

Modeling Gestational Age and Lack of Fit

Mohammad Lutfor Rahman¹, Ema Akter², Zakir Hossain³, Aysha Sultana⁴, Kalyan Das⁵

ABSTRACT

Introduction: Gestational age (GA) refers to age of an unborn baby. Accurate determination of GA is crucial as precise calculation of GA helps reducing post-date labor induction and is particularly useful to assess viability in premature labor and in post-date deliveries.

Aim: The aim of the article was to determine GA by a minimal number of fetal parameters along with greater accuracy.

Materials and methods: This was a prospective cross-sectional study, comprising a total of 229 singleton pregnant mothers enrolled from December 2015 to November 2016 in Ibn Sina Diagnostic and Imaging Center and Ad-Din Hospital in Dhaka, Bangladesh. Multiple linear regression models were fitted with four fetal parameters, namely biparietal diameter (BPD), head circumference (HC), abdominal circumference (AC), and femur length (FL), and lack of fit was tested in each of the cases. In case of having a significant lack of fit, different types of transformation including Box-Cox transformation have been adopted on the variables to improve the adequacy of the model.

Results: Among all the fitted models for gestational age by last menstrual period (GALMP), a model with a single explanatory variable, BPD, was found to be comparatively better than others. The Box-Cox transformation was taken on the dependent variable GALMP with λ that equals 0.2, and lack of fit test was not significant at 1% level (p -value = 0.027). Though the results from comparative models do not vary substantially, emphasis on BPD in GA calculation might suffice while cost of experimentation or screening is a grave concern.

Conclusion: By considering the Box-Cox transformation and proper dealing of outliers, a model for GALMP was obtained without lack of fit where only the BPD appeared to be sufficient as an explanatory variable.

Keywords: Biparietal diameter, Box-Cox transformation, Gestational age, Lack of fit.

Journal of South Asian Federation of Obstetrics and Gynaecology (2021): 10.5005/jp-journals-10006-1902

INTRODUCTION

Gestational age (GA) refers to the age of an unborn baby.¹ Assessment of GA in early detection is valuable in detecting the growth aberration in later stages of pregnancy.² It is specifically useful to assess viability in premature labor and in post-date deliveries.³ Accurate estimates of GA are necessary for both clinical practice and public health concerns. The GA can be computed either as a difference between the first day of the last menstrual period (LMP) and the delivery date (gestational age by last menstrual period—GALMP) or by modern technology-based ultrasonographic measurements (gestational age by ultrasonography—GAUSG).

Estimation of GA by ultrasonographic method is common nowadays.⁴ However, for low- and middle-income countries, estimation of GA through ultrasonography might not be feasible extensively due to limited resources. Often, GA is calculated based on the LMP as it is an easy and low-cost method. However, women with low education and poor economic background are likely to misreport their last menstrual dates. Therefore, the estimation of GA completely based on LMP could entail inaccuracy.

Regression models are widely used for estimating GA.⁵ Mainly four important fetal parameters, namely biparietal diameter (BPD), head circumference (HC), abdominal circumference (AC), and femur length (FL), are used to determine GA in a regression modeling approach. Benson and Doubilet⁶ and Chervenak et al.⁷ argued to use multiple fetal parameters in regression instead of a single second-trimester parameter in GA estimation. Multiple parameters might be useful in case of having growth disruption of any of the fetal parameters due to a fetal syndrome, for instance, FL might be shorter than usual.⁸

¹Institute of Statistical Research and Training, University of Dhaka, Dhaka, Bangladesh

²ICDDR, B, Dhaka, Bangladesh

³Department of Statistics, University of Dhaka, Dhaka, Bangladesh

⁴Bangladesh Institute of Medical Sonography, Dhaka, Bangladesh

⁵Department of Basic and Applied Sciences, NIFTEM, Kundli, Haryana, India

Corresponding Author: Mohammad Lutfor Rahman, Institute of Statistical Research and Training, University of Dhaka, Dhaka, Bangladesh, Phone: +8801946674978, e-mail: lutfor@isrt.ac.bd

How to cite this article: Rahman ML, Akter E, Hossain Z, et al. Modeling Gestational Age and Lack of Fit. *J South Asian Feder Obst Gynae* 2021;13(3):81–85.

Source of support: Nil

Conflict of interest: None

Frequently, clinicians use the unweighted mean of all four biometric parameters, namely BPD, HC, AC, and FL, without proper justification. However, all four parameters are not equally important while predicting GA. Therefore, many authors created regression equations using various combinations of biometric parameters to improve accuracy.⁷ The estimation of regression models is based on some assumptions that need to be examined before making conclusions from the fitted model. Rahman et al.⁹ determined GAUSG and GALMP using linear regression with four fetal parameters in Bangladesh. However, they checked the assumptions of linear regression but did not take into account the lack of fit, particularly for GALMP. None of the earlier studies considered a lack of fit during GA modeling. The current study

has addressed the lack of fit issue to find a better model for GALMP using fetal parameters.

MATERIALS AND METHODS

Data and Variables

The data consisted of 229 singleton pregnant mothers enrolled from December 2015 to November 2016 in Ibn Sina Diagnostic and Imaging Center and Ad-Din Hospital in Dhaka, Bangladesh. Inclusion criteria for participation in the study were certain date of the first day of LMP. Exclusion criteria were multiple gestation, maternal complications (isoimmunization, substance abuse, abnormal glucose tolerance test, and chronic hypertension), and major fetal malformations. Four fetal parameters, namely BPD, HC, AC, FL, were measured through ultrasonographic technology. The response variables were GAUSG and GALMP. Rahman et al.⁹ fitted GAUSG model without having difficulties in satisfying model assumptions, but GALMP model had problems of lack of fit. In this study, the issue of lack of fit has been addressed for better precision in the results, particularly considering GALMP data. Statistical software R (version 4.0.0) has been used for programming and data analysis.

Lack of Fit

The lack of fit test was carried out for ascertaining whether a linear regression model was a good fit for the current data. The test statistic regarding lack of fit can be expressed as follows:

$$F = \frac{MSLF}{MSPE} \quad (1)$$

where MSLF denotes the lack of fit mean square and MSPE denotes the pure error mean square. If a lack of fit test is significant, that implies fitted model might be inadequate. There should have some attempts to discover where and how the inadequacy occurs. On the contrary, if the test is insignificant, there is no doubt about the adequacy of the model on the basis of the test.

Box-Cox Transformation

The Box-Cox transformation helps reducing nonlinearity of the regression function, skewness of the distribution of error variance, and unequal error variances. Initially, Box and Cox¹⁰ considered the power family of transformations on the response as follows:

$$f_{\lambda}(y) = \begin{cases} (y^{\lambda} - 1) / \lambda & \text{when } \lambda \neq 0, \\ \log(y) & \text{when } \lambda = 0 \end{cases} \quad (2)$$

where λ is the transformed response. The objective of the Box-Cox transformations is to ensure that the general assumptions for linear model are more likely to satisfy after the transformation. If the Box-Cox transformation is implemented to the response variable Y in the multiple regression model

$$f_{\lambda}(y) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_{p-1} X_{i,p-1} + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (3)$$

we can estimate λ and other parameters $\beta_0, \beta_1, \dots, \beta_k$ and σ^2 simultaneously by the maximum likelihood method.

The Box-Cox procedure estimates λ automatically as well as the other parameters $\beta_0, \beta_1, \dots, \beta_k$ and σ^2 by the method of maximum likelihood. Further details on the estimation procedure under the Box-Cox transformation can be found in the texts by Neter et al.¹¹ and Yan and Su.¹²

RESULTS

The main idea of analyzing data was to predict the GA by ultrasonography measured fetal parameters using the regression model. Initial assessment showed that there were linear relationships among the response and independent variables. Considering GAUSG and GALMP as responses and BPD, HC, AC, and FL as covariates, we fitted two linear regression models to the data. There was no evidence of lack of fit for modeling GAUSG (p -value = 0.999), whereas the lack of fit test was significant (p -value <0.001) for modeling GALMP. Thus, it was assumed that a transformation of covariates or responses might be helpful to eradicate lack of fit and thereby could provide precise estimates of model parameters.

Transformation of Covariates

We took some transformations of covariates to improve the adequacy of the fitted model. The regression model was modified by altering the nature of the regression function. We did lack of fit test to model GALMP with the main terms, BPD, HC, AC, and FL, and quadratic terms, BPD², HC², AC², and FL², individually. However, none of the models was suitable for predicting GALMP as there was evidence of lack of fit (p -values <0.001). Further, all models with the subset of covariates also shown the existence of lack fit (p -values <0.001).

Transformation of Response Variable

The Box-Cox transformation enables finding the most appropriate transformation of the response variable GALMP. Figure 1 indicates a range of possible values of λ that would be 0–0.3 for the transformation. For these λ values, the response was transformed and performed lack of fit test as shown in Table 1.

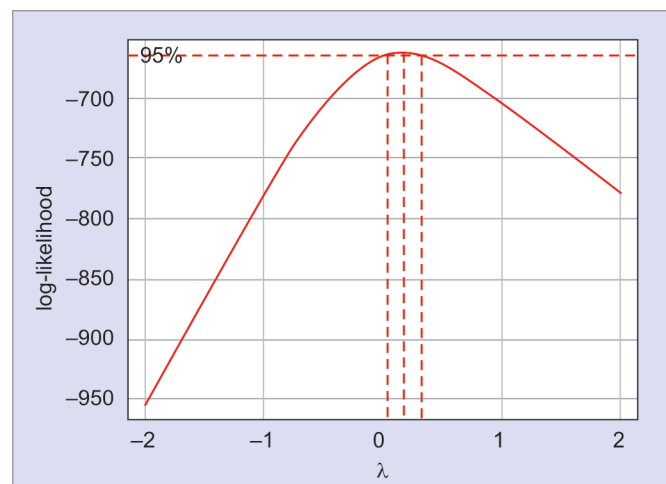


Fig. 1: Plot of λ determination for Box-Cox transformation of GALMP

Table 1: Lack of fit test after the Box-Cox transformation

λ	Regressor(s)	<i>p</i> value	λ	Regressor(s)	<i>p</i> value
0.3	BPD, HC, AC, FL	<0.001	0.0	BPD, FL	<0.001
0.2	BPD, HC, AC, FL	<0.001	0.3	HC, AC	<0.001
0.1	BPD, HC, AC, FL	<0.001	0.2	HC, AC	<0.001
0.0	BPD, HC, AC, FL	<0.001	0.1	HC, AC	<0.001
0.3	BPD, HC, AC	<0.001	0.0	HC, AC	<0.001
0.2	BPD, HC, AC	<0.001	0.3	HC, FL	<0.001
0.1	BPD, HC, AC	<0.001	0.2	HC, FL	<0.001
0.0	BPD, HC, AC	<0.001	0.1	HC, FL	<0.001
0.3	BPD, HC, FL	<0.001	0.0	HC, FL	<0.001
0.2	BPD, HC, FL	<0.001	0.3	AC, FL	<0.001
0.1	BPD, HC, FL	<0.001	0.2	AC, FL	<0.001
0.0	BPD, HC, FL	<0.001	0.1	AC, FL	<0.001
0.3	BPD, AC, FL	<0.001	0.0	AC, FL	<0.001
0.2	BPD, AC, FL	<0.001	0.3	BPD	<0.001
0.1	BPD, AC, FL	<0.001	0.2	BPD	0.027
0.0	BPD, AC, FL	<0.001	0.1	BPD	<0.001
0.3	HC, AC, FL	<0.001	0.0	BPD	<0.001
0.2	HC, AC, FL	<0.001	0.3	HC	<0.001
0.1	HC, AC, FL	<0.001	0.2	HC	<0.001
0.0	HC, AC, FL	<0.001	0.1	HC	<0.001
0.3	BPD, HC	<0.001	0.0	HC	<0.001
0.2	BPD, HC	<0.001	0.3	AC	<0.001
0.1	BPD, HC	<0.001	0.2	AC	<0.001
0.0	BPD, HC	<0.001	0.1	AC	<0.001
0.3	BPD, AC	<0.001	0.0	AC	<0.001
0.2	BPD, AC	<0.001	0.3	FL	<0.001
0.1	BPD, AC	<0.001	0.2	FL	<0.001
0.0	BPD, AC	<0.001	0.1	FL	<0.001
0.3	BPD, FL	<0.001	0.0	FL	<0.001
0.2	BPD, FL	<0.001			
0.1	BPD, FL	<0.001			

AC, abdominal circumference; BPD, biparietal diameter; FL, femur length; HC, head circumference

Model for GALMP

The value of GALMP was obtained by asking the mothers about the start date of menstruation. However, mothers often failed to recall the actual date of menstruation. From Figure 2, we notice that residuals approximately follow the assumption of normality except for some points (say 127th, 201th, and 228th) who are staying away from other observations. These are outliers that might appear due to misreporting menstrual dates by the mothers. We had a new data set by excluding the points numbered with 24, 94, 123, 127, 164, 201, and 228 from the original data set. We performed the lack of fit test for the new data set and obtained results shown in Table 1.

We checked whether there was any significant change ensued in the model for new data. Figure 3 shows response (GALMP) vs covariates excluding outliers and indicates that every covariate was approximately linear with the response. From Table 1, we could fit the model of $GALMP^{0.2}$ to a single explanatory variable BPD with the transformation parameter λ being 0.2 as lack of fit was not significant (p -value = 0.027). Thus, for the data without outliers, the lack of fit was not significant at 1% level meant that

the model was adequate to predict GALMP and the fitted model would be

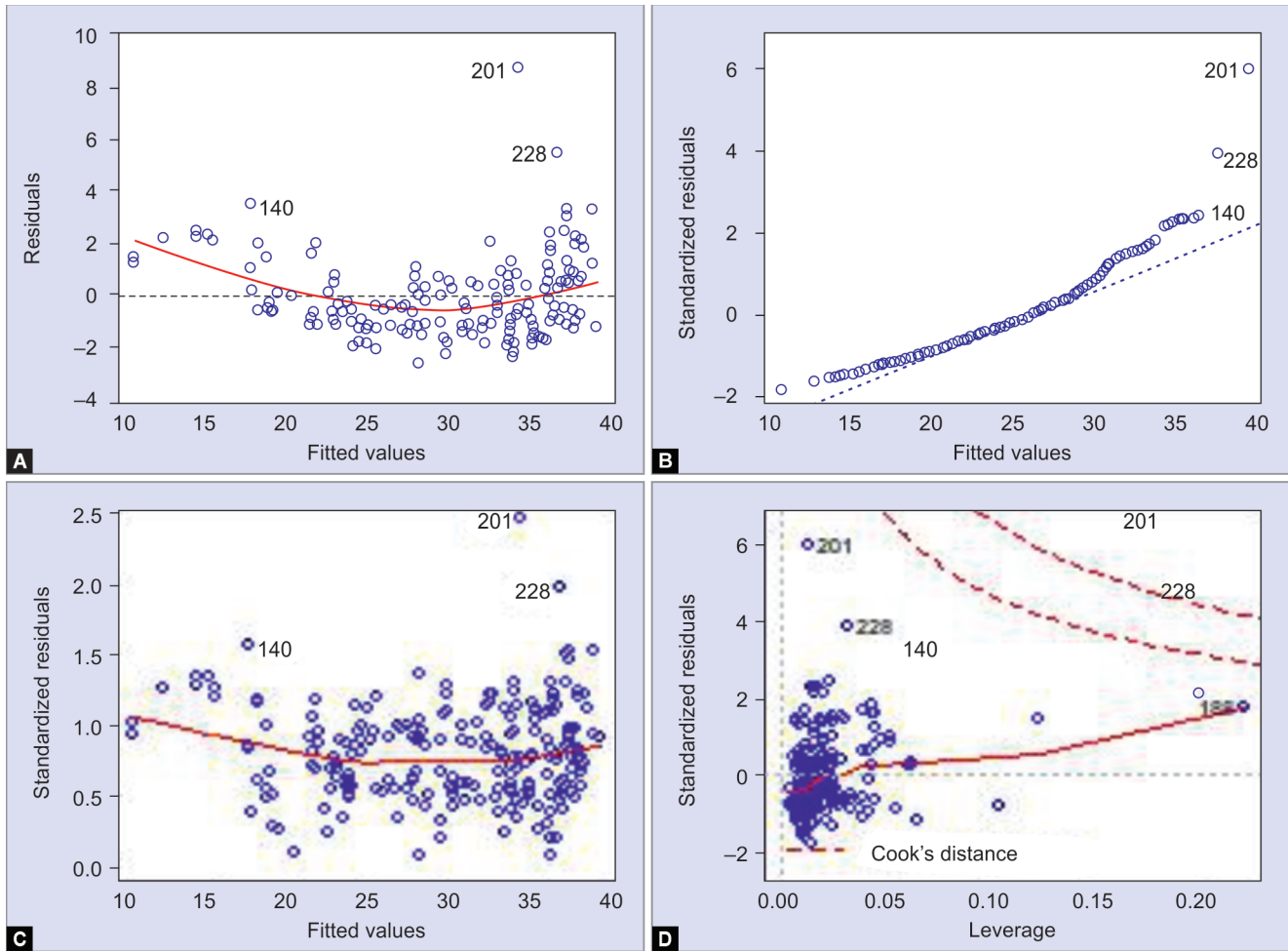
$$GALMP^{0.2} = 1.583440 + 0.005263 \times BPD_i \quad (4)$$

which could be useful for calculating GALMP, specifically for Bangladeshi ethnic community.

DISCUSSION

Precise estimation of GA is crucial for investigation, diagnosis, and treatment of fetus *in vitro*.¹³ This study models LMP-based GA (GALMP) considering lack of fit and fits a simple model of GALMP with a single explanatory variable BPD.

When GAUSG was used as a response in the multiple linear models with four fetal parameters BPD, FL, AC, and HC as explanatory variables, no significant lack of fit was detected.⁹ However, modeling GALMP incurs a significant lack of fit that indicates the fitted model was not adequate for the data. To improve the adequacy of the model, we took a number of transformations including Box-Cox transformation and performed an F test for lack of fit by forming different models based on



Figs 2A to D: Plot of response (GALMP) vs residuals

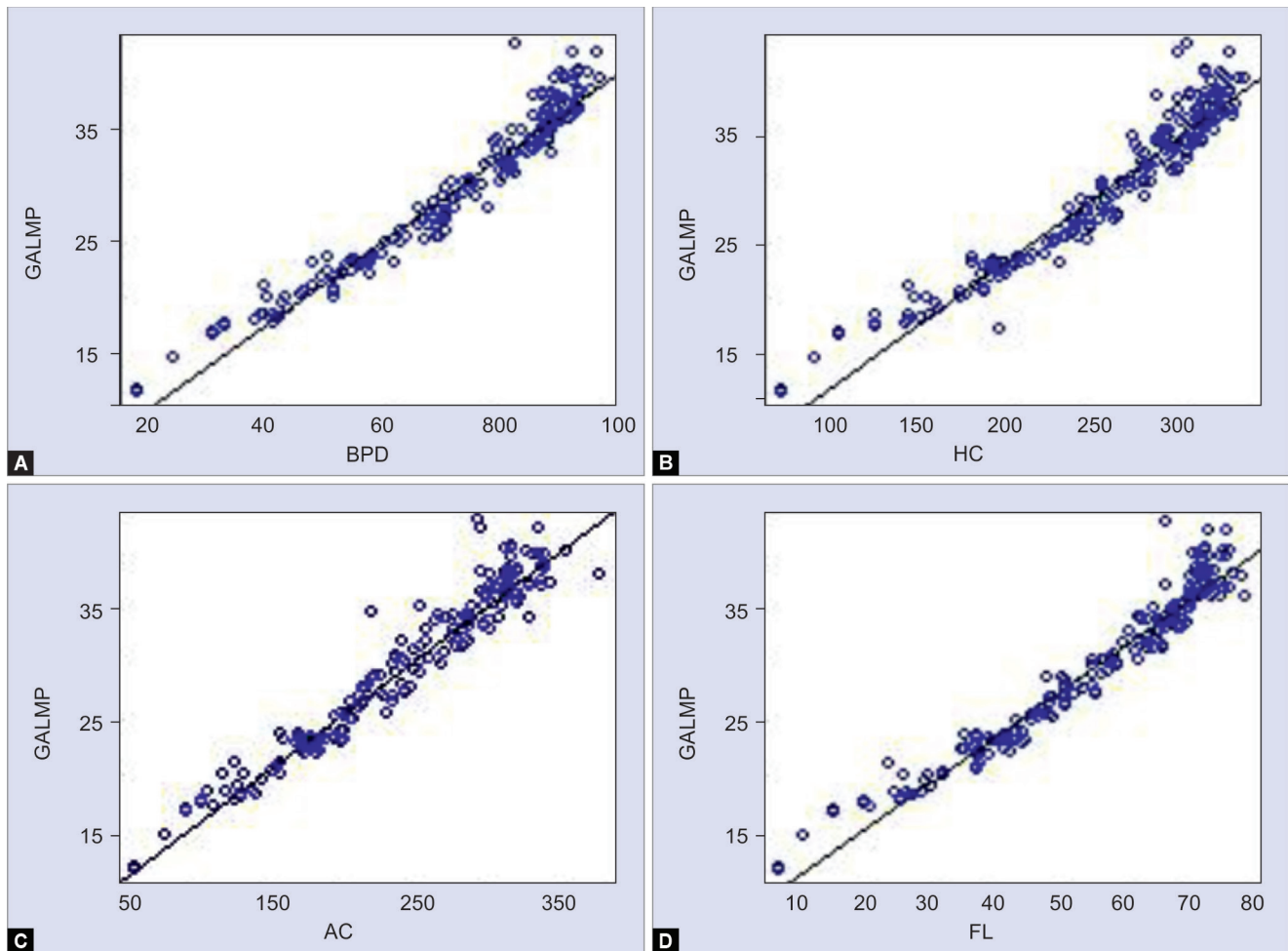
that transformation. We fitted a simple model of GALMP with BPD as an independent variable after Box-Cox transformation on the dependent variable with λ equals to 0.2 and lack of fit was not significant at 1% level (p -value=0.027). Though results from comparative models do not vary substantially (Table 1), yet emphasis on BPD in GA calculations gives a message about not to use redundant parameters while the cost of experimentation or screening is a grave concern. The BPD being a sole predictor of GA is supported by other studies.¹⁴⁻¹⁷ Hadlock et al.¹⁵ described that BPD, out of the four parameters, is the most frequently used parameters in determining the age of the fetus up to 36 weeks. Kurtz et al.¹⁴ emphasized that BPD to be one of the useful criteria to measure GA and to predict the expected date of delivery. Also, Kramer et al.¹⁶ indicated that the BPD was used as “gold standard” to test the validity of GALMP-based estimates.

CONCLUSION

The precise model proposed in this study for GALMP with a single fetal parameter BPD could be used to compute GA in Bangladesh. The other countries with different ethnic communities can follow similar techniques to identify their own models for the sake of calculating GA accurately and economically.

ORCID

- Mohammad L Rahman <https://orcid.org/0000-0001-9100-0476>
- Ema Akter <https://orcid.org/0000-0003-4772-5431>
- Zakir Hossain <https://orcid.org/0000-0003-4128-3097>
- Aysha Sultana <https://orcid.org/0000-0002-9996-0030>
- Kalyan Das <https://orcid.org/0000-0001-9100-0476>



Figs 3A to D: Plot of response (GALMP) vs covariate for data excluding outliers

REFERENCES

- Machado L, Vaclavinkova V, Gibb H. Evaluation of applicability of standard growth curves to healthy native Omani women by fetal biometry at selected gestational ages. *Sultan Qaboos Univ Med J [SQUMJ]* 2000;2(2):97–104. <https://journals.squ.edu.om/index.php/squmj/article/view/1203>.
- Shehzad K, Ali M, Zaidi S. Fetal biometry. *Pak J Med Sci* 2006;22(4):503. https://pjms.com.pk/issues/octdec06/pdf/fetal_biometry.pdf.
- Butt K, Lim K, Bly S, et al. Determination of gestational age by ultrasound. *J Obstet Gynaecol Can* 2014;36(2):171–181. DOI: 10.1016/S1701-2163(15)30664-2.
- Konje J, Abrams K, Bell S, et al. Determination of gestational age after the 24th week of gestation from fetal kidney length measurements. *Ultrasound Obstet Gynecol* 2002;19(6):592–597. DOI: 10.1046/j.1469-0705.2002.00704.x.
- Zaidi S, Shehzad K, et al. Sonographic foetal measurements in a cohort of population of Karachi, Pakistan. *J PMA* 2009;59(4):246–249. <https://www.jpma.org.pk/PdfDownload/1676.pdf>.
- Benson C, Doubilet P. Sonographic prediction of gestational age: accuracy of second- and third-trimester fetal measurements. *AJR* 1991;157(6):1275–1277. DOI: 10.2214/ajr.157.6.1950881.
- Chervenak FA, Skupski DW, Romero R, et al. How accurate is fetal biometry in the assessment of fetal age? *Am J Obstet Gynecol* 1998;178(4):678–687. DOI: 10.1016/S0002-9378(98)70477-6.
- Mailath-Pokorny M, Polterauer S, Worda K, et al. Isolated short fetal femur length in the second trimester and the association with adverse perinatal outcome: experiences from a tertiary referral center. *PLoS One* 2015;10(6):e0128820. DOI: 10.1371/journal.pone.0128820.
- Rahman ML, Sultana A, Das K. Determination of gestational age: a perspective with the Bangladeshi ethnicity. *Eur Sci J* 2018;14(6):375–386. DOI: 10.19044/esj.2018.v14n6p375.
- Box GE, Cox DR. An analysis of transformations. *J R Stat Soc* 1964;26(2):211–243. <http://links.jstor.org/sici?sici=0035-9246%281964%2926%3A2%3C211%3AAAOT%3E2.0.CO%3B2-6>.
- Neter J, Kutner MH, Nachtsheim CJ, et al. *Applied linear statistical models*, vol. 4. Chicago: Irwin; 1996.
- Yan X, Su X. *Linear regression analysis: theory and computing*. World Scientific; 2009. DOI: 10.1142/6986.
- Jaiswal P, Masih WF, Jaiswal S, et al. Assessment of fetal gestational age by ultrasonic measurement of bi-parietal diameter in the southern part of Rajasthan. *Med J Dr DY Patil Univ* 2015;8(1):27. DOI: 10.4103/0975-2870.148839.
- Kurtz AB, Wapner RJ, Kurtz RJ, et al. Analysis of biparietal diameter as an accurate indicator of gestational age. *J Clin Ultrasound* 1980;8(4):319–326. DOI: 10.1002/jcu.1870080406.
- Hadlock FP, Harrist R, Deter RL, et al. Fetal femur length as a predictor of menstrual age: sonographically measured. *Am J Roentgenol* 1982;138(5):875–878. DOI: 10.2214/ajr.138.5.875.
- Kramer MS, McLean FH, Boyd ME, et al. The validity of gestational age estimation by menstrual dating in term, preterm, and postterm gestations. *JAMA* 1988;260(22):3306–3308. DOI: 10.1001/jama.1988.03410220090034.
- Sultana A, Rahman ML. Ultrasonographic estimation of fetal gestational age by biparietal diameter: a cross sectional study in an urban area of Bangladesh. *J Med Sci Clin Res* 2018;6(4):1144–1151. DOI: 10.18535/jmscr/v6i4.187.