Analysis of Bivariate Censored Longitudinal Data: A Case Study

Tingzhi Fang

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science

University of Washington

2018

Committee:

Brian G. Leroux

James P. Hughes

Program Authorized to Offer Degree:

Department of Biostatistics

University of Washington

Abstract

Analysis of Bivariate Censored Longitudinal Data: A Case Study

Tingzhi Fang

Chair of the Supervisory Committee:

Brian G. Leroux

Department of Biostatistics

Laboratory measurements that are below the limit of detection (LOD) are common in continuous longitudinal data in the biomedical sciences. The presence of these left-censored values, or non-detects (NDs), complicates the statistical analysis of such data since improper treatments of NDs may lead to loss of power or biased results. This thesis aims at investigating the effects of high percentages (>50%) of NDs on quantifying the magnitude of dental Bisphenol A (BPA) exposure, measured by urinary BPA concentrations, before and after treatment in children treated with Bisphenol A glycidyl methacrylate (BisGMA) - based dental materials. Data analysis results using the data handled with different statistical methods, including naïve substitution, Paxton's random imputation procedure, modified Paxton's method, and multiple imputation, were compared. In conclusion, our results of all methods suggest that dental treatment using BisGMA-based materials is associated with elevated uBPA concentrations.

1 Introduction

In the biomedical sciences, one of the analysis issues that arises in laboratory-based research is that the measurements may be below the limit of detection (LOD) and thus subject to left-censoring. Such measurements are called non-detects (ND). Because the LOD can vary from laboratory to laboratory depending on the testing methods being implemented, failure to account for the NDs may lead to considerably biased parameter estimates and variance components from a statistical model (Hughes, 1999). Many studies have been done to investigate and compare different statistical approaches to dealing with left-censoring. Deleting the NDs is the simplest approach, but it leads to overestimation by only using the measurements that are above LOD (Baccarelli et al., 2005) and substantial reduction in power especially for datasets with large percentage of NDs (Arunajadai & Rauh, 2012).

The most widely used method in practice is naïve substitution, which assigns a single value, such as LOD, LOD/2, or LOD/$\sqrt{2}$, to replace each of the NDs (Arunajadai & Rauh, 2012; Baccarelli et al., 2005). Nevertheless, naive substitution lacks theoretical foundation (Helsel, 1990) and might give very different results compared with other available methods to handling NDs (Arunajadai & Rauh, 2012).

Maximum Likelihood Estimation (MLE) uses the uncensored observations and the proportion of left-censored observations to compute estimates of summary statistics. It also relies on a distributional assumption that the data come from a specified parametric distribution. A lognormal distribution is commonly assumed with environmental sciences data, while other distributional assumptions could also be considered (Bolks, DeWire, & Harcum, 2014).

Paxton et al. (1997) applied a random imputation procedure that combines the merits of naïve substitution and MLE to their longitudinal left-censored virologic measures data.

Particularly, they first replaced all the NDs with LOD/2 and fitted an appropriate random-effects regression model on the complete data. Then, they used maximum likelihood to fit a normal distribution to the residuals of the regression, and substituted a random number drawn from a specific truncated normal distribution for each of the censored residuals. The new set of residuals were added back to the fitted values of the regression at last to form an imputed dataset. Paxton's random imputation procedure gives less biased estimates for fixed effects as compared with the naïve substitution methods (Hughes, 1999). However, it does not adjust the uncertainty due to censored values.

Hughes (1999) proposed a modified method from the EM algorithm for fitting mixed effects models with normal errors (Dempster, Rubin, & Tsutakawa, 1981; Laird and Ware, 1982) by using a Monte Carlo procedure, the Gibbs sampler. The Monte Carlo expectation maximization (MCEM) method obtains maximum likelihood estimates of parameters of the mixed effects models with right- and/or left-censored correlated data. As is suggested by Fu et al. (2016), the MCEM method provides less biased estimates than other commonly used ad hoc procedures used to adjust for censoring. According to the simulation studies conducted by Hughes (1999), MCEM also performs better than Paxton's method, since MCEM gives less biased fixed effect estimates and variance components. Fortran code that has a R calling routine to implement the MCEM method can be obtained online at: http://faculty.washington.edu/jphughes/pubs.html. One limitation of Hughes' MCEM algorithm is that it does not accept autoregressive error (Jacqmin-Gadda & Thibaut, 2000).

Multiple imputation (MI) is a method that can potentially address the limitations of the Paxton single imputation method. Results from simulation studies from Fu et al. (2016) suggests that MI performs better than naïve substitution and is comparable with MCEM regarding certain

parameter estimates of a mixed effects model. MI replaces each ND with multiple values drawn from an assumed parametric distribution and hence takes account of the uncertainty due to the imputed values. Specifically, MI creates multiple datasets by iterating single imputation for the NDs and then obtains parameter estimates and variance components from analyses of the complete (imputed) data (Rubin, 1987). The pooled estimate and total variance of the final estimate properly reflects the uncertainty due to the NDs, assuming that the underlying distributional assumption is true (Baccarelli et al., 2005).

2 Case Study

This thesis investigates how bivariate left-censoring on laboratory measurements affects the statistical analysis through the Bisphenol A and dental restoration clinical study in children (BPARCS).

2.1 Study Background

Bisphenol A (BPA) is a byproduct of Bisphenol A glycidyl methacrylate (BisGMA)-based dental composite restorative materials and sealants. Previous studies have suggested increased health risks of BPA exposure on brain, endocrine, and reproductive development (Inadera, 2015). However, formal biological toxicity testing was not performed when BisGMA was developed in 1962, and the national dental organizations do not consider the very small quantities of BPA from BisGMA-based materials to have health impacts on humans. BisGMA-based dental materials are widely used among children, and preliminary data suggest dental sealants may account for 17% to 37% of BPA exposure in children (Von Goetz et al., 2010; Sasaki et al., 2005). Besides BisGMA-based dental materials, the use of plastics in general anesthesia (GA) for dental treatment is also a potential source of dental-related BPA exposure. Once the preliminary data is confirmed, a better understanding of dental sources of BPA exposure may be important in understanding health effects of dental treatment in children.

Previous in vitro studies have shown that BPA can leach from resin-based dental materials (Van Landuyt et al., 2011). A systematic review conducted by Marzouk et al. (2018) also concludes that uBPA concentration consistently increases up to 24 hours after dental treatment, and the concentration may remain elevated at 7 days after treatment with different types of resin-based materials. The only prospective study in children (Martin et al., 2005)

indicates that the magnitude of post-treatment urinary BPA (uBPA) concentrations are 2 to 6-fold higher as compared with that of pre-treatment concentrations. Clinical studies suggest that exposure to medical products made from polycarbonate plastic such urinary catheters, gastrointestinal tubes, and cardiopulmonary bypass circuits is associated with increased BPA concentration (Duty et al., 2013; Calafat et al., 2009). Nevertheless, no studies can be found that directly evaluate the BPA exposure from medical products used in anesthesia for dental treatment in children.

A single-site prospective cohort study was conducted with the following aims: (1) quantify the magnitude of dental BPA exposure by measuring uBPA before and after treatment, (2) determine associations between the number of surfaces treated with BisGMA-based dental materials (restorative composites and dental sealants combined) and uBPA, and (3) determine associations between type of anesthesia and uBPA. The study aimed to enroll dental patients aged between 4 and 9 years old (total N = 210), who are treated with either less than four surfaces (N =105) or greater than or equal to four surfaces (N =105) with BisGMA-based dental materials at the University of Washington Center for Pediatric Dentistry. Within each of these two groups, the study aimed to recruit patients who received: no anesthesia (N = 35), nitrous oxide (N = 35); and general anesthesia (N = 35). The study participants had two study visits prior to treatment, and four study visits from 24 hours to 16 weeks after having undergone treatment.

2.2 Data Collection

Data on demographic characteristics were gathered at baseline. Urine samples were collected at each study visit and then evaluated at the Environmental Health Laboratory and Trace Analytical Center at the University of Washington for processing. The samples were delivered to the

laboratory in 4 batches; however, only the first 3 batches containing 942 samples from 186 study participants were available for this analysis. Each sample was divided in two and analyzed before (uBPAb) and after (uBPAa) the enzymatic deconjugation. The difference between uBPAb and uBPAa reflects the amount of BPA (dBPA) in the urine sample. Measurements of uBPA were constrained by the limit of detection (LOD) and the LOD varied by batch. The uBPA concentrations cannot be measured with certainty if they are below the LOD.

2.3 Descriptive Statistics

The LOD is 0.2 ng/ml for uBPAa and uBPAb evaluated in batch 1 and batch 2. The LOD is 0.2 ng/ml for uBPAb and 0.7 ng/ml for uBPAa evaluated in batch 3. The distribution of NDs by visit and batch is shown in the frequency table below. It was observed that earlier batches, especially batch 1, tend to have more samples from earlier visits than the later batches, which implies informative censoring. This will be accounted for by adding batch to the regression analysis model as a fixed effect.

# of NDs \ Visit		Visit 1 (N = 178)	Visit 2 (N = 169)	Visit 3 (N = 157)	Visit 4 (N = 155)	Visit 5 (N = 151)	Visit 6 (N = 132)
Batch 1 (N = 313)	uBPA Before	69	66	52	48	48	9
	uBPA After	42	35	27	28	37	8
Batch 2 (N = 156)	uBPA Before	23	18	22	24	20	40
	uBPA After	15	12	16	19	12	25
Batch 3 (N = 473)	uBPA Before	50	52	42	50	50	55
	uBPA After	43	42	28	36	47	51

Summary statistics for the numbers and percentages of NDs by visit are shown in the tables below. The percentage of NDs for both of the two measurements tend to be the lowest at visit 3.

Visit / # of NDs (%)	Visit 1 (N = 178)	Visit 2 (N = 169)	Visit 3 (N = 157)	Visit 4 (N = 155)	Visit 5 (N = 151)	Visit 6 (N = 132)
uBPA Before	142 (79.8%)	136 (80.5%)	116 (73.9%)	122 (78.7%)	118 (78.1%)	104 (78.8%)
uBPA After	100 (56.2%)	89 (52.7%)	71 (45.2%)	83 (53.5%)	96 (63.6%)	84 (63.6%)

Summary statistics for the numbers and percentages of NDs across batches were also examined. For both uBPA before and after measurements, batch 3 consists of lower percentage of NDs as compared with batch 1 and batch 2, as seen in the tables below.

Batch / # of NDs (%)	Batch 1 (N = 313)	Batch 2 (N = 156)	Batch 3 (N = 473)
uBPA Before	292 (93.3%)	147 (94.2%)	299 (63.2%)
uBPA After	177 (56.5%)	99 (63.5%)	247 (52.2%)

In conducting descriptive analysis, the data was handled by naïve substitution by replacing NDs with LOD/2. Frequency distributions of logBPAb and logBPAa measurements are shown in Figure 1 and Figure 2, respectively. The logBPAb has most measurements being NDs, thus has a large peak at log(0.2/2). The histogram of logBPAa has two peaks at log(0.2/2) and log(0.7/2). The detectable measurements are approximately normally distributed on the log scale for both uBPA before and after. It is also observed that logBPAb values are generally smaller than logBPAa values.

The overall mean is -1.98 (SD = 0.65) for logBPAb and -0.57 (SD = 1.50) for logBPAa. Descriptive statistics for the means and standard deviations (SD) of logBPAb and logBPAa by visit and batch are shown below. The cells with mean value of -2.30 and SD of 0.00 have all measurements being NDs.

Mean(SD) / Visit		Visit 1 (N = 178)	Visit 2 (N = 169)	Visit 3 (N = 157)	Visit 4 (N = 155)	Visit 5 (N = 151)	Visit 6 (N = 132)
Batch 1 (N = 313)	logBPAb	-2.30 (0.00)	-2.27 (0.19)	-2.21 (0.35)	-2.25 (0.27)	-2.24 (0.31)	-2.16 (0.44)
	logBPAa	-1.55 (1.37)	-1.58 (1.18)	-0.58 (1.92)	-1.49 (1.44)	-1.84 (1.07)	-2.30 (0.00)

Batch 2 (N = 156)	logBPAb	-2.27 (0.18)	-2.30 (0.00)	-2.24 (0.29)	-2.26 (0.22)	-2.30 (0.00)	-2.30 (0.00)
	logBPAa	-1.57 (1.17)	-1.97 (0.82)	-1.76 (0.94)	-1.97 (0.80)	-1.60 (1.04)	-1.73 (1.10)
Batch 3 (N = 473)	logBPAb	-1.78 (0.72)	-1.72 (0.89)	-1.62 (0.84)	-1.84 (0.70)	-1.74 (0.85)	-1.86 (0.70)
	logBPAa	-0.22 (0.98)	-0.13 (1.08)	0.54 (1.42)	0.14 (1.29)	-0.21 (1.14)	-0.29 (1.10)

Descriptive statistics for the means and standard deviations (SD) of logBPAb and logBPAa by visit are shown below. The means of logBPAa are greatest at visit 3 and relatively close at the other five visits. The means of logBPAb are close at all visits.

Visit / Mean(SD)	Visit 1 (N = 178)	Visit 2 (N = 169)	Visit 3 (N = 157)	Visit 4 (N = 155)	Visit 5 (N = 151)	Visit 6 (N = 132)
logBPAb	-2.03 (0.57)	-2.00 (0.69)	-1.91 (0.71)	-2.02 (0.59)	-1.94 (0.73)	-2.00 (0.61)
logBPAa	-0.70 (1.40)	-0.65 (1.36)	-0.17 (1.76)	-0.47 (1.63)	-0.77 (1.44)	-0.70 (1.32)

For both before and after measurements, logBPA samples from batch 2 have the lowest mean values, while those from batch 3 have the highest mean values. It is not surprising that logBPAa values using naïve substitution from batch 3 have the highest mean value, because of its higher LOD of 0.7 ng/ml. Summary of descriptive statistics of logBPAb and logBPAa by batch are shown below.

Batch / Mean (SD)	Batch 1 (N = 313)	Batch 2 (N = 156)	Batch 3 (N = 473)
logBPAb	-2.20 (0.39)	-2.22 (0.30)	-1.76 (0.79)
logBPAa	-0.95 (1.71)	-1.46 (1.22)	-0.03 (1.20)

2.4 Scientific Questions and Statistical Models

The primary objective of the study is to quantify the magnitude of change in dBPA (the difference between uBPAb and uBPAa) over time, regardless of the number of surfaces treated with BisGMA-based dental materials and the type of anesthesia. The hypothesis is that children

will have consistently increasing dBPA concentrations up to 24 hours (visit 3), and then the elevated dBPA concentrations are likely to go down at 1 week (visit 4) after last treatment, as compared with pre-treatment (visits 1 and 2). We are especially interested in examining the magnitude of change in dBPA from pre-treatment to visit 3, as well as that from pre-treatment to visit 4. We employed a linear mixed model (LMM) with random intercepts to account for nested clustering by individual ID and family (to account for siblings), as well as fixed effects for visit, batch, and baseline characteristics including demographics (age, gender, race/ethnicity, marital status, height, and weight) and socioeconomic status (household income, education, and insurance). Summary statistics about these demographic characteristics are shown in Table 1. Generalized estimating equations (GEE) were also considered for the correlated data. However, in practice, available GEE software only takes into account a single level of nesting (Stoner, Leroux, & Puumala, 2010). The goal of the inferential analysis is to model the mean dBPA on the log scale, where dBPA is measured by the difference between uBPAb and uBPAa (logdBPA). The LMM, described in R notation, is logdBPA ~ visit + batch + other variables + (1|family/ID). The model estimates the parameters using the maximum likelihood method. We used point estimates and the corresponding 95% confidence intervals to describe the magnitude of effects, as well as hypothesis tests with the standard 0.05 level of statistical significance without adjustment for multiple testing.

For secondary objectives (not shown in this thesis), we will employ LMMs with similar structure but with added fixed effects for number of surfaces treated with BisGMA-based dental materials and type of anesthesia. Specifically, the fitted LMM to assess the association between the amount of BisGMA-based dental materials and uBPA will be: logdBPA ~ visit + batch + dental treatment + other variables + (1|family/ID); and the model to evaluate the effect of types

of anesthesia on uBPA, adjusting for the amount of dental materials, will be: logdBPA ~ visit + batch + dental treatment + anesthesia + other variables + (1|family/ID).

2.5 Statistical Challenges to be Addressed in this Thesis

This thesis aims at addressing the following statistical challenges involved in the data analysis of the BPARCS study. First of all, the data is in a longitudinal bivariate structure. Specifically, two uBPA measurements (uBPAb and uBPAb) were collected for each participant at each of the six study visits. Ideally, the dBPA values used for the final data analysis are the log-transformed differences of uBPAb and uBPAa. However, many of these values cannot be calculated because of the presence of the non-detectable uBPAb and/or uBPAa values and also because the uBPAb value could be larger than the uBPAa value. The joint modeling of longitudinal bivariate outcomes is necessary as the two outcomes are likely to be correlated. Secondly, the data is subject to left censoring with 55.5% of uBPAa measurements being NDs, and 78.3% of uBPAb measurements being NDs. The percentages of NDs in this data are considered to be fairly high, which adds further importance to the identification of an appropriate imputation method. Finally, the data has a hierarchical data structure as the 186 participating children are clustered within 161 families (there are 28 families with siblings and 149 families with a single child who participated in the study).

3 Methods

In this section, each of the four statistical methods, including naïve substitution method, Paxton's random imputation procedure, modified Paxton's method, and multiple imputation, will be applied to the BPARCS data to handle the NDs. It is assumed that the log-transformed uBPA measurements are bivariately normally distributed at each study visit.

3.1 Naïve Substitution Method

For the naïve substitution method, the NDs are replaced by LOD/2. Specifically, the left-censored uBPAb values from all batches are replaced by 0.1, the left-censored uBPAa values evaluated at batch 1 and batch 2 are replaced by 0.1, and the left-censored uBPAa values evaluated at batch 3 are replaced by 0.35. The naïve substitution method was applied on the BPARCS data for conducting descriptive statistics of the NDs, as seen in Section 2.3. Because this data has fairly high percentage of NDs, the naïve substitution method leads to large peaks on the LOD/2 values in the distribution of uBPA measurements (see Figures 1 and 2).

3.2 Paxton's Random Imputation Procedure

Paxton's random imputation procedure accounts for the hierarchical data structure and removes much bias in the fixed effect estimate (Hughes, 1999); however, it ignores the correlation between the uBPAb and uBPAa measurements. The Paxton's procedure was firstly employed to impute the censored logBPAb and logBPAa measurements separately. Particularly, linear mixed models (LMM) with random effects for nested clustering groups, individual and family, as well

as fixed effects for visit number, batch number, deconjugation efficiency, and the demographic characteristics, including education, income, marital status, Hispanic, child race, child gender, child age, weight, and height, were fitted to all data with censored observations being replaced with LOD/2. While one of the logBPAb and logBPAa was the outcome variable, the other logBPA measurement was included in the model as a covariate. The residuals from the LMM models were extracted and fitted to a normal distribution using maximum likelihood. Then, a random number drawn from a univariate truncated normal distribution, where the upper limit was set to the difference of log-transformed LOD and the fitted value from LMM models, was used to substitute for each censored residual. The residuals were added back to the fitted values from the LMM model to generate the imputed dataset. As are shown in Figure 3 and 4, the imputed logBPAb and logBPAa values are more evenly distributed comparing with those values from the naïve substitution method, where NDs were simply replaced with LOD/2. Nevertheless, nearly thirty percent of the imputed data (265 out of 898 observations) had uBPAa being smaller than or equal to uBPAb. Because the dBPA, calculated as the difference between uBPAa and uBPAb, is the outcome variable of interest and we are working on the log scale, a missing value was produced when dBPA was non-positive. The missing logdBPAs were replaced with imputed logBPAa values for further inferential analyses.

3.3 Modified Paxton's Method

In order to generate more accurate imputed values for the censored uBPA measurements and reduce the number of missing logBPA values, some modifications were made on Paxton's procedure to account for the inequality and the bivariate relationship between uBPAa and uBPAb. Specifically, the random numbers used to replace the censored residuals were drawn

from a bivariate truncated normal distribution using Gibbs sampling (Wilhelm & Manjunath, 2010), instead of two univariate truncated normal distributions. Rejection sampling was also considered for generating random samples, however, it is inefficient and impractical when the expected number of iterations is large and the acceptance rate is low. Moreover, the inequality between uBPAa and uBPAb was imposed as a restriction for the random draw of numbers, so that only residuals that made the imputed uBPAa greater than the imputed uBPAb were kept. As a result, the percentage of missing logdBPA was reduced to 6.12% (55 out of 898 observations). The missing values were produced because the ideal inequality was not achieved after 10 random draws from the specified bivariate truncated normal distribution. The missing logBPA values were again replaced with imputed logBPAa values for later analyses. The frequency distributions of imputed logBPAb and after values are presented in Figure 5 and 6. The distributions of imputed values using modified Paxton's method are similar to those using Paxton's random imputation procedure, which are also approximately normally distributed.

3.4 Multiple Imputation Method

When using naïve substitution, Paxton's random imputation procedure, or modified Paxton's method, the statistical estimates do not reflect the uncertainty resulting from the loss of information due to the NDs. To account for the uncertainty, M = 5 imputed datasets were created using the Paxton's random imputation procedure and the modified Paxton's method. The LMM was fitted on each of the 5 imputed datasets, then the parameter estimates, 95% CI, and t values were obtained to explore the sensitivity of the results (Table 5 and 7). The pooled MI estimates were summarized by taking the average over the parameter estimates from all 5 imputed datasets,

and the total MI variances were computed using the Rubin's Rules: $T = \bar{W} + \left(1 + \frac{1}{M}\right) B$, where $\bar{W}$ is the average of the within-imputation variances, and B is the sample variance of the estimates (Graham, Olchowski, & Gilreath, 2007).

4 Results

Data analyses results for the primary objective of the study using naïve substitution, Paxton's random imputation procedure, modified Paxton's method, and multiple imputation are compared and shown below. As described in Section 2.4, LMM accounting for random intercepts for clustering by individual and by family and fixed effects of visit, batch, and demographic characteristics was employed to answer the primary objective of the study. The variance components and exponentiated parameter estimates are reported, along with their 95% confidence and p values.

4.1 Results of Naïve Substitution Method

The results of the LMM model using naïve substitution method are reported in Table 2. Adjusting for batch and demographic characteristics, the dBPA concentrations is estimated to be 1.713 (95%CI: 1.327, 2.212) higher at 24 hours since the last treatment (visit 3), as compared with those pre-treatment at visit 1 and visit 2, in children treated with BisGMA-based dental materials. This result is statistically significant at 0.05 significance level (p value < 0.001). The estimated dBPA concentrations is 1.255 (95% CI: 0.971, 1.622) higher at visit 4, which is 1 week from the last treatment, as compared with pre-treatment measurements. This estimate is not statistically significant. It is worth noting that the standard deviation of random effect for family is 0 (95% CI: 0.000, 0.368), which suggests that there is little variation between families. However, the upper limit of the 95% CI for the standard deviation goes up to 0.368; thus we cannot be certain about the conclusion of no familial variation. The random intercept for family remains included in the LMM for comparison with results from other methods.

4.2 Results of Paxton's Random Imputation Method

The results of the LMM model using Paxton's random imputation method are reported in Table 3. The estimated bPBA concentration is 2.132 (95%CI: 1.453, 3.128) higher for post-treatment samples at visit 3, as compared with pre-treatment samples that are in the same batch and are from children with the same demographic characteristics. This result is statistically significant at 0.05 significance level (p value < 0.001). The estimated dBPA concentration is 1.371 (95% CI: 0.932, 2.015) higher at visit 4, as compared with pre-treatment. This estimate is not statistically significant. The standard deviation of random effect for family is again equal to 0 (95% CI: 0.000, 0.584), which suggests that there is no variation between families.

4.3 Results of Modified Paxton's Random Imputation Method

The results of the adjusted LMM model using modified Paxton's random imputation method are reported in Table 4. The estimated bPBA concentration is 1.964 (95%CI: 1.380, 2.796) higher for post-treatment samples at visit 3, as compared with those pre-treatment at visit 1 and visit 2. This result is statistically significant at 0.05 significance level (p value < 0.001). The estimated difference in mean dBPA concentrations is 1.137 (95% CI: 0.797, 1.622) higher at visit 4, as compared with pre-treatment. This estimate is not statistically significant. The estimated standard deviation of the random intercept for family is 0.093 (0.000, 0.662), which indicates some familial variation. The individual level variation suggested by the model using modified Paxton's method is also higher than those from naïve substitution method and Paxton's random imputation procedure.

4.4 Results of Multiple Imputation Method

The parameter estimates, 95% confidence intervals, and t values for visit 3 from each of the M=5 imputations using Paxton's random imputation procedure are presented in Table 5. The overall results of the multiple imputation using Paxton's random imputation procedure for all of the parameters are reported in Table 6. The combined exponentiated parameter estimate for visit 3 is 2.584 (95%CI: 1.593, 4.194). The overall standard error of visit 3 (0.25) is close to the standard error of that parameter from each of the single imputations, which suggests that M=5 is sufficient for the multiple imputation. The combined exponentiated parameter estimate for visit 4 is 1.415 (95%CI: 0.888, 2.254). The results for visit 3 from each of the M=5 imputations using modified Paxton's method are shown in Table 7. The overall results of the multiple imputation using modified Paxton's method for all of the parameters are reported in Table 8. The combined exponentiated parameter estimate for visit 3 is 1.912 (1.277, 2.862). The overall standard error for visit 3 is 0.21, which is also similar to the standard error from each of the M imputations and therefore confirms our choice of M=5 is sufficient. The combined exponentiated parameter estimate for visit 4 is 1.360 (95%CI: 0.886, 2.088).

5 Summary and Discussion

In this thesis, we compared four statistical approaches to handling left-censoring in longitudinal data, namely, naïve substitution method, Paxton's random imputation procedure, modified Paxton's method, and multiple imputation.

Comparisons of the results of visit 3 and visit 4 from the different methods are shown in Table 9 and Table 10, respectively. Previous studies have shown that BPA is elevated after dental treatment. Our results from all methods suggest that controlling for batch and demographic characteristics, children have significantly increased dBPA 24 hours after dental treatment with BisGMA-based material, which is consistent with previous findings. The parameter estimates for visit 3 vary among different statistical approaches to handling the NDs. Specifically, the LMM using naïve substitution suggests the smallest elevation of dBPA (1.713 higher), while the multiple imputation using Paxton's random imputation procedure suggests the largest elevation of dBPA (2.584) post dental treatment. We cannot be certain which method performs best; nevertheless, the estimate from multiple imputation using Paxton's random imputation procedure is noticeably higher than the estimates from LMMs using other methods, and the corresponding 95% CI for that estimate is also considerably wider than that of visit 3 estimates using other methods. We believe that MI for Paxton's random imputation procedure did not work well for the BPARCS data, and the MI for modified Paxton's method is an improvement of the original Paxton's procedure. The naïve substitution method and the Paxton's random imputation procedure do not properly reflect the variation between the families since the standard deviation of random intercept for family from the LMM fitted on data with NDs replaced by LOD/2 is 0. The modified Paxton's method captures the familial variation, and the standard deviation of random intercept for family from the LMM is 0.093.

The modified Paxton's method accounts for the bivariate structure of the BPARCS data. However, one of the limitations of the modified Paxton's method is that it is based on the distributional assumption that the data come from a bivariate normal distribution. If the assumption is violated, the modified Paxton's method might not perform well.

Other methods that could be applied on this data include robust regression on order statistics (Bolks, DeWire, & Harcum, 2014), generalized Tobit regression (Fu et al., 2016), and Hughes' MCEM (Hughes, 1999). Hughes' MCEM performed better than Paxton's random imputation procedure in the setting of univariately normal distributed data (Hughes, 1999). Nevertheless, the parameter estimates from Paxton's random imputation and Hughes' MCEM were comparable across simulated data with different percentages of censoring, and these two methods clearly outperform naïve substitution as the percentage of censoring increases. The percentage of NDs in our data is higher than 50%, which is the highest percentage of left-censoring simulated in Hughes (1999), so the performance of Paxton's random imputation procedure in comparison with Hughes MCEM method is unknown. Moreover, we modified Paxton's random imputation procedure by changing the distribution from which the censored residuals are drawn and adding the inequality between uBPA before and after measurements. It's unknown whether our modified Paxton's method, and the multiple imputation, would outperform Hughes' MCEM or not. This is a topic for future research.

Acknowledgements

This work was supported in part by National Institute of Dental and Craniofacial Research (NIDCR) grant 1R01DE025229-01. I thank Dr. Christy McKinney for use of the BPARCS study data.

References

1. Inadera, H. (2015). Neurological Effects of Bisphenol A and its Analogues. *International Journal of Medical Sciences, 12*(12), 926–936.
2. Hughes, J. (1999). Mixed Effects Models with Censored Data with Application to HIV RNA Levels. *Biometrics, 55*(2), 625-629.
3. Stoner, J. A., Leroux, B. G., & Puumala, M. (2010). Optimal Combination of Estimating Equations in the Analysis of Multilevel Nested Correlated Data. *Statistics in Medicine*, *29*(4), 464–473.
4. Dempster, A. P., Rubin, D. B., and Tsutakawa, R. K. (1981). Estimation in covariance components models. *Journal of the American Statistics Association 76*, 341-353.
5. Laird, N. M. and Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics 38*, 963-974.
6. Jacqmin-Gadda, H., Thibaut, R., Chene, G., & Commenges, D. (2000). Analysis of left-censored longitudinal data with application to viral load in HIV infection. *Biostatistics,1*(4), 355-368.
7. Bolks, A., DeWire, A., and Harcum, J.B. (2014). Baseline assessment of left-censored environmental data using R. Tech Notes 10. Developed for U.S. Environmental Protection Agency by Tetra Tech, Inc., Fairfax, VA, 28 p.
8. Paxton, W.B., Coombs, R.W., McElrath, M.J., Keefer, M.C., Hughes, J., Sinangil, F., . . . Corey, L. (1997). Longitudinal analysis of quantitative virologic measures in human immunodeficiency virus-infected subjects with ⩾ 400 CD4 lymphocytes: Implications for applying measurements to individual patients. *The Journal of Infectious Diseases*, 175(2), 247-54.
9. Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed□effects models. *Methods in Ecology and Evolution, 4*(2), 133-142.
10. Von Goetz, N., Wormuth, M., Scheringer, M., & Hungerbühler, K. (2010). Bisphenol A: How the Most Relevant Exposure Sources Contribute to Total Consumer Exposure. *Risk Analysis, 30*(3), 473-487.
11. Sasaki, N., Okuda, K., Kato, T., Kakishima, H., Okuma, H., Abe, K., . . . Kubono, K. (2005). Salivary bisphenol-A levels detected by ELISA after restoration with composite resin. *Journal of Materials Science: Materials in Medicine,16*(4), 297-300.
12. Van Landuyt, K.L., Nawrot, T., Geebelen, B., De Munck, J., Snauwaert, J., Yoshihara, K., . . . Van Meerbeek, B. (2011). How much do resin-based dental materials release? A meta-analytical approach. *Dental materials: official publication of the Academy of Dental Materials.* 2011; 27(8): 723-747.
13. Martin, M., Bajet, D., Woods, J., Dills, R., & Poulten, E. (2005). Detection of dental composite and sealant resin components in urine. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology, 99*(4), 429.

14. Duty, S., Mendonca, K., Hauser, R., Calafat, A., Ye, X., Meeker, J., . . . Ringer, S. (2013). Potential sources of bisphenol A in the neonatal intensive care unit. *Pediatrics, 131*(3), 483-9.

15. Calafat, A.M., Weuve, J., Ye, X., Jia, L.T., Hu, H., Ringer, S., Huttner, K., & Hauser, R. (2009). Exposure to bisphenol A and other phenols in neonatal intensive care unit premature infants. *Environmental Health Perspectives, 117*(4), 639-644.

16. Fu, P., Hughes, J., Zeng, G., Hanook, S., Orem, J., Mwanda, O.W., & Remick, S.C. (2016). A comparative investigation of methods for longitudinal data with limits of detection through a case study. *Statistical Methods in Medical Research, 25*(1), 153-166.

17. Helsel, D.R., and Cohn, T.A. (1988). Estimation of descriptive statistics for multiply-censored water-quality data. *Water Resources Research, 24*(12):1997–2004.

18. Rubin, D.B. (1987). Multiple Imputation for Nonresponse in Surveys. *John Wiley*, New York.

19. Marzouk. T., Sathyanarayana, S., Kim, A.S., Seminario, A.L., McKinney, C.M. (2018). A systematic review of exposure to Bisphenol A from Dental Treatment. *Journal of Dental Research, Clinical and Translational Research*, In Press.

20. Helsel, D.R. (1990). Less than obvious-statistical treatment of data below the detection limit. *Environ. Sci. Technol. 24*, 1766–1774.

21. Wilhelm, S., Manjunath, B.G. (2010). tmvtnorm: A package for the truncated multivariate normal distribution. *The R Journal, 2*(2), 25-29

22. Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention science, 8(3),* 206-213.

Tables and Figures

Table 1. Baseline Characteristics for the BPARCS Study Data

Child Race – White, n (%)	436 (48.55)
Hispanic, n (%)	160 (17.82%)
Child Gender – Male, n (%)	435 (48.44%)
Marital Status – Married, n (%)	569 (63.36%)
Weight, kg, median (IQR)	23.9 (20.0 - 29.0)
Height, m, median (IQR)	121.55 (113.40 - 130.00)
Child Age, y, median (IQR)	6 (5 - 8)
Education	
High School or Less, n (%)	180 (20.04%)
Post Graduate, n (%)	249 (27.73%)
Some College or Training, n (%)	469 (52.23%)
Income	
Less than 30000, n (%)	253 (28.17%)
30000 to 75000, n (%)	293 (32.63%)
More than 75000, n (%)	263 (29.29%)
Don't Know, n (%)	89 (9.91%)

Table 2. Results of Naïve Substitution Method
(Parameter estimates and the corresponding 95% CIs are exponentiated)

	Estimate (95% CI)	**P Value**
ID:Family, SD (95% CI)	0.310 (0.000, 0.454)	
Family, SD (95% CI)	0.000 (0.000, 0.368)	
Residual, SD (95% CI)	1.318 (1.253, 1.388)	
Intercept	0.774 (0.122, 4.890)	0.785
Visit 3	1.713 (1.327, 2.212)	< 0.001*
Visit 4	1.255 (0.971, 1.622)	0.083
Visit 5	0.893 (0.689, 1.158)	0.392
Visit 6	0.975 (0.733, 1.295)	0.859
Batch 2	0.609 (0.456, 0.815)	0.001*
Batch 3	2.562 (2.043, 3.228)	< 0.001*
Child race - White	0.978 (0.786, 1.215)	0.838
Hispanic	1.027 (0.767, 1.375)	0.859
Child gender	1.150 (0.936, 1.412)	0.181
Education - post graduate	0.903 (0.639, 1.279)	0.564
Education - some college or training	1.178 (0.878, 1.579)	0.271
Income - 30000 to 75000	0.645 (0.487, 0.852)	0.002*
Income - more than 75000	0.696 (0.497, 0.975)	0.034*
Income – don't know	0.690 (0.467, 1.021)	0.062
Marital status - married	0.905 (0.714, 1.148)	0.408
Weight	1.003 (0.984, 1.022)	0.740
Height	0.995 (0.973, 1.016)	0.620
Child age	1.005 (0.878, 1.150)	0.947

Table 3. Results of Paxton's Random Imputation Procedure
(Parameter estimates and the corresponding 95% CIs are exponentiated)

	Estimate (95% CI)	**P Value**
ID:Family, SD (95% CI)	0.519 (0.000, 0.727)	
Family, SD (95% CI)	0.000 (0.000, 0.584)	
Residual, SD (95% CI)	1.977 (1.880, 2.083)	
Intercept	0.157 (0.009, 2.684)	0.199
Visit 3	2.132 (1.453, 3.128)	< 0.001*
Visit 4	1.371 (0.932, 2.015)	0.109
Visit 5	0.659 (0.447, 0.973)	0.035*
Visit 6	0.775 (0.505, 1.189)	0.242
Batch 2	0.672 (0.433, 1.046)	0.075
Batch 3	2.710 (1.915, 3.868)	< 0.001*
Child race - White	0.841 (0.600, 1.176)	0.309
Hispanic	1.236 (0.788, 1.939)	0.354
Child gender	1.294 (0.943, 1.778)	0.109
Education - post graduate	1.010 (0.591, 1.724)	0.972
Education - some college or training	1.350 (0.858, 2.121)	0.190
Income - 30000 to 75000	0.513 (0.334, 0.789)	0.002*
Income - more than 75000	0.542 (0.284, 0.948)	0.020*
Income – don't know	0.519 (0.323, 0.911)	0.032*
Marital status - married	0.940 (0.652 1.354)	0.737
Weight	1.002 (0.973, 1.031)	0.907
Height	1.008 (0.974, 1.042)	0.659
Child age	0.912 (0.742, 1.123)	0.384

Table 4. Results of Modified Paxton's Method
(Parameter estimates and the corresponding 95% CIs are exponentiated)

	Estimate (95% CI)	P Value
ID:Family, SD (95% CI)	0.552 (0.000, 0.746)	
Family, SD (95% CI)	0.093 (0.000, 0.662)	
Residual, SD (95% CI)	1.819 (1.729, 1.916)	
Intercept	0.903 (0.058 13.884)	0.941
Visit 3	1.964 (1.380, 2.796)	< 0.001*
Visit 4	1.137 (0.797, 1.622)	0.478
Visit 5	0.726 (0.507, 1.040)	0.080
Visit 6	0.730 (0.492, 1.085)	0.119
Batch 2	0.697 (0.461, 1.056)	0.084
Batch 3	3.189 (2.285, 4.487)	< 0.001*
Child race - White	0.833 (0.602, 1.152)	0.266
Hispanic	0.987 (0.639, 1.532)	0.954
Child gender	1.211 (0.892, 1.646)	0.218
Education - post graduate	0.810 (0.483, 1.361)	0.421
Education - some college or training	1.124 (0.725, 1.738)	0.599
Income - 30000 to 75000	0.465 (0.307, 0.704)	< 0.001*
Income - more than 75000	0.619 (0.375, 1.021)	0.059
Income – don't know	0.523 (0.291, 0.936)	0.028*
Marital status - married	0.952 (0.668, 1.356)	0.781
Weight	1.006 (0.978, 1.034)	0.674
Height	0.989 (0.958, 1.022)	0.512
Child age	1.019 (0.833, 1.244)	0.853

Table 5. Results of visit 3 from each of the M = 5 imputations using Paxton's method

	Estimate	Standard Error	t Value
1st	0.847	0.201	4.220
2nd	0.903	0.215	4.204
3rd	0.941	0.229	4.106
4th	0.979	0.243	4.004
5th	1.078	0.250	4.317

Table 6. Overall results of the M = 5 imputations using Paxton's method
(Parameter estimates and the corresponding 95% CIs are exponentiated)

	Estimate (95% CI)	P Value
Intercept	0.338 (0.010, 11.003)	0.541
Visit 3	2.584 (1.593, 4.194)	< 0.001*
Visit 4	1.415 (0.888, 2.254)	0.144
Visit 5	0.776 (0.465, 1.295)	0.331
Visit 6	0.970 (0.531, 1.772)	0.922
Batch 2	0.393 (0.215, 0.715)	0.002*
Batch 3	2.536 (1.649, 3.898)	< 0.001*
Child race - White	0.880 (0.589, 1.315)	0.532
Hispanic	1.082 (0.620, 1.887)	0.783
Child gender	1.374 (0.912, 2.070)	0.129
Education - post graduate	0.788 (0.409, 1.516)	0.475
Education - some college or training	1.426 (0.812, 2.505)	0.217
Income - 30000 to 75000	0.494 (0.288, 0.848)	0.011*
Income - more than 75000	0.529 (0.283, 0.989)	0.046*
Income – don't know	0.452 (0.216, 0.948)	0.036*
Marital status - married	0.831 (0.520, 1.329)	0.440
Weight	1.014 (0.979, 1.050)	0.440
Height	0.995 (0.956, 1.037)	0.817
Child age	0.937 (0.729, 1.205)	0.613

Table 7. Results of visit 3 from each of the M = 5 imputations using modified Paxton's method

	Estimate	Standard Error	t Value
1st	0.593	0.186	3.181
2nd	0.678	0.192	3.529
3rd	0.637	0.199	3.199
4th	0.606	0.200	3.029
5th	0.726	0.207	3.515

Table 8. Overall results of the M = 5 imputations using modified Paxton's method
(Parameter estimates and the corresponding 95% CIs are exponentiated)

	Estimate (95% CI)	P Value
Intercept	0.332 (0.016, 6.711)	0.472
Visit 3	1.912 (1.277, 2.862)	0.002*
Visit 4	1.360 (0.886, 2.088)	0.160
Visit 5	0.768 (0.505, 1.168)	0.218
Visit 6	0.901 (0.572, 1.419)	0.652
Batch 2	0.514 (0.304, 0.867)	0.013*
Batch 3	2.374 (1.570, 3.590)	< 0.001*
Child race - White	0.922 (0.624, 1.361)	0.682
Hispanic	1.115 (0.684, 1.819)	0.663
Child gender	1.262 (0.873, 1.826)	0.216
Education - post graduate	0.987 (0.557, 1.747)	0.963
Education - some college or training	1.441 (0.870, 2.388)	0.156
Income - 30000 to 75000	0.533 (0.340, 0.836)	0.006*
Income - more than 75000	0.590 (0.346, 1.003)	0.051
Income – don't know	0.485 (0.245, 0.960)	0.038*
Marital status - married	0.785 (0.539, 1.142)	0.205
Weight	1.000 (0.969, 1.032)	0.993
Height	0.999 (0.962, 1.037)	0.946
Child age	0.995 (0.789, 1.253)	0.963

Table 9. Comparison of visit 3 results from each of the statistical methods
(Parameter estimates and the corresponding 95% CIs are exponentiated)

	Estimate (95% CI)	**P Value**
Naïve Substitution	1.713 (1.327, 2.212)	< 0.001*
Paxton's Random Imputation	2.132 (1.453, 3.128)	< 0.001*
Modified Paxton's Method	1.964 (1.380, 2.796)	< 0.001*
MI for Paxton's Random Imputation	2.584 (1.593, 4.194)	< 0.001*
MI for Modified Paxton's Method	1.912 (1.277, 2.862)	0.002*

Table 10. Comparison of visit 4 results from each of the statistical methods
(Parameter estimates and the corresponding 95% CIs are exponentiated)

	Estimate (95% CI)	**P Value**
Naïve Substitution	1.255 (0.971, 1.622)	0.083
Paxton's Random Imputation	1.371 (0.932, 2.015)	0.109
Modified Paxton's Method	1.137 (0.797, 1.622)	0.478
MI for Paxton's Random Imputation	1.415 (0.888, 2.254)	0.144
MI for Modified Paxton's Method	1.360 (0.886, 2.088)	0.160

Figure 1. Frequency distribution of logBPAb measurements (NDs replaced by LOD/2)

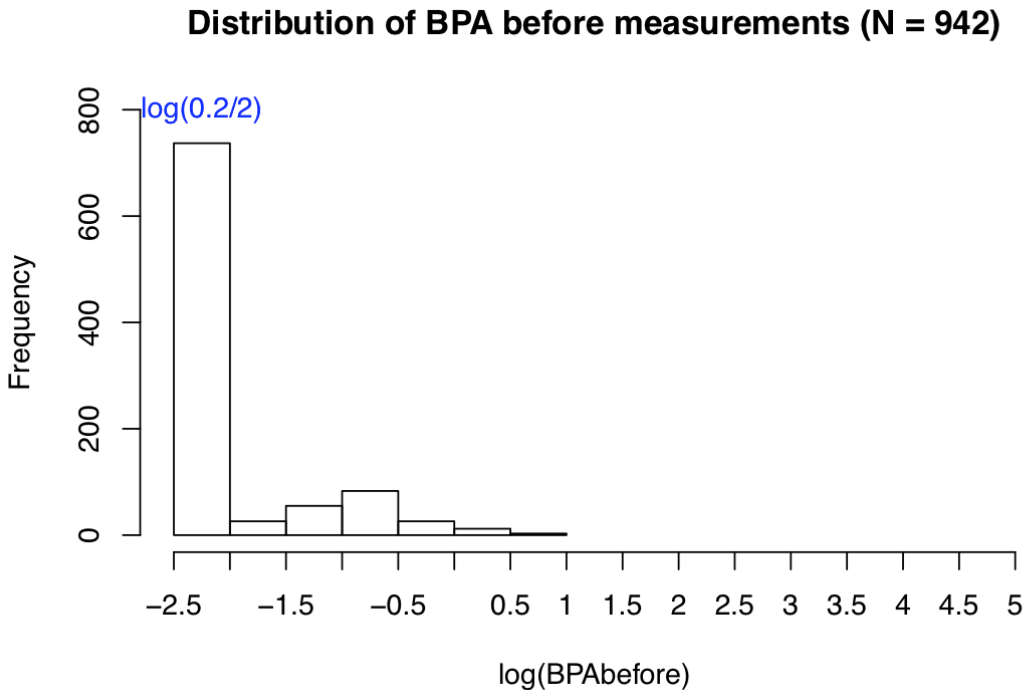

Figure 2. Frequency distribution of logBPAa measurements (NDs replaced by LOD/2)

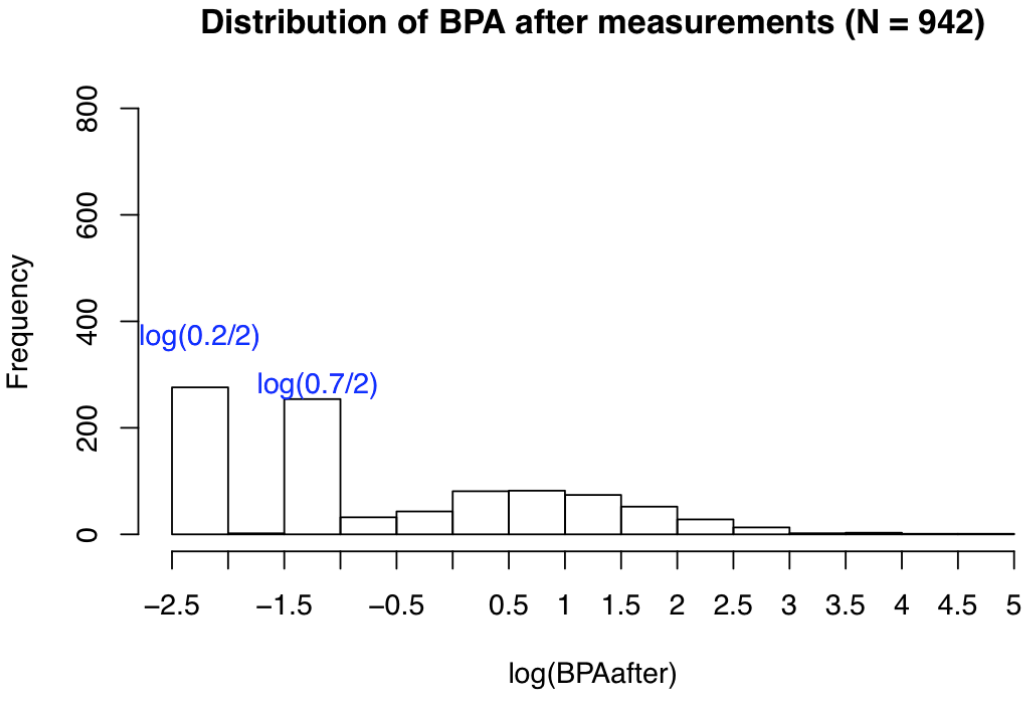

Figure 3. Comparison of logBPAb using the naïve substitution and imputed logBPAb from Paxton's procedure

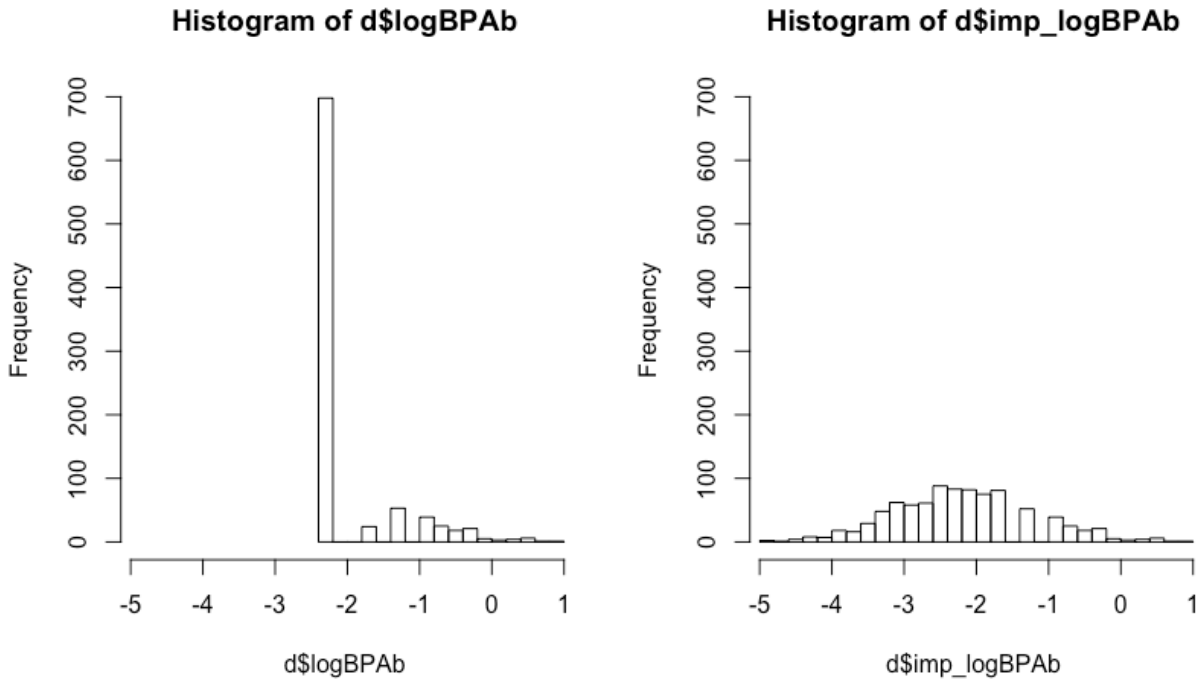

Figure 4. Comparison of logBPAa using the naïve substitution and imputed logBPAa using Paxton's procedure

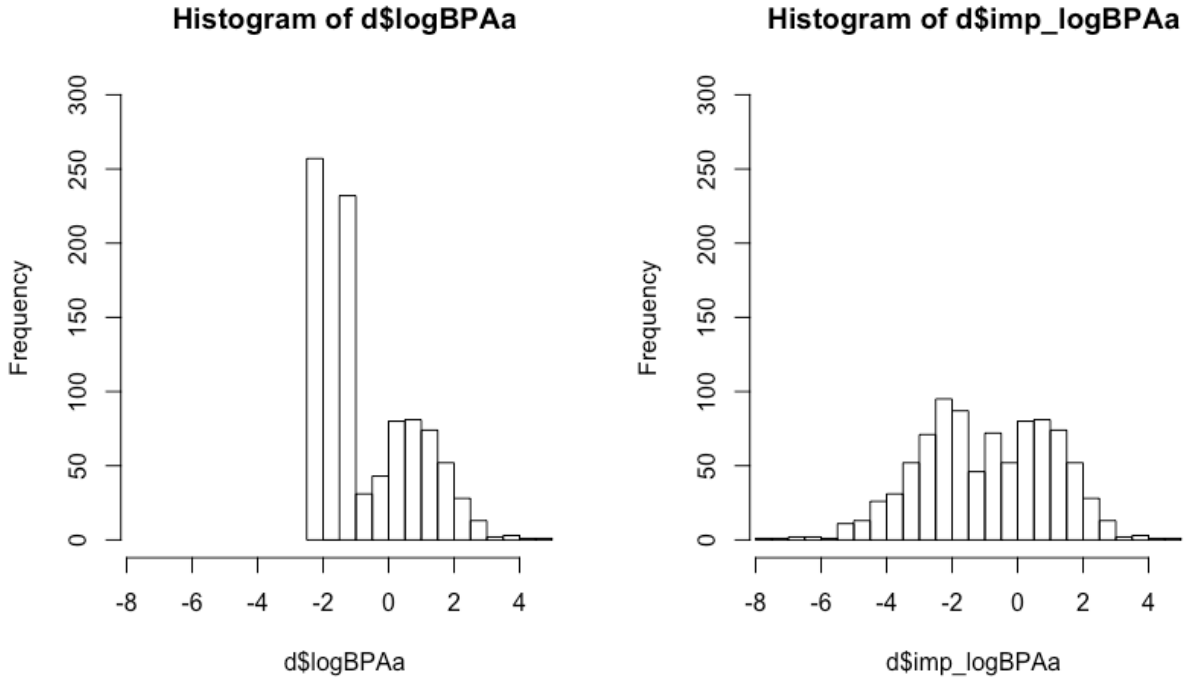

Figure 5. Comparison of logBPAb using the naïve substitution and imputed logBPAb using modified Paxton's method

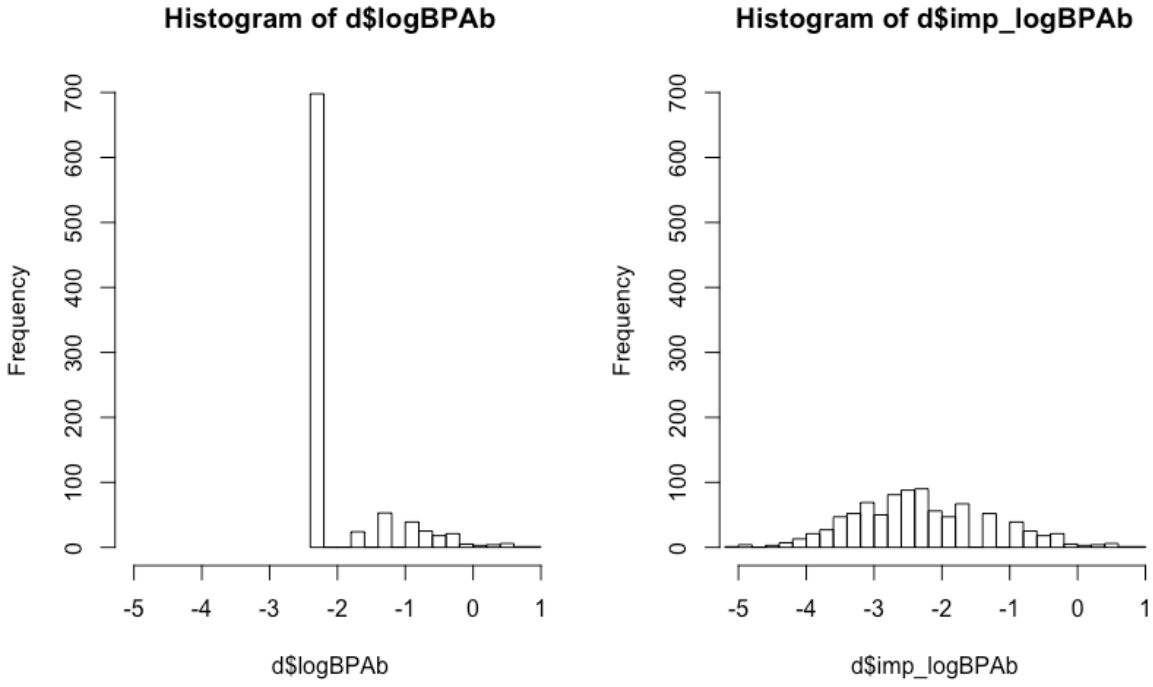

Figure 6. Comparison of logBPAa using the naïve substitution and imputed logBPAa using modified Paxton's method