Yang Lei, Susan Carlson, Lisa N. Yelland, Maria Makrides, Robert Gibson and Byron J. Gajewski

Comparison of dichotomized and distributional approaches in rare event clinical trial design: a fixed Bayesian design

Journal of Applied Statistics, 2017; 44(8):1466-1478

Comparison of Dichotomized and Distributional Approaches in Rare Event Clinical Trial Design: a Fixed Bayesian Design

Yang Lei[1], Susan Carlson[2], Lisa N. Yelland[3], Maria Makrides[4], Robert Gibson[5], Byron J. Gajewski[6*]

1. Department of Biostatistics, The University of Kansas Medical Center, Kansas City, KS 66160, Phone: 319-270-4618, Email: ylei@kumc.edu

2. Department of Dietetics and Nutrition, The University of Kansas Medical Center, Kansas City, KS 66160, Phone: 913-588-5359, Email: scarlson@kumc.edu

3. Women's and Children's Health Research Institute, North Adelaide SA 5006, Australia.School of Population Health, The University of Adelaide, Adelaide SA 5005, Australia, Phone: 618-8313-3215, Email: lisa.yelland@adelaide.edu.au

4. South Australian Health and Medical Research Institute, Adelaide SA 5005, Australia and Discipline of Paediatrics, University of Adelaide, SA 5005, Australia, Phone: 618-8161-6067, Email: maria.makrides@health.sa.gov.au

5. FoodPlus Research Centre, School of Agriculture, Food and Wine, University of Adelaide, SA 5005, Australia, Phone: 618-8313-4333, Email: robert.gibson@adelaide.edu.au

6. Department of Biostatistics, The University of Kansas Medical Center, School of Medicine, Mail Stop 1026, 3901 Rainbow Blvd., Kansas City, KS 66160, Phone: 913-588-1603, Fax: 913-588-0252, E-mail: bgajewski@kumc.edu

*Corresponding author. Email: bgajewski@kumc.edu

Acknowledgements

This study was supported in part by an NIH Clinical and Translational Science Award grant (UL1 TR000001, formerly UL1RR033179) as well as 1R01 HD047315, awarded to the University of Kansas Medical Center. Lisa Yelland was supported by an Australian National Health and Medical Research Council Early Career Fellowship (ID 1052388).

Submitted to *Journal of Applied Statistics*, 10/22/2015

Resubmitted to *Journal of Applied Statistics*, 05/16/2016

Accepted by *Journal of Applied Statistics*, 7/14/2016

ABSTRACT

This research was motivated by our goal to design an efficient clinical trial to compare two doses of docosahexaenoic acid supplementation for reducing the rate of earliest preterm births and/or preterm births. Dichotomizing continuous gestational age data using a classic binomial distribution will result in a loss of information and reduced power. A distributional approach is an improved strategy to retain statistical power from the continuous distribution. However, appropriate distributions that fit the data properly, particularly in the tails, must be chosen, especially when the data are skewed. A recent study proposed a skew-normal method. We propose a three-component normal mixture model and introduce separate treatment effects at different components of gestational age. We evaluate operating characteristics of mixture model, beta-binomial model, and skew-normal model through simulation. We also apply these three methods to data from two completed clinical trials from the USA and Australia. Finite mixture models are shown to have favorable properties in preterm births analysis but minimal benefit for earliest preterm births analysis. Normal models on log transformed data have the largest bias. Therefore we recommend finite mixture model for preterm births study. Either finite mixture model or beta-binomial model is acceptable for earliest preterm births study.

Keyword: Bayesian, Normal mixture model, simulation, Dichotomization, preterm birth

1. INTRODUCTION

In many circumstances, clinical researchers are interested in studying categorized outcomes using cutoff points despite continuous measurements being collected. It has been widely accepted that dichotomizing continuous data prior to analysis results in a loss of information and reduced power [1, 3, 10]. A distributional approach can be used to dichotomize continuous data while retaining the statistical power from the continuous distribution [10]. Peacock et al. [10] described the use of the distributional method and showed the good performance of this parametric approach under standard normal distributional assumptions. Sauzet et al. [12] further discussed the distributional approach when the outcome is skewed and proposed a skew-normal distributional method for dichotomization. They used a logarithm transformation to normalize negatively skewed gestational age data and then applied the skew-normal distributional method under the Frequentist framework. They acknowledged that no satisfactory transformation is available for gestational age data [12]. Mixture models with different components might be a better choice for skewed outcomes such as gestational age, because they allow for greater flexibility in modeling heterogeneous populations [9], which largely explains the skewness of gestational age data.

Our research was motivated by our goal to design an efficient clinical trial to compare two doses of docosahexaenoic acid (DHA) supplementation for reducing the rate of earliest preterm births (ePTB, gestation age<34 weeks) and/or preterm births (PTB, gestational age<37 weeks). Both endpoints have been evaluated in past studies [7]. The United States currently has a PTB rate of 11.4% [6] and babies born preterm are at increased risk of immediate life-threatening health problems, as well as long-term complications and developmental delays [4]. Among preterm infants, those babies who are born the earliest (<34 weeks) are at greatest risk of

complications. Although the overall PTB rates have decreased over time, the ePTB rates in the U.S. have decreased little since 1990 and the overall ePTB rates in the US for 2012 were 3.4% [8]. These births impact overall infant mortality the most and result in much higher hospital costs than uncomplicated births [11]. Docosahexaenoic acid (DHA) supplementation potentially provides a high yield, low risk provocative strategy to reduce early preterm delivery [4]. We designed a Phase III clinical trial (randomized to low or high dose DHA, double-blind) to examine the efficacy of 1000 mg/day DHA supplementation to reduce the probability of earliest preterm births and/or preterm births compared to 200 mg/day, an amount recommended by the FAO/WHO for pregnant and lactating women and currently in many prenatal supplements. Our goal was to identify a powerful design that would provide an efficient estimate of the treatment effect.

Gestational age (GA) data will be measured in completed weeks/days and collected in a continuous form. The two clinically important endpoints of interest are: ePTB (GA<34 weeks) and PTB (GA<37 weeks). The traditional analysis approach is to dichotomize the continuous gestational age data using these cutoff points and to compare the probabilities of binary outcomes, using a chi-square test for example. Distributional methods compare the proportions below the cutoff points in continuous distributions [4, 10]. Sauzet et al. [12] proposed a skew-normal method and used normal distribution on the logarithmic transformed data. We propose a three-component normal mixture model and apply the distributional approach directly. The aim of this study is to compare these three statistical methods under a fixed Bayesian design framework for a very rare endpoint (ePTB) and a less rare endpoint (PTB).

The remainder of this article is arranged as follows. In section 2, we describe three statistical models using the pre-dichotomizing and distributional methods separately. In section 3, we

provide the simulation details under a fixed Bayesian clinical trial design framework and compare these three statistical methods in several realistic outcome scenarios. In section 4, we apply these three methods to data from two completed clinical trials, one in the USA and one in Australia. The results from the real data analysis are examined and compared. In section 5, we discuss the observations from the simulations and real data analysis and further investigate the rationale of these observations. In section 6, we discuss the limitations of this study. In section 7, we draw conclusions from our analysis and give suggestions to future studies.

2. STATISTICAL MODELS

Let $\boldsymbol{Y}_j = (Y_{j1,\ldots}, Y_{jn_j})$ denote the continuous data of gestational age, where *j* denotes the treatment group assignment (*j=c* for participants in the control group and *j=t* for participants in the treatment group) and n_j denotes the sample size in the j^{th} treatment group in a two-armed randomized clinical trial design. Let p_j denote the probability of ePTB or PTB in the j^{th} treatment group.

The first method considered involves dichotomizing the data prior to modeling. We propose a beta-binomial model to simplify a Bayesian inference of $P(p_c > p_t|data)$, denoting the posterior probability that control has a higher ePTB/PTB rate than treatment. Because the endpoints considered are rare, using a uniform prior or a beta (1,1) prior might induce non-negligible bias. We therefore assume a very weak prior of p_j as beta (0.01, 0.01). Furthermore, the posterior mode is close to a classical Frequentist approach (i.e., Maximum Likelihood Estimator). Let $X_j = \sum_{i=1}^{n_j} I\left(Y_{ji} < 34\ or\ 37\right)$, where $I(x < y) = \begin{cases} 1, x < y \\ 0, x \geq y \end{cases}$ and n_j is the sample size in the j^{th} treatment group (i=1,...n_j). The distribution of X_j is assumed to follow a binomial

distribution: $X_j|p_j \sim binomial(n_j, p_j)$. The posterior distribution of $p_j|X_j \sim beta\ (X_j + 0.01, n_j - X_j + 0.01\ )$.

The second method considered is a distributional approach, where we will apply the transformation recommended by Sauzet et al. [12]. First we take a logarithmic transformation of (45-GA) to normalize the data because we expect GA in weeks to be <45 and is negatively skewed, $Z_{ji} = \log(45 - Y_{ji}), i = 1, \ldots n_j$, and then assume $Z_{ji} \overset{i.i.d}{\sim} N(\mu_j, \sigma_j^2)$. Since the logarithmic transformation is a continuous and monotonic transformation, this does not affect the proportion below a cut-point [12]. The proportions of GA below 34 and 37 are translated into the proportions greater than log(45-34) =2.3979 and log(45-37) =2.0794 in the normal distribution $N(\mu_j, \sigma_j^2)$. We use non-informative conjugate priors for the parameters in the normal distribution: $N(0, 100^2)$ for μ_j and *Gamma (*0.001, 0.001) for $\frac{1}{\sigma_j^2}$. The posterior probability of ePTB or PTB $(p_j|\boldsymbol{Z}_j)$ is calculated as $p_j|\boldsymbol{Z}_j = \int_{2.3979\ or\ 2.0794}^{\infty} \Phi(y|\mu_j, \sigma_j^2)$ dy, where $\Phi(y|\mu_j, \sigma_j^2)$ is a normal density function with posterior variance $\sigma_j^2|\boldsymbol{Z}_j \sim IG(0.001 + \frac{n_j}{2}, 0.001 + \frac{1}{2}\sum_{i=1}^{n_j}(Z_{ji} - \bar{Z}_j)^2)$, and posterior mean $\mu_j|\boldsymbol{Z}_j, \sigma_j^2 \sim Normal\ (\frac{100^2 n_j \bar{Z}_j}{100^2 n_j + \sigma_j^2}, \frac{100^2 \sigma_j^2}{100^2 n_j + \sigma_j^2})$.

The third method considered is another distributional approach using the finite normal mixture model. Peacock et al. [10] showed the good performance of the parametric approach under traditional normal distributions. We extend this approach here and propose a finite mixture normal model to allow for population heterogeneity. In this method, we apply a three-component normal mixture model derived from the North Carolina Detailed Birth Record (NCDBR) database with 336,129 observations in the final analysis: a three-component mixture of N(39.59, 0.96), N(38.26, 2.48), and N(33.29, 13.23) [14]. The 95% CIs for the parameter estimates in this

model show these estimates are reliable in this registry data. The first component has a mean of 39.59 (39.58, 39.61), and variance of 0.96 (0.95, 0.97). The second component has a mean of 38.26 (38.20, 38.32) and variance of 2.48 (2.42, 2.54). The third component has a mean of 33.29 (33.07, 33.51) and variance of 13.23 (12.78, 13.67) [14]. Although we used fixed parameter estimates from a U.S. registry data, this model has unprecedented advantages in gestational age data analysis or clinical trial design, even for a different population. Firstly, the parameter estimates are derived from a huge registry data thus is representative and has generalizablility. Secondly, a three-component mixture normal model has its own flexibility to model similar but not exactly the same gestational age data from a different population by allowing various component weights. Thirdly, the three components are realistic and interpretable. The three components represent low, medium, and high-risk groups for PTB separately. We assume a unity prior for Δ_j (j=c,t), the mixture weights in the j^{th} treatment group, and the three-component normal mixture model can be written as: $f(Y_{ji}|\Delta_j) = \Delta_{1j}\phi\left(Y_{ji}\middle|39.59, 0.96\right) + \Delta_{2j}\phi\left(Y_{ji}\middle|38.26, 2.48\right) + \Delta_{3j}\phi\left(Y_{ji}\middle|33.29, 13.23\right)$, where $\phi\ (y|\mu, \sigma^2)$ denotes the density of y in a normal distribution with mean μ and variance σ^2, and $\Delta_{1j}, \Delta_{2j}, and\ \Delta_{3j}$ denote the mixture weights in the j^{th} treatment group, with $\Delta_{1j} + \Delta_{2j} + \Delta_{3j} = 1$. In this method, the posterior probability of ePTB or PTB ($p_j|\boldsymbol{Y}_j$) is calculated as: $p_j|\boldsymbol{Y}_j = \int_{-\infty}^{34\ or\ 37} f\left(y\middle|\Delta_j\right)$ dy. A more general approach would be to let each component's mean and variance be freely modeled. However, we found our approach was flexible and appropriate for our focus on the lower tail. More flexible models allowing components' means and variances to vary are considered for future analysis, especially when more data are available.

3. FIXED BAYEISAN CLINICAL TRIAL DESIGN AND SIMULATION STUDY

A previous Phase III trial comparing 600 mg DHA per day and placebo, Kansas University DHA Outcome Study [KUDOS], found an 85% reduction in ePTB with DHA supplementation [2]. Another Australian trial, DHA to Optimize Mother Infant Outcome [DOMInO] trial, which compared 800 mg DHA per day and placebo, found a 50% reduction in ePTB with DHA supplementation [7]. In both trials, ePTB was a secondary outcome [2, 7]. The primary aim of the current proposed Phase III randomized, double-blind trial is to test the hypothesis that ePTB and/or PTB is reduced by 1000 mg of DHA per day compared to 200 mg DHA per day. We performed a simulation study based on realistic response scenarios to investigate the operating characteristics of this fixed Bayesian clinical trial design.

3.1 Simulation Methods

We simulated gestational age data using different true values of mixture weights ($\Delta_{\mathbf{0}}$) with resulting probabilities of ePTB or PTB close to probability scenarios we observed from our clinical trials [2]. In the beta-binomial model, we used simulation to generate the posterior distribution of $p_j|X_j \sim beta\left(X_j + 0.01, n - X_j + 0.01\right)$ for both treatment and control groups and calculated the probability of $p_c|X_c > p_t|X_t$. In the finite mixture model, we used Markov Chain Monte Carlo (MCMC) to generate posterior distributions of $\Delta_{\boldsymbol{j}}$ and the posterior probability $p_j|\boldsymbol{Y}_j$ was calculated as: $p_j|\boldsymbol{Y}_j = \int_{-\infty}^{34\ or\ 37} f\left(y|\Delta_{\boldsymbol{j}}\right)$ dy. In the logarithmic transformation method, we used Gibbs sampling to generate posterior distributions of $\mu_{jp}\ and\ \sigma_{jp}^2$. The posterior probability $p_j|\boldsymbol{Z}_j$ was calculated as $\int_{2.3979\ or\ 2.0794}^{\infty} \Phi\left(y|\mu_{jp}, \sigma_{jp}^2\right)$ dy. If $\Pr(p_c > p_t|data) > \delta$, we counted this as a trial success. The posterior mean of $p_j|data$, $\widehat{p_j} = \mathrm{E}(p_j|data)$, was saved for each simulation in each of the three models. In all models, the expected estimated probability of ePTB or PTB, $E(\widehat{p_j})$ was calculated as the average of $\widehat{p_j}$ across

simulations. V_j, the sample variance of $\widehat{p_j}$, was calculated as $\frac{\sum_{j=1}^{S}(\widehat{p_j}-E(\widehat{p_j}))^2}{S-1}$ for each treatment group, where S was the number of simulations. The MSE of $E(\widehat{p_j})$ was calculated as $bias^2 + sample\ variance = (E(\widehat{p_j}) - P_{j,0})^2 + V_j$, with $P_{j,0}$ denoting the true probability of ePTB or PTB in the j^{th} treatment group.

To mimic situations for ePTB in future trials, we simulated 5 scenarios with varying treatment effects: no effect (3 vs. 3%, difference=0), very small (3 vs. 2%, difference=1%), small (3 vs. 1%, difference=2%), medium (3 vs. 0.5%, difference=2.5%) and large (4 vs. 1%, difference=3%) based on our previous clinical trial results [2]. To mimic situations for PTB in future trials, we simulated another 5 scenarios: no treatment effect (8 vs. 8%, difference=0), very small (8 vs. 7%, difference=1%), small (8 vs. 6%, difference=2%), medium (8 vs. 5%, difference=3%) and large (8 vs. 4%, difference=4%) based on results from our previous clinical trial [2]. In the null scenarios where the treatment effect was 0, we identified the δ values which made the average success rate across simulations approximately equal to 0.05, $\mathrm{P}(\Pr(p_c > p_t|data) > \delta) \approx 0.05$. δ values can vary in different statistical methods. This ensured the type I error rate was about 5%. In other scenarios, $\mathrm{P}(\Pr(p_c > p_t|data) > \delta)$ was used to calculate the power of the tests.

We compared the simulated trial operating characteristics, (bias, power and MSE) across the three models for both ePTB and PTB. These were based upon 1000 simulations and 600 subjects in each group because our designed trial has a sample size around 1200. The δ value was 0.95 for both ePTB and PTB simulations, in the beta-binomial model and the finite mixture model. In the logarithmic transformation model, δ was 0.999 for ePTB simulations and 0.997 for PTB simulations. The δ value was variable in simulation studies to ensure the Type I error rate is controlled at 5% level for each model. The purpose is to compare model performance after

controlling for Type I error. Using the same δ value for log-transformed model will boost the Type I error for this method. All methods were implemented in R 3.1.1 and Openbugs.

3.2 Simulation Results

In the simulation study of probability of ePTB (<34 weeks), the beta-binomial model had lower bias compared to the finite mixture model and the logarithmic transformation model in all scenarios. The MSE in the finite mixture model was consistently lower than in the beta-binomial model and logarithmic transformation model in the control group and slightly higher than in the beta-binomial model in the last three scenarios in the treatment group (Table 1). Figure 1 shows the comparisons of bias, variance, MSE and power across the three models. In the null scenario, the type I error rate was 0.048 in the beta-binomial model, 0.054 in the finite mixture normal model, and 0.053 in the logarithmic transformation model (Table 1). The power for the finite mixture model was slightly higher than the beta-binomial model in other scenarios, but the difference was small (Figure1). The logarithmic transformation model had the largest bias and lowest power (Table 1).

In the simulation study of probability of PTB (<37 weeks), the beta-binomial model continued to have lower bias compared to the finite mixture normal model and the logarithmic transformation model. The difference in MSE between the finite normal mixture model and the beta-binomial model was larger than that in the ePTB simulations (Table 1 and Table 2). The logarithmic transformation model again had the largest bias and largest MSE (Table 2). In the null treatment effect scenario, the type I error rate was 0.054 in the beta-binomial model, 0.05 in the finite mixture normal model, and 0.051 in the logarithmic transformation model (Table 2). The power for the finite mixture model was higher than the beta-binomial model in small to large difference scenarios, with differences as large as 7.5% when the true effect was likely (8 vs. 5%)

(Table 2). Figure 2 shows the comparisons of bias, variance, MSE and power across the three models.

These simulation results demonstrated that although the bias from the finite mixture method was larger than that from the pre-dichotomizing method, the parameter estimates from the finite mixture method had desirable properties such lower variance. In ePTB simulation, the finite mixture model did not appear to be more desirable than the beta-binomial model. However, the finite mixture model improved the power and MSE in PTB analysis. The logarithmic transformation method has the largest bias and highest MSE. In a word, the logarithmic transformation model appeared to be inferior to the finite mixture model. The bias in the log-transformed model is not driven by δ value but the fact that this model cannot model the distribution of gestation age very well.

4. APPLICATION TO REAL DATA

To illustrate the use of the three models in real data, we reanalyzed the gestational age data from an Australia based clinical trial and a USA based clinical trial.

4.1 DOMInO Trial

The DOMInO trial was a double-blind, multicenter, randomized controlled trial conducted in five Australian maternity hospitals. The trial included 2399 women who were less than 21 weeks' gestation with singleton pregnancies and who were recruited between October 31, 2005, and January 11, 2008 [7]. This study compared fish oil capsules (providing 800 mg/d of DHA) or matched vegetable oil capsules without DHA. Gestational age data were available for 2367 (1183 in control and 1184 in treatment) participants in this study.

We looked at the posterior summary statistics of the posterior component probabilities in the control and treatment groups from the finite mixture model (Table 3). Compared to the control

group, the posterior probability of the first component (low risk of PTB) increased from 0.783 to 0.813 and the posterior probability of the third component (high risk of PTB) decreased from 0.04 to 0.022. The posterior probability of the second component decreased from 0.177 to 0.165. Convergence diagnostics were checked to ensure the convergence of posterior samples.

In Table 4, we show the calculated and estimated probability of ePTB and PTB and the standard deviation of the estimated probabilities. In this analysis, we found the benefits of the finite mixture model were not clear in ePTB but the standard deviation was slightly smaller in the finite mixture model in PTB analysis. The estimated proportions for the log transformation model are quite different to the raw data (Table 4). Since we don't know the true parameter value, we won't be able to calculate bias and MSE.

4.2 KUDOS Trial

KUDOS was a Phase III, randomized, double-blind, placebo-controlled clinical trial involving 299 women [2]. This study compared participants in the placebo group ($n_1 = 145$) and participants who received 600 mg/day DHA ($n_2 = 154$) in the second and third trimester during pregnancy from 2001 to 2006 in the University of Kansas Hospital [2].

The posterior summary statistics of the mixture weights were summarized in Table 3. Compared to the DOMInO trial, the difference in the three component probabilities between treatment and control groups was much larger (Table 3). The mixture weight of the third component (high risk of PTB) decreased dramatically from 0.089 in the control group to 0.029 in the intervention group. The weight of the second component increased from 0.073 to 0.196. Both indicate the improvement in the intervention group. Convergence diagnostics were checked to ensure the convergence of posterior samples.

In Table 4, we show the calculated and estimated probability of ePTB and PTB and the standard deviation of the estimated probabilities. Again in this analysis, we found the advantages of the finite mixture model compared to beta-binomial model were not very clear for ePTB but the standard deviation was smaller in the finite mixture model for PTB. Both the DOMInO and KUDOS data were consistent with the simulation studies and showed that the benefits of the finite mixture model were questionable for ePTB but might exist in PTB analysis. The logarithmic transformation model produced quite different results compared to the other two models, which may be due to the bias in this method observed in the simulation study.

5. DISCUSSION

We aimed to investigate the properties of pre-dichotomizing and distributional approaches using a three-component normal mixture model and a logarithmic transformation model. The three-component normal mixture model has been demonstrated to be identifiable and superior to two-component mixture models while avoiding the poor mixing in models with four or more components [14]. The Bayesian framework provides us with a convenient tool to compare distributional approaches and the pre-dichotomizing method.

In the simulation study, we used a weak beta prior for the beta-binomial model to ensure the bias was negligible and the estimates were close to the Frequentist approach. As a result, the bias from the finite mixture model was greater than that from the beta-binomial model. However, the finite mixture model had lower variance in all scenarios (Figure 1 and Figure 2). In the ePTB analysis where the endpoint was very rare, the power of the finite mixture model was only slightly higher than the beta-binomial model and the benefits of the finite mixture model were relatively small. The benefits of the finite mixture model were more apparent in the PTB analysis where the endpoint was less rare. In this case, the variance and hence the MSE were lower in the

finite mixture model compared with the other methods. The power in the finite mixture model was higher except in the very small difference scenario, where it is slightly lower than the beta-binomial model. The logarithmic transformation model had the largest bias and MSE.

In real data analysis, both DOMInO and KUDOS trial data demonstrated that the finite mixture model was not superior in ePTB analysis. The finite mixture model had lower standard deviation compared to the beta-binomial model for PTB in both datasets. The logarithmic transformation model produced quite different results in both analyses. These findings confirmed previous findings that the logarithmic transformation was not satisfactory for GA data [12].

Further investigating the three-component mixture model facilitates understanding of our observations in the simulation study and real data analysis. The three mixture components are: N(39.59, 0.96), N(38.26, 2.48), and N(33.29, 13.23). The mixture weights are about 70-80% for the first component, 10-20% for the second component, and less than 10% for the third component. The three components have different means and standard deviations (heteroscedastic). Therefore it is not straightforward to describe the exhibition of the mixture distribution. However, we can still speculate the mixture exhibition from the three mixture components and the mixture weights. The first two components have close means and different standard deviations. Distribution mixing these two will display high kurtosis with a sharper peak and heavier tails than a single distribution [5]. The third component is sufficiently separated from the first two components. The difference in the means between the second and third components is greater than two times the standard deviation of the second component. Mixing of these two could form a bimodal distribution [13]. Since the mixture weights of the first two components are dominant and the standard deviation of the third component is large, the exhibition could have a long left tail with a small peak on the tail.

Based on the exploration of the finite mixture model, we can obtain an intuitive explanation of our observations. In the ePTB analysis we used GA<34 as a cutoff. Given the exhibition of the mixture model, the area below 34 was mainly captured by the third component of the distribution. In the PTB analysis we used GA<37 as a cutoff and the area below 37 was comprised of the second and the third components, while the influence of the first component was trivial. Therefore in the ePTB analysis, the finite mixture model did not appear to be much better than the beta-binomial model in terms of power because most of the information we needed to make inference on the probability of ePTB was captured by one mode in a bimodal exhibition. In the PTB analysis, the information to make inference on the probability of PTB was captured by two components and the finite mixture model captured the information from the trend of the two components and retained the power from the continuous distribution. Gestation age analysis is a single example in real life where we care about dichotomized outcomes while continuous data are collected. This study showed the cutoff value and the exhibition of the distribution were important to understand the mechanism of gaining power from a continuous distribution. This conclusion can be generalized to other studies in which the outcome is dichotomized while data are collected in a continuous form.

6. LIMITATIONS

There are a few assumptions we have made to pursue this study. Firstly, we used the parameter estimates of the normal components from the North Carolina Detailed Birth Record (NCDBR) database and applied them to different populations. We assumed these component parameters were valid in different populations and they appeared to be fine in this study as the estimated probabilities are quite close to the true data. Although the finite mixture model has certain flexibility to allow component weights to vary, the parameter estimates or even the

formation of the mixture model could change in other populations if the population is extremely different. Secondly, we assumed there was no measurement error in the gestational age data. Gestational age data were obtained from medical records but we do not have a technique to test the measurement error in the current study. If the measurement error was large, it could blur the boundary of ePTB and PTB.

7. CONCLUSION

In studies where endpoints are collected as continuous variables but clinicians are interested in studying dichotomized outcomes, a pre-dichotomizing or distributional approach could be used for analysis. In general, a distributional approach that fits the data well retains information and power from the continuous distribution, while a dichotomizing method is close to the traditional Frequentist approach and may result in less bias. The benefits of a distributional method depend on model fit, cutoff values, and the exhibition of the continuous distribution. Meticulous investigation of the distributions is necessary, especially in rare endpoint analysis where retaining statistical power is more important. In our clinical trial designs for gestational age data, we recommend the finite mixture normal model if the endpoint is PTB (<37 weeks) since this is a more powerful design and beta-binomial model if the endpoint is ePTB (<34 weeks) since the power from these two designs are close and beta-binomial model has less bias.

Conflict of Interest

There is no conflict of interest.

Reference

[1] D.G. Altman, and P. Royston, The cost of dichotomising continuous variables, BMJ (Clinical research ed) 332 (2006), pp. 1080. Available at <Go to ISI>://MEDLINE:16675816.

[2] S.E. Carlson et al., DHA supplementation and pregnancy outcomes, The American journal of clinical nutrition 97 (2013), pp. 808-15. Available at <Go to ISI>://MEDLINE:23426033.

[3] B.A. Deyi, A.S. Kosinski, and S.M. Snapinn, Power considerations when a continuous outcome variable is dichotomized, Journal of biopharmaceutical statistics 8 (1998), pp. 337-52. Available at <Go to ISI>://MEDLINE:9598427.

[4] B.J. Gajewski et al., Commensurate Priors on a Finite Mixture Model for Incorporating Repository Data in Clinical Trials, (In review).

[5] N. Gridgeman, A Comparison of Two Methods of Analysis of Mixtures of Nomnal Distributions1, Technometrics 12 (1970), pp. 823-833.

[6] J.N. House, U.S. Preterm Birth Rate Still Too High: Q&A with Jennifer L. Howse, PhD, the March of Dimes, in Culture of Health, NewPublicHealth, Robert Wood Johnson Foundation, 2014.

[7] M. Makrides et al., Effect of DHA supplementation during pregnancy on maternal depression and neurodevelopment of young children: a randomized controlled trial, Jama 304 (2010), pp. 1675-83. Available at <Go to ISI>://MEDLINE:20959577.

[8] J.A. Martin et al., Division of Vital Statistics Births: Final Data for 2012, National Vital Statistics Reports 15 (2013). Available at http://www.cdc.gov/nchs/data/nvsr/nvsr62/nvsr62_09.pdf.

[9] G. McLachlan, and D. Peel, Finite Mixture Models ed, Wiley Series in Probability and Statistics, New York, 2000.

[10] J.L. Peacock et al., Dichotomising continuous data while retaining statistical power using a distributional approach, Statistics in medicine 31 (2012), pp. 3089-103. Available at <Go to ISI>://MEDLINE:22865598.

[11] R.B. Russell et al., Cost of hospitalization for preterm and low birth weight infants in the United States, Pediatrics 120 (2007), pp. e1-9. Available at <Go to ISI>://MEDLINE:17606536.

[12] O. Sauzet, M. Ofuya, and J.L. Peacock, Dichotomisation using a distributional approach when the outcome is skewed, BMC medical research methodology 15 (2015), pp. 40. Available at <Go to ISI>://MEDLINE:25902850.

[13] M.F. Schilling, A.E. Watkins, and W. Watkins, Is human height bimodal?, The American Statistician 56 (2002), pp. 223-229.

[14] S.L. Schwartz, A.E. Gelfand, and M.L. Miranda, Joint Bayesian analysis of birthweight and censored gestational age using finite mixture models, Statistics in medicine 29 (2010), pp. 1710-23. Available at <Go to ISI>://MEDLINE:20575047.

Table 1: Simulated Trial Operating Characteristics for Probability of ePTB (GA<34 weeks)

Scenarios	Method	Bias		MSE×10^5		Power
		Control	Tx	Control	Tx	
No difference (3 vs. 3%)	Beta-Binomial Model	.00009	.00006	4.62	4.78	.048
	Finite Mixture Model	.00144	.00119	4.02	4.17	.054
	Logarithmic Transformation Model	.01983	.01954	58.56	59.21	.053
Very Small (3 vs. 2%)	Beta-Binomial Model	.00037	.00005	4.53	3.39	.275
	Finite Mixture Model	.00099	.00123	4.08	3.13	.286
	Logarithmic Transformation Model	.02051	.01493	53.46	28.14	.164
Small (3 vs. 1%)	Beta-Binomial Model	.00034	.00028	5.03	1.69	.845
	Finite Mixture Model	.00160	.00171	4.63	1.82	.857
	Logarithmic Transformation Model	.02051	.00713	53.46	7.19	.58
Medium (3 vs. 0.5%)	Beta-Binomial Model	.00043	.00004	4.69	0.89	.983
	Finite Mixture Model	.00156	.00160	4.20	1.00	.985
	Logarithmic Transformation Model	.01983	.00347	58.58	2.51	.794
Large (4 vs. 1%)	Beta-Binomial Model	.00010	.00010	6.57	1.51	.984
	Finite Mixture Model	.00120	.00143	5.47	1.55	.989
	Logarithmic Transformation Model	.02554	.00667	89.75	10.72	.863

Power: average success rate across simulations, P(Pr($p_c > p_t|data$) $> \delta$), $\delta = 0.95$ for Beta-binomial and finite mixture model, $\delta = 0.999$ for logarithmic transformation model

Table 2: Simulated Trial Operating Characteristics for Probability of PTB (GA<37 weeks)

Scenarios	Method	Bias		MSE×10^5		Power
		Control	Tx	Control	Tx	
No difference (8 vs. 8%)	Beta-Binomial Model	.00056	.00003	12.3	11.6	.054
	Finite Mixture Model	.00235	.00224	10.3	10.1	.05
	Logarithmic Transformation Model	.03265	.03286	175	178	.051
Very Small (8 vs. 7%)	Beta-Binomial Model	.00002	.00016	12.6	11.4	.164
	Finite Mixture Model	.00184	.00239	10.4	9.59	.163
	Logarithmic Transformation Model	.03266	.02859	175	151	.129
Small (8 vs. 6%)	Beta-Binomial Model	.00070	.00024	12.1	9.05	.378
	Finite Mixture Model	.00149	.00218	9.90	7.47	.418
	Logarithmic Transformation Model	.03266	.02134	175	98	.343
Medium (8 vs. 5%)	Beta-Binomial Model	.00051	.00006	12.2	7.86	.693
	Finite Mixture Model	.00224	.00225	10.8	6.33	.768
	Logarithmic Transformation Model	.03266	.01439	175	56	.687
Large (8 vs. 4%)	Beta-Binomial Model	.00025	.00008	12.9	6.64	.908
	Finite Mixture Model	.00224	.00246	11.5	4.98	.952
	Logarithmic Transformation Model	.03223	.00789	157	19	.94

Power: average success rate across simulations, P(Pr($p_c > p_t|data) > \delta$), $\delta = 0.95$ for Beta-binomial and finite mixture model, $\delta = 0.997$ for logarithmic transformation model

Table 3: Posterior summary statistics for mixture weights in finite mixture model in DOMInO and KUDOS trial (10000 simulations)

		Control		Treatment	
		mean	std	mean	std
DOMInO	Δ_1	.783	.022	.813	.021
	Δ_2	.177	.023	.165	.021
	Δ_3	.040	.007	.022	.006
KUDOS	Δ_1	.838	.048	.775	.060
	Δ_2	.073	.048	.196	.063
	Δ_3	.089	.027	.029	.018

Δ_1: posterior probability of component 1, N(39.59, 0.96)
Δ_2: posterior probability of component 2, N(38.26, 2.48)
Δ_3: posterior probability of component 3, N(33.29, 13.23)

Table 4: DOMIinO and KUDOS Data analysis: calculated and estimated probability of GA less than certain cutoff, standard deviation, and variance (10000 simulations)

	End Point	Data		Beta-Binomial Model				Finite Mixture Model				Log-Transformation Model			
		Pc0	Pt0	Pc	Pt	SDc	SDt	Pc	Pt	SDc	SDt	Pc	Pt	SDc	SDt
DOMInO	<34 wks	.023	.011	.023	.011	.004	.003	.024	.013	.004	.003	.009	.004	.001	.001
	<37 wks	.072	.055	.072	.055	.008	.007	.075	.057	.007	.006	.099	.068	.007	.007
KUDOS	<34 wks	.048	.007	.048	.007	.018	.007	.052	.018	.016	.011	.016	.004	.006	.002
	<37 wks	.09	.065	.09	.065	.024	.02	.094	.069	.022	.017	.136	.069	.023	.016

Outcome: probability of GA less than a certain amount of time
Pc0: the calculated probability in the data in control group
Pt0: the calculated probability in the data in treatment group
Pc: the estimated probability in the control group
Pt: the estimated probability in the treatment group
SDc: standard deviation in the control group
SDt: standard deviation in the treatment group

List of Figures