Micronutrient deficiencies and developmental delays among infants: evidence from a cross-sectional survey in rural China

Renfu Luo,1 Yaojiang Shi,2 Huan Zhou,3 Ai Yue,2 Linxiu Zhang,1 Sean Sylvia,4 Alexis Medina,5 Scott Rozelle5

ABSTRACT

Objectives: Research increasingly indicates the importance of the nutritional programming that occurs in the first 2–3 years of life. Quality nutrition during this brief window has been shown to have large and significant effects on health and development throughout childhood and even into adulthood. Despite the widespread understanding of this critical window, and the long-term consequences of leaving nutritional deficiencies unaddressed, little is known about the status of infant nutrition in rural China, or about the relationship between infant nutrition and cognitive development in rural China.

Design, setting and participants: In April 2013 and October 2013, we conducted a survey of 1808 infants aged 6–12 months living in 351 villages across 174 townships in nationally designated poverty counties in rural areas of southern Shaanxi Province, China.

Main outcome measures: Infants were administered a finger prick blood test for haemoglobin and assessed according to the Bayley Scales of Infant Development. They were also measured for length and weight. Caregivers were administered a survey of demographic characteristics and feeding practices.

Results: We found that 48.8% of sample infants were anaemic, 3.7% were stunted, 1.2% were underweight and 1.6% were wasted. Approximately 20.0% of the sample infants were significantly delayed in their cognitive development, while just over 32.3% of the sample infants were significantly delayed in their psychomotor development. After controlling for potential confounders, infants with lower haemoglobin counts were significantly more likely to be delayed in both their cognitive (p<0.01) and psychomotor development (p<0.01).

Conclusions: The anaemia rates that we identify in this study classify anaemia as a ‘severe’ public health problem according to the WHO. In contrast, there is virtually no linear growth failure among this population. We find that low haemoglobin levels among our sample population are associated with significant cognitive and psychomotor delays that could eventually affect children’s schooling performance and labour force outcomes.

Trial registration number: ISRCTN44149146.

INTRODUCTION

Despite China’s rapid growth, many residents in inland provinces are still plagued by poverty. One of the outcomes of this poverty is that up to 40% of primary school children in poor regions suffer from micronutrient deficiencies.1 In China, the most common form of micronutrient deficiency among school-aged children is iron deficiency.2 Iron deficiency anaemia in the first years of life has been related to many adverse consequences. In the short term, infants with iron deficiency anaemia are at higher risk of cognitive, social and emotional delays.3–11 In the longer run, it can negatively affect school performance and behaviour, reduce overall educational attainment and negatively affect work outcomes into adulthood; these

Strengths and limitations of this study

The key strengths of this study are its population-based sampling technique, large sample size (over 1800 children) and rigorous child development testing, all of which increase confidence in the validity and generalisability of our findings.

The primary study limitation is that we were unable to conduct full blood panel testing for nutritional deficiencies, and are therefore limited to considering haemoglobin as our sole indicator of micronutrient deficiency. We acknowledge that (1) low haemoglobin levels alone do not definitively indicate iron deficiency and (2) other micronutrient deficiencies such as zinc, folate or vitamin A may be driving our correlational results.

Since our study is correlational in nature, we are limited in the conclusions we may draw from our results. Future work is needed to refine the results presented here, and to better identify whether a causal link exists between infant nutrition and early child development in rural China.
consequences are irreversible even if the anaemia is corrected in later childhood.\textsuperscript{12-15} Many of these consequences can be avoided, however, simply through proper dietary intake in the first years of life; studies have shown that regular iron supplementation can reduce anaemia rates by as much as 50%.\textsuperscript{16} Previous work in China has shown that interventions of vitamin supplements and school lunch programmes reduce anaemia prevalence and raise test scores among primary school children.\textsuperscript{17-19}

Little is known, however, about micronutrient deficiencies among one of the most vulnerable population groups in China. Research increasingly indicates the importance of the nutritional programming that occurs in the first 2–3 years of life.\textsuperscript{20-22} Quality nutrition during this brief window has been shown to have large and significant effects on health and development throughout childhood, and even to affect outcomes in adulthood.\textsuperscript{12-14}

Despite the widespread understanding of this critical window and the long-term consequences of leaving nutritional deficiencies unaddressed, little is known about the status of infant nutrition or the relationship between infant nutrition and cognitive development in rural China. Although researchers have attempted to measure the prevalence of nutritional deficiencies in China overall, evidence on the situation in rural China is mixed. The Chinese Food and Nutrition Surveillance System found anaemia prevalence among infants aged 6–12 months in rural areas to be around 28% in 2010.\textsuperscript{2} Other, more geographically focused studies have found anaemia prevalence ranging from 22.6% (in Guangxi)\textsuperscript{23} to 58.2% (in Gansu)\textsuperscript{24} among the same age group. Moreover, most of these studies have been based on relatively small sample sizes, limited sample areas or both. To the best of our knowledge, while there has been one study on the impact of mothers’ nutrition on early infant development in China,\textsuperscript{25} no studies have reported on the status of infant development or examined the correlation between infant micronutrient deficiencies and development among infants aged 6–12 months in China.

The goal of this paper is to provide an overview of infant nutrition in rural China, and to explore the link between nutritional status and child development. More specifically, we measure the prevalence of micronutrient deficiencies (haemoglobin (Hb) concentrations/anaemia) and macronutrient deficiencies (stunting, underweight and wasting) among infants aged 6–12 months living in nationally designated poverty counties in rural Shaanxi Province. In addition, we examine correlations between Hb concentrations, and infants’ cognitive and motor development, using the Bayley Scales of Infant Development (BSID).

METHOD

Sample selection

Our study was conducted in 2013, in 11 nationally designated poverty counties located in southern Shaanxi Province. From each of these 11 counties, all townshipshas were enrolled in the study. If a village had fewer than five infants in our desired age range, we randomly selected an additional village in the same township for inclusion in the study, and continued to randomly select additional villages until five infants per township had been found. The sample villages were all revisited in October 2013, and a new cohort of infants (also aged 6–12 months) was surveyed at that time. Subsequently, overall, our study included 1808 infants in 351 villages across 174 townships.

All surveys and tests of nutritional status and cognitive development were administered on the same day for each household.

Data collection

With the assistance of trained nurses from Xi’an Jiaotong Medical School, we collected Hb concentrations from all participating infants and their mothers. Hb concentrations were measured onsite using a HemoCue Hb 201+ finger prick system. The nurses also measured the length and weight of each infant, according to WHO recommendations.\textsuperscript{26}

Teams of trained enumerators collected socioeconomic data from all households participating in the study. Each infant’s primary caregiver was identified, and administered a detailed survey on parental and household characteristics, including each child’s gender and birth order, maternal age and education, and whether the family was receiving minimum living standard guarantee payments (a poverty indicator). The infant’s age was obtained from his or her birth certificate. The primary caregiver was individually identified by each family as the individual most responsible for the infant’s care (typically the child’s mother or grandmother).
All infants were also administered the BSID Version I, an internationally scaled test of infant and toddler cognitive and motor development. This test is well recognised in the psychological literature and is listed by the American Psychiatric Association as a way to diagnose certain developmental disorders. The test was formally adapted to the Chinese language and environment in 1992, and scaled according to an urban Chinese sample.

Following the example of other published studies that use the BSID to assess infant development in China, it was this officially adapted version of the test that was used in this study. The test has an inter-rater reliability of 0.99 for each of the two subindices, the Mental Development Index (MDI) and the Psychomotor Development Index (PDI). (Both indices are described in more detail below.) The test-retest reliability is high, at 0.82 for MDI and 0.88 for PDI. The parallel forms reliability is also high, at 0.85 for MDI and 0.87 for PDI, indicating that the test scores are consistent when there is a variation in the methods or instruments used in the test.

All BSID enumerators attended a week long training course on how to administer the BSID, including a 2.5 day experiential learning programme in the field. The test was administered one-on-one in the household using a set of standardised toys and a detailed scoring sheet. The BSID takes into consideration each infant’s age in days, as well as whether he or she was a premature birth. These two factors, combined with the infant’s performance on a series of tasks using the standardised toy kit, contribute to the establishment of two independent, internationally standardised scores: MDI, which evaluates memory, habitation, problem solving, early number concepts, generalisation, classification, vocalisations and language to produce a measure of cognitive development; and PDI, which evaluates gross muscle groups (rolling, crawling and creeping, sitting, standing, walking, running and jumping) and fine motor manipulation to produce a measure of psychomotor development.

This study represents one of the largest administrations of the BSID ever conducted in China and, to the best of our knowledge, the only administration of the BSID ever conducted in rural communities in China’s nationally designated poverty counties.

Ethical approval
All participating caregivers gave their oral consent for both their own and their infant’s involvement in the study. Children who were found to have severe anaemia were referred to the local hospital for treatment.

Statistical analysis
Anaemia status was assessed based on finger blood analysis for Hb. Following internationally accepted standards, anaemia was defined as Hb<110 g/L, moderate anaemia was defined as Hb>70 g/L but <100 g/L, and severe anaemia was defined as Hb<70 g/L.

Physical indicators of length and weight were compared with the 2006 WHO child growth standards, to calculate length-for-age (LAZ), weight-for-age (WAZ) and weight-for-length z-scores (WLZ). We followed internationally recognised cut-offs to consider children whose LAZ, WAZ, or WLZ fall more than two SDs below the international mean to be stunted, underweight or wasted, respectively.

Cognitive development (MDI) and psychomotor development (PDI) scores were determined using the BSID protocol. Both indices are scaled to have an expected mean of 100 and a SD of 16. Scores on each index can range between 50 and 150. Mild impairment for each index is defined as 70≤MDI<80 and 70≤PDI<80, while moderate or severe impairment for each index is defined as MDI<70 and PDI<70. Infants failing to achieve the minimum MDI or PDI score (50) were assigned a score of 49.

Statistical analyses were performed using STATA 12.0. p Values below 0.05 were considered statistically significant. Linear regression and logistic regression were used for multivariate analyses as appropriate. We included the following variables as potential confounders in the multivariate analysis: gender, age, whether the child was born prematurely, birth order, whether the child’s mother was identified as the primary caregiver, maternal educational level and age, and whether the child’s family received minimum living standard guarantee payments. All multiple linear regressions adjust for county fixed effects. SEs account for clustering at the village level. Lowess curves reflecting the relationship between Hb concentration and BSID score were estimated using the loess procedure in SAS.

Role of the funding source
The study sponsors had no role in study design; in the collection, analysis, or interpretation of data; in the writing of the report; or in the decision to submit the paper for publication.

RESULTS
Table 1 shows the demographic characteristics of the sample. Slightly over half of the sample infants were male (53.2%). Around 4.6% of sample infants were premature and 62.6% were first-order births. The mother was the primary caregiver for 79.6% of the infants in the sample. The majority of the mothers (81.1%) had completed fewer than 9 years of schooling; 49.9% were aged over 25 years. About a one-quarter (24.3%) of sample families reported receiving minimum living standard guarantee payments, a form of government welfare for the lowest income families nationwide.

Hb concentration, anaemia and physical development among rural Chinese infants
Hb concentrations were normally distributed with a mean of 109.1 g/L (table 2). In our sample, 48.8% of...
infants had Hb concentrations below 110 g/L, indicating anaemia. A total of 363 infants (19.4%) had Hb concentrations below 100 g/L, classifying them as moderately or severely anaemic. There was considerable variation in anaemia prevalence across the sample counties, ranging from 33.3% to 69.7%.

Table 2 further shows indicators of the infants’ physical development, as measured by LAZ, WAZ and WLZ. The data show that 3.7% of infants in our sample are stunted, 1.2% are underweight and 1.6% are wasted.

MDI and PDI among rural Chinese infants
BSID measures were available for all 1808 infants. Two per cent of the sample infants failed to achieve the minimum score (50) on the MDI and were assigned a score of 49. No infants scored below 50 on the PDI. Table 3 shows the MDI scores for all 1808 sample infants. The mean MDI score for the sample was 96.7, significantly lower than the expected mean of 100 (p<0.001). The SD was 17.0. The BSID test results show that 6.9% had an MDI score below 70, thereby classifying them as moderately or severely impaired in their cognitive development. Around 13.1% of infants had an MDI score between 70 and 80, which indicates mild cognitive impairment. In total, 20.0% had MDI scores lower than 80.

Table 3 also shows the PDI score for all 1808 sample infants. The mean PDI score for the sample was 90.0, significantly lower than both the observed MDI score and the expected mean PDI score of 100 (p<0.001). The SD was 17.2. The data show that 13.1% of sample infants were moderately or severely impaired in their psychomotor development (PDI<70) and 19.2% were mildly impaired (70≤PDI<80).

Multivariate analysis
The multivariate analyses show a significant positive association between Hb concentrations and both MDI and PDI scores after controlling for potential confounding variables (table 4). More precisely, our analysis shows that a 10 g/L rise in infant Hb concentration is associated with a 0.9 point rise in the MDI score and a 1.0 point rise in the PDI score. Logistic analyses using three commonly used BSID score cut-offs are largely consistent with this result for both MDI and PDI (see online supplementary appendix table S1).

Figure 1 shows a lowess plot of the relationship between Hb concentrations and BSID scores. Both panels show a positive relationship between Hb concentrations, and MDI and PDI scores. This positive relationship is approximately linear and evident both below and above the WHO cut-off for anaemia (110 g/L).
**DISCUSSION**

We observed that 48.8% of infants aged 6–12 months in low-income areas of rural China are anaemic. This is nearly twice as high as the 2011 average for East and Southeast Asia. Nearly 20% of the sample suffers from moderate or severe anaemia. Stunting, underweight and wasting, however, are much less prevalent in this population, ranging from only 1.2–3.7%. Such low rates are indicative of a population with virtually no linear growth failure, a notable finding for China, especially given that, as recently as 2012, stunting prevalence among children under 5 years of age in rural areas was observed to be around 20%. This contrast suggests that it is now the quality of infant diets, rather than the quantity of food, that is lacking in rural China today.

**Table 3** Cognitive and psychomotor development of sample infants in rural Shaanxi Province (n=1808)

<table>
<thead>
<tr>
<th></th>
<th>Mean/percent (95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MDI score</strong></td>
<td>96.7±17.0 (96.0 to 97.5)</td>
</tr>
<tr>
<td><strong>Cognitive impairment</strong></td>
<td></td>
</tr>
<tr>
<td>Moderate or severe (MDI&lt;70)</td>
<td>6.9 (125) (5.7 to 8.1)</td>
</tr>
<tr>
<td>Mild (70≤MDI&lt;80)</td>
<td>13.1 (237) (11.6 to 14.7)</td>
</tr>
<tr>
<td>Any (MDI&lt;80)</td>
<td>20.0 (362) (18.2 to 21.9)</td>
</tr>
<tr>
<td><strong>PDI score</strong></td>
<td>90.0±17.2 (89.2 to 90.8)</td>
</tr>
<tr>
<td><strong>Psychomotor impairment</strong></td>
<td></td>
</tr>
<tr>
<td>Moderate or severe (PDI&lt;70)</td>
<td>13.1 (236) (11.5 to 14.6)</td>
</tr>
<tr>
<td>Mild (70≤PDI&lt;80)</td>
<td>19.2 (348) (17.4 to 21.1)</td>
</tr>
<tr>
<td>Any (PDI&lt;80)</td>
<td>32.3 (584) (30.1 to 34.5)</td>
</tr>
</tbody>
</table>

Data are presented as mean±SD or % (n) for categorical variables.

MDI, Mental Development Index.

**Table 4** Association between Hb concentration and BSID scores (n=1808)

<table>
<thead>
<tr>
<th></th>
<th>MDI score*</th>
<th>p Value</th>
<th>PDI score</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hb concentration (g/L)</strong></td>
<td>Coefficient (95% CI)</td>
<td>p Value</td>
<td>Coefficient (95% CI)</td>
<td>p Value</td>
</tr>
<tr>
<td>7</td>
<td>0.09 (0.03 to 0.15)</td>
<td>&lt;0.01</td>
<td>0.10 (0.04 to 0.16)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Sex (female=1)</td>
<td>1.31 (–0.16 to 2.78)</td>
<td>0.08</td>
<td>0.44 (–1.05 to 1.93)</td>
<td>0.56</td>
</tr>
<tr>
<td>Infant age (6 months is base)</td>
<td>(months) 7</td>
<td>9.05 (5.56 to 12.53)</td>
<td>&lt;0.01</td>
<td>6.20 (3.30 to 9.10)</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>7.79 (4.38 to 11.20)</td>
<td>&lt;0.01</td>
<td>4.06 (1.09 to 7.03)</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>6.69 (3.17 to 10.21)</td>
<td>&lt;0.01</td>
<td>–0.65 (–3.89 to 2.58)</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>5.92 (2.34 to 9.49)</td>
<td>&lt;0.01</td>
<td>–7.69 (–11.0 to –4.35)</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>4.78 (1.45 to 8.12)</td>
<td>&lt;0.01</td>
<td>–2.04 (–5.36 to 1.28)</td>
</tr>
<tr>
<td>Is the infant premature? (yes=1)</td>
<td>–1.70 (–6.10 to 2.69)</td>
<td>0.45</td>
<td>0.76 (–2.71 to 4.23)</td>
<td>0.67</td>
</tr>
<tr>
<td>Birth order of infant (second or higher=1)</td>
<td>–0.83 (–2.79 to 1.14)</td>
<td>0.41</td>
<td>–1.37 (–3.41 to 0.67)</td>
<td>0.19</td>
</tr>
<tr>
<td>Mother is primary caregiver (yes=1)</td>
<td>0.55 (–1.49 to 2.59)</td>
<td>0.60</td>
<td>1.40 (–0.85 to 3.65)</td>
<td>0.22</td>
</tr>
<tr>
<td>Maternal educational level (more than 9 years=1)</td>
<td>0.96 (–0.93 to 2.85)</td>
<td>0.32</td>
<td>2.41 (0.51 to 4.30)</td>
<td>0.01</td>
</tr>
<tr>
<td>Maternal age (more than 25 years=1)</td>
<td>–1.19 (–3.12 to 0.73)</td>
<td>0.22</td>
<td>–0.94 (–2.86 to 0.98)</td>
<td>0.34</td>
</tr>
<tr>
<td>Family receives minimum living standard guarantee (yes=1)</td>
<td>–0.42 (–2.16 to 1.32)</td>
<td>0.64</td>
<td>–0.71 (–2.58 to 1.17)</td>
<td>0.46</td>
</tr>
</tbody>
</table>

*In addition to the covariates presented, the multiple linear regressions also adjust for county fixed effects. SEs account for clustering at the village level.

BSID, Bayley Scales of Infant Development; Hb, haemoglobin; MDI, Mental Development Index; PDI, Psychomotor Development Index.
Our data further show that the high prevalence of anaemia that we observe is significantly correlated with high rates of cognitive and psychomotor delays, as measured by infant performance on the BSID. Around 20% of the sample suffered from some degree of cognitive delay, while nearly one-third (32.3%) of the sample suffered from some degree of psychomotor delay. These numbers are quite a bit higher than what would be expected in a normal population. In the urban Chinese population against which the BSID scores were normalised, only 11.7% of children would be expected to have cognitive delays, and only 11.0% would be expected to have psychomotor delays. In total, 39.2% of the sample suffered from one or both types of developmental delay, compared with an expected 22.7% in a healthy Chinese population. Moreover, children with lower Hb counts were significantly more likely to suffer from either type of delay.

The graph of the relationship between Hb concentrations and BSID scores demonstrates that the association between nutrition and child development exists both below and above the 110 g/L mark, the traditional cut-off point used to identify an infant as having anaemia. In other words, even for those infants the WHO considers to be non-anaemic, higher Hb levels are associated with better cognitive and psychomotor development. While our data cannot prove causality, these results merit careful consideration, since they point to a link between Hb and child development even for children traditionally considered to be “healthy”.

We note that our results are not sensitive to our treatment of infants failing to reach the minimum MDI score. As described in the Results section, in our sample, 31 infants, or 2% of the sample infants, failed to achieve the minimum score (50) on the MDI and were assigned a score of 49. We included these measures in the analyses, reasoning that their inclusion helps contribute to better overall explanatory power. Since some researchers prefer to exclude ‘failing’ observations from the analyses, we also ran all of our analyses with these 31 failing observations excluded from the analysis (n=1777, not shown), and the results were not statistically different from the results shown in this paper.

The strengths of this study include its population-based sampling technique, large sample size and rigorous child development testing, all of which increase confidence in the validity and generalisability of our findings. A primary weakness is that we were unable to conduct full blood panel testing for nutritional deficiencies, and are therefore limited to considering Hb as our sole indicator of micronutrient deficiency. We acknowledge that (1) low Hb levels alone do not definitively indicate iron deficiency and (2) other micronutrient deficiencies such as zinc, folate, or vitamin A may be driving our correlational results. While iron deficiency is not necessarily the only cause of anaemia in our sample areas, evidence suggests that other potential causes (such as helminthic infections) are unlikely to be the driving factor behind low levels of Hb among our sample population. Indeed, several studies report that 85–95% of anaemia in China is caused by iron deficiency.

While our study highlights the important link between infant nutrition and early child development, we point out that other factors may also be at play. For example, studies elsewhere have indicated the importance of parental stimulation as a contributing factor to child development. Maternal mental health and the home environment may also play a role. Since our study is correlational in nature, future work is needed to refine the results presented here and better identify whether a causal link exists between infant nutrition and early child development in rural China.

Research in context

Overall, our findings indicate that low Hb levels are a serious problem in poor areas of rural China, affecting a large proportion of Chinese infants living in these areas, and officially categorising anaemia as a ‘severe’ public health problem according to the WHO. The severity of this problem in a country that has already reached middle-income status is alarming, especially in the face of recent evidence that anaemia among this age group can be effectively treated with a low-cost daily iron supplement. Moreover, we find that the low Hb levels that we observe in this study are associated with significant cognitive and psychomotor delays that could eventually affect children’s schooling performance and labour force outcomes. The Chinese government should take steps to ensure that these infants have access to a more balanced diet, and should do so before irreversible damage is done to this vulnerable population.
submitted version of the manuscript and accept accountability for all aspects of the work.

Funding International Initiative for Impact Evaluation (3ie) (grant number PW2.03.CH.1E), UBS Optimus Foundation (grant number 6140), China Medical Board (grant number 13-158), Bank of East Asia, HJ Heinz Company Foundation, Huqiao Foundation and Nobelles.

Competing interests None declared.

Ethics approval This study received ethical approval from the Stanford University Institutional Review Board (IRB) (Protocol ID 25734) and from the Sichuan University Ethical Review Board (Protocol ID 2013005–01).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Statistical code and datasets are available from corresponding author on emailed request.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

REFERENCES


Supplemental Appendix: Sample Size Calculations

The total sample size for this study was determined by power calculations for a larger, interventional study not reported in this paper. The larger interventional study was designed as a cluster-randomized controlled trial with two intervention groups and a control group.

In order to determine sample size we performed power calculations using Optimal Design software, developed by a group of researchers at the University of Michigan. Specifically, we calibrated the model as appropriate for a Cluster Randomized Trial with a continuous outcome and clustering at the township level. (For details about the math behind the calculation, please visit the Optimal Design website at http://sitemaker.umich.edu/group-based/optimal_design_software.)

The power needed to detect a difference in outcome variables between the treatment and control groups in a Cluster Randomized Trial with a continuous outcome depends on 6 factors:

a.) the number of babies per township (n);
b.) the number of townships (J);
c.) the intra-township correlation in outcome variable (ρ);
d.) the minimum effect size that we would expect to be able to detect (δ); and
e.) the proportion of variation in the true township-level post-intervention outcome variable explained by township-level pre-intervention outcome variable ($R^2$).

Using estimates from a survey conducted by the School of Public Health at Sichuan University, we assumed 10 infants will be born per township over a six month period, and $ρ=0.10$.

Based on a Chinese study of the nutritional supplement packets that were ultimately used as part of the project interventions (citation available upon request), we predicted a minimum effect size of 0.25. We supposed that we would be able to follow-up with 9 of 10 infants per township after one year, with $R^2=0.5$.

With the above assumptions on the parameters, using the Cluster Randomized Trials setting in the Optimal Design software, we found that with regard to Intervention 1, for one pair of treatment-control comparison, with about 80 townships, we could achieve power=0.81 at the conventional significance level ($α=0.05$). We would then randomly assign 40 townships into Intervention Group 1 and 40 townships into the Control Group. We could also achieve power=0.81 when we compare the difference between Intervention Group 1 and Intervention Group 2 at the conventional significance level of $α=0.05$. (This is because the difference of minimum detectable effect (MDE) between these two intervention groups was also 0.25 S.D.)

For the second evaluation survey, the power calculations were slightly different, though the end result is the same. In this case, we desired a minimum effect size of 0.3 and 0.6 for Intervention Groups 1 and 2, respectively, in the second follow-up survey. We supposed that we could only follow-up 6 of 10 infants aged 24-30 months per township, with $R^2=0.5$. 
With these new assumptions on the parameters, using the Cluster Randomized Trials setting in the Optimal Design software, we found that with regard to Intervention 1, for one pair of treatment-control comparison, with about 80 townships, we could achieve power=0.84 at the conventional significance level ($\alpha=0.05$). We would then randomly assign 40 townships into Intervention Group 1 and 40 townships into the Control Group. We could also achieve power=0.84 when we compared the difference between Intervention Groups 1 and 2 at the conventional significance level of $\alpha=0.05$. (This is because the difference of MDE between these two intervention groups is also 0.25 S.D.)

Overall then, we required at least 40 townships in the Control Group, 40 townships in Intervention Group 1, and 40 townships in Intervention Group 2 in order to observe differences among the three groups. We rounded up to 50 townships per arm to allow for the possibility of attrition, giving a total of 150 townships. If there were no attrition, we welcomed the additional power that a larger sample would provide.
### Appendix Table 1: Association between Hb concentration and BSID scores, using three commonly used score cutoffs (n=1808)

<table>
<thead>
<tr>
<th></th>
<th>MDI &lt;100 (Yes=1)</th>
<th></th>
<th>MDI &lt;85 (Yes=1)</th>
<th></th>
<th>MDI &lt;80 (Yes=1)</th>
<th></th>
<th>PDI &lt;100 (Yes=1)</th>
<th></th>
<th>PDI &lt;85 (Yes=1)</th>
<th></th>
<th>PDI &lt;80 (Yes=1)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Hb concentration (g/L)</td>
<td>-0.002 (0.00; 0)</td>
<td>0.02</td>
<td>-0.00 (0.00; 0)</td>
<td>0.47</td>
<td>-0.00 (0.00; 0)</td>
<td>0.29</td>
<td>-0.00 (0.00; 0)</td>
<td>0.01</td>
<td>-0.002 (0.00; 0)</td>
<td>0.03</td>
<td>-0.00 (0.00; 0)</td>
<td>0.29</td>
</tr>
<tr>
<td>Sex (Female=1)</td>
<td>-0.03 (0.07; 0.01)</td>
<td>0.18</td>
<td>-0.03 (0.06; 0.01)</td>
<td>0.12</td>
<td>-0.01 (0.05; 0.02)</td>
<td>0.44</td>
<td>-0.03 (0.07; 0.02)</td>
<td>0.23</td>
<td>0.00 (0.04; 0.04)</td>
<td>0.90</td>
<td>-0.01 (0.04; 0.02)</td>
<td>0.44</td>
</tr>
<tr>
<td>Infant age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 months</td>
<td>-0.18 (-0.27; -0.09)</td>
<td>&lt;0.01</td>
<td>-0.18 (-0.26; -0.10)</td>
<td>&lt;0.01</td>
<td>-0.14 (-0.21; -0.08)</td>
<td>&lt;0.01</td>
<td>-0.06 (-0.14; 0.03)</td>
<td>0.17</td>
<td>-0.21 (-0.29; -0.12)</td>
<td>&lt;0.01</td>
<td>-0.14 (-0.21; -0.08)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>8 months</td>
<td>-0.15 (-0.25; -0.06)</td>
<td>&lt;0.01</td>
<td>-0.15 (-0.23; -0.08)</td>
<td>&lt;0.01</td>
<td>-0.12 (-0.18; -0.05)</td>
<td>&lt;0.01</td>
<td>-0.03 (-0.12; 0.06)</td>
<td>0.54</td>
<td>-0.19 (-0.28; -0.11)</td>
<td>&lt;0.01</td>
<td>-0.18 (-0.18; -0.05)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>9 months</td>
<td>-0.07 (-0.12; 0)</td>
<td>0.12</td>
<td>-0.09 (-0.17; -0.01)</td>
<td>0.04</td>
<td>-0.09 (-0.16; -0.02)</td>
<td>0.01</td>
<td>-0.07 (-0.02; 0.15)</td>
<td>0.11</td>
<td>0.03 (-0.06; 0.12)</td>
<td>0.58</td>
<td>-0.09 (-0.16; 0)</td>
<td>0.01</td>
</tr>
<tr>
<td>10 months</td>
<td>-0.08 (-0.17; 0.01)</td>
<td>0.10</td>
<td>-0.08 (-0.16; 0.00)</td>
<td>0.06</td>
<td>-0.08 (-0.14; 0.01)</td>
<td>0.02</td>
<td>0.15 (0.06; 0.23)</td>
<td>&lt;0.01</td>
<td>0.21 (0.12; 0.30)</td>
<td>&lt;0.01</td>
<td>-0.08 (0.14; 0.01)</td>
<td>0.02</td>
</tr>
<tr>
<td>11 months</td>
<td>-0.08 (-0.17; 0)</td>
<td>0.05</td>
<td>-0.03 (-0.11; 0.05)</td>
<td>0.44</td>
<td>-0.02 (-0.09; 0.04)</td>
<td>0.50</td>
<td>0.04 (-0.04; 0.12)</td>
<td>0.33</td>
<td>0.06 (-0.03; 0.15)</td>
<td>0.17</td>
<td>-0.02 (-0.09; 0.04)</td>
<td>0.50</td>
</tr>
<tr>
<td>Is the infant pregnant?</td>
<td>0.00 (0.01; 0.11)</td>
<td>0.08</td>
<td>0.08 (-0.00; 0.17)</td>
<td>0.05</td>
<td>0.04 (-0.03; 0.11)</td>
<td>0.26</td>
<td>-0.14 (-0.06; 0.04)</td>
<td>0.40</td>
<td>-0.07 (-0.17; 0.03)</td>
<td>0.17</td>
<td>-0.03 (-0.03; 0.11)</td>
<td>0.26</td>
</tr>
<tr>
<td>Birth order of infant</td>
<td>0.06 (0.00; 0.12)</td>
<td>0.04</td>
<td>0.04 (-0.01; 0.09)</td>
<td>0.15</td>
<td>0.01 (-0.03; 0.05)</td>
<td>0.63</td>
<td>-0.04 (-0.04; 0.07)</td>
<td>0.57</td>
<td>0.00 (-0.05; 0.06)</td>
<td>0.87</td>
<td>0.00 (-0.03; 0.05)</td>
<td>0.63</td>
</tr>
<tr>
<td>Mother is primary caregiver?</td>
<td>-0.04 (0.10; 0.02)</td>
<td>0.17</td>
<td>0.01 (-0.03; 0.06)</td>
<td>0.56</td>
<td>-0.01 (-0.01; 0.07)</td>
<td>0.15</td>
<td>-0.06 (-0.06; 0.05)</td>
<td>0.82</td>
<td>-0.03 (-0.09; 0.02)</td>
<td>0.23</td>
<td>-0.01 (-0.01; 0.07)</td>
<td>0.15</td>
</tr>
<tr>
<td>Maternal educational level (more than 9 years=1)</td>
<td>-0.02 (-0.08; 0.03)</td>
<td>0.43</td>
<td>0.01 (-0.04; 0.05)</td>
<td>0.72</td>
<td>0.01 (-0.05; 0.03)</td>
<td>0.67</td>
<td>-0.06 (-0.11; -0.01)</td>
<td>0.03</td>
<td>-0.06 (-0.11; 0.00)</td>
<td>0.06</td>
<td>-0.01 (-0.05; 0.03)</td>
<td>0.67</td>
</tr>
<tr>
<td>Maternal age (more than 25 years=1)</td>
<td>0.01 (-0.04; 0.07)</td>
<td>0.46</td>
<td>-0.00 (-0.05; 0.04)</td>
<td>0.96</td>
<td>-0.01 (-0.03; 0.05)</td>
<td>0.53</td>
<td>0.01 (-0.04; 0.06)</td>
<td>0.68</td>
<td>0.06 (0.01; 0.11)</td>
<td>0.03</td>
<td>0.01 (-0.03; 0.05)</td>
<td>0.53</td>
</tr>
<tr>
<td>Family receives Minimum Living Standard Guarantee (Yes=1)</td>
<td>0.04 (-0.01; 0.09)</td>
<td>0.13</td>
<td>-0.01 (-0.05; 0.04)</td>
<td>0.80</td>
<td>-0.01 (-0.05; 0.03)</td>
<td>0.66</td>
<td>0.02 (-0.03; 0.08)</td>
<td>0.34</td>
<td>0.03 (-0.02; 0.07)</td>
<td>0.30</td>
<td>-0.01 (-0.05; 0.03)</td>
<td>0.66</td>
</tr>
</tbody>
</table>

*In addition to the covariates presented, the multiple logistic regressions also adjust for county fixed effects. Standard errors account for clustering at the village level. Marginal effects show the change in probability when the independent variable increases by one unit.*