Weill Cornell Medicine

NewYork-Presbyterian

Identifying risk factors associated with low birth weight

Weill Cornell Medicine NewYork-Presbyterian

Background

What is low birth weight?

Those who were born with the weight under 2500 grams are called low birth weight children.

What is the negative impacts of low birth weight?

- Intelligence and social adaptability
- The mortality rate is higher

Objective

To identify potential risk factors that may have an association with low birth weight.

Weill Cornell Medicine NewYork-Presbyterian

Variables

Variable	Meaning
Age	Age of mother
Lwt	Weight at the last menstrual period
Race	A factor for race: 1(white),2 (black) and 3 (other)
Smk	Indicator of smoking status:1=smoker, 0=non-smoker
Ptl	history of premature labor: 0=none, 1=one, etc
Ht	History of hypertension: 1=yes, 0=no
UI	Presence of uterine irritability: 1=yes, 0=no
Ftv	Number of physician visits during the first trimester 0=None, 1=one, 2=two, etc

Analysis Plan

1. Demographic analysis
2. Univariate analysis in logistic regression
3. Multivariate analysis in logistic regression
4. Model checking
5. Conclusion
6. Discussion

Weill Cornell Medicine NewYork-Presbyterian

Demographic Analysis

The descriptive analysis to see whether the distribution of the variables are different in two groups.

Continuous variable: mean and standard deviation

Normal: t-test

Non-normal: Wilcoxon rank-sum test

Categorical variable:

Contingency table and Chi-square test

Table 1. demographic analysis

Effect	Low birth weight (1)	Normal birth weight (0)	P-value
Count	59	130	
Age (mean(sd))	23.66 (5.58)	22.31 (4.51)	0.247
Weight (mean(sd))	122.14 (26.56)	133.3 (31.72)	0.013
Race			0.08
White (1)	23	73	
Black (2)	11	15	
Other (3)	25	42	
Smoking Status			0.03
Smoker (1)	30	44	
Non-smoker (0)	29	86	
Premature labor			0.0002
Yes (1)	18	12	
No (0)	41	118	
Hypertension			0.036
Yes (1)	7	5	
No (0)	52	125	
Uterine irritability			0.02
Yes (1)	14	14	
No (0)	45	116	
Physician visits			0.133
Yes (1)	23	66	
No (0)	36	64	

Weill Cornell Medicine NewYork-Presbyterian

Univariate analysis in logistic regression

1. Regroup the history of premature labor and the times of the physician visits at the first three month of pregnancy

2. Apply Deviance, Wald test and AIC to preliminarily select the variables

3. The odds ratio and the confidence interval can give us a peek on furthering multivariate analysis

Univariate Analysis
Model selection for 3 models

Table 2. univariate analysis

Effect	Odds ratio	Confidence Interval	p value	Change in Deviance	AIC
age	0.95	0.893-1.011	0.1047	2.76	235.912
weight	0.986	0.974-0.998	0.0227	5.981	232.691
Race2 vs 1	2.327	0.938-5.772	0.0854	5.01	235.662
Race3 vs 1	1.889	0.955-3.735			
smk	2.022	1.081-3.783	0.0276	4.867	233.805
ptl	4.317	1.916-9.726	0.0004	12.774	225.898
ht	3.365	1.021-11.088	0.0461	4.022	234.65
UI	2.578	1.139-5.834	0.0231	5.076	233.596
ftv	0.62	0.331-1.159	0.1339	2.587	236.394

Multivariate analysis in logistic regression

Selection criteria:

- Deviance
- Wald Chi-square test
- AIC

Deviance:

1. Compare the deviance of one-variable logistic model with null model and find variables that significantly reduce the deviance. Set the threshold for the significance level in deviance to be 2.07 χ^2 *(1,0.15).*

2. Compute the increase in the deviance when each variable on its own is omitted from the set, and retain the variables whose omission induces significant increase

3. Add back non-important variables to check whether to lower the deviance.

4. Interaction: Brought the lowest in deviance will be retained

Multivariate Analysis
Model selection: Deviance

Table 4. main factor model using deviance selection

Parameter	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	joint test Chi-Square	Pr > ChiSq
Intercept	0.6688	0.3954	2.8612	0.0907		
race 2	0.3656	0.3076	1.4131	0.2345	7.2994	0.026
race 3	0.3292	0.2631	1.566	0.2108		
smk	0.4653	0.1959	5.644	0.0175	5.644	0.0175
ptl	0.5952	0.221	7.257	0.0071	7.257	0.0071
ht	0.685	0.3214	4.5428	0.0331	4.5428	0.0331
UI	0.4197	0.227	3.4173	0.0645	3.4173	0.0645

Those variables whose omission induces significant increase in the deviance

Interaction between variables(Deviance)

Table 6. multivariate analysis using deviance selection

Effect	Coefficient	Odds Ratio	Confidence Interval	P-value
Ptl	0.6344	3.556	1.448-8.736	0.0057
Ht	0.7098	4.135	1.091-15.68	0.0368
UI	0.5276	2.872	1.122-7.354	0.0278
Race 1 & Smk 1	0.2721	17.849	2.189-145.546	0.4521
Race 2 & Smk 0	0.0697	14.578	1.472-144.356	0.8955
Race 2 & Smk 1	1.5194	62.128	5.573-692.658	0.0132
Race 3 & Smk 0	0.4317	20.938	2.581-169.882	0.2272
Race 3 & Smk 1	0.3169	18.665	1.744-199.778	0.5913

Wald Chi-Square test

1. Backward Selection

2. Set the threshold: $p=0.15$

3. In each step, only the variable not significant with largest p-value will be dropped.

4. The significant interaction term will be retained in the model for further study

Wald Chi-Square test method

Table 7. multivariate analysis using wald test selection

Effect	Coefficient	Odds Ratio	Confidence Interval	P-value
lwt	-0.0189	0.981	0.967-0.996	0.0112
smk	0.4672	2.546	1.130-5.737	0.0242
ptl	0.6039	3.346	1.367-8.191	0.0082
ht	0.9954	7.321	1.775-30.191	0.0059
race=1 & ui=1	-1.0616	1.137	0.287-4.058	0.0860
race=2 & ui=0	-0.1707	2.773	0.922-8.346	0.7213
race=2 & ui=1	2.4970	39.931	2.668-597.604	0.0282
race=3 & ui=0	-0.4531	2.090	0.797-5.481	0.2649
race=3 & ui=1	0.3781	4.798	1.161-19.837	0.5255

Akaike information criterion(AIC)

1. Backward selection

2. Variables with the largest AIC will be dropped in each step

AIC

The history of premature labor, the presence of uterine irritability, and the weight of the mother are identified to have a significant association with low birth weight

$$Logit(\frac{P}{1-P} = \beta_0 + \beta_1 lwt + \beta_2 ptl + \beta_3 ui + beta_4 age + \beta_5 ht + \beta_6 age \times ht$$

Table 8. multivariate analysis using AIC selection

Effect	Coefficient	Odds Ratio	Confidence Interval	P-value
lwt	-0.0147	0.985	0.972-0.999	0.0336
ptl	0.7141	4.172	1.710-10.179	0.0017
ui	0.3423	1.983	0.812-4.844	0.1331

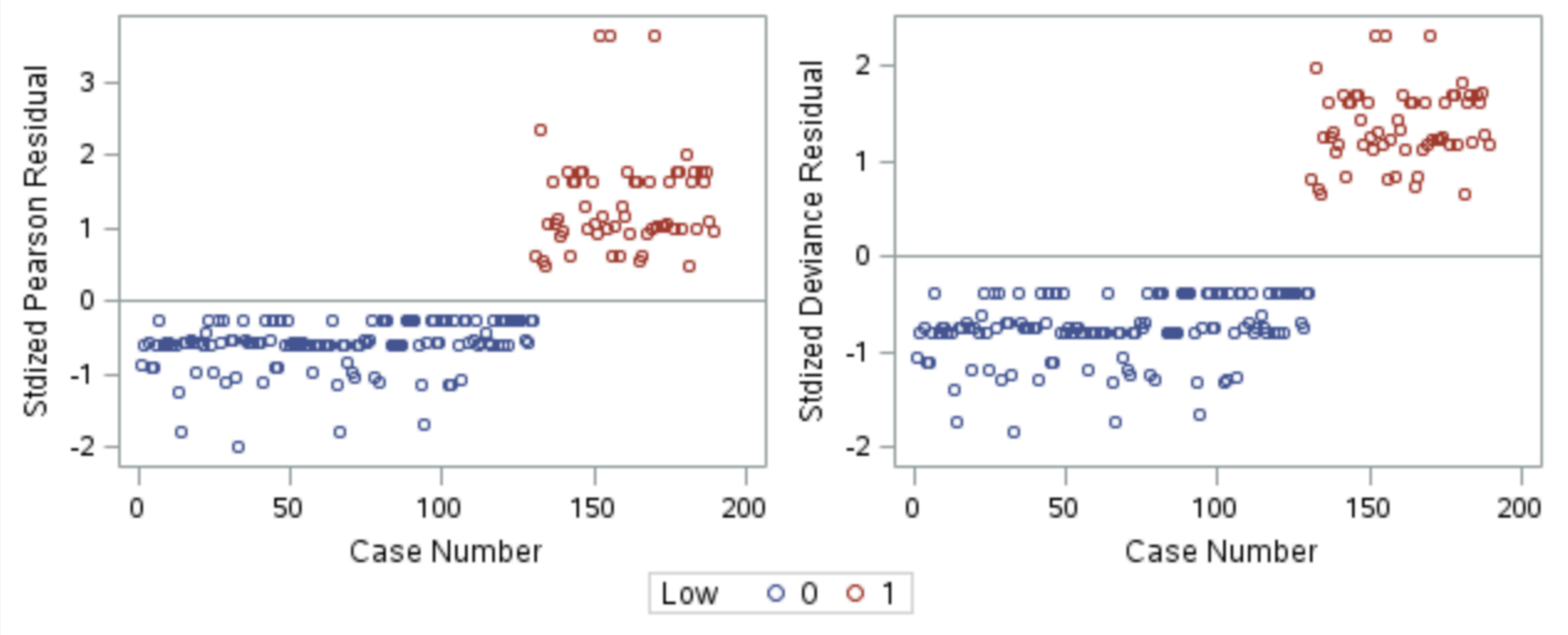
Model checking

To evaluate the goodness-of-fit of the model and to identify the outliers

- Hosmer-Lemeshow goodness of fit test: unreplicated data for each pattern of covariate
  - The observation whose delta deviance larger than 4 is identified as outlier and should be deleted. χ^2 (1,0.05)
- Deviance goodness of fit test: replicated data for each pattern of covariate
  - Observation whose standardized deviance residuals larger than 2 and smaller than -2 should be deleted

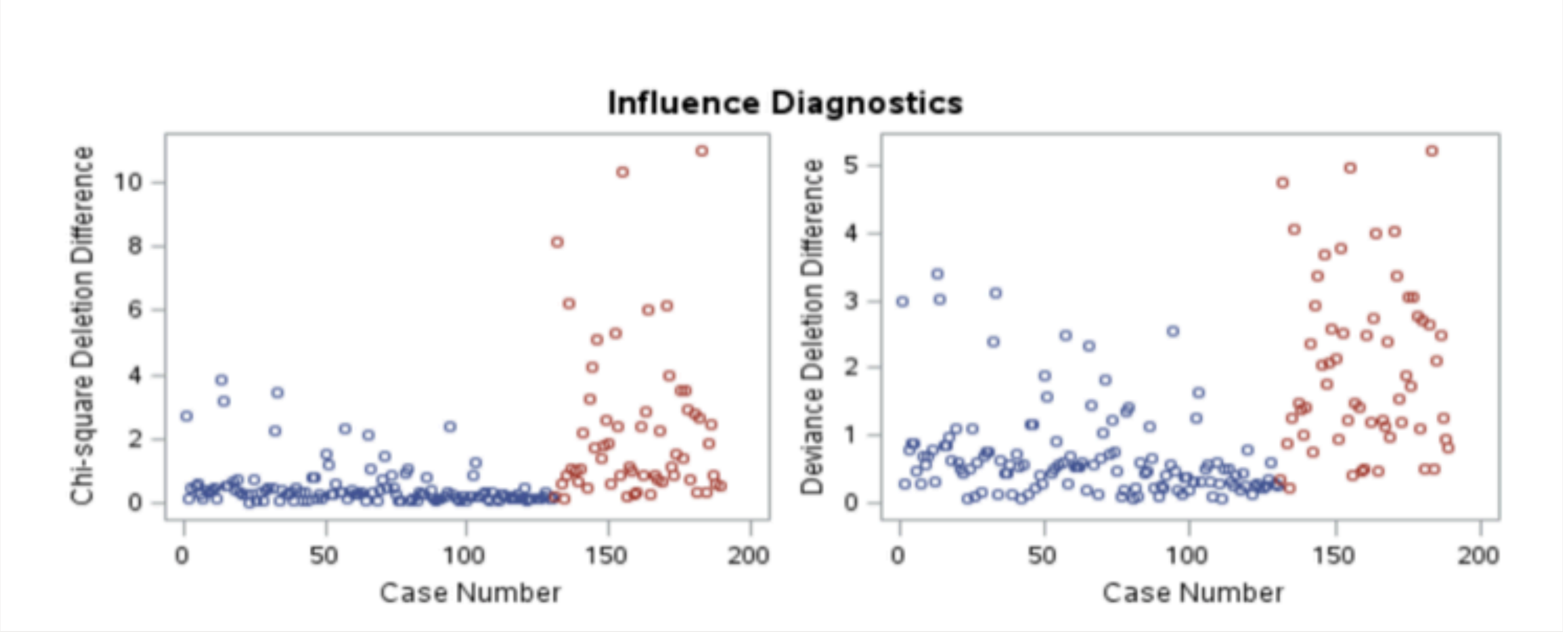
Model checking for deviance

- Replicated data
- The p-value of the Deviance test is 0.56
- Does not lack goodness-of-fit.
- In Figure 1, the observations whose deviance standardized residuals lager than 2 or less than -2 are deleted.



Model checking for Wald selection

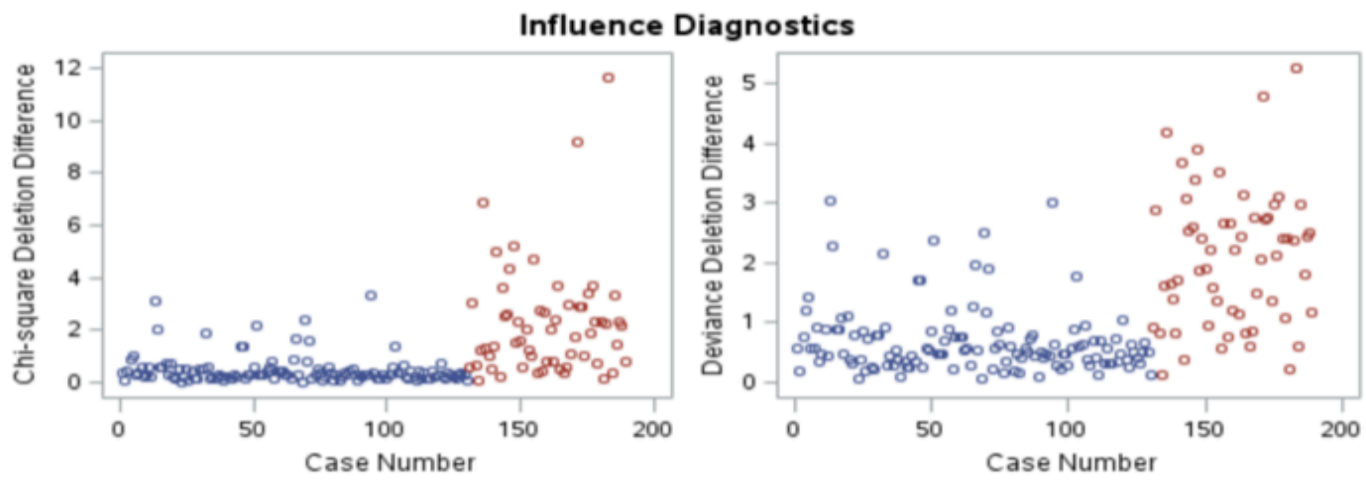
- Unreplicated data
- The p-value of the Hosmer-Lemeshow test is 0.6107
- Does not lack goodness-of-fit.
- The delta deviance of cases is shown in Figure 2, and the observations whose value lager than 4 are deleted.



Model checking for AIC

- Unreplicated data
- The p-value of the Hosmer-Lemeshow test is 0.46
- Does not lack goodness-of-fit.
- The delta deviance of cases is shown in Figure 2, and the observations whose value lager than 4 are deleted.

Figure 3. delta deviance of cases for model selected by AIC



Conclusion

- Discuss the risk factors of low birth weight

- the history of premature labor, the history of hypotension, the presence of uterine irritability are risk factors that increase the odds of delivering a low birth weight baby in each of the model.

- Deviance works best, because it covers the fewest variables and all of them are significant with a smallest deviance

Discussion

1. More data and details are needed
2. Lack of considerations of other factors

- socioeconomic disadvantages
- different educational levels of women

3. More articles and more statistical models are needed to select the best model

Reference

Dennis.J &Mollborn.S (2014) Young maternal age and low birth weight risk: An exploration
Of racial/ethnic disparities in the birth outcomes of mothers in the United States. *NIH Public Access Author Manuscript 50(4)*. 625-634

Ediriweera.DS &Dilina.N & Perera.U& Flores.F &Samita.S (2017) Risk of low birth weight on adulthood hypertension-evidence from a tertiary care hospital in a South Asian country, SriLanka: a retrospective cohort study BMC Public Health 17. 358

Kleinman.JC&Kessel.SS. (1987) Racial differences in low birth weight.Trends and risk factors.*The New England Journal of Medicine. 317(12)*.749-753

Kleinman JC & Madans JH(1985) The effects of maternal smoking, physical stature, and educational attainment on the incidence of low birth weight. *American Journal of Epidemiology 121(6)* 843-855

Kramer (1987). Reviews Analyses. *Bulletin of the World Health Organization.65(5)*.663-737

Meis PJ,Ernest JM& Moore ML(1987) Causes of low birth weight births in public and private patients. *American Journal of obstetrics and gynecology 156(5)* 1165-1168

Su.D(2013) The study of risk factors on low birth weight. *China Medical Herald.28.60-62*