Quantifying and Exploiting the Age Dependence in the Effect of Supplementary Food for Child Undernutrition

Milinda Lakkam¹, Stefan Wager², Paul H. Wise³, Lawrence M. Wein⁴*

¹ Institute for Computational and Mathematical Engineering, Stanford University, Stanford, California, United States of America, ² Statistics Department, Stanford University, Stanford, California, United States of America, ³ School of Medicine, Stanford University, Stanford California, United States of America, ⁴ Graduate School of Business, Stanford University, Stanford, California, United States of America

Abstract

Motivated by the lack of randomized controlled trials with an intervention-free control arm in the area of child undernutrition, we fit a trivariate model of weight-for-age z score (WAZ), height-for-age z score (HAZ) and diarrhea status to data from an observational study of supplementary feeding (100 kCal/day for children with WAZ < −2.5) in 17 Guatemalan communities. Incorporating time lags, intention to treat (i.e., to give supplementary food), seasonality and age interactions, we estimate how the effect of supplementary food on WAZ, HAZ and diarrhea status varies with a child’s age. We find that the effect of supplementary food on all 3 metrics decreases linearly with age from 6 to 20 mo and has little effect after 20 mo. We derive 2 food allocation policies that myopically (i.e., looking ahead 2 mo) minimize either the underweight or stunting severity – i.e., the sum of squared WAZ or HAZ scores for all children with WAZ or HAZ < 0. A simulation study based on the statistical model predicts that the 2 derived policies reduce the underweight severity (averaged over all ages) by 13.6–14.1% and reduce the stunting severity at age 60 mo by 7.1–8.0% relative to the policy currently in use, where all policies have a budget that feeds 4% of children. While these findings need to be confirmed on additional data sets, it appears that in a low-dose (100 kCal/day) supplementary feeding setting in Guatemala, allocating food primarily to 6–12 mo infants can reduce the severity of underweight and stunting.

Introduction

With over 3 M deaths per year of children under 5 years attributable to undernutrition [1] and the level of food aid far less than required [2,3], it is vital to optimally allocate food to the appropriate children in the appropriate amounts. To address this problem in a rigorous manner requires knowledge about three key aspects [4]: (i) the evolution of weight and height (and perhaps disease) of children under 5 years in the absence of food aid; (ii) the impact that weight and height (and perhaps other factors such as age, sex and disease) have on morbidity and mortality; and (iii) the impact of supplementary or therapeutic food on weight and height (and perhaps other factors such as disease).

A bivariate statistical model of weight and height can be constructed from a longitudinal data set [4] to address (i), and logistic regression [4,5] or proportional hazards [6] models can estimate the impact of anthropometric measurements on mortality to address (ii). The ideal approach to (iii) would use longitudinal measurements of weight and height from a large randomized control trial with a control arm that does not receive any food. There appears to be only one study that possesses these characteristics [7] and no studies involving supplementary feeding [8]. As a result, very little is known about how the impact of food on weight and height is influenced by a child’s age, sex and pre-intervention weight, height and presence of disease, and the amount of therapeutic or supplementary food consumed. Furthermore, the ethical issues inherent in using a control group that does not receive food suggests that there may not be any large randomized control trials with intervention-free arms performed in the future.

Consequently, the most promising approach to making headway on aspect (iii) is to jointly estimate aspects (i) and (iii) – i.e., fit a statistical model of weight and height in the presence of nonrandom food aid – using observational or dose-response studies of nutrition programs that do not include intervention-free control arms. In this study, we make a first step in this direction and analyze data from a nutrition program in Guatemala, where every child under 5 years in 17 villages has his weight, height and diarrhea status (the number of days in the previous week with diarrhea) measured every 2 mo, and children with weight-for-age z score (WAZ) less than −2.5 receive supplementary food of ≥ 100 kCal/day. We use the WAZ = −2.5 threshold to cast the nutrition program as a natural experiment – where children with WAZ just above −2.5 (who do not receive treatment) act as a control group for the children with WAZ just below −2.5 (who do receive treatment) – and analyze it via a regression discontinuity design [9], which is a design where the assignment of an intervention is based on the value of an observed covariate lying below (or above) a fixed threshold. By comparing observations that are just above and just below WAZ = −2.5, the regression discontinuity design allows us to estimate the average treatment effect. In econometrics,
statistics and related fields, regression discontinuity designs have become the preferred approach for evaluating the effects of intervention in a setting such as ours, where randomization is not possible [9]. More specifically, a generalized additive model (GAM) [10] is used to predict future (i.e., 2 mo hence) WAZ, height-for-age z score (HAZ), and diarrhea status (because deaths from wasting do not occur in this setting, we focus on WAZ and HAZ instead of weight-for-height z score, WHZ) in terms of current and past values of (WAZ, HAZ, diarrhea) along with seasonal and age-interaction terms. This statistical model allows us to isolate how age influences the impact that supplementary food has on WAZ, HAZ, and diarrhea status. We use these statistical results to derive 2 allocation policies that minimize the underweight (or stunting) severity, i.e., the sum of squared expected WAZ (or HAZ) scores among children with WAZ < 0 (or HAZ < 0), 2 mo into the future. We build a simulation model based on our GAM that compares the derived allocation policies to the allocation policy that is currently used in the nutrition program in Guatemala, and to the counterfactual of no supplementary food.

Materials and Methods

Data

Our data set tracks 2125 children (1047 boys, 1078 girls) from 17 communities participating in a nutrition program in Guatemala from November 2009 to April 2012. Every 2 months, children from birth to age 60 mo had their WAZ, HAZ – both measured according to World Health Organization (WHO) child growth standards [11] – and the number of days with diarrhea in the previous week recorded. The nutrition program provided supplemental food to all children with WAZ < −2.5. Children over 12 mo received ≈100 kCal/day of Incaparina (a supplement made from maize and soy flours) and sugar, and children between 6–12 mo received half this amount and – for children with inadequate or no breastfeeding despite promoter efforts to support breastfeeding until at least 1 year of age – 2 400-gm cans of powdered infant formula (Nestle NAN Pro 1 or 2, made of whey and casein milk proteins) per mo. The Stanford Institutional Review Board, which includes members with medical and social science expertise, and the Review Committee of the Hospital Obras Sociales de Monsignor Gregorio Schaffer, San Lucas Toliman reviewed and approved this study prior to study initiation. Informed consent in Spanish and/or the local indigenous language was obtained from a parent/guardian for all children at enrollment. All patient information was anonymized and de-identified prior to analysis.

Generalized Additive Model

For boys and girls, we build separate GAMs for the evolution of \( \{W_t, H_t, D_t\} \) for age \( t \) varying from 1 to 60 mo, which are the WAZ, HAZ and the number of days with diarrhea in the previous week for a child of age \( t \) mo (see Table 1 for a description of most of the variables). Define the Intention To Treat (ITT) indicator variable \( I_t = 1 \) if \( W_t < -2.5 \) and \( I_t = 0 \) otherwise, and let \( M_t \) be the month of the year when a child is of age \( t \). Let \( t = (t - 31.76)/15.74 \) be the standardized age, where 31.76 mo and 15.74 mo are the mean and standard deviation of the ages among all entries in the data set.

An analysis of GAM results with varying amounts of history suggests that we need to maintain 4 mo of history; i.e., if we are trying to estimate a child’s \( (W_{t+2}, H_{t+2}, D_{t+2}) \) then we need to use \( (W_t, H_t, D_t) \) and \( (W_{t-2}, H_{t-2}, D_{t-2}) \). To ease the interpretation of our GAM coefficients, we let the response variables be

<p>| Table 1. The estimated parameters for boys from the GAM. |</p>
<table>
<thead>
<tr>
<th>Variable Description</th>
<th>Est.</th>
<th>S.E.</th>
<th>( \beta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.230</td>
<td>0.018</td>
<td>0.000</td>
</tr>
<tr>
<td>( W_t )</td>
<td>-0.095</td>
<td>0.007</td>
<td>0.000</td>
</tr>
<tr>
<td>( H_t )</td>
<td>-0.095</td>
<td>0.007</td>
<td>0.000</td>
</tr>
<tr>
<td>( W_{t-2} )</td>
<td>-0.236</td>
<td>0.011</td>
<td>0.000</td>
</tr>
<tr>
<td>( H_{t-2} )</td>
<td>-0.090</td>
<td>0.010</td>
<td>0.000</td>
</tr>
<tr>
<td>( W_t - W_{t-2} )</td>
<td>-0.084</td>
<td>0.006</td>
<td>0.000</td>
</tr>
<tr>
<td>( H_t - H_{t-2} )</td>
<td>-0.062</td>
<td>0.005</td>
<td>0.000</td>
</tr>
<tr>
<td>( W_t \times age )</td>
<td>0.007</td>
<td>0.008</td>
<td>0.000</td>
</tr>
<tr>
<td>( H_t \times age )</td>
<td>0.031</td>
<td>0.003</td>
<td>0.000</td>
</tr>
<tr>
<td>( W_t \times \text{inc. age} )</td>
<td>0.022</td>
<td>0.011</td>
<td>0.000</td>
</tr>
<tr>
<td>( H_t \times \text{inc. age} )</td>
<td>0.011</td>
<td>0.008</td>
<td>0.000</td>
</tr>
<tr>
<td>( W_t \times \text{inc. age} )</td>
<td>0.004</td>
<td>0.004</td>
<td>0.000</td>
</tr>
<tr>
<td>( H_t \times \text{inc. age} )</td>
<td>0.014</td>
<td>0.004</td>
<td>0.000</td>
</tr>
<tr>
<td>( W_t \times \text{inc. age} )</td>
<td>0.014</td>
<td>0.004</td>
<td>0.000</td>
</tr>
<tr>
<td>( H_t \times \text{inc. age} )</td>
<td>0.013</td>
<td>0.004</td>
<td>0.000</td>
</tr>
</tbody>
</table>

For each variable, we give the estimated coefficient (Est.), the standard error (S.E.) and the \( \beta \) value.
(W_{i+2} = W_i, H_{i+2} = H_i, D_{i+2}). We fit W_{i+2} - W_i and H_{i+2} - H_i by least squares and model D_i by a Poisson random variable. We let the link functions be the identity function for W_{i+2} - W_i and H_{i+2} - H_i and the natural logarithm for D_{i+2}. The predictor variables include the current values (W_i, H_i, D_i) and their age-interaction terms (W_i t, H_i t, D_i t), the most recent increments ((W_{i+2} - W_i)(H_{i+2} - H_i)), and their age interaction terms ((W_{i+2} - W_i), (H_{i+2} - H_i), t) as well as seasonal functions (f_0^t_W(M_i), f_0^t_H(M_i), f_0^t_D(M_i)) for the month of the year, age functions (g^t_W(t), g^t_H(t), g^t_D(t)) in the absence of supplementary food, and age functions (g^t_W(t) I_t, g^t_H(t) I_t, g^t_D(t) I_t) due to the ITT. The seasonal functions are modeled as periodic cubic splines and the age functions are modeled as cubic smoothing splines. We denote the sample covariance matrix calculated from the residuals of the GAM by $\Sigma$. The 3 GAM equations are stated with their resulting coefficients in §1 of File S1.

To make the most use of the available data without imputing any missing data, we include every set of 3 consecutive data set when they were 6 mo. At each 2-mo time step, the values any missing data, we include every set of 3 consecutive

**Proposed Allocation Policies**

Because of the difficulty of solving a dynamic program with a system state that is the joint population-wide probability density function (PDF) of (W_i, H_i, D_i), we derive 2 derived policies – referred to as the derived WAZ policy and the derived HAZ policy – that estimate every child’s (W_{i+2} H_{i+2}, D_{i+2}) given all the information known at age t, and minimizes the sum of $E[W_{i+2}]^2$ (or $E[H_{i+2}]^2$) over all children with $W_i < 0$ (or $H_i < 0$), subject to a constraint on the amount of supplementary food allocated. As in [12], this objective captures moderate and mild forms of undernutrition and also incorporates the fact that more severe undernutrition leads to more severe consequences; however, we use $E[W_{i+2}]^2$ rather than $E[W_i^2]$ so that the proposed policy is relatively simple.

Let $c_t = 0.5$ if $t \leq 6, 12$ and $c_t = 1$ if $t > 12$, which represents the normalized amount of supplementary food consumed by a child at age t. The allocation policy that solves this constrained optimization problem prioritizes all children with $W_i < 0$ by ($E[W_{i+2} I_t = 1]^2 - E[W_{i+2} I_t = 0]^2)/c_t$, (with smaller values getting higher priority), and allocates the food until it is gone ($\S2$ of File S1).

**Simulation Model**

The simulation model tracks 1 M children of each sex from age 6 to 60 mo. A child’s initial (W_i, H_i, D_i) at 6 mo is randomly sampled with replacement from all children of the same sex in the data set when they were 6 mo. At each 2-mo time step, the values of (W_{i+2}, H_{i+2}, D_{i+2}) are calculated using the GAM and then adding an additive trivariate noise term that is normally distributed with mean zero and covariance matrix $\Sigma$ (and rounding the value of $D_{i+2}$ to the nearest integer).

We consider the 5 policies described in Table 2. All policies except the no food policy allocate the same amount of food, which is dictated by the amount of food allocated under the current policy (i.e., $W_i < -2.5$). This food budget is 1.2 (0.7) units of food per boy (girl) over a child’s 27 feeding opportunities, which corresponds to 4.7% of children being fed at any given time. We do a 1-dimensional search for thresholds (for the quantities being prioritized in Table 2) for the simple WAZ policy and the 2 derived policies so that the amount of food allocated by these policies is equal to this food budget. The thresholds for the simple WAZ policy are $-1.8$ for boys (i.e., a boy under 2 yr receives food if $W_i < -1.8$) and $-1.7$ for girls, the thresholds for the derived HAZ policy are $-0.082$ for boys (i.e., a boy receives food if ($E[W_{i+2} I_t = 1]^2 - E[W_{i+2} I_t = 0]^2)/c_t < -0.082$) and $-0.138$ for girls, and the thresholds for the derived HAZ policy are $-0.44$ for boys and $-0.56$ for girls.

For each policy and various values of $\theta$, we report $P(H_i < \theta)$ and $P(W_i < \theta)$ for all t $\leq 6, 60$ mo. We also compute 2 variants of the severity index considered in [12], the average of $\sum_{t=6}^{60} (W_{i+2} I_{W_i < 0})^2$ over all children and the average of $H_{i+2} I_{H_i < 0}$, where $I_{[\_]}$ is the indicator function of the event $\_$. These quantities are the average squared shortfalls below the reference median (i.e., zero) for WAZ averaged over all ages and for HAZ at age 60 mo. We refer to these measures as the underweight severity and the stunting severity, respectively.

**Results**

**GAM Results**

The GAM parameters are very similar for boys (Table 1) and girls (Table 1 of File S1), and the statistically significant parameters have the following interpretations. WAZ exhibits global mean reversion (i.e., heavier children tend to lose weight and lighter children tend to gain weight), which is more pronounced in younger children, and local mean reversion (i.e., children who have recently lost weight tend to gain it back and vice versa), which is more pronounced in older children. Taller children tend to gain weight although children who recently grew taller are less likely to gain weight. Current diarrhea status has no impact on weight gain overall, although older boys are more likely to gain weight after a diarrhea episode while younger boys are more apt to lose weight (this effect is not observed in girls). HAZ also exhibits global mean reversion that is more prominent in younger children and local mean reversion that is more prominent in older children, and the magnitude of global mean reversion for HAZ is larger than it is for WAZ. Heavier children are more apt to grow taller, an effect that is more evident in younger children. Current diarrhea status and recent weight gain are not associated with height gain. Diarrhea is reinforcing; i.e., current diarrhea tends to lead to future diarrhea. Lighter children and children who recently gained weight both tend to get more diarrhea, and the former effect is more pronounced in older girls than younger girls.

Current height of boys has no overall impact on getting diarrhea, although taller girls are more likely to get diarrhea than shorter girls. Older tall children are more likely to get diarrhea than younger tall children.

Turning to the spline functions, which are also very similar for boys (Fig. 1) and girls (Fig. 1 of File S1), we see that low WAZ and high diarrhea prevalence are on very similar annual cycles (with November - April being the most severe months), while HAZ has a more rapid increase and lags slightly behind WAZ, peaking during August - November (Figs. 1a-c). In the absence of intervention, fatering in WAZ and HAZ occurs during the first 2 yr, at which point WAZ stabilizes and HAZ increases for boys (Figs. 1d-e); for girls, WAZ does not stabilize until $\approx 40$ mo and HAZ begins to increase at 30 mo. The prevalence of diarrhea decreases with age (Fig. 1f). The effect of ITT (i.e., supplementary food) has a similar pattern for all 3 variables: the impact is beneficial (i.e., higher WAZ and HAZ, lower diarrhea) and linearly decreasing between 6 and 20 mo, and thereafter has little or no effect (with the exception that ITT leads to an increase in diarrhea for older girls). ITT’s effect on HAZ is twice as large as its effect on WAZ.
Table 2. Food allocation policies.

<table>
<thead>
<tr>
<th>Name of Policy</th>
<th>Children Receiving Food</th>
<th>Severity Index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Underweight</td>
</tr>
<tr>
<td>No Food</td>
<td>None</td>
<td>1.87</td>
</tr>
<tr>
<td>Current</td>
<td>( W_{t+2} &lt; -2.5 )</td>
<td>1.84</td>
</tr>
<tr>
<td>Simple WAZ</td>
<td>Under 2 yr and prioritized by ( W_t )</td>
<td>1.77</td>
</tr>
<tr>
<td>Derived WAZ</td>
<td>Prioritized by ( (E[W_{t+2}]</td>
<td>I_t=1)^2 - E[W_{t+2}]</td>
</tr>
<tr>
<td>Derived HAZ</td>
<td>Prioritized by ( (E[H_{t+2}]</td>
<td>I_t=1)^2 - E[H_{t+2}]</td>
</tr>
</tbody>
</table>

A description of the food allocation policies and their severity indices for boys, which measure the average of squared shortfalls below the reference median (i.e., zero) for WAZ averaged over all 28 measured ages (underweight) and for HAZ at age 60 mo (stunting).

do:10.1371/journal.pone.0099632.t002

The covariance parameters are very similar for boys and girls (§1 of File S1), where weight gain and height gain are positively correlated, weight gain and diarrhea are negatively correlated, and height gain and diarrhea are the least correlated. The adjusted \( R^2 \) for \( \{W_{t+2} - W_{t}, H_{t+2} - H_{t}, D_{t}\} \) are 0.151, 0.259 and 0.041 for boys, and 0.226, 0.309 and 0.052 for girls.

Food Allocation Results

The results for boys (Figs. 2–5) and girls (Figs. 2–5 of File S1) are very similar, aside from the fact that girls are less underweight than boys, and hence have a smaller food budget. Given the strong influence of age on the effect of intervention (Figs. 1g–i), it is not surprising that the two derived policies allocate much of the food to children < 12 mo (Fig. 2). Consequently, the two derived policies reduce the far left tail of the HAZ and WAZ distributions between 6–12 mo (Figs. 3–4). Although some of this gain is dissipated through the next 4 years, these 2 policies still dominate the other policies throughout the first 5 years. The current policy incurs a significant increase in extreme underweight (WAZ < -3) during 12–30 mo because it fails to feed enough children < 12 mo who are in danger of faltering. A similar outcome – but at WAZ < -2.5 – is incurred by the simple WAZ policy for children during 20–24 mo.

The severity indices (Table 2 for boys, Table 2 of File S1 for girls) reveal that the derived HAZ policy achieves slightly smaller underweight and stunting severity indices than the derived WAZ policy, although both policies perform well on both measures relative to the other policies, and achieve a 13.6–14.1% reduction in WAZ severity and a 7.1–8.0% reduction in HAZ severity relative to the current policy. The performance gap between the 2 derived policies and the simple WAZ policy is 2–3 times larger than the gap between the simple WAZ policy and the no food policy. In our simulations, the current policy increases the HAZ severity index by 1.1% relative to the no food policy for boys, and increases it by 0.1% for girls. An examination of the stunting PDFs for the 5 policies (Fig. 5) reveals that the current policy reduces the number of children with \( H_{60}\in[-3.5,-1] \), but increases the number of children with \( H_{60}\leq -4 \). This increase in the extreme tail occurs because there is a small group of children with WAZ values that are consistently <-2.5 and with very low HAZ levels, and ITT further lowers their HAZ levels during months 18–36 and 46–60 (although even at the lowest points in Fig. 1 h of \( \sim 0.08 \) at 25 mo and \( \sim 0.1 \) at 60 mo, children continue to grow taller under ITT). In contrast, the 2 derived policies reduce the number of children with \( H_{60} < -2 \) and move them to the \( H_{60} > -2 \) region.

Sensitivity Analysis

We perform a sensitivity analysis for boys to assess the robustness of our main result that the 2 derived policies outperform the current policy. We perform the following procedure 10^3 times: take a random sample with replacement of 1047 boys (which is the original sample size in the data set) from our data set, re-estimate the statistical model (i.e., get new values for the parameters in Table 1 and the functions in Fig. 1), use the statistical model to simulate 10^3 boys under the current policy and the derived WAZ policy (for all runs, the derived WAZ policy is based on the results in Table 1 and Fig. 1, not on the re-estimated model), and compute the percentage reduction in the two severity indices that the derived WAZ policy achieves relative to the current policy. The PDFs of the 10^3 percentage reductions (Figs. 6a–b) show that the derived WAZ policy achieves a mean percentage reduction of 14.7% for underweight severity and 7.6% for stunting severity relative to the current policy, and the percentage of times that the current policy performs better than the derived WAZ policy is 4.7% for underweight severity and 0.2% for stunting severity.

We perform an identical sensitivity analysis to assess whether the current policy reduces the underweight and stunting severities relative to the no food policy. We find that the current policy reduces underweight severity by 1.4% on average and reduces it 79.7% of the time, whereas the current policy increases stunting severity by 1.2% on average and increases it 90.6% of the time (Figs. 6c–d).

Discussion

Statistical Results

The similarity between some of our statistical results and the existing literature is reassuring, and provides some confidence in the remaining statistical results that have no counterpart in the literature. The patterns of growth faltering in Figs. 1d-e are similar to the patterns seen elsewhere [13], as are the decrease in diarrhea prevalence with age (Fig. 1f and [14]). Similarly, the seasonal patterns in Figs. 1a–c are not unlike those found in Malawi [15], and – as is typical in Latin America [16] – stunting is much more prevalent than underweight. Rearranging equations (13)-(16) in the Supporting Material of [4] shows that the model there, which is based on data from Bwamanda in the Democratic Republic of Congo [17], also exhibits global and local mean reversion; global mean reversion in HAZ data is also observed in [15].

Turning to the cross terms, we find that high weight leads to a future increase in HAZ, but that recent weight gain does not; the latter finding agrees with correlations derived in [15]. An analysis

PLOS ONE | www.plosone.org
of WHZ (rather than WAZ) and HAZ find that both low WHZ and a recent drop in WHZ lead to future reductions in HAZ [18].

Our statistical results provide limited evidence to support the fact that diarrhea can reduce nutritional status. The lack of significance of the $D_t$ coefficient in predicting future WAZ or HAZ gain may be because most of the catch-up growth occurs before the next visit. In addition, the coefficient of $D_{tt}$ in the WAZ equation for boys is consistent with the observation that infants take longer to catch up in weight than older children [19]. Although our model predicts that underweight children are more apt to get diarrhea, which is consistent with earlier results [20,21], the same is not true of stunted children (in fact, taller girls are more likely to get diarrhea than shorter girls in our model); some of the anticipated effect (i.e., stunted children are more likely to get diarrhea) may be partially masked by the age dependence (i.e., the value of the $H_{tt}$ coefficient) and the 2-mo follow-up period. In any case, diarrhea is very difficult to predict and the only factors that play a strong role are seasonality, being underweight, and currently having diarrhea, which collectively suggest environmental and microbiome causes.

Our most intriguing statistical result is the strong dependence of age on the impact of intervention (Figs. 1g–i). Considerable
evidence has been amassed about the importance of nutrition in the first 1000 days (from conception to 24 mo) [1,22], and Figs. 1g-i can be viewed as a refinement of this principle. Our age-dependent results are similar to those in an earlier study in Guatemala [23], which found that – after controlling for diarrhea, previous height and weight, sex, home diet and socioeconomic status – height gain and weight gain due to supplemental food decreased throughout the first 3 years of life and had little effect after age 3. A meta-analysis in Web Appendix 4 of [24] suggests that linear growth during 6–12 mo is larger than during 12–18 mo in populations with average per capita income >$1/day, but is smaller for incomes <$1/day. Fig. 1 h is not inconsistent with this observation because the income in Guatemala is >$1/day.

The magnitude of the seasonality in Figs. 1a-b and of the global and local mean reversion (Table 1) implies that approaches that ignore these factors can lead to misleading estimates – and systematic overestimates in the case of global mean reversion – of the increase in weight and height that is due to supplementary feeding.

Food Allocation Results

The 2 derived policies anticipate growth faltering and exploit the age dependence in the impact of intervention by front-loading much of the intervention on infants (Fig. 2), who gain a substantial head start relative to the current policy (Figs. 3–4). Although some of the gap disappears by the time children reach age 5 years (perhaps partly due to global mean reversion), much of the improvement remains intact at age 5. The derived policies are fairly simple to implement, although they do require knowledge of a child’s height and weight from his previous visit.

Due to the very limited amount of food distributed by the current policy (≈4% of children receive food at any point in time) and the fact that many more children are stunted than underweight, the improvements achieved by the 2 derived policies are modest when measured by stunting (HAZ < −2), but are more pronounced for severe stunting (HAZ < −3) and in the extreme left tail of the HAZ distribution (Fig. 4).

Because of the reinforcing cross terms – i.e., large HAZ leads to future increases in WAZ and large WAZ leads to future increases in HAZ – and the similar age dependence in how intervention increases WAZ and HAZ, the 2 derived policies have very similar performance with respect to reducing the severity of underweight and stunting.

There has been increasing focus on the lifetime effects of stunting [25], and in Latin America, where severe wasting is rare and stunting is prevalent, reducing stunting is one of the primary goals of supplementary feeding programs. To this end, we note that our model predicts that the current policy – because it allows some underweight, stunted children to become very severely...
stunted (Fig. 5b) – has a slightly higher severity of stunting than the no food policy (Fig. 6d). This result is driven by the fact that ITT leads to a slight decrease in HAZ in our model during months 18–36 and 46–60 (Fig. 1h); it is not clear whether this effect is real (albeit small), or is caused by overshooting of the cubic spline in Fig. 1h. In contrast, despite the wide confidence intervals in Figs. 1g–i, our sensitivity analysis (Figs. 6a–b) suggests that our main result – i.e., that the 2 derived policies outperform the current policy – is somewhat robust.

There has also been concern about the long-term cardiovascular risks associated with catch-up growth in early childhood [26]. While this remains an important issue, particularly for rapid weight gain among infants born at low birth weight, the immediate mortality and developmental risks associated with serious childhood malnutrition would seem to warrant monitored nutritional supplementation.

Limitations of Analysis

Aside from the obvious shortcoming that our analysis is based on data from a non-randomized study, there are 2 major limitations in our data set. The first is that health promoters in the Guatemala nutrition program had discretion to feed sick children, and hence ITT did not exactly coincide with the actual receipt of food. We use ITT so as not to be biased by the behavior

---

**Figure 4. For boys, the left tails, P[HAZ < \( \hat{A} \)] for \( \theta \) equals (a) –2, (b) –2.5, (c) –3, (d) –3.5, (e) –4, (f) –4.5 vs. age under the various policies.**

doi:10.1371/journal.pone.0099632.g004
of health promoters. Note that if the health promoters were not given discretion to feed sick children, there may have been an incentive to alter WAZ data (e.g., change $-2.3$ to $-2.6$ so as to feed a sick child). The second limitation in our data set is the amount of missing data. While children were supposed to see a health promoter every 2 mo, the actual time between consecutive visits in the data set was 1 mo, 2 mo, 3 mo, 4 mo and $>4$mo with probabilities 0.694, 0.638, 0.122, 0.121 and 0.024, respectively.

Figure 5. For boys, (a) the PDF for HAZ at age 60 mo, and (b)-(e) the difference in PDFs for HAZ at age 60 mo between 2 policies. doi:10.1371/journal.pone.0099632.g005
With respect to estimating the impact of ITT and its age interaction, the primary concern with our assumption that consecutive visits are exactly 2 mo is if ITT causes children to visit more or less frequently relative to no ITT. To assess this issue, we re-ran the GAM with the same predictor variables as before but now with the time until the next visit as the response variable. The overall mean effect of ITT was very close to zero (see Fig. 6 of File S1, which is the analogue to Figs. 1g-i in the main text) and the 3 p-values for the cubic spline were 0.91, 0.88 and 0.96.

There are other aspects that could be investigated: estimating the variance of ITT, the interaction of diarrhea and z scores (e.g., a $D_{t}W_{z}$ interaction term could address whether diarrhea lowers future weight gain only when current WAZ is low [19]), and whether ITT is a function of $D_{t}$ [19]. However, we note that our qualitative conclusions would not change even if diarrhea was omitted from the model. Also, the data used here are not well suited to estimate the influence that pre-intervention (WAZ, HAZ) levels have on the impact of intervention because of the lack of intervention-free children with low WAZ and because of the presence of global mean reversion, which may have a similar effect (Fig. 2 in [27]).

The age dependence in the impact of treatment should not be extrapolated to other settings, other doses or other undernutrition thresholds without performing similar statistical analyses of other longitudinal data sets. It should be noted that the community health system within which the nutritional surveillance and supplementation program is embedded also provides basic preventive and therapeutic health services, including routine antihelminthic medication for children greater than 2 years of age and when needed, referral to local and regional, clinical facilities. In addition, identification of malnutrition and the provision of supplementation may elevate the nutritional needs of the child in question and result indirectly in a relative reallocation of some family resources to the child. This in turn could modify the impact of supplementation on growth in different cultural or resource settings. It may be the case that 100 KCal/day is an insufficient dose – particularly after taking into account the possibility of food sharing within families – to improve the nutritional status of children over 2 years old, but that larger doses are sufficient. Similarly, even at the current dose, our findings should not be extrapolated to thresholds other than WAZ = −2.5. It is possible that age-specific treatment effects differ by level of undernutrition. In particular, children >20 mo may respond more to supplementary feeding when their initial undernutrition levels are more severe, because of the developmental delays associated with undernutrition. If so, regression discontinuity around a more
extreme threshold might have yielded flatter estimated age functions in Figs. 1g–i.

Conclusion

Fitting a trivariate statistical model of WAZ, HAZ and diarrhea to a longitudinal data set from a nutritional program in Guatemala generates results that reinforce and quantify relationships that are already known, and identify new relationships, including a strong age dependence in the impact of supplementary food on WAZ, HAZ and diarrhea, with the benefits decreasing from 6 mo to 20 mo old and staying at a small level between 20 and 60 mo old. Our findings highlight the dangers of ignoring seasonality and global and local mean reversion when estimating the impact of supplementary food. The food allocation policies derived here perform better than the current policy: by allocating much of the food to infants, the derived policies reduce the underweight severity by 13.6–14.1% and the stunting severity by 7.1–8.0% relative to the current policy, despite using a budget that allocates food to only 4.7% of the children at any time under the current policy. This result should not be extrapolated to other geographical regions, to doses other than 100 Kcal/day, or to undernutrition thresholds other than WAZ = −2.5. To confirm this result, the derived policies should be compared to the current policy or the simple WAZ policy in a randomized study. More generally, given the paucity of randomized controlled trials with intervention-free control groups, our statistical approach has the potential to uncover new knowledge from observational studies pertaining to therapeutic and supplementary food for children, and our optimization approach allows this knowledge to be leveraged to improve nutritional status in a budget-constrained setting.

Supporting Information

File S1 Supporting Material. Gives the regression model equations, the derivation of the proposed food allocation policies, and the results for girls. (PDF)

Acknowledgments

We thank Mike Baicchi and Jeremy Goldhaber-Fiebert for helpful conversations. We also wish to thank Dr. Rafael Tun, Jos Vicente Macario Cosigu, and Dominga Pic Salazar for coordination of the community health worker and nutrition program.

Author Contributions

Conceived and designed the experiments: SW LMW. Performed the experiments: ML SW. Analyzed the data: ML SW LMW. Contributed reagents/materials/analysis tools: PHW. Wrote the paper: LMW.

References

The GAM is given in §1 and the derived allocation policies are constructed in §2. Tables 1-2 and Figs. 1-5, which are the results for girls, and Fig. 6, which contains the results of a sensitivity analysis, are referred to in the main text.

1 Regression Model

With the parameter estimates included, the GAM equations for boys are given by

\[
E[W_{t+2} - W_t] = -0.230 - 0.095W_t + 0.014W_t \hat{t} - 0.256(W_t - W_{t-2}) - 0.090(W_t - W_{t-2})\hat{t} \\
+ 0.038H_t + 0.006H_t \hat{t} - 0.020(H_t - H_{t-2}) - 0.007(H_t - H_{t-2})\hat{t} \\
+ 0.003D_t + 0.014D_t \hat{t} + f^W_M(M_t) + g^W(t) + g^W_I(t)I_t, \tag{1}
\]

\[
E[H_{t+2} - H_t] = -0.907 + 0.124W_t - 0.068W_t \hat{t} - 0.000(W_t - W_{t-2}) + 0.021(W_t - W_{t-2})\hat{t} \\
- 0.177H_t + 0.092H_t \hat{t} - 0.205(H_t - H_{t-2}) - 0.022(H_t - H_{t-2})\hat{t} \\
- 0.003D_t + 0.007D_t \hat{t} + f^H_M(M_t) + g^H(t) + g^H_I(t)I_t, \tag{2}
\]

\[
\ln E[D_{t+2}] = -0.751 - 0.184W_t - 0.056W_t \hat{t} + 0.183(W_t - W_{t-2}) + 0.084(W_t - W_{t-2})\hat{t} \\
- 0.007H_t + 0.068H_t \hat{t} + 0.063(H_t - H_{t-2}) - 0.041(H_t - H_{t-2})\hat{t} \\
+ 0.102D_t - 0.021D_t \hat{t} + f^D_M(M_t) + g^D(t) + g^D_I(t)I_t. \tag{3}
\]

The sample covariance matrix calculated from the residuals of equations (1)-(3) is

\[
\Sigma = \begin{pmatrix}
0.10 & 0.03 & -0.04 \\
0.03 & 0.19 & 0.01 \\
-0.04 & 0.01 & 1.54
\end{pmatrix}. \tag{4}
\]
The GAM equations for girls are

\[
E[W_{t+2} - W_t] = -0.178 - 0.101W_t + 0.029W_t\hat{t} - 0.249(W_t - W_{t-2}) - 0.071(W_t - W_{t-2})\hat{t} \\
+ 0.033H_t + 0.004H_t\hat{t} - 0.006(H_t - H_{t-2}) + 0.019(H_t - H_{t-2})\hat{t} \\
+ 0.003D_t + 0.002D_t\hat{t} + f_W^W(M_t) + g^W(t) + g_I^W(t)I_t ,
\]

\[(5)\]

\[
E[H_{t+2} - H_t] = -0.765 + 0.118W_t - 0.070W_t\hat{t} + 0.003(W_t - W_{t-2}) - 0.009(W_t - W_{t-2})\hat{t} \\
- 0.188H_t + 0.106H_t\hat{t} - 0.193(H_t - H_{t-2}) - 0.029(H_t - H_{t-2})\hat{t} \\
- 0.003D_t + 0.006D_t\hat{t} + f_H^H(M_t) + g^H(t) + g_I^H(t)I_t ,
\]

\[(6)\]

\[
\ln E[D_{t+2}] = -0.660 - 0.232W_t - 0.092W_t\hat{t} + 0.148(W_t - W_{t-2}) + 0.073(W_t - W_{t-2})\hat{t} \\
+ 0.115H_t + 0.068H_t\hat{t} + 0.032(H_t - H_{t-2}) + 0.060(H_t - H_{t-2})\hat{t} \\
+ 0.132D_t + 0.017D_t\hat{t} + f_D^D(M_t) + g^D(t) + g_I^D(t)I_t .
\]

\[(7)\]

The sample covariance matrix calculated from the residuals of equations (5)-(7) is

\[
\Sigma = \begin{pmatrix}
0.10 & 0.03 & -0.05 \\
0.03 & 0.19 & 0.00 \\
-0.05 & 0.00 & 1.48
\end{pmatrix}.
\]

\[(8)\]

2 The Derived Allocation Policies

To construct a policy that is easy to use and understand, we make 4 simplifying assumptions. First, we restrict ourselves to myopic policies, i.e., policies that consider only WAZ severity 2 mo in advance (all our arguments in this section use WAZ although they also hold if we replace WAZ by HAZ). Suppose we are making food allocation decisions at time \(\tau\). We have \(N\) children indexed by \(i = 1, \ldots, N\), and let \(W_{i,\tau+2}\) be child \(i\)'s WAZ score at time \(\tau + 2\) and
Let the decision variables be $I_{i\tau} = 1$ if child $i$ receives food at time $\tau$ and $I_{i\tau} = 0$ otherwise. At time $\tau$ and conditional on our decision $I_{i\tau}$, this child’s expected contribution to the WAZ severity at time $\tau + 2$ is $E[W_{i,\tau+2}^2 | I_{i\tau}]I_{\{W_{i,\tau+2} < 0\}}$. We make 2 more assumptions to simplify this expression: we replace $E[W_{i,\tau+2}^2 | I_{i\tau}]$ by $E[W_{i,\tau+2} | I_{i\tau}]^2$ (or, equivalently, we assume that there is no randomness in predicting $W_{i,\tau+2} | I_{i\tau}$ at time $\tau$) and we assume that children’s WAZ scores do not change sign between times $\tau$ and $\tau + 2$. While this latter assumption may be violated, it only impacts our policy if the food budget $B$ is sufficiently large that children with WAZ near 0 might receive food (which is not the case in our setting or in most practical settings). These 2 assumptions allow us to approximate $E[W_{i,\tau+2}^2 | I_{i\tau}]I_{\{W_{i,\tau+2} < 0\}}$ by $E[W_{i,\tau+2} | I_{i\tau}]^2 I_{\{W_{i\tau} < 0\}}$. Our fourth assumption will be introduced later.

Let $c_{i\tau} = 0.5$ if child $i$’s age is $\in [6, 12)$ mo at time $\tau$ and $c_{i\tau} = 1$ if child $i$’s age is $\in [12, 60]$ mo at time $\tau$. Then our optimization problem at time $\tau$ is

$$\min_{I_{i\tau} \in \{0, 1\}} \sum_{i=1}^{N} E[W_{i,\tau+2} | I_{i\tau}]^2 I_{\{W_{i,\tau+2} < 0\}}$$

subject to $\sum_{i=1}^{N} c_{i\tau} I_{i\tau} \leq B$. \hspace{1cm} (9)

Define $\delta_{i\tau} = (E[W_{i,\tau+2} | I_{i\tau} = 1]^2 - E[W_{i,\tau+2} | I_{i\tau} = 0]^2)I_{\{W_{i\tau} < 0\}}$, which is the change in child $i$’s contribution to WAZ severity due to food, where the children with more negative values generate a greater reduction in severity. If $c_{i\tau}$ equaled 1 for all $i$ and $\tau$, then the optimal solution to (9)-(10) would rank children by $\delta_{i\tau}$ and give food to the $B$ children with the lowest $\delta_{i\tau}$ values; this can be shown with a simple interchange argument (i.e., a proof by contradiction, where one assumes that a different child with a higher value of $\delta_{i\tau}$ receives some food instead of a child with a lower value of $\delta_{i\tau}$). However, because $c_{i\tau} = 0.5$ for children under 12 mo, it is optimal to make 2 lists of children ranked by $\delta_{i\tau}$, where list A has children with $c_{i\tau} = 0.5$ and list B has children with $c_{i\tau} = 1$, and then allocate the food by comparing the sum of the $\delta_{i\tau}$ values of the next 2 children in list A to the $\delta_{i\tau}$ value of the next child in list B; see Algorithm 1 for details.
Algorithm 1 Optimal policy

Set $I_{i\tau} = 0 \forall i$ and total amount of food given = 0
Create two lists of children, $A = \{i : c_{i\tau} = 0.5\}$ and $B = \{i : c_{i\tau} = 1\}$
Sort A and B in the increasing order of $\delta_{i\tau}$, and let $A(1)$, $A(2)$ and $B(1)$ denote the first two children in list A and the first child in list B

while total amount of food allocated is less than $B$ do
    if $\delta_{A(1)\tau} + \delta_{A(2)\tau} > \delta_{B(1)\tau}$ then
        set $I_{B(1)\tau} = 1$; i.e., give food to the first child in list B
        increase the total amount of food allocated by 1
    else
        set $I_{A(1)\tau} = 1$; i.e., give food to the first child in list A
        increase the total amount of food allocated by 0.5
    end if
    Delete the child who received food from the list
end while

if total amount of food allocated is greater than $B$ then
    Instead of the last child who got food (who is from list B), give food to the first child in list A
end if

Algorithm 1 is awkward to implement. Hence, we introduce our fourth assumption, which is that $N$ is large enough so that the WAZ scores are dense in the real line and its distribution is independent of age. It follows that the heuristic policy (see Algorithm 2), which simply ranks all children by $\delta_{i\tau}/c_{i\tau}$ and allocates the food to children with smaller (i.e., more negative) values, achieves nearly the same performance as the optimal solution.

Algorithm 2 Heuristic policy

Set $I_{i\tau} = 0$ for all $i$, total amount of food given = 0
Sort all the children in the increasing order of $\delta_{i\tau}/c_{i\tau}$
Let $n = \max\{j | \sum_{i=1}^{j} c_{i\tau} \leq B\}$
Set $I_{i\tau} = 1 \ \forall i \leq n$
Tables

Table 1: The estimated parameters for girls from the GAM

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Estimating $W_{t+2} - W_t$</th>
<th>Estimating $H_{t+2} - H_t$</th>
<th>Estimating $D_{t+2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Est.</td>
<td>S.E.</td>
<td>p</td>
</tr>
<tr>
<td>Intercept</td>
<td></td>
<td>-0.178</td>
<td>0.018</td>
<td>0.000</td>
</tr>
<tr>
<td>$W_t$</td>
<td>WAZ</td>
<td>-0.101</td>
<td>0.006</td>
<td>0.000</td>
</tr>
<tr>
<td>$W_t \times age$</td>
<td>WAZ × age</td>
<td>0.029</td>
<td>0.006</td>
<td>0.000</td>
</tr>
<tr>
<td>$W_t - W_{t-2}$</td>
<td>WAZ increment</td>
<td>-0.249</td>
<td>0.011</td>
<td>0.000</td>
</tr>
<tr>
<td>$(W_t - W_{t-2})t$</td>
<td>WAZ incr. × age</td>
<td>-0.071</td>
<td>0.010</td>
<td>0.000</td>
</tr>
<tr>
<td>$H_t$</td>
<td>HAZ</td>
<td>0.033</td>
<td>0.005</td>
<td>0.000</td>
</tr>
<tr>
<td>$H_t \times age$</td>
<td>HAZ × age</td>
<td>0.004</td>
<td>0.005</td>
<td>0.376</td>
</tr>
<tr>
<td>$H_t - H_{t-2}$</td>
<td>HAZ increment</td>
<td>-0.006</td>
<td>0.009</td>
<td>0.514</td>
</tr>
<tr>
<td>$(H_t - H_{t-2})t$</td>
<td>HAZ incr. × age</td>
<td>0.019</td>
<td>0.007</td>
<td>0.008</td>
</tr>
<tr>
<td>$D_t$</td>
<td>Diarrhea</td>
<td>0.003</td>
<td>0.003</td>
<td>0.335</td>
</tr>
<tr>
<td>$D_t \times age$</td>
<td>Diarrhea × age</td>
<td>0.002</td>
<td>0.003</td>
<td>0.549</td>
</tr>
</tbody>
</table>

For each variable, we give the estimated coefficient (Est.), the standard error (S.E.) and the p-value.

Table 2: Food allocation policies

<table>
<thead>
<tr>
<th>Name of Policy</th>
<th>Children Receiving Food</th>
<th>Severity Index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Underweight</td>
</tr>
<tr>
<td>No Food</td>
<td>None</td>
<td>1.62</td>
</tr>
<tr>
<td>Current</td>
<td>$W_t &lt; -2.5$</td>
<td>1.60</td>
</tr>
<tr>
<td>Simple WAZ</td>
<td>Under 2 yr and prioritized by $W_t$</td>
<td>1.55</td>
</tr>
<tr>
<td>Derived WAZ</td>
<td>Prioritized by $(E[W_{t+2}</td>
<td>I_t = 1])^2 - E[W_{t+2}</td>
</tr>
<tr>
<td>Derived HAZ</td>
<td>Prioritized by $(E[H_{t+2}</td>
<td>I_t = 1])^2 - E[H_{t+2}</td>
</tr>
</tbody>
</table>

A description of the food allocation policies and their severity indices for girls, which measure the average of squared shortfalls below the reference median (i.e., zero) for WAZ averaged over all 28 measured ages (underweight) and for HAZ at age 60 mo (stunting).
Figure 1: The estimated spline functions (and 95% confidence intervals) for girls from the GAM. The seasonal functions (a) $f^W_{{t}}(M_t)$, (b) $f^H_{{t}}(M_t)$ and (c) $f^D_{{t}}(M_t)$ for WAZ and HAZ increments and diarrhea. The age functions in the absence of supplementary food (d) $g^W(t)$, (e) $g^H(t)$ and (f) $g^D(t)$ for WAZ and HAZ increments and diarrhea. The age functions due to ITT (g) $g^W_{t}(t)$, (h) $g^H_{t}(t)$ and (i) $g^D_{it}(t)$ for WAZ and HAZ increments and diarrhea.
**Figure 2:** The proportion of girls by age who receive food under the various allocation policies.
Figure 3: For girls, the left tails, \( P(WAZ < \theta) \) for \( \theta \) equals (a) -2, (b) -2.5, (c) -3 vs. age under the various policies.
Figure 4: For girls, the left tails, $P(\text{HAZ} < \theta)$ for $\theta$ equals (a) -2, (b) -2.5, (c) -3, (d) -3.5, (e) -4, (f) -4.5 vs. age under the various policies.
Figure 5: For girls, (a) the PDF for HAZ at age 60 mo, and (b)-(e) the difference in PDFs for HAZ at age 60 mo between 2 policies.
Figure 6: For boys, the estimated spline function (and 95% confidence interval) for the age-dependent impact of ITT in the GAM, where the response variable is the time until the next visit for a child of age $t$ mo ($T_i$).