A Study on the Effect of Maternal Risk Factors on Low Birth Weight

Johan Carlo A. Ilagan

18 October 2024

Executive Summary

This study investigates the relationship between low birth weight in newborns and maternal characteristics using logistic regression. The dataset contains 187 observations and includes eight explanatory variables - maternal age, weight, race, smoking status, history of previous premature labours, hypertension, uterine irritability, and the number of physician visits during pregnancy.

Descriptive statistics were used for initial data exploration. Three logistic regression models were developed: 1. A full model with all predictors 2. A model selected using the `stepAIC()` function 3. A manually reduced model excluding statistically insignificant variables

Model comparison using ANOVA indicated the third model as the most effective. The final model showed that: - Maternal age had a negative association with the log-odds of low birth weight. - Race, smoking status, and history of hypertension were positively associated with increased risk of low birth weight.

These findings may provide insight into maternal risk factors associated with adverse birth outcomes.

1. Introduction

Early identification of low birth weight is crucial for providing timely medical intervention, as it significantly affects both maternal health and the newborn's development. Low birth weight is commonly associated with short-term complications such as respiratory distress and neurological issues, as well as long-term consequences like vision and hearing impairments and stunted growth (Stanford Medicine Children's Health, n.d.). The World Health Organization (WHO) defines low birth weight as a birth weight of less than 2.5 kilograms (WHO, n.d.).

Over the years, numerous studies have investigated potential maternal and environmental risk factors for low birth weight. Anil et al (2020) and Diabelková et al (2022) highlighted the role of preterm birth, while the Australian Institute of Health and Welfare (AIHW, n.d.) identified maternal age and drug use as significant contributing factors. Wubutu et al (2021) employed descriptive statistics and multiple linear regression to explore correlations with maternal age and weight. Arabzadeh et al (2024) used the meta-umbrella R package to focus specifically on substance use. Meanwhile, Adugna and Worku (2022) demonstrated that hypertension, body mass index (BMI), and preterm delivery were significant predictors.

Multivariable logistic regression is a commonly applied method in such studies. It allows for the modelling of a binary outcome - in this case, low versus normal birth weight - based on multiple predictors (Kalan et al, 2021). The general form of logistic regression model is given by:

$$log(\frac{\pi_i}{1-\pi_i}) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n$$

In this analysis, we aim to explore the relationship between low birth weight and various maternal characteristics using a logistic regression framework. The report outlines the statistical methodology, presents the results of model fitting and selection, and concludes with a discussion of the findings and their implications.

2. Methodology

The dependent variable in this study is binary, coded as 1 if the birth weight is less than 2.5 kg (low birth weight) and 0 otherwise. Given this binary outcome, logistic regression for binomial data was used to model the relationship between low birth weight and various maternal factors.

The analysis began with descriptive statistics, including the minimum, maximum, mean, median, and standard deviation for each variable. Exploratory data analysis was conducted through visualizations to assess potential patterns or associations between low birth weight and each independent variable.

An initial logistic regression model was fitted using all available maternal variables. No interaction terms were included, as there was no prior evidence to suggest interaction effects between the predictors. Following this, model selection was performed to identify statistically significant predictors. Variables with p-values greater than the significance threshold ($\alpha = 0.05$) were iteratively removed to arrive at the final reduced model.

The final model was used to estimate the log-odds of low birth weight based on the retained predictors. All analyses and visualizations were conducted in the R statistical environment.

3. Results

For descriptive analysis, categorical variables were summarized using frequency counts, while continuous variables were described using the minimum, median, mean, maximum, and standard deviation. The `summary()` and `sd()` functions were used to compute the following statistics:

'low': 'No'= 129, 'Yes' = 58

'age': Min = 14.0; Median = 22.0; Mean = 23.1; Max = 36.0; SD = 5.07

'lwt': Min = 80.0; Median = 121.0; Mean = 129.9; Max = 250.0; SD = 30.73

'race': 'White' = 95, 'Black' = 26, 'Other' = 66

'smoke': 'No' = 114, 'Yes' = 73

'ptl': Min = 0.0; Median = 0.0; Mean = 0.1925; Max = 3.0; SD = 0.49

'ht': 'No' = 175, 'Yes' = 12

'ui': 'No' = 160, 'Yes' = 27

'ftv': Min = 0.0; Median = 0.0; Mean = 0.7968; Max = 6.0; SD = 1.06

'bwt': Min = 1021; Median = 2977; Mean = 2946; Max = 4593; SD = 698.65

From the descriptive statistics, it is evident that the categorical variables show imbalanced distributions - particularly `low`, where 129 cases are classified as "No" and only 58 as "Yes". Additionally, the large standard deviation for `bwt` suggests a wide spread in birth weight values. Notably, both `ptl` and `ftv` have medians of 0, but their means are non-zero, indicating the presence of outliers or skewed distributions.

To explore potential associations, boxplots were generated for the continuous predictors (`age`, `lwt`, `ptl`, `ftv`) against low birth weight status, while mosaic plots were used for categorical variables (`race`, `smoke`, `ht`, `ui`). The boxplots did not show strong visual separation, suggesting the need for further analysis. However, the mosaic plots suggested possible relationships between low birth weight and `smoke`, `ht`, and `ui` (see Figure 1).

An initial logistic regression model (`M1`) was fitted using all predictors without interaction terms. The model was then refined using the `stepAIC()` function from the MASS package in a backward elimination procedure to form `M2`. Based on p-values, a further manual reduction was performed to remove statistically insignificant variables (`ptl`, `ui`), resulting in the final model `M3`.

To determine the best-fitting model, an ANOVA test was used to compare `M1`, `M2`, and `M3`. Based on the test statistics, `M3` showed a significantly better fit and was therefore selected as the final model. The resulting logistic regression model for the log-odds of low birth weight is:

$$logit(\widehat{\pi}) = -0.018(lwt) + 1.26(race2) + 0.863(race3) + smoke1 + 1.757(ht1)$$

To assess model adequacy, a residuals vs. fitted values plot was generated (see Figure 2). The plot shows two horizontal bands, one corresponding to cases with low birth weight (1) and one without (0), with no evident violations of model assumptions.

A deviance goodness-of-fit test was also conducted. The p-value obtained was 0.00014, indicating the the model fit is significantly better than the null model, and confirming `M3` as the preferred model.

4. Discussion

In the final model (M3), the significant predictors of low birth weight were maternal weight (`lwt`), race, smoking status (`smoke`), and history of hypertension (`ht`). These variables were retained after stepwise model selection and manual reduction based on statistical significance ($\alpha = 0.05$).

The final logistic regression equation is:

$$logit(\widehat{\pi}) = -0.018(\text{lwt}) + 1.26(\text{race2}) + 0.863(\text{race3}) + \text{smoke1} + 1.757(\text{ht1})$$

Th effects of each predictor on the log-odds of low birth weight are summarized below: 1. **Maternal weight(`lwt`)**: The negative coefficient (-0.018) suggests that as the maternal weight increases, the likelihood of low birth weight decreases. Specifically for every one-unit increase in weight, the log-odds of low birth weight decreases by 0.018 units.

2. **Race**: Compared to the baseline group (White), mothers identified as Black (`race2`) have an increased log-odds of 1.26 for delivering a low birth weight baby, and those classified as Other (`race3`) have an increase of 0.863. This indicates that race is a meaningful predictor in the model.
3. **Smoking Status (`smoke1`)**: A coefficient of approximately 1 implies that smoking during pregnancy substantially increases the likelihood of low birth weight. The odds ratio is $e^1 \approx 2.72$, meaning the smoking mothers are about 2.7 times more likely to deliver a low birth weight baby compared to non-smokers.
4. **History of Hypertension (`ht1`)**: The coefficient of 1.757 translates to an odds ratio of $e^{1.757} \approx 5.79$. This suggests that mothers with a history of hypertension are nearly 5.8 times more likely to deliver a low birth weight baby.

These findings align with previous studies in the literature, reinforcing known risk factors such as smoking, hypertension, and racial disparities in maternal health. By identifying statistically significant predictors, this model contributes to a growing body of research aimed at reducing the incidence of low birth weight. These insights can inform clinical risk assessments and public health interventions targeting vulnerable populations.

Appendix

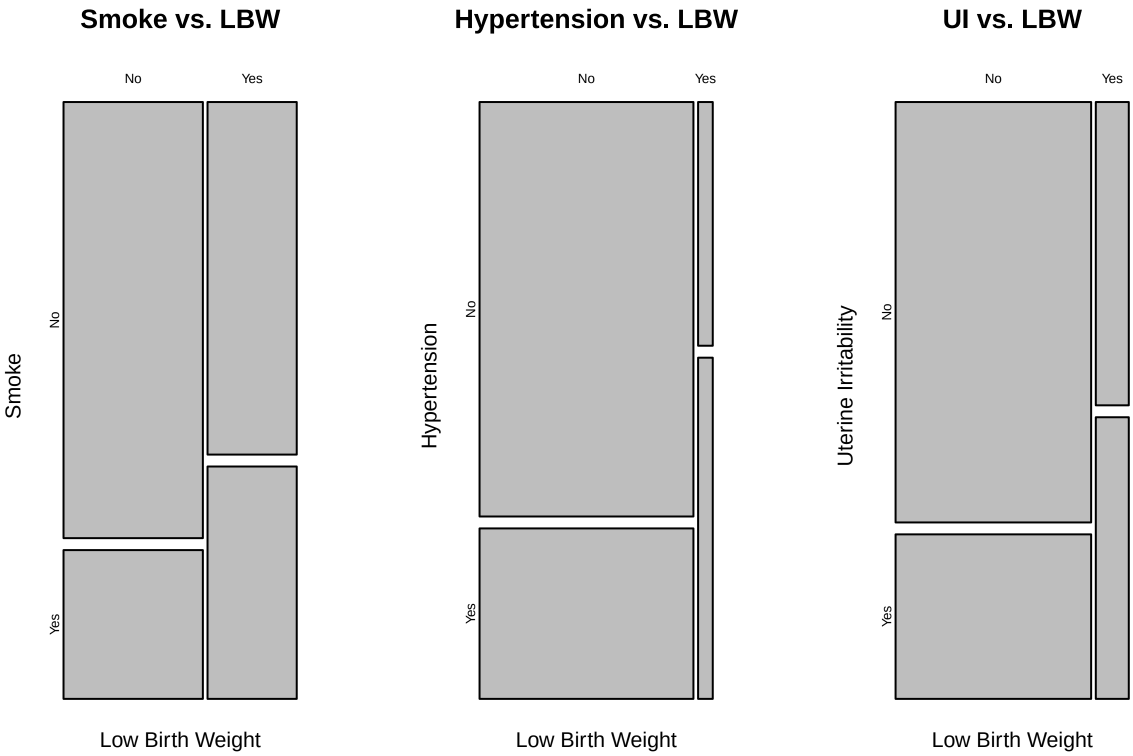

Figure 1. Mosaic Plots of `smoke`, `ht`, and `ui` against `low`

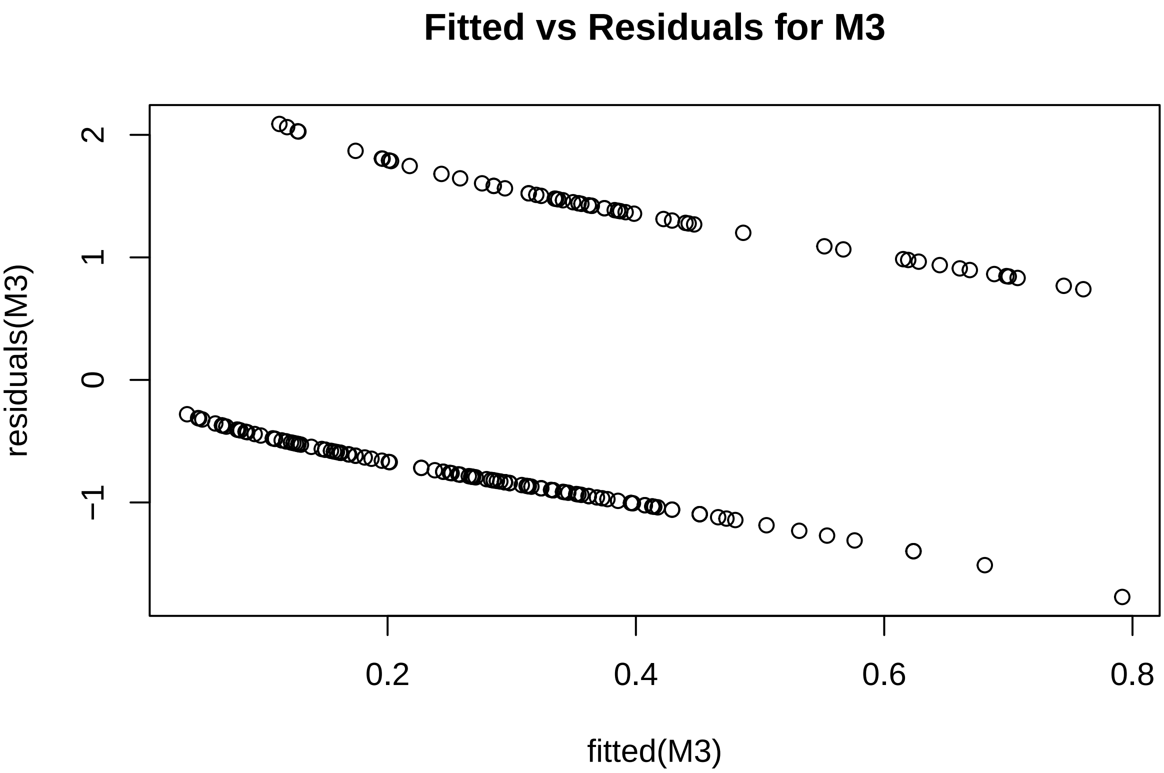

Figure 2. Fitted Values vs Residuals for M3

References

Adugna, D. G., & Worku, M. G. (2022). Maternal and neonatal factors associated with low birth weight among neonates delivered at the University of Gondar comprehensive specialized hospital, Northwest Ethiopia. Frontiers in Pediatrics, 10. https://doi.org/10.3389/fped.2022.899922

Arabzadeh, H., Doosti-Irani, A., Kamkari, S., Farhadian, M., Elyasi, E., & Mohammadi, Y. (2024). The maternal factors associated with infant low birth weight: an umbrella review. BMC Pregnancy and Childbirth, 24(1). https://doi.org/10.1186/s12884-024-06487-y

Australian Institute of Health and Welfare. (2020). Australia's children, Birthweight. Australian Institute of Health and Welfare. https://www.aihw.gov.au/reports/children-youth/australias-children/contents/health/birthweight

Diabelková, J., Rimárová, K., Urdzík, P., Dorko, E., Houžvičková, A., Andraščíková, Š., Drabiščák, E., & Škrečková, G. (2022). Risk factors associated with low birth weight. Central European Journal of Public Health, 30, S43–S49. https://doi.org/10.21101/cejph.a6883

K. C., A., Basel, P. L., & Singh, S. (2020). Low birth weight and its associated risk factors: Health facility-based case-control study. PLoS ONE, 15(6), e0234907. https://doi.org/10.1371/journal.pone.0234907

Kalan, M. E., Jebai, R., Zarafshan, E., & Bursac, Z. (2020). Distinction between two statistical terms: multivariable and multivariate logistic regression. Nicotine & Tobacco Research, 23(8). https://doi.org/10.1093/ntr/ntaa055

Stanford Medicine Children's Health. (n.d.). Low Birth Weight. Stanford Medicine Children's Health. https://www.stanfordchildrens.org/en/topic/default?id=low-birth-weight-90-P02382

World Health Organization. (2023). Low birth weight. Www.who.int. https://www.who.int/data/nutrition/nlis/info/low-birth-weight

Wubetu, A. D., Amare, Y. E., Haile, A. B., & Degu, M. W. (2021). Newborn Birth Weight and Associated Factors Among Mother-Neonate Pairs in Public Hospitals, North Wollo, Ethiopia. Pediatric Health, Medicine and Therapeutics, 12, 111–118. https://doi.org/10.2147/phmt.s299202