

# Foetal ultrasound measurement imputations based on growth curves versus multiple imputation chained equation (MICE)

Kelly K. Ferguson<sup>1</sup>  | Youfei Yu<sup>2</sup> | David E. Cantonwine<sup>3</sup> | Thomas F. McElrath<sup>3</sup> | John D. Meeker<sup>4</sup> | Bhramar Mukherjee<sup>2</sup>

<sup>1</sup>Epidemiology Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

<sup>2</sup>Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI, USA

<sup>3</sup>Division of Maternal-Fetal Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

<sup>4</sup>Department of Environmental Health Sciences, School of Public Health, University of Michigan, Ann Arbor, MI, USA

## Correspondence

Bhramar Mukherjee, Departments of Biostatistics and Epidemiology, University of Michigan School of Public Health, Ann Arbor, MI, USA.  
Email: bhramar@umich.edu

## Funding information

NIEHS, Grant/Award Number: P30ES017885, P42ES017198, R01ES018872 and ZIA103321

## Abstract

**Background:** Ultrasound measures are valuable for epidemiologic studies of risk factors for growth restriction. Longitudinal measurements enable investigation of rates of change and identification of windows where growth is impacted more acutely. However, missing data can be problematic in these studies, limiting sample size, ability to characterise windows of vulnerability, and in some instances creating bias. We sought to compare a parametric linear mixed model (LMM) approach to multiple imputation in this setting with multiple imputation by chained equation (MICE) methodology.

**Methods:** Ultrasound scans performed for clinical purposes were abstracted from women in the LIFECODES birth cohort ( $n = 1003$ ) if they were close in time to three study visits (median 18, 26, and 35 weeks' gestation). We created imputed datasets using LMM and MICE and calculated associations between demographic factors and ultrasound parameters cross-sectionally and longitudinally. Results were compared with a complete-case analysis.

**Results:** Most participants had ultrasounds at 18 weeks' gestation, and ~50% had measurements at 26 and 35 weeks; 100% had birthweight. Associations between demographic factors and ultrasound measures were similar in magnitude, but more precise, when either imputed datasets were used, compared with a complete-case analysis, in both the cross-sectional or longitudinal analyses.

**Conclusions:** MICE, though ignoring the non-linear features of the trajectory and within subject correlation, is able to provide reasonable imputation of foetal growth data when compared to LMM. Because it simultaneously imputes missing covariate data and does not require specification of variance structure as in LMM, MICE may be preferable for imputation in this setting.

## KEYWORDS

foetal growth, Imputation, longitudinal modelling, MICE, statistical methodology, ultrasound

## 1 | INTRODUCTION

Ultrasound measurements collected in pregnancy are the best available tool for diagnosing growth restriction prior to delivery and

have proven useful for research studies investigating mechanisms of foetal growth restriction and risk factors. Longitudinal measurements allow for investigation of rates of change in growth, identification of tangential points when growth becomes perturbed, and

**TABLE 1** Gestational age and foetal growth measurements by visit in the original (N = 1003), LMM imputed (N = 897), and MICE imputed (N = 1003) study populations

	N missing (%)	Original mean (SD)	LMM mean (SD)	MICE mean (SD)
<b>Visit 1</b>				
Gestational age (wk)	86 (8.6)	10.6 (2.1)		10.6 (2.1)
Crown-rump length (mm)	86 (8.6)	41.2 (21.0)		41.1 (21.0)
<b>Visit 2</b>				
Gestational age (wk)	111 (11.1)	18.1 (1.1)	18.1 (1.1)	18.1 (1.1)
Abdominal circumference (mm)	112 (11.2)	126.7 (14.7)	126.5 (14.4)	126.4 (14.5)
Head circumference (mm)	111 (11.1)	148.5 (14.7)	148.3 (14.3)	148.3 (14.5)
Femur length (mm)	111 (11.1)	26.5 (3.7)	26.5 (3.6)	26.5 (3.7)
Estimated foetal weight (grams)	112 (11.2)	254.5 (60.2)	253.6 (58.9)	253.2 (59.0)
<b>Visit 3</b>				
Gestational age (wk)	560 (55.8)	26.4 (1.9)	26.4 (1.9)	26.3 (1.9)
Abdominal circumference (mm)	560 (55.8)	221.6 (23.7)	222.1 (23.5)	220.6 (22.7)
Head circumference (mm)	560 (55.8)	246.9 (22.8)	247.2 (21.7)	246.5 (21.6)
Femur length (mm)	560 (55.8)	49.4 (5.4)	49.5 (5.2)	49.2 (5.3)
Estimated foetal weight (grams)	601 (59.9)	1049.8 (266.0)	1056.6 (293.0)	995.0 (273.9)
<b>Visit 4</b>				
Gestational age (wk)	417 (41.6)	35.2 (1.7)	35.2 (1.7)	35.1 (1.8)
Abdominal circumference (mm)	417 (41.6)	317.5 (25.2)	317.5 (23.7)	318.0 (24.8)
Head circumference (mm)	428 (42.7)	316.5 (13.8)	316.5 (13.0)	316.5 (13.8)
Femur length (mm)	417 (41.6)	67.8 (4.2)	67.7 (4.0)	67.5 (4.6)
Estimated foetal weight (grams)	417 (41.6)	2684.5 (524.8)	2619.3 (508.8)	2684.3 (517.8)
<b>Delivery</b>				
Final gestational age (wk)	1 (0.1)	38.5 (2.3)	38.6 (2.3)	38.5 (2.3)
Birthweight (grams)	0 (0)	3233.6 (631.6)	3241.9 (623.1)	3233.6 (631.6)

The sample size reduces because for LMM method we only used patients who have complete covariates and at least one gestational age observed whereas MICE imputes the growth parameters and covariates sequentially.

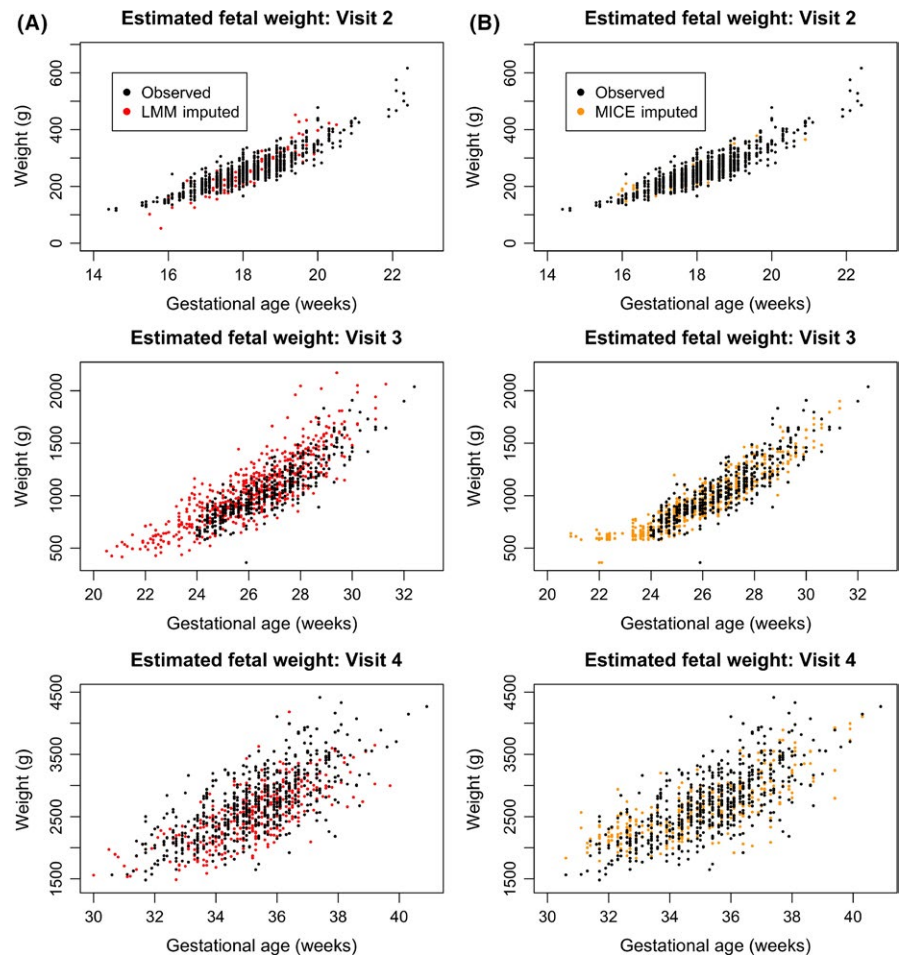
the examination of multiple parameterisations of growth (eg, head size in addition to weight). These facets add significant understanding beyond what was capable with the examination of birthweight alone.

Numerous pregnancy cohorts incorporate longitudinal ultrasound information in order to assess research questions pertaining to foetal growth.<sup>1-4</sup> A common problem in these studies is missingness in ultrasound measurements at one or more time points. Complete case analyses will have reduced precision and may incur bias if the missing data mechanism depends on the missing growth features themselves. Thus, imputation is particularly valuable under this context.

One approach to model this data is to create linear mixed effects models (LMM) with polynomial or spline terms for gestational age that capture non-linearity in the growth trajectory and to then extract predicted values for individuals from these models, either at all time points or at missing time points. This is common for parametric analysis of the growth curve, although less commonly used as a tool for imputation. While appropriate for use in this setting, LMM can be computationally intensive and

require careful specification of the mean and variance structure. In addition, LMM cannot simultaneously handle imputation of missing covariate data.

An alternative approach may exist in the Multiple Imputation by Chained Equation (MICE) methodology. MICE imputes missing values in a dataset under conditionally specified models using a Bayesian sampling framework, and can be implemented easily using the R package *mice*.<sup>5,6</sup> MICE has the advantage of imputing both covariate and foetal growth data during the imputation process, with a sequential conditional imputation strategy cycling iteratively through univariate imputation models. MICE is applied to impute a variety of types of missing data but, to our knowledge, has not been examined in the context of missing foetal growth measurements. Presumably, this is due to the perception that the methods developed for creating foetal growth curves more accurately predict missing growth measurements as a trajectory over time compared to algorithms implemented in MICE, which does not utilise temporal information in an explicit way. However, the comparative performance of these two different methods has not been examined quantitatively.



**FIGURE 1** Observed versus imputed measurements for A, LMM compared to B, MICE imputed measurements of estimated foetal weight

## 2 | METHODS

LIFECODES is a prospective birth cohort conducted at Brigham and Women's Hospital (BWH) in Boston, Massachusetts, USA. Women are eligible if they are at least 18 years of age, are recruited prior to 15 weeks' gestation, and plan to deliver at BWH.<sup>7</sup> The present analysis includes data on pregnancies from singleton livebirths delivered 2006-2008 ( $n = 1003$ ).

Gestational age at delivery is calculated by combining self-reported date of last menstrual period with early first trimester ultrasound dating.<sup>7</sup> All pregnancies at BWH receive a second trimester (~18 weeks) ultrasound scan for assessment of anthropometry, including abdominal and head circumference, femur length, and estimated foetal weight (EFW; calculated from the formula of Hadlock).<sup>8</sup> In addition, many pregnancies have additional ultrasounds performed outside of the standard window. For the present analysis, ultrasound scans were abstracted if they were close in time to three study visits (Visits 2-4; median 18, 26, and 35 weeks' gestation). Birthweight was recorded at delivery for all pregnancies. All foetal growth measurements were imputed and modeled in their raw form (ie, not z-scored).

First, we fitted LMM for each anthropometric parameter using the `lme` function in the `nlme` package<sup>9</sup> in R with model selection based on the Akaike Information Criterion (AIC; additional details

in Data S1). We created 50 datasets with imputations for missing measurements based on the fitted model parameters. Gestational age for the missing measurement was randomly drawn from a normal distribution.

Second, we created another set of 50 imputed datasets using the `mice` package in R (Data S2).<sup>5,6</sup> Foetal growth variables used in the imputation procedure for our dataset along with their missing percentages are listed in Table 1. Additionally, missing covariates were imputed including: health insurance provider ( $n = 30$ ); body mass index (BMI,  $n = 14$ ); alcohol use in pregnancy ( $n = 20$ ); and infant sex ( $n = 1$ ). Other covariates without missingness were included as well, consistent with those used in LMM.

We compared the imputation methods by examining empirical associations between growth measurements and factors known to be associated with foetal growth,<sup>7</sup> including: maternal age, BMI, infant sex, smoking, and preeclampsia. For the cross-sectional analysis, we examined all factors in one model in relation to each foetal growth measurement (eg, EFW) separately using linear regression. For the longitudinal analysis, we utilised the same growth curve models that were fitted for the LMM imputation approach as our analysis model. Finally, we examined rate of change in growth by day of gestation by including an interaction term between gestational age and each co-factor in LMM. We performed these analyses using the original data (un-imputed,

**TABLE 2** Adjusted<sup>a</sup> associations between pregnancy factors and estimated foetal weight capturing cross-sectional associations at visit 3 (median 26 wk gestation), longitudinal associations across pregnancy, and rate of change in growth by pregnancy factors per day

	Original data $\beta$ (95% confidence interval)	LMM Imputed $\beta$ (95% confidence interval)	MICE Imputed $\beta$ (95% confidence interval)
Cross-sectional			
Maternal age (y)	1.9 (-0.7, 4.5)	1.8 (-0.3, 3.8)	1.2 (-0.6, 3.1)
Maternal BMI (kg/m <sup>2</sup> )	4.7 (2.5, 6.8)	6.1 (4.3, 7.9)	4.0 (2.4, 5.6)
Infant sex	-2.5 (-27.3, 22.2)	-41.1 (-60.3, -21.9)	-9.2 (-27.4, 8.9)
Maternal smoking	-9.8 (-61.5, 41.9)	-45.0 (-88.8, -1.2)	5.9 (-35.1, 46.8)
Preeclampsia	3.1 (-38.3, 44.6)	-22.3 (-59.6, 15.0)	-6.0 (-37.7, 25.8)
Longitudinal			
Maternal age (y)	0.6 (-2.1, 3.3)	0.5 (-2.1, 3.0)	1.2 (-1.1, 3.4)
Maternal BMI (kg/m <sup>2</sup> )	5.4 (3.1, 7.7)	4.1 (1.8, 6.4)	5.1 (3.1, 7.1)
Infant sex	-42.7 (-67.8, -17.5)	-40.9 (-65.1, -16.7)	-40.7 (-62.0, -19.4)
Maternal smoking	-43.7 (-100, 12.9)	-34.5 (-89.7, 20.7)	-26.2 (-74.2, 21.8)
Preeclampsia	-9.9 (-57.7, 38.0)	-8.2 (-55.2, 38.8)	-13.1 (-54.5, 28.2)
Rate of change by day <sup>b</sup>			
Maternal Age $\times$ GA	0.4 (0.23, 0.58)	0.4 (0.2, 0.5)	0.3 (0.2, 0.5)
Maternal BMI $\times$ GA	0.6 (0.39, 0.75)	0.5 (0.4, 0.7)	0.5 (0.3, 0.7)
Infant gender $\times$ GA	-5.5 (-7.4, -3.5)	-5.3 (-7.0, -3.6)	-4.9 (-6.8, -3.0)
Maternal smoking $\times$ GA	-6.8 (-11.2, -2.4)	-6.9 (-10.9, -3.0)	-4.9 (-9.3, -0.6)
Preeclampsia $\times$ GA	-1.1 (-5.1, 3.0)	-0.5 (-4.0, 3.0)	-0.4 (-4.2, 3.4)

<sup>a</sup>All models additionally adjusted for maternal race, health insurance provider.

<sup>b</sup>A separate model with covariate by time interaction was fitted for each covariate of interest. In particular, the model for the growth parameter  $Y_{ij}$  is specified as  $Y_{ij} = \beta_0 + \beta_1 t_{ij} + \beta_2 t_{ij}^2 + \beta_3 X_i + \beta_4 X_i \cdot t_{ij} + \gamma C_i + b_{0i} + b_{1i} t_{ij} + \epsilon_{ij}$ , where  $t_{ij}$  denotes gestational age of  $i$ -th subject at  $j$ -th visit,  $X_i$  denotes the covariate of interest (eg, maternal age),  $C_i$  denotes the rest of the covariates of  $i$ -th subject.

complete-case analysis), LMM-imputed data, and MICE-imputed data for comparison. Methods for pooling estimates from imputed datasets are described in Data S3 and were the same for LMM and MICE.

### 3 | RESULTS

Characteristics of the study population have been presented previously.<sup>7</sup> Only ~10% of participants were missing measurements from Visit 2 and ~56% and 40% from Visits 3 and 4, respectively (Table 1). All participants had birthweight recorded at delivery.

Following imputation, mean and standard deviation of growth measurements by either approach were similar to those observed in the original data (Table 1, Figures S1 and S2 in Data S4), except for EFW, where imputed LMM data were generally greater in magnitude than observed measurements from Visit 3 and were more varied (Figure 1). Differences in imputed measurements were generally normally distributed and similar across the duration of pregnancy (Figure S3 in Data S4).

Table 2 shows a condensed set of associations between maternal and pregnancy factors and EFW by imputation method for: 1) Cross-sectional analyses (Visit 3 only); 2) Repeated measures analysis; and 3) Rate of change analysis. Expanded results with

additional visits (cross-sectional) and other ultrasound parameters (all methods) are presented in Tables S1-S4 in Data S4. We expected to see similar effect estimates with more precision in both imputed models, and this was largely the case. For Visit 3, confidence intervals for almost all associations were narrowed in LMM- or MICE-imputed models compared to those observed in the original data, with tighter intervals observed in LMM-imputed models. In longitudinal models, precision was generally, but not always, improved in LMM- and MICE-imputed models. In models of weight, precision was best in MICE-imputed models. Finally, in rate of change analyses, we observed very similar interaction terms and precision in the original data as compared to LMM- and MICE-imputed data (Table 2).

### 4 | COMMENT

Ultrasound measures of foetal growth during gestation have the potential to shed light on the mechanisms and causes of growth restriction and overgrowth. These data are more precise, can capture individual anthropometric measurements (eg, head circumference), and allow for calculation of rates of change in pregnancy or specific windows of vulnerability.<sup>10-13</sup> While some studies have utilised imputation for ultrasound data, most

perform complete-case analyses or methods that handle drop-outs when they are missing at random.<sup>3,14,15</sup> This often limits our power to characterise windows of susceptibility. Acknowledging that MICE is a suitable approach in this setting could extend its use and the knowledge to be gained from analyses using ultrasound data.

The choice to impute over using complete case-analyses or missing outcome data also depends on the amount of missingness in the dataset and whether missingness is dependent on the observed variables (missing at random) or unobserved outcome (not missing at random). Our improved precision rarely identified associations that would have gone unnoticed in complete-case results, which may be a result of low proportions of missing data. We cannot comment on the bias incurred under each approach without a gold-standard. However, both imputation methods can handle data that are missing at random. We posit that using an imputed dataset may be less biased in this particular setting (ie, when ultrasounds are collected for clinical rather than research purposes) because some measurements may be collected specifically because the clinician anticipates undergrowth of the foetus. Finally, it should be clear that while imputation is appropriate for epidemiologic research for exploring associations, it is not designed to be used for individual-level prediction.

## 5 | CONCLUSION

In a comparison of two approaches for imputing missing ultrasound measurements of foetal growth, we found the MICE method performed as well as the LMM approach, with results from both imputed datasets showing improved precision compared to complete-case analyses. Because of the relative ease of implementing MICE, since it does not require specification of non-linear terms to capture the growth trajectory or specification of the variance-covariance structure, and because it can simultaneously handle imputation of missing covariate data, MICE can be applied in future work utilising ultrasound data of foetal growth with missing measurements. Future longitudinal studies of foetal growth may consider MICE as an imputation approach, although caution should be taken with extension to other populations.

## ACKNOWLEDGEMENTS

This work was funded by the National Institute of Environmental Health Sciences, National Institutes of Health (ZIA103321, R01ES018872, P42ES017198, P30ES017885). Support for K. Ferguson was provided by the Intramural Research Program of the National Institute of Environmental Health Sciences.

## ORCID

Kelly K. Ferguson  <http://orcid.org/0000-0001-8467-3250>

## REFERENCES

1. Aguilar I, Garcia-Esteban R, Iñiguez C, et al. Prenatal exposure to traffic-related air pollution and ultrasound measures of fetal growth in the INMA Sabadell cohort. *Environ Health Perspect*. 2010;118:705.
2. Grewal J, Grantz KL, Zhang C, et al. Cohort profile: NICHD fetal growth studies-singletons and twins. *Int J Epidemiol*. 2018;47:25-251.
3. Nguyen PH, Addo OY, Young M, et al. Patterns of fetal growth based on ultrasound measurement and its relationship with small for gestational age at birth in rural Vietnam. *Paediatr Perinat Epidemiol*. 2016;30:256-266.
4. American College of Obstetricians and Gynecologists. ACOG practice bulletin. Diagnosis and management of preeclampsia and eclampsia. Number 33, January 2002. American College of Obstetricians and Gynecologists. *Int J Gynaecol Obstet*. 2002;77:67.
5. Bondarenko I, Raghunathan T. Graphical and numerical diagnostic tools to assess suitability of multiple imputations and imputation models. *Stat Med*. 2016;35:3007-3020.
6. van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. *J Stat Softw*. 2011;45:1-67.
7. Cantonwine DE, Ferguson KK, Mukherjee B, et al. Utilizing longitudinal measures of fetal growth to create a standard method to assess the impacts of maternal disease and environmental exposure. *PLoS One*. 2016;11:e0146532.
8. Hadlock FP, Harrist RB, Martinez-Poyer J. In utero analysis of fetal growth: a sonographic weight standard. *Radiology*. 1991;181:129-133.
9. Pinheiro J BD, DebRoy S, Sarkar D; R Core Team. nlme: linear and nonlinear mixed effects models. 2017; R package version 3.1-131, <https://CRAN.R-project.org/package=nlme>. Accessed 07/09/18.
10. Manzano-Salgado CB, Casas M, Lopez-Espinosa M-J, et al. Prenatal exposure to perfluoroalkyl substances and birth outcomes in a Spanish birth cohort. *Environ Int*. 2017;108:278-284.
11. Zhang C, Hediger ML, Albert PS, et al. Association of maternal obesity with longitudinal ultrasonographic measures of fetal growth: findings from the NICHD fetal growth studies-singletons. *JAMA Pediatr*. 2017;172:24-31.
12. Zheng T, Zhang J, Sommer K, et al. Effects of environmental exposures on fetal and childhood growth trajectories. *Ann Glob Health*. 2016;82:41-99.
13. Jaddoe VW, Verburg BO, De Ridder M, et al. Maternal smoking and fetal growth characteristics in different periods of pregnancy: the generation R study. *Am J Epidemiol*. 2007;165:1207-1215.
14. Casas M, Valvi D, Ballesteros-Gomez A, et al. Exposure to bisphenol A and phthalates during pregnancy and ultrasound measures of fetal growth in the INMA-Sabadell cohort. *Environ Health Perspect*. 2016;124:521.
15. van den Hooven EH, Pierik FH, de Kluizenaar Y, et al. Air pollution exposure during pregnancy, ultrasound measures of fetal growth, and adverse birth outcomes: a prospective cohort study. *Environ Health Perspect*. 2012;120:150.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Ferguson KK, Yu Y, Cantonwine DE, McElrath TF, Meeker JD, Mukherjee B. Foetal ultrasound measurement imputations based on growth curves versus multiple imputation chained equation (MICE). *Paediatr Perinat Epidemiol*. 2018;32:469-473. <https://doi.org/10.1111/ppe.12486>