SGA infants in each quintile of the risk score, respectively.

lead to possible nutrition counseling according to the maternal physique. The predictive ability of the risk score including the EFW during mid-gestation for SGA infants in this study was similar to that of other prediction models<sup>10,14</sup>. McCowan et al. reported that the C-statistic of the prediction model for SGA infants, created using clinical data and ultrasound variables collected at 15 and 20 weeks of gestation, respectively, was 0.73<sup>10</sup>. Erkamp et al. also reported that the C-statistic of the prediction model for SGA infants using maternal characteristics and the EFW during the second trimester was 0.72 (95% CI: 0.70–0.74)<sup>14</sup>.

Considering maternal complications, SGA (as a surrogate of FGR) requires careful surveillance with modalities such as the EFW, fetal doppler velocimetry, and cardiotocography at tertiary institutions<sup>7</sup>. If pregnant women are considered at a high-risk of delivering SGA infants by the prediction model, a detailed fetal ultrasonography will be needed. Fetal ultrasonography includes confirmation of congenital morphological abnormality, measuring maternal uterine artery pulsatility index for evaluation of placental dysfunction, close follow-up of EFW and fetal abdominal circumference for evaluation of fetal growth velocity. If fetal growth deteriorates, evaluation of fetal doppler velocimetry, including umbilical artery pulsatility index, middle cerebral artery pulsatility index, and flow of ductus venosus in combination with cardiotocography or biophysical profile scoring will be needed. In Japan,





**Figure 4.** Decision curve analysis of the risk scores for predicting SGA infants. “All” (solid grey line) indicates that all subjects are considered to be at a high risk of delivering SGA infants. “None” (dashed black line) indicates that no subjects are considered to be at a high risk of delivering SGA infants.

Cut-off (Minimum risk score of quintile 5)	TPR (Sensitivity) (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Positive LR (95% CI)	Negative LR (95% CI)	NB
<b>Early gestation</b>							
Risk score = 5 (Predicted probability of SGA infants = 0.094)	0.401 (0.373–0.430)	0.797 (0.791–0.803)	0.122 (0.112–0.133)	0.950 (0.946–0.953)	1.98 (1.83–2.14)	0.75 (0.72–0.79)	0.007
<b>Mid-gestation</b>							
Model 1: Risk score = 4 (Predicted probability of SGA infants = 0.095)	0.447 (0.418–0.476)	0.790 (0.784–0.797)	0.131 (0.120–0.141)	0.953 (0.949–0.957)	2.13 (1.98–2.29)	0.70 (0.66–0.74)	0.008
Model 2: Risk score = 4 (Predicted probability of SGA infants = 0.110)	0.432 (0.403–0.461)	0.841 (0.835–0.846)	0.161 (0.147–0.174)	0.954 (0.951–0.958)	2.71 (2.51–2.92)	0.68 (0.64–0.71)	0.011

**Table 3.** Discrimination and the NB based on the lowest risk score of quintile 5 as the cut-off value in each risk score. Abbreviations: CI, confidence interval; LR, likelihood ratio; NB, net benefit; NPV, negative predictive value; PPV, positive predictive value; SGA, small for gestational age; TPR, true positive rate.

pregnant women are commonly managed at midwife homes or at primary, secondary, and tertiary institutions<sup>28</sup>. Obstetric medical institutions are increasingly becoming more centralized due to a shortage of obstetricians in Japan and a decrease in the number of medical institutions providing perinatal care<sup>29</sup>. Therefore, division of roles is increasingly practiced in tertiary institutions and institutions that manage low-risk pregnancies; the latter need to determine when to transfer pregnant women at a high risk of delivering SGA infants to tertiary institutions. Here, our prediction model may provide an early opportunity for the recognition of such women, allowing sufficient time for decision-making on their transfer. However, the NB for each of the risk scores was low in this study. Therefore, it is necessary to create a prediction model for SGA infants that increases NB more than the current model. Additionally, our model (in particular, model 2 during mid-gestation) for predicting preterm SGA infants should be updated though recalibration in the future. We would like to propose the use of a risk score for predicting SGA infants, as a supportive rather than a mandatory tool, at prenatal checkups. The decision to use our risk score for predicting SGA infants should be taken by the medical institution concerned.

The strength of this study is that many variables that were used for creating a prediction model for SGA infants were collected prospectively in a large sample size cohort study. Conversely, the limitations of the study are as follows. First, external validation of the prediction model was not performed. Other prediction models for SGA infants, which were constructed in other countries, have been evaluated for external validation<sup>30</sup>. Although several characteristics in this study (including the maternal age; proportions of primipara, preterm births, and low-birth-weight infants; mean gestational age at delivery; and infant birth weight) were similar to those used in the Japan Environment and Children’s Study (a nationwide birth cohort study in Japan), we will perform an external validation in the near future<sup>31</sup>. Second, it is unknown whether low maternal birth weight was attributable

to SGA or preterm births, because information on the gestational week when the subjects were born in was not collected. Third, data on other predictors of SGA infants, including prenatal ultrasonographic findings on abnormal cord insertion site, and abnormal cord coiling which leads to impaired cord blood flow were not collected in this study<sup>32,33</sup>. Furthermore, neither the maternal uterine artery pulsatility index nor biochemical markers such as the pregnancy-associated plasma protein-A were measured in this study<sup>34,35</sup>. Moreover, because the placental growth factor data were available for only a small number of subjects in this study, this parameter could not be incorporated into the prediction model. Therefore, we could not evaluate predictive performance of other prediction models of SGA infants for subjects in this study. In addition, we also could not compare predictive performance among risk scores in this study and other prediction models. However, these parameters are not routinely measured in clinical practice in Japan. Therefore, our prediction model may be useful in environments where provisions for such skill-intensive techniques and special measurement systems are not available. Riskin-Mashiah, S. et al. reported that fasting plasma glucose (FPG) in the first trimester was a predictor of infant birth weight<sup>36</sup>. However, 86.0% of data on FPG in the first trimester was missing in this study. In addition, the percentage of missing data on family history of DM, a risk factor of gestational diabetes mellitus, was 37.3%. Therefore, neither FPG in the first trimester nor family history of DM could be incorporated into the prediction model due to the high proportion of missing data in this study<sup>37</sup>. Although other parameters of fetal ultrasonography, including biparietal diameter (BPD), abdominal circumference (AC), femur length (FL), and fetal congenital anomalies may improve the predictive ability of our prediction model for SGA infants, there is a high proportion of missing data for at least one of BPD, AC, and FL in this study. Furthermore, neither prenatal ultrasonographic findings on fetal congenital anomalies nor doppler assessment of umbilical artery were recorded in this study. Therefore, we could not include BPD, AC, FL, fetal congenital anomalies, and doppler assessment of umbilical artery in the prediction model. For a sensitivity analysis, the discrimination performance of SGA infants without major congenital anomalies using the prediction model was evaluated. The discrimination performance of SGA infants without major congenital anomalies was similar to that of all SGA infants. The C-statistics and tenfold cross-validated C-statistics of the risk score during early gestation were 0.659 (95% CI: 0.643–0.676) and 0.659 (95% CI: 0.642–0.677), respectively. In model 1 during mid-gestation, the C-statistics and tenfold cross-validated C-statistics were 0.677 (95% CI: 0.661–0.694) and 0.673 (95% CI: 0.661–0.694), respectively. In model 2 during mid-gestation, the C-statistics and tenfold cross-validated C-statistics were 0.723 (95% CI: 0.708–0.738) and 0.723 (95% CI: 0.707–0.739), respectively.

The cause of hypertensive disorders of pregnancy (HDP), especially preeclampsia, is thought to be impaired uterine spiral artery remodeling followed by angiogenic imbalance. As a result, the insufficient uteroplacental perfusion leads to SGA (as a surrogate of FGR)<sup>38</sup>. As low-dose aspirin treatment for pregnant women at high risk of preeclampsia would decrease the risk of delivery of SGA infants, construction of a prediction model for SGA infants with HDP (i.e., preeclampsia) in early gestation will be needed in Japan in the future<sup>39</sup>.

In conclusion, our prediction model for SGA infants, particularly during mid-gestation, may aid in the detection of pregnant women at a high risk of delivering SGA infants in Japan. Further studies for its external validation and improvement of its predictive ability are necessary.

## Methods

**Study design and participants.** This study was part of the TMM BirThree Cohort Study, an ongoing prospective cohort study. The TMM BirThree Cohort Study, one of the several cohort studies conducted by the TMM, aimed to 1) monitor the damage to health status due to the Great East Japan Earthquake, 2) study the early diagnosis and treatment of diseases, and 3) perform molecular-epidemiological studies to examine the associations between genetic and environmental factors and diseases<sup>40</sup>.

The TMM BirThree Cohort Study recruited pregnant women (and their family members), whose expected date of confinement was later than February 1, 2014, from obstetric clinics and hospitals in the Miyagi and Iwate Prefectures in Japan from July 19, 2013 to March 31, 2017. Timing of consent to participate in the TMM BirThree Cohort Study was from the whole period of pregnancy to one month after delivery. Written informed consent was obtained from all participants. The study protocol was approved by the Institutional Review Board of the Tohoku University Graduate School of Medicine (approval number: 2013–1–103–1). The study has been conducted in accordance with the Declaration of Helsinki, the Ethical Guidelines for Human Genome/Gene Analysis Research, and all other applicable guidelines<sup>41,42</sup>. The details and cohort profile of the TMM BirThree Cohort Study have been described previously<sup>40,43,44</sup>.

For easy use in clinical settings, we developed a prediction model based on risk scores to predict the delivery of SGA infants, and an internal validation was performed using the data of all the enrolled subjects. This study was also described based on the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement<sup>45,46</sup>. The official TRIPOD checklist is shown in the supplementary information<sup>46</sup>.

**Candidate explanatory variables of the risk score for predicting SGA infants.** The maternal age, height, BMI, weight gain, general medical history and medical history during previous pregnancy, conception method, smoking status, alcohol drinking, clinic blood pressure (BP), standard deviation (SD) of the EFW during mid-gestation, and maternal birth weight were considered as the candidate explanatory variables of the risk score<sup>10,15–17,19,21,34,47–51</sup>. Details on the data collection of the candidate explanatory variables are presented in the supplementary information.

There is a potential for extramatern survival at  $\geq 22$  weeks of gestation; furthermore, reference ranges of birth weights are available for Japanese infants from  $\geq 22$  weeks of gestation onward<sup>52</sup>. Therefore, we considered that a prediction model for SGA infants at  $< 22$  weeks of gestation would be clinically significant for fetal surveillance at

the prenatal checkup. Furthermore, the pregnancy period was divided into two periods separated by 18 weeks of gestation, because Japanese reference ranges for the EFW are unavailable for < 18 weeks of gestation<sup>53</sup>. Therefore, we constructed the risk scores to predict SGA infants in the following two gestational periods that were less than 22 weeks: 1) early gestation: 11 weeks, 0 days to 17 weeks, 6 days and 2) mid-gestation: 18 weeks, 0 days to 21 weeks, 6 days.

**Definition of SGA infants.** There is a difference in mean infant birth weight between different races<sup>22</sup>. Since 1980, the mean infant birth weight has decreased and the proportion of LBW infants has increased rapidly in Japan<sup>54</sup>. Additionally, owing to the prevalent use of customized infant birth weight percentiles among health practitioners in clinical practice in Japan, we used it to define SGA infants in this study. Data on the infant birth weight, parity, delivery week, and sex were obtained from the medical records, because the infant birth weight percentile in Japan is customized based on these parameters<sup>52,55</sup>. SGA infants were defined as infants whose birth weight was in the < 10<sup>th</sup> percentile.

**Statistical analyses.** Continuous variables of the maternal and neonatal characteristics in this study were expressed as means  $\pm$  SDs or median (interquartile range), as appropriate. Categorical variables were also expressed as numbers (percentages).

First, we explored the candidate explanatory variables that were associated with SGA infants using a univariate logistic regression model. When the risk score during early gestation was constructed, the maternal age, maternal height, maternal pre-pregnancy BMI, BMI during early gestation, maternal weight gain between pre-pregnancy BW and the initial BW during early gestation, parity, conception method, maternal birth weight, BP during early gestation, smoking status, alcohol consumption, and medical histories of diabetes mellitus (DM), systemic lupus erythematosus (SLE) and/or antiphospholipid syndrome (APS), chronic kidney diseases (CKD), hyperthyroidism, and hypothyroidism were included in the univariate logistic regression model. When the risk score during mid-gestation was constructed, the maternal BMI during mid-gestation, maternal weight gain between initial BW during early gestation and initial BW during mid-gestation, BP during mid-gestation, and SD value of the EFW during mid-gestation were included in a univariate logistic regression model in addition to the explanatory variables that were considered during early gestation. Furthermore, risk scores without EFW during mid-gestation (model 1) and with EFW during mid-gestation (model 2) were created. Explanatory variables that showed a two-sided *P*-value of < 0.20 in the univariate logistic regression model were included in a multiple logistic regression model. Furthermore, the maternal age during early or mid-gestation and height, which were parameters related to SGA infants in previous studies, were also included in the multiple logistic regression model, even when the two-sided *P*-value was  $\geq$  0.20 in the univariate logistic regression model<sup>15,16</sup>. If the variance inflation factor was greater than 2.5, multicollinearity among several explanatory variables was suspected, and the explanatory variables were either combined or one variable was chosen to decrease the multicollinearity<sup>56</sup>. In the multiple logistic regression model, explanatory variables that contributed to the prediction of SGA infants were selected when each two-sided *P*-value was < 0.05. The sample size in this study satisfied the condition that the number of SGA infants per explanatory variable was 10 or more to avoid overfitting in the multiple logistic regression model<sup>57</sup>.

The regression coefficients, rather than the odds ratio, were divided by the smallest absolute value of the regression coefficients among the selected explanatory variables and then rounded to an integer score<sup>58</sup>. Next, the risk score for predicting SGA infants was calculated by summing the integer score of each explanatory variable.

To assess the discrimination performance of each risk score for the SGA infants, we created receiver operating characteristics (ROC) curves and calculated the C-statistics (also known as the area under the ROC curves). The C-statistics were interpreted as follows: 0.5 (no discrimination); > 0.5 and < 0.7 (poor discrimination);  $\geq$  0.7 and < 0.8 (acceptable discrimination);  $\geq$  0.8 and < 0.9 (excellent discrimination); and  $\geq$  0.9 (outstanding discrimination)<sup>59</sup>. For the internal validation of each risk score, ten-fold cross validation was conducted. Furthermore, a calibration plot using a restricted cubic spline function with four knots was created to assess the calibration<sup>60</sup>. Calibration evaluates the concordance between the predicted probability and observed proportion of the SGA infants. To evaluate the clinical utility of the risk scores for SGA infant prediction by a net benefit (NB), we conducted a decision curve analysis (DCA)<sup>61–63</sup>.

To evaluate the differences in the predictive abilities of the different risk scores for SGA infants, we evaluated the differences in the C-statistics and reclassification. Differences in the C-statistics were determined using the Delong's test<sup>64</sup>. Reclassification was evaluated by calculating the net reclassification improvement (NRI) and integrated discrimination improvement (IDI)<sup>65–67</sup>. Since there is no established meaningful risk category of SGA infants, we evaluated the continuous NRI (cNRI), rather than the category-based NRI. The overall cNRI evaluates the upward and downward changes in the predicted risk of SGA infants by changing the reference model to the new model. The overall cNRI was calculated as the sum of the event cNRI and nonevent cNRI. The event cNRI indicated the net proportion of subjects who delivered SGA infants and were correctly predicted to have a higher risk of delivering SGA infants. The nonevent cNRI indicated the net proportion of subjects who did not deliver SGA infants and were correctly predicted to have a lower risk of delivering SGA infants. The IDI was calculated by the following formula: IDI = change in the average predicted probability of SGA infants for subjects who delivered SGA infants between the two models—change in the average predicted probability of SGA infants for subjects who had not delivered SGA infants between the two models. When differences in the C-statistics, cNRI, and IDI among the risk scores during early gestation, model 1, and model 2 during mid-gestation were compared, a two-sided *P*-value of < 0.0167 (0.05/3) by the Bonferroni's correction was considered as statistically significant. In addition, we also compared the decision curves of the risk scores.

Considering that it is unknown whether infants are SGA or not until they are born, we also explored cut-off values with the predicted probability of SGA infants. Thus far, there is no established threshold probability to

categorize pregnant women as having a high risk of delivering SGA infants, although the other study reported several cut-off values according to the predicted probability of SGA infants<sup>30</sup>. Therefore, several cut-off values of the risk scores were considered to define a high risk of SGA infant delivery in this study. After the subjects were divided into quintiles based on the distribution of each risk score, the minimum risk scores of the fifth quintile in each risk score were set as the cut-off values. Furthermore, the risk scores closest to threshold probabilities of 0.05, 0.10, 0.15, and 0.20 (i.e., 5%, 10%, 15%, and 20%) were set as the cut-off values for each risk score. The numbers that must be tested corresponding to the threshold probabilities of 0.05, 0.10, 0.15, and 0.20 were 20, 10, 7.7, and 5, respectively. The risk scores which had the maximum Youden index (i.e., sensitivity + specificity—1) were also set as the cut-off values<sup>68</sup>. Then, the true positive rate (i.e., sensitivity), specificity, positive predictive value, negative predictive value, positive likelihood ratio (LR), and negative LR were calculated in each cut-off of the risk score<sup>69</sup>. Furthermore, we calculated each NB when each cut-off risk score was set.

In the sensitivity analysis, we constructed a risk score after multiple imputations by a chained equation (MICE), as described in the Supplementary Material<sup>70</sup>. As an additional analysis, we evaluated the model performance, calibration, and clinical utility of each risk score for predicting preterm and term SGA infants. We also compared the model performance between different risk scores for predicting preterm and term SGA infants.

Statistical software used in the statistical analysis are described in the supplementary information.

## Data availability

The datasets analyzed in this study are available from the corresponding author on reasonable request.

Received: 27 June 2021; Accepted: 16 May 2022

Published online: 26 May 2022

## References

- Gardosi, J., Madurasinghe, V., Williams, M., Malik, A. & Francis, A. Maternal and fetal risk factors for stillbirth: Population based study. *BMJ* **346**, f108. <https://doi.org/10.1136/bmj.f108> (2013).
- Sylvestre, G., Fisher, M., Westgren, M. & Divon, M. Y. Non-reassuring fetal status in the prolonged pregnancy: The impact of fetal weight. *Ultrasound Obstet. Gynecol.* **18**, 244–247. <https://doi.org/10.1046/j.1469-0705.2001.00460.x> (2001).
- Takeuchi, A. *et al.* Neurodevelopment in full-term small for gestational age infants: A nationwide Japanese population-based study. *Brain Dev.* **38**, 529–537. <https://doi.org/10.1016/j.braindev.2015.12.013> (2016).
- Leite, D. F. B. & Cecatti, J. G. Fetal growth restriction prediction: How to move beyond. *Sci. World J.* **2019**, 1519048. <https://doi.org/10.1155/2019/1519048> (2019).
- Lindqvist, P. G. & Molin, J. Does antenatal identification of small-for-gestational age fetuses significantly improve their outcome?. *Ultrasound Obstet. Gynecol.* **25**, 258–264. <https://doi.org/10.1002/uog.1806> (2005).
- Verljijsdonk, J. W., Winkens, B., Boers, K., Scherjon, S. & Roumen, F. Suspected versus non-suspected small-for-gestational age fetuses at term: perinatal outcomes. *J. Matern. Fetal. Neonatal. Med.* **25**, 938–943. <https://doi.org/10.3109/14767058.2011.600793> (2012).
- Lees, C. C. *et al.* ISUOG Practice Guidelines: Diagnosis and management of small-for-gestational-age fetus and fetal growth restriction. *Ultrasound Obstet. Gynecol.* **56**, 298–312. <https://doi.org/10.1002/uog.22134> (2020).
- Haragan, A. & Himes, K. Accuracy of ultrasound estimated fetal weight in small for gestational age and appropriate for gestational age grown periviable neonates. *Am. J. Perinatol.* **35**, 703–706. <https://doi.org/10.1055/s-0037-1617433> (2018).
- Stephens, K. *et al.* Comparing the relation between ultrasound-estimated fetal weight and birthweight in cohort of small-for-gestational-age fetuses. *Acta Obstet. Gynecol. Scand.* **98**, 1435–1441. <https://doi.org/10.1111/aogs.13645> (2019).
- McCowan, L. M. *et al.* Clinical prediction in early pregnancy of infants small for gestational age by customised birthweight centiles: Findings from a healthy nulliparous cohort. *PLoS ONE* **8**, e70917. <https://doi.org/10.1371/journal.pone.0070917> (2013).
- Lesmes, C., Gallo, D. M., Panaiotova, J., Poon, L. C. & Nicolaides, K. H. Prediction of small-for-gestational-age neonates: Screening by fetal biometry at 19–24 weeks. *Ultrasound Obstet. Gynecol.* **46**, 198–207. <https://doi.org/10.1002/uog.14826> (2015).
- McCowan, L. M. *et al.* Prediction of small for gestational age infants in healthy nulliparous women using clinical and ultrasound risk factors combined with early pregnancy biomarkers. *PLoS ONE* **12**, e0169311. <https://doi.org/10.1371/journal.pone.0169311> (2017).
- Allotey, J. & Thangaratinam, S. Prognostic models need to look beyond fetal size. *BJOG* **126**, 485. <https://doi.org/10.1111/1471-0528.15564> (2019).
- Erkamp, J. S. *et al.* Second and third trimester fetal ultrasound population screening for risks of preterm birth and small-size and large-size for gestational age at birth: A population-based prospective cohort study. *BMC Med.* **18**, 63. <https://doi.org/10.1186/s12916-020-01540-x> (2020).
- Kyozuka, H. *et al.* The effect of maternal age at the first childbirth on gestational age and birth weight: The Japan environment and children's study (JECS). *J. Epidemiol.* **29**, 187–191. <https://doi.org/10.2188/jea.JE20170283> (2019).
- Ogawa, K. *et al.* Association of shorter height with increased risk of ischaemic placental disease. *Paediatr. Perinat. Epidemiol.* **31**, 198–205. <https://doi.org/10.1111/ppe.12351> (2017).
- Tamura, N. *et al.* different risk factors for very low birth weight, term-small-for-gestational-age, or preterm birth in Japan. *Int. J. Environ. Res. Public Health*, <https://doi.org/10.3390/ijerph15020369> (2018).
- Suzuki, K., Shinohara, R., Sato, M., Otawa, S. & Yamagata, Z. Association between maternal smoking during pregnancy and birth weight: An appropriately adjusted model from the Japan environment and children's study. *J. Epidemiol.* **26**, 371–377. <https://doi.org/10.2188/jea.JE20150185> (2016).
- Iwama, N. *et al.* Maternal clinic and home blood pressure measurements during pregnancy and infant birth weight: The BOSHI study. *Hypertens. Res.* **39**, 151–157. <https://doi.org/10.1038/hr.2015.108> (2016).
- Morisaki, N. *et al.* Pre-pregnancy BMI-specific optimal gestational weight gain for women in Japan. *J. Epidemiol.* **27**, 492–498. <https://doi.org/10.1016/j.je.2016.09.013> (2017).
- Goldstein, R. F. *et al.* Gestational weight gain across continents and ethnicity: Systematic review and meta-analysis of maternal and infant outcomes in more than one million women. *BMC Med.* **16**, 153. <https://doi.org/10.1186/s12916-018-1128-1> (2018).
- Shiono, P. H., Klebanoff, M. A., Graubard, B. I., Berendes, H. W. & Rhoads, G. G. Birth weight among women of different ethnic groups. *JAMA* **255**, 48–52 (1986).
- Morisaki, N., Kawachi, I., Oken, E. & Fujiwara, T. Parental characteristics can explain Why Japanese women give birth to the smallest infants in the United States. *Paediatr. Perinat. Epidemiol.* **30**, 473–478. <https://doi.org/10.1111/ppe.12308> (2016).
- McCowan, L. M. *et al.* Spontaneous preterm birth and small for gestational age infants in women who stop smoking early in pregnancy: Prospective cohort study. *BMJ* **338**, b1081. <https://doi.org/10.1136/bmj.b1081> (2009).

25. McCowan, L. M., Figueras, F. & Anderson, N. H. Evidence-based national guidelines for the management of suspected fetal growth restriction: Comparison, consensus, and controversy. *Am. J. Obstet. Gynecol.* **218**, S855–S868. <https://doi.org/10.1016/j.ajog.2017.12.004> (2018).
26. Pringle, P. J. *et al.* The influence of cigarette smoking on antenatal growth, birth size, and the insulin-like growth factor axis. *J. Clin. Endocrinol. Metab.* **90**, 2556–2562. <https://doi.org/10.1210/jc.2004-1674> (2005).
27. Andersen, M. R., Simonsen, U., Ulbjerg, N., Aalkjaer, C. & Stender, S. Smoking cessation early in pregnancy and birth weight, length, head circumference, and endothelial nitric oxide synthase activity in umbilical and chorionic vessels: An observational study of healthy singleton pregnancies. *Circulation* **119**, 857–864. <https://doi.org/10.1161/CIRCULATIONAHA.107.755769> (2009).
28. Sugawara, J. *et al.* Regional Birth Outcomes after the 2011 Great East Japan Earthquake and Tsunami in Miyagi Prefecture. *Prehosp Disaster Med* **33**, 215–219. <https://doi.org/10.1017/S1049023X18000183> (2018).
29. Matsumoto, M. *et al.* Selection and concentration of obstetric facilities in Japan: Longitudinal study based on national census data. *J. Obstet. Gynaecol. Res.* **41**, 919–925. <https://doi.org/10.1111/jog.12663> (2015).
30. Meertens, L. *et al.* External validation and clinical usefulness of first-trimester prediction models for small- and large-for-gestational-age infants: A prospective cohort study. *BJOG* **126**, 472–484. <https://doi.org/10.1111/1471-0528.15516> (2019).
31. Michikawa, T. *et al.* Baseline profile of participants in the Japan environment and children's study (JECS). *J Epidemiol* **28**, 99–104. <https://doi.org/10.2188/jea.JE20170018> (2018).
32. O'Quinn, C., Cooper, S., Tang, S. & Wood, S. Antenatal diagnosis of marginal and velamentous placental cord insertion and pregnancy outcomes. *Obstet. Gynecol.* **135**, 953–959. <https://doi.org/10.1097/aog.0000000000003753> (2020).
33. Hasegawa, J. Ultrasound screening of umbilical cord abnormalities and delivery management. *Placenta* **62**, 66–78. <https://doi.org/10.1016/j.placenta.2017.12.003> (2018).
34. Papastefanou, I. *et al.* Competing-risks model for prediction of small-for-gestational-age neonate from biophysical and biochemical markers at 11–13 weeks' gestation. *Ultrasound Obstet. Gynecol.* **57**, 52–61. <https://doi.org/10.1002/uog.23523> (2021).
35. Royal College of Obstetricians and Gynaecologists. The Investigation and Management of the Small-for-Gestational-Age Fetus. *Green-top guideline* (2014).
36. Riskin-Mashiah, S., Younes, G., Damti, A. & Auslender, R. First-trimester fasting hyperglycemia and adverse pregnancy outcomes. *Diabetes Care* **32**, 1639–1643. <https://doi.org/10.2337/dc09-0688> (2009).
37. Solomon, C. G. *et al.* A prospective study of pregravid determinants of gestational diabetes mellitus. *JAMA* **278**, 1078–1083 (1997).
38. Chaemsaihong, P., Sahota, D. S. & Poon, L. C. First trimester preeclampsia screening and prediction. *Am. J. Obstet. Gynecol.* **226**, S1071–S1097 e1072. <https://doi.org/10.1016/j.ajog.2020.07.020> (2022).
39. Roberge, S. *et al.* The role of aspirin dose on the prevention of preeclampsia and fetal growth restriction: systematic review and meta-analysis. *Am J Obstet Gynecol* **216**, 110–12. <https://doi.org/10.1016/j.ajog.2016.09.076> (2017).
40. Kuriyama, S. *et al.* The Tohoku Medical Megabank Project: Design and mission. *J Epidemiol* **26**, 493–511. <https://doi.org/10.2188/jea.JE20150268> (2016).
41. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA* **310**, 2191–2194. <https://doi.org/10.1001/jama.2013.281053> (2013).
42. Ministry of Education, Culture, Sports, Science and Technology, Ministry of Health, Labour and Welfare, Ministry of Economy, Trade and Industry. *Ethical Guidelines for Human Genome Gene/ Gene Analysis Research*. Tokyo: Ministry of Education, Culture, Sports, Science and Technology, Ministry of Health, Labour and Welfare, Ministry of Economy, Trade and Industry 2001.
43. Sugawara, J. *et al.* Maternal baseline characteristics and perinatal outcomes: The tohoku medical megabank project birth and three-generation cohort study. *J. Epidemiol.* <https://doi.org/10.2188/jea.JE20200338> (2020).
44. Kuriyama, S. *et al.* Cohort Profile: Tohoku medical megabank project birth and three-generation cohort study (TMM BirThree Cohort Study): Rationale, progress and perspective. *Int. J. Epidemiol.* **49**, 18–19m. <https://doi.org/10.1093/ije/dyz169> (2020).
45. Collins, G. S., Reitsma, J. B., Altman, D. G. & Moons, K. G. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD statement. *Ann Intern Med* **162**, 55–63. <https://doi.org/10.7326/M14-0697> (2015).
46. TRIPOD Checklist: Prediction Model Development <https://www.tripod-statement.org/resources/>. [accessed 1 January 2021].
47. Kuhle, S. *et al.* Comparison of logistic regression with machine learning methods for the prediction of fetal growth abnormalities: A retrospective cohort study. *BMC Pregnancy Childbirth* **18**, 333. <https://doi.org/10.1186/s12884-018-1971-2> (2018).
48. Fischer-Betz, R. & Specker, C. Pregnancy in systemic lupus erythematosus and antiphospholipid syndrome. *Best Pract. Res. Clin. Rheumatol.* **31**, 397–414. <https://doi.org/10.1016/j.berh.2017.09.011> (2017).
49. Pelkonen, S. *et al.* Perinatal outcome of children born after frozen and fresh embryo transfer: The Finnish cohort study 1995–2006. *Hum. Reprod* **25**, 914–923. <https://doi.org/10.1093/humrep/dep477> (2010).
50. Turunen, S. *et al.* Maternal hyperthyroidism and pregnancy outcomes: A population-based cohort study. *Clin. Endocrinol. (Oxf)* **93**, 721–728. <https://doi.org/10.1111/cen.14282> (2020).
51. Derakhshan, A. *et al.* Association of maternal thyroid function with birthweight: A systematic review and individual-participant data meta-analysis. *Lancet Diabetes Endocrinol.* **8**, 501–510. [https://doi.org/10.1016/s2213-8587\(20\)30061-9](https://doi.org/10.1016/s2213-8587(20)30061-9) (2020).
52. Itabashi, K. *et al.* New normative birthweight among Japanese infants according to gestational week at delivery. *Acta Paediatr. Jpn.* **114**, 1271–1293 (In Japanese) (2010).
53. Okai, T. Standard values of ultrasonic measurements in Japanese fetuses. *Journal of Medical Ultrasonics* **30**, J415–440 (In Japanese) (2003).
54. Kato, N., Sauvaget, C., Yoshida, H., Yokoyama, T. & Yoshiike, N. Factors associated with birthweight decline in Japan (1980–2004). *BMC Pregnancy Childbirth* **21**, 337. <https://doi.org/10.1186/s12884-021-03819-0> (2021).
55. Uehara, R., Miura, F., Itabashi, K., Fujimura, M. & Nakamura, Y. Distribution of birth weight for gestational age in Japanese infants delivered by cesarean section. *J. Epidemiol.* **21**, 217–222. <https://doi.org/10.2188/jea.je20100123> (2011).
56. Katz, M. H. *Multivariable analysis. A practical guide for clinicians and public health researchers*. Third edn, (Cambridge University Press, 2011).
57. Peduzzi, P., Concato, J., Kemper, E., Holford, T. R. & Feinstein, A. R. A simulation study of the number of events per variable in logistic regression analysis. *J. Clin. Epidemiol.* **49**, 1373–1379. [https://doi.org/10.1016/s0895-4356\(96\)00236-3](https://doi.org/10.1016/s0895-4356(96)00236-3) (1996).
58. Moons, K. G., Harrell, F. E. & Steyerberg, E. W. Should scoring rules be based on odds ratios or regression coefficients?. *J. Clin. Epidemiol.* **55**, 1054–1055. [https://doi.org/10.1016/s0895-4356\(02\)00453-5](https://doi.org/10.1016/s0895-4356(02)00453-5) (2002).
59. David W. Hosmer, J., Lemeshow, S., Sturdivant, R. X. *Applied Logistic Regression*. Third edn, (Wiley, 2013).
60. Van Calster, B. *et al.* A calibration hierarchy for risk models was defined: from utopia to empirical data. *J. Clin. Epidemiol.* **74**, 167–176. <https://doi.org/10.1016/j.jclinepi.2015.12.005> (2016).
61. Vickers, A. J. & Elkin, E. B. Decision curve analysis: A novel method for evaluating prediction models. *Med. Decis. Making* **26**, 565–574. <https://doi.org/10.1177/0272989X06295361> (2006).
62. Van Calster, B. *et al.* Reporting and interpreting decision curve analysis: A guide for investigators. *Eur. Urol.* **74**, 796–804. <https://doi.org/10.1016/j.eururo.2018.08.038> (2018).
63. Vickers, A. J., van Calster, B. & Steyerberg, E. W. A simple, step-by-step guide to interpreting decision curve analysis. *Diagn. Progn. Res.* **3**, 18. <https://doi.org/10.1186/s41512-019-0064-7> (2019).
64. DeLong, E. R., DeLong, D. M. & Clarke-Pearson, D. L. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics* **44**, 837–845 (1988).

65. Pencina, M. J., D'Agostino, R. B., Sr., D'Agostino, R. B., Jr. & Vasan, R. S. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat. Med.* **27**, 157–172; discussion 207–112, doi:<https://doi.org/10.1002/sim.2929> (2008).
66. Pencina, M. J., D'Agostino, R. B. Sr. & Steyerberg, E. W. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat. Med.* **30**, 11–21. <https://doi.org/10.1002/sim.4085> (2011).
67. Zhou, Z. R. *et al.* In-depth mining of clinical data: the construction of clinical prediction model with R. *Ann. Transl. Med.* **7**, 796. <https://doi.org/10.21037/atm.2019.08.63> (2019).
68. Youden, W. J. Index for rating diagnostic tests. *Cancer* **3**, 32–35. [https://doi.org/10.1002/1097-0142\(1950\)3:1%3c32::aid-cnrcr2820030106%3e3.0.co;2-3](https://doi.org/10.1002/1097-0142(1950)3:1%3c32::aid-cnrcr2820030106%3e3.0.co;2-3) (1950).
69. McGee, S. Simplifying likelihood ratios. *J. Gen. Intern. Med.* **17**, 646–649. <https://doi.org/10.1046/j.1525-1497.2002.10750.x> (2002).
70. Michael, O. & Bohdana, R. *Clinical Trials with Missing Data, A Guide for Practitioners.*, (Wiley, 2014).

## Acknowledgements

This work was supported by the Japan Agency for Medical Research and Development (AMED), Japan [grant number: JP20km0105001], AMED Birthday [grant number: JP21gk0110039], and Japan Society for the Promotion of Science (JSPS) KAKENHI [grant number: JP19K18659]. We thank all the study participants, participating medical institutions, and staff involved in the TMM BirThree Cohort Study. We also thank all the past and present members of the Tohoku Medical Megabank Organization at the Tohoku University (present members are listed at <https://www.megabank.tohoku.ac.jp/english/a200601/>). We would also like to thank Editage (<https://www.editage.com/>) for their support in the revision of the manuscript.

## Author contributions

N.I. obtained the funding, performed the statistical analysis, and drafted the manuscript. T.O., K.M., A.N., F.U., T.O., F.M., H.M., and S.K. contributed to funding acquisition and designed and conducted the BirThree cohort study. M.I. contributed to the design, conducted the BirThree cohort study, and supervised the drafting of the manuscript. T.O., K.M., A.N., F.U., T.O., F.M., H.M., and S.K. supervised the drafting of the manuscript. T.H., M.S., J.S., and N.Y. contributed to the acquisition of data and supervised the drafting of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-12892-0>.

**Correspondence** and requests for materials should be addressed to N.I.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022