

# Modelling weight of a Newborn based on baby's characteristics for low birth weight Babies

**M. Abdollahian<sup>a</sup>, S. Nuryani<sup>b</sup> and D. Anggraini<sup>c</sup>**

*School of Mathematical and Geospatial Sciences RMIT University, Melbourne Australia<sup>a</sup>*

*Ulin Hospital (RSUD Ulin) Banjarmasin Indonesia<sup>b</sup>*

*Department of Mathematics Lambung Mangkurat University, Banjararu, Indonesia<sup>c</sup>*

*Email: [mali.abdollahian@rmit.edu.au](mailto:mali.abdollahian@rmit.edu.au)*

**Abstract:** Neonatal mortality rate (NMR) is an increasingly important public health issues in many developing countries. Neonatal death now accounts for about two-thirds of the eight million infant deaths that occur globally each year. It is well-documented that low birth weight (LBW) is the most significant factor influencing NMR. This paper deploys multi-regression models to identify the significant factors for forecasting the weight of the LBW babies. The model explores the relationship between weights, baby's characteristics, gestation age and mother pre-pregnancy BMI. The results indicate that 65.9% of the variations in the weight of the LBW babies can be explained by baby's characteristics such as the height, head and chest circumferences, the gestation age and mother's BMI. The proposed model was then used to estimate the recorded weights together with their corresponding 95% confidence and predication interval. Analysis of the prediction errors shows that the mean prediction error for the recorded data is one gram. The research is based on a case study in Indonesia intended to improve the mortality rate.

**Keywords:** *Multivariate simulation, multivariate regression, neonatal mortality rate*

## 1. INTRODUCTION

Low birth weight is a major public health problem in the world, contributing substantially both to infant mortality and to childhood handicap. This study was conducted to provide community-based data on maternal health and predictors of the weight for low birth weight newborn in Indonesia. The study was undertaken among pregnant women attending the Banjarmasin Clinic in Indonesia during 2010-2011. A total of 199 pregnant women between the ages of 16 and 42, who attended the clinic for antenatal care or routine follow up were included in the study. For each patient; age, body mass index BMI, gestation age at the time of delivery, hemoglobin level before (in the third semester of pregnancy) and after (towards the end after consumption of Vitamin C and Sulfas Ferroses), systolic and diastolic blood pressure, baby's weight, height, chest and head circumference were measured. Out of the 199 deliveries 10 had a low birth weight. Low birth weight is defined as a birth weight of less than 2500 grams and is a well-documented risk factor for neonatal mortality (Allen et al (2003), Chang et al (1992), Mathews et al (2002), Chen et al (2010), Copper et al (1995), Maddah (2005)). It is also well documented that, reduced-size-at-birth infants, who include low birth weight infants, are at greater risk of having reduced educational capacity, school performance, and intellectual development than are infants of normal birth weight (Bennett et al (1999), Lagerstorm et al (1991), Rich-Edwards et al (1997, 1999), Rondo et al (2003), Thorsdottir et al (1998)). Many authors have investigated the relationship between newborns weight and characteristics of mothers and babies (Brown et al (2002), Campbell et al (1997), Chang et al (1992), Chen et al (2010), Frederick et al (2008), Goldfrey et al (2000), Graves et al (2006), Kaminski et al (2006), Kulmala et al (2001), Moshaddeque et al (1994), Paneth (1995), Weisman et al (2010)).

In this study, the information on the above characteristics for individual babies was obtained from the records. The distribution of all these characteristics for the low birth weight babies has been investigated. It was observed that all characteristics follow normal distribution with a p-value of  $\geq 0.01$ . Multivariate normal distribution based on the observed means and standard deviations was used to obtain 955 simulated sets of data for the LBW babies. Multi-regression analysis was first deployed to find the relationship between weight for the low weight birth babies and all the above characteristics. The regression model was then used to assess the effect of only BMI and other babies' characteristics on the weight of LBW babies.

The regression model including data from all routine investigations explained 95.55% of the variance in the newborn weights, suggesting that routine antenatal measurements had a power to predict the size of term live-born babies. However, our investigation in this paper is associated to the LBW babies only. This paper is organized in the following manner. Time series and distribution analysis are discussed in section 2. A review of the multi-regression analysis and multi-normal simulation are presented in section 3. Discussion based on Simulated and application example with real clinical data is presented in section 4. This is followed by the conclusion in section 5.

## 2. TIME SERIES AND DISTRIBUTION ANALYSIS

### 2.1 Time series plots

Time series plots are used to evaluate patterns in data over time. It is also used to investigate whether the patterns on different characteristics have uniform or corresponding pattern. The time series plot for the weight of LBW babies together with their mother's BMI and other babies' characteristics is presented in Figure 1.

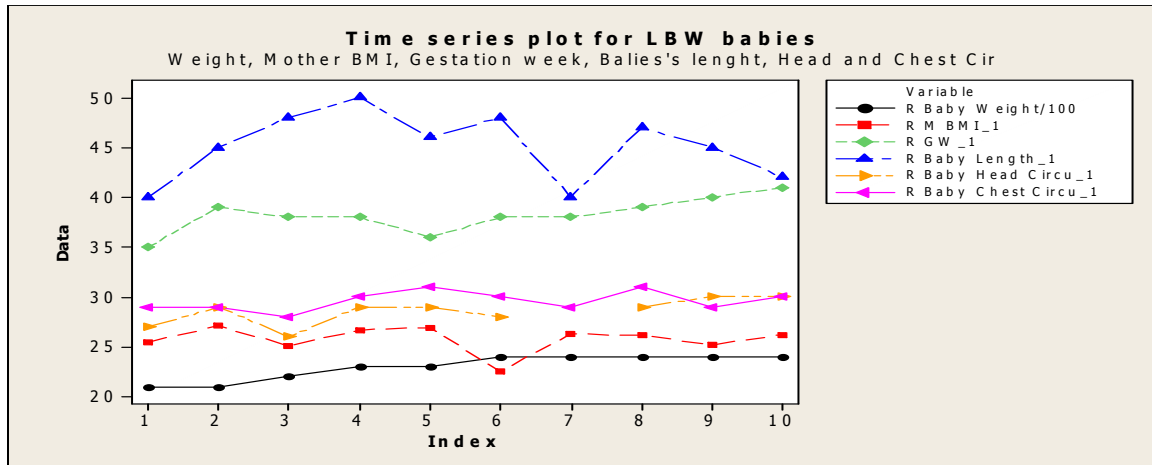


Figure 1: Time series plot for LBW data

Figure 1 clearly shows that the trend in the weight of LBW babies is closely related to the gestation age, height, head and chest circumference but is opposite to the trend in the Mother's BMI, i.e., as BMI increases the weight of the LBW baby decreases.

## 2.2 Normality Test

To carry out statistical analysis one often requires normal data. The normal probability plot is a graphical technique for assessing whether or not a data set is approximately normally distributed. The statistical package Minitab 16 is used to fit the commonly used normal distribution function to the individual characteristic of the low birth weight data. The probability plot is presented in Figure 2.

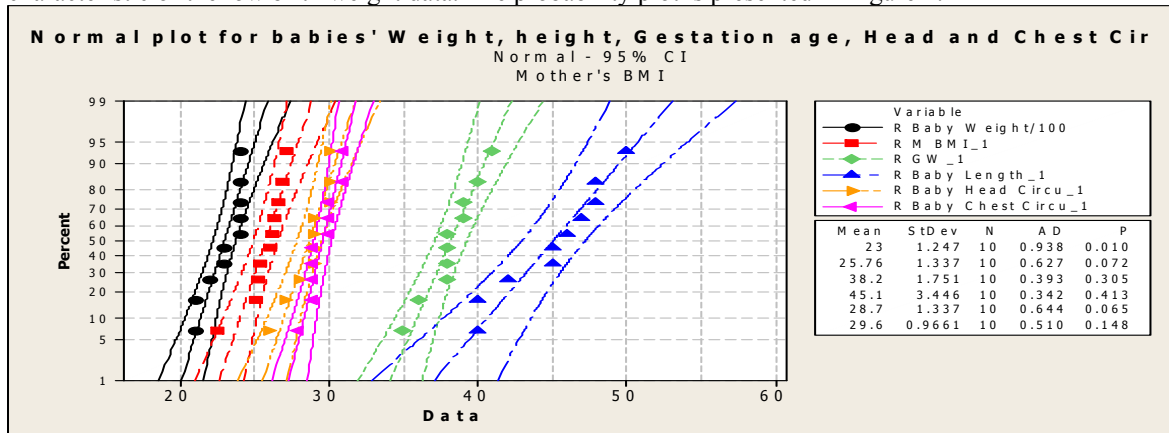


Figure 2: Normal probability plot of the recorded LBW data

Figure 2 shows that the p-value corresponding to all characteristics is greater than 0.01, therefore, we can conclude that the recorded follows normal distribution.

The multi-normal distribution based on the mean vector and the covariance matrix of the recorded LBW data is deployed to obtain 955 simulated data sets. The Normal probability plot for the simulated data together with the corresponding p-value for the individual characteristics is given in Figure 3.

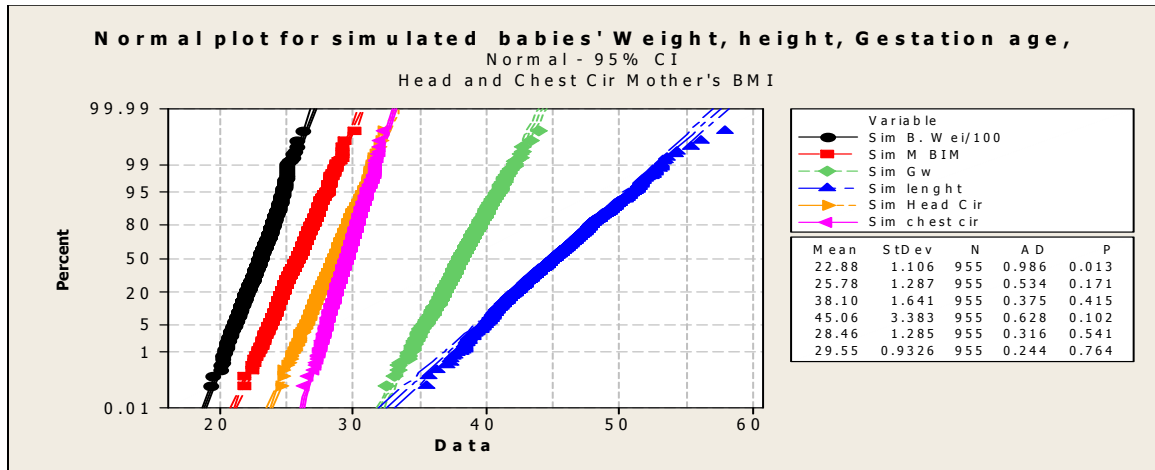


Figure 3: Normal probability plot of the simulated LBW data.

Figure 3 indicates that the p-value corresponding to all characteristics is greater than 0.01, therefore, the 955 simulated data follow normal distribution.

### 3. MULTI-REGRESSION MODEL

Regression is an appropriate approach used to accurately model the relationship between a scalar variable  $y$  and one or more explanatory variables denoted by  $X$ . It is assumed in regressions that the residuals (observed minus predicted values) are distributed normally. Even though most tests (specifically the  $F$ -test) are quite robust with regard to violations of this assumption, it is always a good idea, before drawing final conclusions, to review the distributions of the major variables of interest and test the correlation between explanatory variables.

Given a data set  $\{y_i, x_{i1}, x_{i2}, \dots, x_{ip}\}$  of  $n$  statistical units, a linear regression model assumes that the relationship between the dependent variable  $y_i$  and the  $p$ -vector of regressors  $x_i$  is linear. This relationship is modelled through a so-called “disturbance term”  $\epsilon_i$ , an unobserved random variable that adds noise to the linear relationship between the dependent variable and regressors. Thus the model takes the form

$$y_i = \beta_1 x_{i1} + \dots + \beta_p x_{ip} + \epsilon_i \quad i = 1, \dots, n. \quad (1)$$

In this study we are interested to find the relationship between; the baby’s weight for low birth weight babies (dependent variable) and mother’s pre-pregnancy BMI, gestation age at the time of delivery, baby’s height, chest and head circumference (independent variables). Prior to fitting the model one should investigate the correlation between the explanatory variables. Table 1 provides the correlation coefficients between these variables.

Table 1: Correlations: Gestation Age, Baby’s Height, Head and Chest Circumference and Mother’s BMI

	GW	Baby Height	Baby Head Cir
Baby Height	0.107		
p-value	0.769		
Baby Head Circumference	0.550	-0.234	
p-value	0.099	0.516	
Baby Chest Circumference	-0.013	0.247	0.413
p-value	0.971	0.492	0.236
M BMI	0.030	-0.189	0.395
p-value	0.934	0.601	0.258
Baby Chest Circumference			
M BMI	0.157		
p-value	0.665		

Table 1 shows that the p-values corresponding to the correlation between the independent variables are not significant, therefore there is no correlation between the independent variables.

### 3.1. Fitting the Regression Model and Estimating the Forecasting Error

Using the actual observed data, the following multi- regression model with the  $R^2 = 72.9\%$  was obtained. Therefore, 72.9% of the variation in the weight of LBW babies is explained by the mother's BMI and other characteristics of the baby.

$$\text{Baby Weight}/100 = 2.6 + 0.141\text{GW} + 0.018\text{Height} + 0.551\text{Head Cir} + 0.375 \text{Chest Cir} - 0.494 \text{M BMI} \quad (2)$$

Using the 955 multi-normal simulated data we obtained the following regression model with the  $R^2 = 65.9\%$ .

$$\text{Baby Weight}/100 = 5.14 + 0.183\text{Gw} - 0.0268\text{Height} + 0.345\text{Head Cir} + 0.424\text{chest cir} - 0.401 \text{M BIM} \quad (3)$$

The correlation coefficient for the regression model given in (3) is  $r = 0.8$  which indicates that the model can be used to predict the weight for the LBW babies. To assess the efficacy of the proposed model, we have deployed the model to predict the observed weight in the recorded data. The predicted values, together with the standard error of the prediction, 95% confidence and predicted intervals and the prediction error are given in table 2.

Table 2: Predicted Values for the recorded baby's weight based on the proposed multi regression model.

	Fit	SE Fit	95% CI	95% PI	Prediction error
1	21.8474	0.0625	(21.7247, 21.9701)	(20.3706, 23.3242)	-0.8474
2	22.4769	0.0430	(22.3925, 22.5613)	(21.0028, 23.9510)	-1.4769
3	21.5688	0.0655	(21.4402, 21.6974)	(20.0915, 23.0461)	0.4312
4	22.7714	0.0536	(22.6661, 22.8766)	(21.2959, 24.2468)	0.2286
5	22.8210	0.0580	(22.7073, 22.9348)	(21.3450, 24.2971)	0.1790
6	24.1158	0.0734	(23.9717, 24.2599)	(22.6370, 25.5945)	-0.1158
7	23.0684	0.0802	(22.9110, 23.2258)	(21.5883, 24.5484)	0.9316
8	23.6623	0.0654	(23.5339, 23.7907)	(22.1850, 25.1395)	0.3377
9	23.7407	0.0663	(23.6105, 23.8708)	(22.2633, 25.2181)	0.2593
10	24.0260	0.0675	(23.8936, 24.1585)	(22.5484, 25.5037)	-0.0260

Table 3: Summary statistics for the forecasting error of the proposed model

N	Mean	Standard error of Mean	Standard deviation
10	-0.010	0.217	0.686

Table 3 shows that the proposed regression model can predict the true value of the newborn weight for LBW babies with the mean error of  $-0.01/10 = -0.001$  Kg (1 gram, the weight used in the model is gr/100). The standard error of the mean is 0.217. Therefore we can claim that the accuracy of the model is significant and precise.

## 4. DISCUSSION

The study was conducted among women enrolled in a maternity clinic in Banjarmasin Indonesia. One of the objectives in this research experiment was to investigate the relationship between baby's weight (for the low birth weight babies) and baby's characteristics such as height, chest and head circumference, gestation age as well as mother's pre-pregnancy BMI. The study has proposed a model to predict weight of the LBW babies based on the above characteristics. The independent effect of a number of variables on the baby's weight was calculated by using multi-regressions. The slope (the beta-coefficient) that shows the amount of change in the dependent variable (baby's weight) for one unit change in an independent variable, such as; BMI and other variables together with the Pearson correlation coefficient were used to identify the most effective independent variables for predicting the baby's weight. All the analyses were performed using the MINTAB 16 statistical software package.

The slopes and their corresponding p-values in the proposed regression model show that the most effective variables are baby's chest and head circumference( slop= 0.424 and 0.345 respectively) and mother's pre-

pregnancy BMI (slope= -0.401). However the least effective variables are gestation age (slope=0.183) and height (slope= -0.0268).

The proposed regression model has a correlation coefficient of 80%. Therefore, one can claim that the relationship between weight and the other five independent variables is strong enough to predict the newborn's weight. The model was then used to predict the observed weight data to assess its predicting accuracy. Results presented in tables 2 and 3 shows that the estimated weight using the proposed model is very close to the actual recorded weight for the LBW babies with the mean error prediction of one gram.

## 5. CONCLUSION

Low birth weight is an increasingly prevalent factor in the Maternal Mortality Rate (MMR). Therefore many studies have attempted to identify the sources of variation in the newborn weight. In this study multi-regression model is used to assess the independent effects of the mother's BMI, gestation age at the time of delivery, baby's height, head and chest circumference on the newborn's weight for the low birth weight babies.

Nine hundred and fifty five simulated data based on the means and standard deviations of the recorded low birth weight data were used to estimate the parameters of the model. The results show that for the low birth weight babies, there is a statistically significant relationship between the newborn's weight and mother's BMI, gestation age and baby's height, head and chest circumference. The proposed multi-regression model based on these independent variables has a correlation coefficient of 80%.

The proposed model was then used to estimate the recorded weights together with their corresponding 95% confidence and predication interval. Analysis of the prediction errors shows that the mean prediction error for the recorded data is one gram. Therefore one can conclude that the proposed multi-regression model is capable of accurately predicating the weight for the low birth weight babies based on the characteristics of the mother and the baby. The model is based on the simulated data using the sample measurements and in future would be tested on a larger set of observed data. The outcome of this research suggests that medical practitioners may be able to reduce the mortality rate due to the low birth weight by using ultrasonic measurements of these characteristics in the third semester of the pregnancy and apply suitable preventive intervention to ensure that the weight would increase enough to guarantee a birth weight greater than 2.5 kg.

## Acknowledgement

The authors would like to thank the medical practitioners and staff in the maternity clinic in Banjarmasin Indonesia for their effort in collecting this data and providing them to us for this research.

## REFERENCES

- Allen, M, Mor, J. (2003). US birth weight/gestational age-specific neonatal mortality: 1995-1997 rates for whites, Hispanic, and blacks. *Pediatrics*, 111, 61-6.
- Bennett, BB. (1999). Shoulder dystocia: An obstetric emergency. *Obstet Gynecol Clin North Am* 26, 445-58.
- Brown, JE., Murtaugh, MA., Jacobs, Jr DA., and Margellos, HC. (2002). Variation in newborn size according to pregnancy weight change by trimester, *Am J Clin Nutr*, 76, 205-9. Printed in USA. American Society for Clinical Nutrition.
- Campbell, S, Wilkin, D. (1975). Ultrasonic measurement of fetal abdominal circumference in the estimation of fetal weight. *Br J Obstet Gynecol*, 82, 689-97.
- Chang, TC, Robson, SC, Boys, RJ, Spencer, JA. (1992). Prediction of the small for gestational age infant: Which ultrasonic measurement is the best?. *Obstet Gynecol*, 80, 1030-7.
- Chen, et al., (2010). Pre-pregnancy body mass index, gestational weight gain, and pregnancy outcomes in China. *Int. J. Gynaecol. Obstet.* 109, 41-44.
- Copper, RL, Dubard, MB, Goldenberg, RL, Oweis, AL. (1995). The relationship of maternal attitude toward weight gain to weight gain during pregnancy and low birth weight. *Obstet. Gynecol*, 85(4), 590-595.
- Frederick, IO, William, MA, Sales AE, Martin, DP, Killien, M. (2008). Pro-pregnancy body mass index, gestational weight gain and other maternal characteristics in relation to infant weight. *Maternal child Health J.*, 12, 557-567.
- Goldfrey, K.M. & Barker, D.J. (2000). Fetal nutrition and adult disease. *Am J Clin Nutr*, Vol. 71 (5 Suppl), 1344S-52S.

- Graves, BW, Dejoy, SA, Heath, A, Pekow, P. (2006). Maternal body mass index, delivery route, and induction of labor in a midwifery caseload. *J. Midwifery, Womens Health*, 51(4), 254-259.
- Kaminski, M., Goujard, J. and Rumeau-Rouquette, C. (1973). Prediction of Low Birth weight and Prematurity by a Multiple Regression Analysis with Maternal Characteristics Known since the Beginning of the Pregnancy. *International Journal of Epidemiology*, Vol. 2. No. 2, 195-203.
- Kulmala, Teijia, Vaahtera, M., Ndekha, M., Koivis, A., Cullinan, T. (2001). Gestational Health and Predictors of Newborn Weight Amongst Pregnant Women in Rural Malawi. *African Journal of Reproductive Health / La Revue Africaine de la Santé Reproductive*, Vol. 5, No. 3, 99-108.
- Lagerstrom, M.; Bremme, K.; Eneroth, P. & Magnusson, D. (1991). School performance and IQ-test scores at age 13 as related to birth weight and gestational age. *Scand J Psychol.* Vol. 32 (4), 316 – 24.
- Maddah, M. (2005). Pregnancy weight gain in Iranian women attending cross-sectional study of public health centers in Rasht. *Midwifery*, 21(4), 365-370.
- Mathews, TJ, MacDorman, MF, Menacker, F. (2002). Infant mortality statistics from the 1999 period linked birth/death data set. *Natl Vital Stat Rep*, 50, 1-28.
- Metgud, CS., Vijaya, A., Naik, A., Maheshwar, D. (2012). Factors Affecting Birth Weight of a Newborn – A Community Based Study in Rural Karnataka, India. *PLoS ONE*, [www.plosone.org](http://www.plosone.org) Volume 7 | Issue 7 | e40040.
- Moshaddeque, Hossain, M., Mostafa, H., Herbert, L. (1994). Association Between Birth Weight and Birth Arm Circumference of Neonates in Rural Egypt: *Indian J Pediatr*, 61, 81-87.
- Paneth, NS. (1995). The Problem of Low Birth Weight. *The Future of Children, Low Birth Weight*. Vol. 5, No. 1, 19-34.
- Phillips, DI. (1998). Birth weight and the future development of diabetes: *A review of the evidence. Diabetes Care*, 21, 150-5.
- Rich-Edwards, JW, Stampfer, MJ, Manson, JE, Rosner, B, Hankinson, SE, Colditz, GA, et al. (1977). Birth weight and risk of cardiovascular disease in a cohort of women followed up since 1976. *BMJ*, 315, 396-400.
- Rich-Edwards, JW, Colditz, GA, Stampfer, MJ, Willett, WC, Gillman, MW, Hennekens, CH, et al. (1999). Birth weight and the risk for type 2 diabetes mellitus in adult women. *Ann Intern Med*, 130, 278-84.
- Rondo, PHC., Ferreira, RF., Nogueira, F, Ribeiro, MCN., Lobert, H. (2003). Maternal psychological stress and distress as predictors of low birth weight, prematurity and intrauterine growth retardation. *European Journal of Clinical Nutrition*, 57, 266–272.
- Thorsdottir, I. and Birgisdottir, BE. (1998) Different weight gain in women of normal weight before pregnancy: postpartum weight and birth weight. *Obstet Gynecol*, 92(3), 377-383.
- Weisman, CS, Hillemejer, MM., Downs, DS., Chuang, CH., Dyer, AM. (2010). Preconception predictors of weight gain during pregnancy: prospective findings from the Central Pennsylvania women's Health Study. *Womens Health Issues*, 20(2), 126-132.