Treatment Effects of Maternal Micronutrient Supplementation Vary by Percentiles of the Birth Weight Distribution in Rural Nepal\textsuperscript{1,2}

Joanne Katz,*\textsuperscript{1,3} Parul Christian,* Francesca Dominici,\textsuperscript{1} and Scott L. Zeger\textsuperscript{1}

\textsuperscript{*Department of International Health, \textsuperscript{1}Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205-2103

ABSTRACT Certain antenatal micronutrient supplements increased birth weight by 40–70 g in rural Nepal. The effect was estimated by calculating the mean difference in birth weight between control and treatment groups, which assumes a constant treatment effect across the birth weight distribution. By estimating differences (and CI) in birth weight between treatment and control groups as a nonlinear, smooth function of the percentiles of the birth weight distribution, we can examine whether the shape of the birth weight distribution for a treatment group is different from that of the control group. Supplementation groups were folic acid, folic acid and iron, folic acid and iron and zinc, and a multiple micronutrient supplement all with vitamin A, compared with the control group of vitamin A alone. The shape of the birth weight distribution in the multiple micronutrient group was the same as that of the control group; however, the location of the distribution had shifted. The folic acid and iron group had fewer infants in the lower tail of its distribution but a similar proportion in the upper tail compared with the control group. The biologic pathways affecting intrauterine growth may vary by micronutrients such that some may confer a benefit among the most vulnerable infants, whereas others may have a more constant effect across the birth weight distribution. Future analytic approaches to estimating benefits of maternal supplementation on birth weight should examine whether there is a constant or variable treatment effect across the distribution of birth weight. J. Nutr. 136: 1389–1394, 2006.

KEY WORDS: • birth weight • micronutrients • pregnancy • infant mortality • Nepal

Many studies use birth weight as an outcome of nutritional or other interventions in pregnancy because it is a measure of intrauterine growth retardation in the absence of gestational age data and also a strong predictor of early infant survival. The usual analytic approaches to the estimation of treatment effects involve calculating the mean difference in birth weight between control and treatment groups. This assumes that treatment shifts the distribution of birth weights by a constant amount. Alternately, researchers examine the relative risk of low or very low birth weight but this compares only the left tails of the 2 distributions, and may not adequately capture changes in the middle or upper end of the distribution.

Previously, we published the effects of daily antenatal supplementation with different combinations of micronutrients on birth weight (1). In that community-based, randomized trial, despite the apparently smaller adjusted birth weight increase with folic acid + iron alone (37 g) compared with a multiple micronutrient that included folic acid + iron (64 g), the percentage reduction in low birth weight (<2500 g) was very similar with folic acid + iron (16%) compared with the multiple micronutrients (14%), suggesting that the birth weight distributions of the treatment groups may have had different shapes from that of the control group, especially at the lower end of the distributions. Furthermore, the multiple micronutrient supplement increased the proportion of babies with birth weight >2500 g (7.7% compared with 5.3% in the control group), an increase not observed with folic acid + iron supplementation (6.0% compared with 5.3%) (1). Thus, one issue that may be relevant to randomized trials of nutritional interventions is whether treatments change the shape of the birth weight distribution in a way that may not be reflected when comparing means of treatment and control distributions.

In this paper, we explore these findings further and describe an analytic approach to addressing this issue in trials of nutritional interventions to improve birth weight. We use data from our community-randomized trial that examined the effect of different maternal micronutrient supplements on birth weight in rural Nepal to illustrate this approach.

MATERIALS AND METHODS

The design, methods, and results of the randomized trial were described previously (1–4). Briefly, 426 communities in Sarlahi...
district, Nepal, were randomly assigned to receive 1 of 5 different maternal supplements. From December 1998 through April 2001, all married women of childbearing age who were not already pregnant or breast-feeding an infant <9 mo old were visited every 5 wk and asked if they had experienced menses in the past 5 wk. If they had not, they were given a urine-based pregnancy test. If pregnant, they were enrolled in the trial and supplemented daily with caplets of identical appearance containing vitamin A alone as the control group (1000 µg retinyl acetate), vitamin A + folic acid (400 µg), vitamin A + folate acid + iron (60 mg ferrous fumarate), vitamin A + folic acid + iron + zinc (30 mg zinc sulfate), or a multiple micronutrient supplement that included the same quantities of vitamin A, iron, folate acid, and zinc, along with vitamin D (10 µg cholecalciferol), vitamin E (10 mg) thiamin (1.6 mg), riboflavin (1.8 mg), niacin (20 mg nicotinamide), vitamin B-6 (2.2 mg pyridoxine hydrochloride), vitamin B-12 (2.6 µg), vitamin C (100 mg ascorbic acid), vitamin K (65 µg menadione), copper (2.0 mg cupric oxide), and magnesium (100 mg magnesium oxide).

The trial received ethical approval from the Committee on Human Research of the Johns Hopkins Bloomberg School of Public Health and the Nepal Health Research Council.

Pregnant women were interviewed at the time of enrollment and maternal height, weight, age, date of last menstrual period, parity, smoking history, and other characteristics were recorded. The main outcomes of the study were birth weight and infant survival. Because 95% of births occurred in the home, a female supplement distributor resident in the village reported the birth to a supervisor who dispatched an anthropometrist to the home to obtain "birth weight." The aim was to weigh the infant as soon after birth as possible. Weights were measured using a digital (Seca 727) infant weighing scale, accurate to 2 g. We enrolled a total of 4096 pregnancies that resulted in 4130 live-born infants; ~80% of weights were obtained within 72 h.

Because weight is measured at variable times after birth, we wished to construct a statistical model that provided a better estimate of true birth weight using the age of the infant at the time of measurement. In addition, ~7% of infants were missing weight measurements; 34% of these infants died very soon after birth (2.4% of all those missing birth weight), before the anthropometrist was able to reach the household. These infants are likely to have been smaller at birth than those who survived beyond the time they were weighed. We developed a measurement error model that allowed us to estimate the weight of the baby at birth for measurements made at variable times after birth and to impute missing measurements. Exploratory analysis of the relation between birth weight and time of the measurement indicated that birth weight dropped ~50 g between 24 and 48 h, was constant between 48 and 72 h, and increased linearly after 72 h. We assumed the expected birth weight for baby i at time t_i had a main effect for the treatment and the vital status, was modeled as a natural cubic spline with knots placed at 24, 48, and 72 h for weights taken between birth and 72 h, varied linearly with time after 72 h and with number of cigarettes smoked, but varied nonlinearly with gestational age and maternal age, weight and height (modeled as natural splines).

Specifically, let \( t_i \) and \( W_{i0} \) be the time and the corresponding birth weight measurement for infant i and let \( N_{ib} \) be an indicator of vital status (alive or died) at time \( t_i \). The following model was fit:

\[
E[W_{i0} | t_i, N_{ib}, T_{ri}, x_i] = \beta_0 + \beta_T t_i + \beta_R N_{ib} + s(t_i, knots = (24, 48, 72)) + \beta_num.cig + s(\text{gest.age}, 3) + s(\text{maternal.weight}, 3) + s(\text{maternal.height}, 3) + s(\text{maternal.age}, 3),
\]

where \( x_i \) is the vector of covariates for the ith woman (gestational age, maternal weight in the first trimester, maternal height, maternal age, and number of cigarettes smoked in the 7 d before the first trimester interview). \( T_{ri} \) was the type of supplement the ith infant’s mother received, and \( s(x, \lambda) \) was a cubic spline with \( \lambda \) df (3 in this case). The gestational age of the infant was based on date of the last menstrual period from the first trimester interview, or if not available, the date of first positive pregnancy test was used. The imputation assumes that those with missing birth weight had weights similar to those with the same covariates who did have birth weights measured. Treatment group was one covariate in this model because treatment affected birth weight. Including treatment in the model would prevent the imputed values from being biased by the treatment effect (5).

Missing birth weights were imputed using a multiple imputation method (6). Specifically, let \( \hat{W}_{i0} = E[W_{i0} | t_i, N_{ib}, T_{ri}, x_i] \) be the predicted birth weight at time \( t_i \) conditioned on the vital status of the infant, supplement group, and maternal covariates. Let \( \epsilon_i \) be the estimated residual variance of the regression model. Imputed data sets \( n = 50 \) were created by sampling \( \hat{W}_{i0,j} \) from a normal distribution with mean \( \hat{W}_{i0,j} \) and SD \( \epsilon_i \) for \( j = 1 \ldots 50 \). If the weight was taken after 72 h, then the birth weight was recalibrated by taking \( \hat{W}_{i0} = \hat{W}_{i0,j} \). A small percentage of infants (4%) had missing values for some of the mother’s covariates. We imputed missing data in the mother’s characteristics with the corresponding mean in the population. Figure 1 denotes the observed (circles), predicted (triangles) and imputed (squares) birth weights at time \( t_i \) for the babies that were measured after (A) and before (B) the 72 h interval.

We first calculated the treatment effect at each percentile by comparing the weight for the smallest infant in one of the treatment groups with the expected birth weight for baby i at time t_i; the pre-
that of the weight for the smallest infant in the control group. We continued this all the way across the percentiles through to the largest infants. This was done for each of the 4 treatment groups compared with the control group. We then smoothed these treatment effects across the percentiles and estimated the treatment effect by taking the mean of these smoothed values.

Specifically, we assumed that the treatment effect was a smooth function of the percentiles of the birth weight distribution:

\[
TE(p) = Q_T(p) - Q_C(p) = s(p, \lambda),
\]

where \( Q_T(p) \) and \( Q_C(p) \) are the quantile functions of the birth weight distribution for treatment and control groups, respectively, and \( s \) is a natural cubic spline of the percentile \( p \) with \( \lambda \), which we have chosen to equal 4. To estimate \( TE(p) \), we did the following: 1) calculated the percentiles \( p_i = i/(n+1) \) with \( n = 750 \) (the smallest number of infants across treatment groups); 2) calculated the corresponding empirical quantiles of the birth weights \( q_{Ti}, q_{Ci} \) under the treatment and control groups, respectively; and 3) smoothed the difference between quantiles \( q_{Ti} - q_{Ci} \) across the percentiles \( p_i \).

Note that if we set \( P = 0.5 \), then our approach reduces to the common method of estimating a treatment effect by comparing the central values of the birth weight distribution (means or medians) between the treatment and control groups. The smoothing of quantile differences across percentiles to improve estimation of the mean difference between 2 outcomes was introduced by Dominici and Cope (7) for continuous and skewed data. This approach was implemented for estimating mean medical expenditures between diseased and nondiseased patients (8). In this paper, we modified this idea for a continuous and symmetric outcome to estimate percentile-specific treatment effects. For more details of this approach, see Dominici et al. (9).

To account for the uncertainty in the imputation of the missing values, we repeated steps 1–3 separately for 50 imputed data sets. We then calculated the percentile-specific treatment effect and its corresponding total statistical variance by using standard multiple imputation methods (6). More specifically, let \( \overline{TE}(p) \) and \( \overline{TV}(p) \) be the point estimate and the bootstrap variance of \( TE(p) \) for the \( j \)-imputed data set, respectively. For each \( p_i \), the overall estimate of the treatment effect and its total variance, denoted by \( \overline{TE}(p), \overline{TV}(p) \), were obtained as follows:

\[
\overline{TE}(p) = (1/J) \sum_{j=1}^{J} \overline{TE}_j(p),
\]

\[
\overline{TV}(p) = \overline{WV}(p) + (1/1+J/JBW(p)),
\]

where

\[
\overline{WV}(p) = (1/J) \sum_{j=1}^{J} \overline{V}(p)
\]

\[
BW(p) = (1/(J-1)) \sum_{j=1}^{J} (\overline{TE}_j(p) - \overline{TE}(p))^2
\]

where

\[
\overline{TE}(p) = \sum_{j=1}^{J} \overline{TE}_j(p).
\]

To test whether there was a constant treatment effect across the distribution, we performed a permutation test. Specifically, for \( J = 1, \ldots, 500 \), we randomly resampled the birth weights to the 2 treatment groups (each of the 4 active treatment groups against control) and calculated the test statistics \( T_m = \sum_{i=1}^{n} (s^i\(p, \lambda \) - \( s_m)) \), where \( s = \sum_{i=1}^{n} s^i\(p, \lambda \) \). The 1-sided \( P \)-value was calculated as the probability that \( T_m \) exceeded the observed test statistics \( T_{obs} = \sum_{i=1}^{n} (s^i\(p, \lambda \) - \( s \)) \) where \( s = \sum_{i=1}^{n} s^i\(p, \lambda \) \).

Because the correlation of birth weight within clusters that were randomized to different treatments was essentially zero, these analyses did not take into account the cluster randomization (1).

**RESULTS**

Variables that predicted birth weight in a measurement error model were treatment group, infant survival, time of infant weight measurement, gestational age at birth, maternal height, weight, and age, and number of cigarettes smoked in the 7 d before the first trimester interview (Table 1, Fig. 1A and B, Fig. 2A–D). The \( r^2 \) for this model was 0.54. Infants who died were predicted to weigh 340 g less than those who survived. The predicted birth weight was 11.7 g lower for each cigarette smoked in the 7 d before the first trimester interview. The predicted effect of the multiple micronutrient supplement on birth weight was 75 g, and the iron + folic acid group had an increase of 43 g compared with the group taking vitamin A alone (Table 1).

The treatment effects were estimated as smooth functions of the percentiles of the birth weight distribution \( TE(p) \) with the corresponding 95% bootstrap confidence bands for each treatment group (Fig. 3A–D). The circles denote the difference in the empirical quantile functions between the 2 groups \( q_{Ti} - q_{Ci} \) as a function of the percentiles. As previously reported, there was a mean treatment effect for the folic acid + iron and multiple micronutrient groups (1). The 1-sided \( P \)-values of the permutation tests for the hypothesis of a constant treatment effect across percentiles of the birth weight distribution were 0.10 for folic acid + iron, 0.96 for multiple nutrients, 0.04 for folic acid + iron + zinc, and 0.60 for folic acid. Therefore, for folic acid + iron and for folic acid + iron + zinc, the treatment effect is significantly different across the distribution of birth weight at the 5 and 10% level, respectively.

More specifically, we found the following: 1) folic acid + iron increased birth weight for babies smaller than \(~2800 \) g with a decline in the treatment effect for the larger babies; 2) the multiple micronutrients increased birth weight across the entire distribution of weights; 3) folic acid + iron + zinc increased birth weight only for babies in the middle of the birth weight distribution (between 2400 and 2900 g approximately); and 4) folic acid alone did not significantly increase the birth weight across the entire distribution.

**DISCUSSION**

Our analysis demonstrates that the standard approach of estimating a mean difference in a continuous outcome between a treatment and control group may not adequately capture the effect of nutritional supplementation on birth weight. The ability to assess whether the treatment effect varies across the

**TABLE 1**

<table>
<thead>
<tr>
<th>Variable†</th>
<th>Estimate</th>
<th>SE</th>
<th>( t ) Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>768.86</td>
<td>239.91</td>
<td>3.20</td>
</tr>
<tr>
<td>Folic acid</td>
<td>-0.96</td>
<td>21.78</td>
<td>-0.04</td>
</tr>
<tr>
<td>Folic acid + iron</td>
<td>43.01</td>
<td>21.76</td>
<td>1.98</td>
</tr>
<tr>
<td>Folic acid + iron + Zinc</td>
<td>8.84</td>
<td>21.47</td>
<td>0.41</td>
</tr>
<tr>
<td>Multiple micronutrients</td>
<td>74.65</td>
<td>21.00</td>
<td>3.55</td>
</tr>
<tr>
<td>Death</td>
<td>-340.39</td>
<td>34.81</td>
<td>-9.78</td>
</tr>
<tr>
<td>Number of cigarettes</td>
<td>-11.7</td>
<td>2.17</td>
<td>-5.39</td>
</tr>
<tr>
<td>Maternal height (cm)</td>
<td>2.83</td>
<td>1.53</td>
<td>1.85</td>
</tr>
</tbody>
</table>

† Also adjusted for time of measurement (h), gestational age (wk), maternal age (y) and maternal weight (kg).
distribution of the outcome may provide insights into the mechanism by which the treatment affects the outcome, and ideas to explain why a surrogate outcome (such as birth weight) may not reflect the effect of treatment on a primary outcome of interest (mortality). Such an approach could also be applied more generally to continuous outcomes in trials of nutritional interventions such as hemoglobin or Z-scores for weight-for-height or height-for-age. From a public health perspective, this approach can also help identify whether a targeted, rather than a universal supplementation program would be more effective and efficient in achieving a nutritional goal for a population.

Our data demonstrated that the mean treatment effect with iron folic acid was half that of the mean multiple micronutrient effect but in environments such as rural Nepal, it may be more important to affect the lower than the upper part of the birth weight distribution. In fact, affecting the upper part of the distribution may be harmful to the mother and infant. In our study, the multiple micronutrient supplement was associated with a slightly elevated risk of early infant mortality [relative risk (RR): 1.07, 95% CI: 0.75, 1.58], especially among term births (RR: 1.74, 95% CI: 1.00, 3.04) (2). This was despite the significant 14% reduction in low birth weight. The risk of birth asphyxia as a cause of neonatal mortality also appeared to be higher in the group administered the multiple micronutrient supplement. On the other hand, folic acid + iron was associated with an overall 21% reduction (95% CI: 48% reduction to increase of 20%), and ~47% reduction in 3-mo mortality (95% CI: 8, 70%) among preterm births. Given an improvement in birth weight at the lower end of the distribution, this intervention may have produced improved survival overall, whereas the multiple micronutrient appeared to have no effect on survival because deaths averted in the smaller infants were negated by higher mortality at the upper end of the distribution. In a recent study examining a multiple micronutrient supplement relative to a control supplement of iron + folic acid in Janakpur, Nepal, there was a higher mean birth weight of 77 g (95% CI: 24, 130) in the multiple micronutrient group compared with the iron + folic acid group, but higher neonatal mortality in the multiple micronutrient group (RR: 1.53, 95% CI: 0.72, 3.2) (10). Pooled analyses of the Sarlahi

![FIGURE 2](https://academic.oup.com/jn/article-abstract/136/5/1389/4670045) The predicted values of birth weight plotted against time of measurement using a spline with knots at 24, 48, and 72 h (A). The deviations in predicted values from the average birth weight are plotted against gestational age at birth (B), maternal age (C), and first trimester maternal weight (D), also using natural splines.
and Janakpur studies found higher neonatal mortality (RR: 1.52 (1.03, 2.25) among those in the multiple micronutrient group relative to the folic acid group (2,11).

When examined cross-sectionally, our data also found evidence at a population level for a decline in weight of $50\text{ g}$ between 24 and 48 h after birth. This is comparable to the drop in weight seen in a population of newborns in Bangladesh (12). The weights rise thereafter, and from 72 h through 3 mo of age, weights can be modeled as a linear function of age in this population. For those who died before a weight measurement could be obtained, or who survived but were missing weight measurements due to migration, birth weights can be imputed using a variety of maternal and other covariates that are strong predictors of birth weight. Using recalibrated birth weights for those measured at varying times after birth, and imputed weights for those with missing weights, the estimated effect of maternal supplementation was slightly larger for all treatment groups than when the treatment effects were estimated without imputation (Fig. 3 and Table 2). The CI for the treatment effects using imputed values were wider, reflecting the uncertainty associated with the imputed birth weights. Although there is an assumption in imputation that the infants missing birth weight would have a similar weight to infants with similar covariates who were not missing weight, there are more infants missing weight because they died soon after birth, and such infants might have been smaller than those who survived long enough to be weighed. However, we do have birth weights for some infants who later died, and this imputation model does use those infants to help predict the birth weight of those who died before we could weigh them. The birth weight differential predicted from the model between those who died and those who survived is $300\text{ g}$, although the babies we did weigh who

FIGURE 3 The estimated treatment effects shown as a function of the percentiles of the birth weights in the control group (A, iron + folic acid; B, multiple micronutrients; C, folic acid + iron + zinc; D, folic acid). The x-axis show the percentiles of the birth weights in the control group. The zero line indicates no treatment effect and a negative treatment effect implies a reduction in birth weight associated with the treatment in that range of the birth weight distribution. The central solid black line is the smoothed treatment effect, with upper and lower 95% bootstrapped confidence bands. The dotted line on the right denotes the estimated crude treatment effects and 95% CI. The solid line on the right denotes the estimated treatment effect and 95% bootstrap CI obtained by averaging the treatment effect across the percentiles making use of the imputed values for birth weight.
TABLE 2

Crude and imputed mean difference and 95% CI in birth weight between treatment and control groups

<table>
<thead>
<tr>
<th>Treatment group</th>
<th>Crude mean difference (CMD)</th>
<th>Imputed mean difference (IMD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Folic acid</td>
<td>-7.1 (-61.0, 46.8)</td>
<td>16.9 (-54.5, 88.3)</td>
</tr>
<tr>
<td>Folic acid + iron</td>
<td>60.0 (4.1, 115.9)</td>
<td>71.1 (0.4, 141.9)</td>
</tr>
<tr>
<td>Folic acid + iron + zinc</td>
<td>0.3 (-57.1, 57.8)</td>
<td>26.3 (-42.4, 95.0)</td>
</tr>
<tr>
<td>Multiple micronutrients</td>
<td>63.5 (9.0, 118.0)</td>
<td>84.6 (15.6, 153.6)</td>
</tr>
</tbody>
</table>

Later died may not adequately capture this difference in birth weight. It is also possible that babies who were large relative to the size of their mothers may have died so soon after birth that they were also not weighed. This may be true given a reverse J-shaped risk profile for mortality of these infants by birth weight (13). Thus, there may be some bias, but the direction of the bias is uncertain. However, only 2.4% of missing birth weights were imputed because the child had died before weight could be obtained; therefore, it is likely that the bias from imputing weights for these infants on overall birth weight is small. Given the constraints in obtaining birth weight in settings in which births occur at home, future studies or programs are likely to face the same situation with missing birth weight. Our measurement error model for imputing the missing birth weights and predicting the weights "at birth" for measurements made after the 72 h may be relevant for future studies that examine intervention effects on birth weight.

Because of the strong association between birth weight and survival, which may be an inverse J-shape in this population (13), understanding the effect of a treatment on different parts of the birth weight distribution may provide more insight into the effect of the treatment on both the surrogate (birth weight) and primary outcomes (survival) (14). Such analyses can help explain contradictions in treatment effects on surrogate and primary outcomes, and may help identify the extent to which a treatment effect acts directly on survival or through its effect on birth weight. Direct analyses of treatment effects on mortality by birth weight distribution are currently underway (9).

LITERATURE CITED