Clin Pediatr Endocrinol 2019; 28(3), 81–89

Original Article

Identifying a risk score for childhood obesity based on predictors identified in pregnant women and 1-year-old infants: An analysis of the data of the Hokkaido Study on Environment and Children's Health

Yasuaki Saijo[1], Yoshiya Ito[2], Eiji Yoshioka[1], Yukihiro Sato[1], Machiko Minatoya[3, 4], Atsuko Araki[3], Chihiro Miyashita[3], and Reiko Kishi[3]

[1]*Department of Social Medicine, Asahikawa Medical University, Sapporo, Japan*

[2]*Faculty of Nursing, Japanese Red Cross Hokkaido College of Nursing, Sapporo, Japan*

[3]*Center for Environmental and Health Sciences, Hokkaido University, Sapporo, Japan*

[4]*Faculty of Health Sciences, Hokkaido University, Sapporo, Japan*

Abstract. This study aimed to construct a childhood obesity risk index based on predictors identified in pregnant women and 1-yr-old infants. The primary outcome was an identified obesity index of > 20% at 6–8 yr of age. Of a total sample size of 6,846 mother-child pairs, 80% and 20% were randomly allocated to the derivation and validation cohorts, respectively. For the derivation cohort, univariate and multivariate logistic regression analyses of data were conducted to identify the final predictors to determine the childhood obesity risk score algorithm. These included pre-pregnancy body mass index (BMI), child's gender, smoking during pregnancy, education, and obesity index at one yr of age. The ß coefficients for categories of predictor variables were each divided by the smallest value among them. The quotient was rounded off to the integer and assigned to the risk score, and a value of zero was assigned to reference categories. A total risk score was calculated for each individual. A cutoff point ≥ 16 had 22.2% and 21.8% positive predictive values in the derivation and validation cohorts, respectively. In conclusion, the childhood obesity risk score algorithm was constructed based on generic predictors that can be easily obtained from maternal and child health handbooks.

Key words: childhood, obesity, risk score, pregnancy, body mass index

Introduction

Since 1975, obesity has approximately tripled worldwide (1), and the increase in the number of children who are overweight and obese has become an important public health issue in developed and developing countries (2). According to the 2017 Annual Report of School Health Statistics by the Ministry of Education, Culture, Sports, Science, and Technology, the rate of overweight students in Japan increased

Received: February 19, 2019
Accepted: May 10, 2019
Corresponding author: Reiko Kishi, MD, PhD, MPH, Department of Public Health Sciences, Hokkaido University Graduate School of Medicine, North 15 West 7, Kita-ku, Sapporo 060-8638, Japan.
E-mail: rkishi@med.hokudai.ac.jp

until the 2000s; however, it has been gradually decreasing. Because childhood obesity can cause various illnesses in adulthood, including cardiovascular diseases (2), preventive measures are required.

Effective treatment of obesity is not well established; therefore, early prevention is better than treatment later in life (3). Identifying children at a high risk of becoming overweight or obese can emphasize the development of preventive measures. A recent systematic review reported several childhood risk factors for becoming overweight and obese using factors determined during pregnancy (4); however, the included studies were conducted in North America and Europe, and rates of obesity varied across countries. Japan has relatively low overweight and obesity rates (5). Therefore, Japan must have unique childhood obesity risk factors and scores.

To construct the risk score, easily comprehensible predictors obtained from maternal and child health handbooks should be included, and the risk score algorithm developed should be easy to use during standard infant health checks. Accordingly, this study aimed to construct a childhood obesity risk score based on predictors identified in pregnant women and 1-yr-old infants included in a Japanese population-based birth cohort for the Hokkaido Study on Environment and Children's Health (6).

Methods

Participants

Details regarding the Hokkaido Study on Environment and Children's Health are described elsewhere (6–8). Participants included women in early pregnancy (< 13 wk gestational period) who visited any of the 37 hospitals (including three university hospitals) or their associated clinics evenly distributed throughout the Hokkaido Prefecture, from February 2003 to March 2012. The cohort still exists for a future follow up, and the current study included data

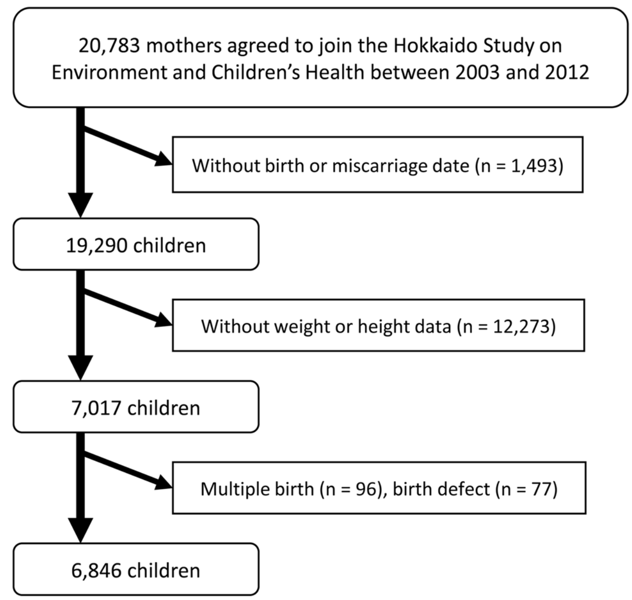

Fig. 1. Flow diagram of the selection of study participants.

up to September 2018. Figure 1 shows a flow chart of the study participants. A total of 20,783 expectant mothers were initially enrolled. Details of their pre-pregnancy height and weight, self-administered questionnaires during early pregnancy, and medical records at delivery were acquired. The following mother-child pairs were excluded from the study: 1,493 with missing data reflecting their birth or miscarriage date, 12,273 with missing height and weight data at 6–8 yr of age, 96 with multiple births, and 77 with birth defects (9); therefore, data of 6,846 mother-child pairs were analyzed.

We obtained written informed consent from all mothers. The study protocol was approved by the institutional ethics boards of the Hokkaido University Center for Environmental and Health Sciences (reference number 14; March 22, 2012) and the Hokkaido University Graduate School of Medicine (May 31, 2003).

Predictor variables

On the basis of previous childhood obesity risk score studies (4), predictor variables that could be determined easily during the 18-mo infant health check were selected. Information regarding pre-pregnancy height and weight, birth order (first-born child or younger), and education (high school diploma or lower or higher education) were obtained from a self-administered baseline questionnaire during early pregnancy. Pre-pregnancy body mass index (BMI) was calculated as weight (kg) divided by height (m^2), and was categorized as < 18.5, 18.5 to < 25, 25 to < 30, and ≥ 30 kg/m^2. Infant birth weight (< 2,500, 2,500 to < 3,000, 3,000 to < 3,500, and ≥ 3,500 g) and gender (male or female) were obtained from medical records at birth. Positive smoking status was defined as smoking during pregnancy or quitting after gestational week 16 based on a questionnaire administered when the child was four months old. Feeding (formula versus breastfeeding or combination feeding) was based on a questionnaire administered when the child was 1 yr; however, if data were missing, a report of continued breastfeeding from a questionnaire administered when the child was 2 yr was used. Furthermore, the questionnaire at 2 yr was used to obtain the height and weight of the infants when they were between 1 and 2 yr. The obesity index (relative body weight) was calculated (10) and categorized as < 10%, 10% to < 20%, or ≥ 20%.

Primary outcome variable

The height and weight of the child when they were between 6- and 8-yr-old were obtained from a questionnaire that was administered when they were seven years old (for some cases, 8-yr-old height and weight data were used because questionnaires were returned late), and the obesity index was calculated (10). Childhood obesity was defined as an obesity index of ≥ 20% (11).

Statistical analyses

The evaluated predictor variables were the pre-pregnancy BMI (four categories), birth weight (four categories), child's gender, smoking during pregnancy, education level, birth order, formula feeding, obesity index at 1 yr of age (three categories), age of the mother at delivery, gestational period, and the primary outcome (obesity index of ≥ 20%). Overall, 25.5% of participants had missing data for at least one variable, which were replaced using multiple imputations (25 imputed datasets) based on the assumption that data were missing at random. There were no missing data for the age of the mother at the time of delivery, and only one missing value for the gestational period.

The cohort was divided into two groups: 80% of the sample size was randomly selected as the derivation cohort to develop the obesity risk algorithm, and the remaining 20% was selected as the validation cohort.

In the derivation cohort, a univariate logistic regression analysis was used to obtain the crude odds ratios (ORs) of the predictor variables for childhood obesity; those with P-values < 0.20 were subsequently included in a mutually adjusted multivariate logistic regression analysis.

The method used to develop and validate the obesity risk prediction algorithm has been described previously (12). A risk score algorithm was constructed using the β coefficients of predictor variables in the mutually adjusted multivariate logistic regression model. The β coefficients for the categories of the predictor variables were each divided by the smallest value among them. The quotient was rounded off to the closest integer and assigned to the appropriate risk score, and a value of zero was assigned to reference categories. After the integer values were assigned to each of the variable categories, a total risk score was calculated for each mother-child pair within the derivation cohort.

Based on a 3-point risk score obtained in the derivation cohort, two cutoff points were identified based on the finding that scores > 16 and > 22 were associated with a > 10% and > 25% obesity prevalence, respectively (Supplementary Table 1: online only). Then, the total risk score was calculated within the validation cohort, and the obesity prevalence for each cutoff point (16 and 22)

Table 1. Participant characteristics

	n = 6,846	%
Experienced childhood obesity	581	8.5
Child's gender: Female	3,360	49.1
Mother's BMI (kg/m^2)		
< 18.5	1,138	16.6
18.5–24.9	4,873	71.2
25.0–29.9	500	7.3
≥ 30	135	2.0
Missing	200	2.9
Smoked during pregnancy		
Yes	501	7.3
No	5,784	84.5
Missing	561	8.2
Education		
High school or lower	2,744	40.1
Higher education	3,951	57.7
Missing	151	2.2
Birth weight (g)		
≤ 2,499	475	6.9
2,500–2,999	2,649	38.7
3,000–3,499	2,915	42.6
≥ 3,500	802	11.7
Missing	5	0.1
Birth order		
First-born	2,825	41.3
Younger (second-born or later)	3,580	52.3
Missing	441	6.4
Feeding type		
Formula feeding	1,808	26.4
Breast- or combination feeding	4,794	70.0
Missing	244	3.6
Obesity Index in 1-yr-old infants (%)		
< 10	5,016	73.3
10–19.9	736	10.8
≥ 20	99	1.4
Missing	995	14.5

BMI, body mass index.

was obtained. In both cohorts, predictive metrics of sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were determined, and the discrimination of raw scores was evaluated using the area under the receiver operating curve (AUROC).

All analyses were conducted using IBM SPSS Statistics 25.0 for Windows (SPSS Inc., Chicago, IL, USA).

Results

The characteristics of mother–child pairs are reported in Table 1. Overall, 8.5% of children were obese at 6–8 yr of age.

Results of the univariate logistic regression analyses in the derivation cohort are reported in Table 2. Formula feeding (versus breastfeeding or combination feeding) and first-born (versus

Table 2. Results of a univariate logistic regression analysis for childhood obesity in the derivation cohort (n = 5,488)

	OR	95% CI	P-value
Child's gender: Female (versus male)	1.32	1.09 to 1.60	0.004
Mother's BMI (kg/m^2)			
< 18.5	Ref		
18.5–24.9	3.49	2.25 to 5.41	< 0.001
25.0–29.9	11.48	7.08 to 18.61	< 0.001
≥ 30	16.90	9.34 to 30.57	< 0.001
Smoked during pregnancy (versus non-smoking)	1.47	1.06 to 2.03	0.020
Education high school or lower (versus ≥ higher education)	1.21	1.00 to 1.47	0.049
Birth weight (g)			
≤ 2,499	Ref		
2,500–2,999	1.25	0.99 to 1.58	0.331
3,000–3,499	1.65	1.31 to 2.07	0.027
≥ 3,500	2.17	1.35 to 3.48	0.002
Second-born or later (versus first-born)	1.13	0.92 to 1.37	0.245
Formula feeding (versus breast- or combination feeding)	1.14	0.92 to 1.40	0.236
Obesity index in 1-yr-old infants (%)			
< 10	Ref		
10–19.9	4.28	3.37 to 5.44	< 0.001
≥ 20	9.76	6.14 to 15.53	< 0.001

Missing values were replaced using multiple imputation (25 imputed datasets). CI, confidence interval; OR, odds ratio; Ref, reference.

second-born or younger child) had P-values ≥ 0.2; therefore, these two variables were excluded from the multivariate model.

In an initial multivariate logistic regression analysis in the derivation cohort, data regarding the birthweight categories were not linear (the odds ratio [OR] for category 3 was lower than that for category 2), and ORs for all birthweight categories were low (Supplementary Table 3: online only); therefore, this variable was excluded. Results of the final multivariate logistic regression analysis in the derivation cohort are reported in Table 3.

Childhood obesity risk of low-, moderate-, and high-risk categories in the derivation cohort are reported in Table 4. Scores ≤ 15 (low risk), 16–25 (moderate risk), and ≥ 26 (high risk) had childhood obesity risks of 4.9%, 18.4%, 36.3%, respectively (each 3-point childhood obesity risk is reported in Supplementary Table 1).

In the validation cohort, scores ≤ 15 (low risk), 16–25 (moderate risk), and ≥ 26 (high risk) had childhood obesity risks of 5.2%, 18.7%, 34.4%, respectively (Table 5) (Supplementary Table 2: online only).

Predictive metrics in both cohorts, including sensitivity, specificity, PPV, and NPV at a cutoff point ≥ 16, and AUROC of the risk scores are reported in Table 6. The sensitivity of the validation cohort was slightly higher; however, results were very similar between cohorts.

Discussion

This study aimed to construct a childhood obesity risk score using predictors identified in pregnant women and 1-yr-old infants. Using pre-pregnancy BMI (4 categories), child's gender (two categories), smoking during pregnancy (two categories), education (two categories), and

Table 3. Results of a multivariate logistic regression analysis for childhood obesity in the derivation cohort (n = 5,488)

	OR	95% CI lower limit		95% CI upper limit	P-value	β	Integer score
Child's gender							
Male	Ref						0
Female	1.33	1.08	to	1.63	0.006	0.284	4
Mother's BMI (kg/m^2)							
< 18.5	Ref						0
18.5–24.9	3.56	2.22	to	5.73	< 0.001	1.271	18
25.0–29.9	10.37	6.15	to	17.50	< 0.001	2.339	33
≥ 30	17.69	9.35	to	33.44	< 0.001	2.873	41
Smoked during pregnancy							
No	Ref						0
Yes	1.21	0.94	to	1.55	0.140	0.189	3
Education level							
High school or lower	Ref						0
Higher education	1.21	0.98	to	1.49	0.071	0.191	3
Obesity index (%)							
< 10	Ref						0
10–19.9	3.47	2.72	to	4.42	< 0.001	1.243	18
≥ 20	9.69	5.98	to	15.70	< 0.001	2.271	32

Missing values were replaced using multiple imputation (25 imputed datasets). BMI, body mass index; CI, confidence interval; OR, odds ratio; Ref, reference.

Table 4. Predicted risk score, risk category, and corresponding percentage of childhood obesity risk at 6–8 yr of age in the derivation cohort (n = 5,488)

Risk score	Childhood obesity (%)	Risk category
≤ 15	4.9%	Low risk
16–25	18.4%	Moderate risk
≥ 26	36.3%	High risk

Missing values were replaced using multiple imputation (25 imputed datasets).

Table 5. Predicted risk score, risk category, and corresponding percentage of childhood obesity risk at 6–8 yr of age in the validation cohort (n = 1,358)

Risk score	Childhood obesity (%)	Risk category
≤ 15	5.2%	Low risk
16–25	18.7%	Moderate risk
≥ 26	34.4%	High risk

Missing values were imputed by using multiple imputation (25 imputed datasets)

Table 6. Diagnostic measures for assessing the predictive capability of the risk score for childhood obesity at 6–8 yr of age in the derivation and validation cohorts

Diagnostic measure	Derivation cohort	Validation cohort
Sample size (n)	5,488	1,358
AUROC	71.3	68.8
Sensitivity [a]	54.2	49.7
Specificity [a]	82.2	83.8
PPV [a, b]	22.2	21.8
NPV [a, b]	95.1	94.8

Missing values were replaced using multiple imputation (25 imputed datasets). [a] Based on risk score threshold ≥ 16. [b] Based on an 8.5% prevalence of obesity in the derivation cohort, and 8.3% in the validation cohort. AUROC, area under the receiver operating curve; NPV, negative predictive value; PPV, positive predictive value.

obesity index at one year of age (three categories), a risk score algorithm was constructed to predict childhood obesity with a 22% PPV at a cutoff point of ≥ 16. To our knowledge, there have been no previously reported childhood obesity risk scores in Japan.

As previously mentioned, a systematic review of childhood overweight and obesity risk scores using predictors identified during pregnancy was reported from studies conducted in North America and Europe (4). Although one study evaluated genetic predictors (13), most studies used generic anthropometric measures, socioeconomic indicators, smoking status during pregnancy, and feeding type. In this study, generic predictors obtained during a standard health check for 1.5-yr-old infants were included. Some studies have directly used regression equations as predictor algorithms (14–16). In comparison, here, the β coefficients were converted into integers to facilitate the use of the risk score among mothers and public health nurses, according to a previously reported method (12).

In Japan, the overweight and obesity prevalence is lower than that reported in North America and Europe (5). In addition, adult obesity is defined as having a BMI higher than 25, rather than overweight, and there are no criteria for being classified as overweight in childhood; classification begins with mild obesity (11). In the results of the study this methodology is based on (12), 23.4% of participants were overweight at three years of age, according to the International Obesity Task Force (17). In contrast, in the present study, only 8.5% of children were obese at 6–8 yr of age. Because PPV is influenced by the prevalence of an outcome, the PPV in this study was only 22% compared to 37% in the previous study. In addition, because noninvasive preventive measures for childhood obesity are required, a higher false-positive rate is permissible.

The sensitivity and specificity of the derivation and validation cohorts were 54.9% and 49.7%, and 82.2 and 83.8, respectively, although they had different cutoff points. The typical AUROC range for a prediction model is 0.6–0.85, with a result between 0.70–0.80 denoting adequate discriminatory power and 0.80–0.90 considered to be excellent (18). In this study, the AUROCs of the derivation and validation cohorts were 71.3% and 68.8%, respectively. Because the β coefficients were changed to integer values for easier use, the AUROC was lower than the c-statistic (0.74) of the final multivariate logistic regression model.

Cases with a high-risk score, particularly those ≥ 16, must receive public health interventions to prevent obesity. A comprehensive approach that considers diet, physical activity, psychological health, and environmental factors are required to prevent childhood obesity (2). This can be implemented through consultation with a nurse and nutritionist during standard health checks for 1.5-yr-old infants, along with local government initiatives to provide an environment that promotes healthy diet and exercise in nurseries, kindergartens, and schools.

This study had several limitations. First, because this study aimed to develop a childhood obesity risk score algorithm, the cause and effect of each predictor variable were not considered. Instead, predictor variables were selected on the basis of their practical use in maternal and child health. Second, because of a large volume of missing obesity data for children aged 6–8 yr, participants included in the analyses do not reflect the entire cohort population. The excluded mother-child pairs differed in terms of their smoking behaviors, education levels, and birth weights (Supplementary Table 4: online only), and the childhood obesity prevalence was lower among the included mother-child pairs. This suggests the PPV may be slightly higher in the general population. Third, birth weight was measured in the hospital, whereas other measures reflecting the height and weight of mothers and children were self-reported. However, mothers were asked to refer to a standard health check or school-based physical measurement when reporting the height and weight of their child. Fourth, small for their gestational age (SGA) is related to later obesity (19). However, instead of selecting SGA which needs complicated calculation to be used as a predictor variable, we selected birth weight, which is an easily comprehensible predictor and can be obtained from maternal and child health handbooks.

The strength of this study was the use of prospective population-based birth cohort data, including the collection of predictor variables prior to determining the outcome, which reduces the risk of recall bias.

Conclusions

A childhood obesity risk score algorithm was developed based on a Japanese population-based birth cohort of mother and child pairs. Predictors, including factors for both pregnant women and 1-year-old infants, were pre-pregnancy BMI (four categories), child's gender (two categories), smoking during pregnancy (two categories), education level (two categories), and obesity index at one year of age (three categories). A 22% PPV was achieved at a cutoff point ≥ 16. Further studies are required to validate and improve this algorithm.

Conflict of interests: The authors declare that they have no conflict of interest.

Acknowledgments

This research was supported by AMED under Grant Number 18gk0110032h001.

References

1. World Health Organisation. Obesity and overweight. WHO. 2018 (http://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight).
2. Lee EY, Yoon KH. Epidemic obesity in children and adolescents: risk factors and prevention. Front Med 2018;12: 658–66. [Medline] [CrossRef]
3. Pandita A, Sharma D, Pandita D, Pawar S, Tariq M, Kaul A. Childhood obesity: prevention is better than cure. Diabetes Metab Syndr Obes 2016;9: 83–9. [Medline] [CrossRef]
4. Canfell OJ, Littlewood R, Wright OR, Walker JL. Clinical relevance and validity of tools to predict infant, childhood and adulthood obesity: a systematic review. Public Health Nutr 2018;21: 3135–47. [Medline] [CrossRef]
5. Afshin A, Forouzanfar MH, Reitsma MB, Sur P, Estep K, Lee A, *et al*. GBD 2015 Obesity

Collaborators Health effects of overweight and obesity in 195 countries over 25 years. N Engl J Med 2017;377: 13–27. [Medline] [CrossRef]
6. Kishi R, Araki A, Minatoya M, Hanaoka T, Miyashita C, Itoh S, *et al*. Members of The Hokkaido Study on Environment and Children's Health. The Hokkaido birth cohort study on environment and children's health: cohort profile-updated 2017. Environ Health Prev Med 2017;22: 46. [Medline] [CrossRef]
7. Kishi R, Sasaki S, Yoshioka E, Yuasa M, Sata F, Saijo Y, *et al*. Hokkaido Study on Environment and Children's Health. Cohort profile: the Hokkaido study on environment and children's health in Japan. Int J Epidemiol 2011;40: 611–8. [Medline] [CrossRef]
8. Kishi R, Kobayashi S, Ikeno T, Araki A, Miyashita C, Itoh S, *et al*. Members of the Hokkaido Study on Environment and Children's Health. Ten years of progress in the Hokkaido birth cohort study on environment and children's health: cohort profile--updated 2013. Environ Health Prev Med 2013;18: 429–50. [Medline] [CrossRef]
9. Ito K, Hanaoka T, Tamura N, Sasaki S, Miyashita C, Araki A, *et al*. Association between maternal serum folate concentrations in the first trimester and the risk of birth defects: the Hokkaido study of environment and children's health. J Epidemiol 2019;29: 164–71. [Medline] [CrossRef]
10. Ito Y, Fujieda K, Okuno A. Weight-for-height charts for Japanese children based on the year 2000 Report of School Health Statistics Research. Clin Pediatr Endocrinol 2016;25: 77–82. [Medline] [CrossRef]
11. Kubo T. Common approach to childhood obesity in Japan. J Pediatr Endocrinol Metab 2014;27: 581–92. [Medline] [CrossRef]
12. Weng SF, Redsell SA, Nathan D, Swift JA, Yang M, Glazebrook C. Estimating overweight risk in childhood from predictors during infancy. Pediatrics 2013;132: e414–21. [Medline] [CrossRef]
13. Seyednasrollah F, Mäkelä J, Pitkänen N, Juonala M, Hutri-Kähönen N, Lehtimäki T, *et al*. Prediction of adulthood obesity using genetic and childhood clinical risk factors in the cardiovascular risk in young Finns study. Circ Cardiovasc Genet 2017;10: e001554. [Medline] [CrossRef]
14. Steur M, Smit HA, Schipper CM, Scholtens S, Kerkhof M, de Jongste JC, *et al*. Predicting the risk of newborn children to become overweight later in childhood: the PIAMA birth cohort study. Int J Pediatr Obes 2011;6: e170–8. [Medline] [CrossRef]
15. Santorelli G, Petherick ES, Wright J, Wilson B, Samiei H, Cameron N, *et al*. Developing prediction equations and a mobile phone application to identify infants at risk of obesity. PLoS One 2013;8: e71183. [Medline] [CrossRef]
16. Robson JO, Verstraete SG, Shiboski S, Heyman MB, Wojcicki JM. A risk score for childhood obesity in an urban Latino cohort. J Pediatr 2016;172: 29–34.e1. [Medline] [CrossRef]
17. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. BMJ 2000;320: 1240–3. [Medline] [CrossRef]
18. Lee YH, Bang H, Kim DJ. How to establish clinical prediction models. Endocrinol Metab (Seoul) 2016;31: 38–44. [Medline] [CrossRef]
19. Hong YH, Chung S. Small for gestational age and obesity related comorbidities. Ann Pediatr Endocrinol Metab 2018;23: 4–8. [Medline] [CrossRef]