

Nutritional Impact of Supplementation in the INCAP Longitudinal Study: Analytic Strategies and Inferences^{1,2}

JEAN-PIERRE HABICHT,*³ REYNALDO MARTORELL† AND JUAN A. RIVERA‡

*Division of Nutritional Sciences, Cornell University, Ithaca, NY 14853-6301, †Department of International Health, The Rollins School of Public Health of Emory University, Atlanta, GA 30322, and ‡Centro de Investigaciones en Salud Pública, Instituto Nacional de Salud Pública, 62508 Cuernavaca, Morelos, México

ABSTRACT From 1969 to 1977 a supplementation trial was conducted in Guatemala to ascertain the effects on physical and behavioral outcomes of improved nutrition in pregnant women and in preschool children. This paper reviews different strategies to analyze the effect of the intervention on physical growth. One strategy compares outcomes in two villages that were randomly allocated to receive Atole, a supplement containing high amounts of protein and energy, with values in two other villages that received Fresco, a beverage containing no protein and little energy. Both supplements contained micronutrients. This comparison of village means gives a probability significance statement ($P < 0.005$) that the difference in growth was because of the supplement intervention, although it does not specify the aspect of the intervention that caused the effect. Complementary strategies increase the credibility that the effect of the supplement was nutritional. Thus, analysis of the dose response with increasing supplement intake within the villages excludes the possibility that the above findings were the result of knowing which villages received which supplement (i.e., measuring biases). A greater effect in those most likely to respond nutritionally also increases the credibility that the mechanism was nutritional. In studying other behavioral and biomedical impacts of this supplementation intervention, analyses for credibility should always be included. *J. Nutr.* 125: 1042S-1050S, 1995.

INDEXING KEY WORDS:

- analytic methods • pregnancy
- supplementation • growth • Guatemala

The study population, the experimental design and the methods used in the INCAP longitudinal study (1969-77) are described elsewhere (Martorell et al. 1995). In summary, two kinds of supplements were distributed in a central refectory in four villages, in

midmorning and midafternoon, to any villager who attended. Supplement consumption was recorded for infants, children <7 y of age and pregnant and lactating women. From 1969-1977, two of the villages received a high-protein, high-energy supplement called "Atole" and the other two were given a no-protein, low-energy supplement called "Fresco." The villages were paired by size (i.e., large and small) and allocation within pairs to the supplements was random. From 1971 to 1977, the supplements had the same concentrations of specified micronutrients.

Growth and behavior, the key outcomes in children, were assessed periodically during the preschool period, as were measurements of potential modifiers or confounders such as morbidity, home diet and socioeconomic status. Maternal nutrition and health information also was collected periodically during pregnancy and lactation.

This paper presents different approaches to the analysis of the impact of supplementation on physical growth. These approaches may be divided into those that took advantage of the randomized design and

¹ Presented in the symposium on Nutrition in Early Childhood and its Long-Term Functional Significance, FASEB, April 6, 1992, Anaheim, CA. Published as a supplement to *The Journal of Nutrition*. Guest editors for this supplemental publication were Reynaldo Martorell, The Rollins School of Public Health of Emory University, Atlanta, GA, and Nevin Scrimshaw, The United Nations University, Boston, MA.

² The INCAP longitudinal study was supported by contract No. HD-5-0640 from the National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland. Additional sources of support were from the Agency for International Development, Washington, DC (AID-TA-C/1224) and a grant from Rockefeller Foundation (73030-E7352). The follow-up study was supported by RO1 grant HD-22440.

³ To whom correspondence should be addressed: Division of Nutritional Sciences, 210 Savage Hall, Cornell University, Ithaca, NY 14853-6301.

those that sought other ways to control for confounding. The analyses selected as examples are tests of the effect of the supplement type on growth in length by three years of age, the effect of different amounts of maternal supplementation during pregnancy on birth weight and the effect of supplementation on improving weight in wasted children.

Except for the effect on growth presented in Table 1, all the findings discussed have been published elsewhere and are appropriately cited.

Using the randomized design

The analysis. To take advantage of the randomized design, one must use village as the unit of analysis because villages and not individuals were randomized. Most publications describing differences in outcomes between Atole and Fresco villages use the child or mother as the unit of analysis. Thus, the error term used to test these differences has many-fold more degrees of freedom and therefore will result in greater statistical significance than in analyses using village as the unit of analysis. These analyses do not give the statistical significance relating the treatment to the difference and instead only provide the statistical significance that the difference is not due to chance. The difference might be due to intrinsic village differences and not due to the treatments themselves. For example, the children of Espíritu Santo, the small Fresco village, had smaller head circumferences than the children in the other villages. When all the children from the Fresco villages are compared with all the children in the Atole villages, the systematically lower values of Espíritu Santo exaggerate the statistical significance of the difference between groups compared with using village as the units of comparison.

Martorell et al. (1982) developed an approach using the consistency of the response to supplementation across the two village sizes and two genders, (i.e., four sex-size groups for each treatment). This paper showed that the lengths of 3-y olds who had lived in Fresco villages their entire lives after the supplementation program began did not differ significantly ($P \geq 0.05$) in any of the four sex-village size groups compared with 3-y-old children measured before the study in the same groups. The range of change was -0.7 cm to 1.1 cm with a mean of 0.45 cm. In contrast, the change in the Atole villages relative to baseline values was statistically significant for all four sex-size groups ($P < 0.05$). The range of this secular change was 2.5 – 3.6 cm with a mean of 2.90 cm. Of course, the statistical probability of the change within each of the Atole sex-village size groups is not the probability that this was caused by Atole. However, the consistency of the changes across the Atole groups compared with the negligible changes observed in the Fresco groups over time makes the inference that the Atole improved growth credible.

TABLE 1

Length¹ of 3-y-old children before and after supplementation by village size and type of supplement

	Large villages		Small villages	
	Atole	Fresco	Atole	Fresco
After ²	86.70	84.00	85.95	84.35
Before ³	83.45	83.30	83.40	84.15
Change	3.25	0.70	2.55	0.20
Difference in change	2.55		2.35	

Overall difference in change: mean = 2.45 ± 0.10 , t -test = 24.50, $P < 0.005$ (Two-tailed probability, $df = 2$).

¹ Means of sex-specific data calculated from Table 3 in Martorell et al. (1982).

² Born between 1969 and 1973.

³ Measured in 1968

A more rigorous statistical test can be made of the above-mentioned changes in Atole (A) and Fresco (F) villages by using village as the unit of analysis (Table 1). This analysis is true to the randomized design and deals with potential intrinsic differences between villages within each pair of similar sized villages by incorporating them into the statistical probability statement.

According to this analysis, the difference in net change (Atole minus Fresco) in the large villages was 2.55 cm and in the small villages it was 2.35 cm (Table 1). The mean of these differences is 2.45 ± 0.10 cm (mean \pm SD). Even though the standard deviation only has 2 deg of freedom, the t -test is 24.5 with a two-tailed probability of $P < 0.005$. It is well known that the probability statement, $P < 0.005$, means that there is only one chance in a thousand that this difference could be due to chance. What is less well understood is that such a probability statement, except in a randomized design, does not exclude the likelihood, often strong, that the difference is due to something other than the intervention. Only a randomized design incorporates the potential effects of confounding factors into the probability statement. Thus, one can infer, with very little chance of being wrong ($P < 0.005$), that the difference in growth between Atole and Fresco villages was due to difference in the interventions and not to chance or to confounding.

The probability of the t -test shown above is for a two-tailed test. However, there is such a clear expectation that the effect of Atole will be beneficial compared with Fresco that it may be more appropriate to use a one-tailed test. In this case, P would be < 0.0025 .

Potential biases. It is generally well understood that the statistical significance of the above impact cannot be due to initial village differences because these are included in the error term of the test statistic. Similarly, differential changes that occurred among the

villages during the period of supplementation also are included in the error term so long as these are not associated with the supplementation.

Also, the effect of the intervention on growth cannot be explained by self-selection to ingest the supplement. A repeated criticism of the study is that children who came for supplementation may have had parents who were more concerned about child health and nutrition and thus, would have grown and performed better anyway. However, this self-selection hypothesis also would predict that the village mean growth would remain unchanged. This, as seen above, was not the case. Therefore, these and similar factors associated with ingestion of supplementation within a village could not affect the comparison across Atole and Fresco groups as presented above. Even differential self-selection where, for instance, the better off children in the Atole villages and the worst off children in the Fresco villages ingested the supplement, would not bias the results in Table 1. Thus, self-selection for ingestion of the supplement within the villages cannot introduce, by itself, biases into the analyses performed appropriately for the randomized design.

The causal statistical significance for an effect of the intervention is impressive, both in its statistical significance and in its exclusion of other factors related to the villages and to those who ingested the supplement. It is important, however, to remember that it does not specify what aspect of the intervention is responsible for the effect. Anything done in the villages that was associated with the supplement could have caused the effect seen. This is why care was taken to spread the INCAP presence equally across the villages through designing and implementing all interventions similarly in all villages, and through rotation of all personnel (Martorell et al. 1995).

One effect associated with supplementation across villages that could not be excluded is the effect of knowing the kind of supplementation a village received. The villagers were, for all practical purposes, "blinded" to this fact because of the distances and the lack of communication among the villages. However, the measurers could not be "blinded". All field workers knew that both supplements were good for mothers and children, so one might expect them not to have been biased. Nevertheless, this possibility must be excluded as described below when discussing the dose response to supplementation.

Another measurement effect that could be associated with supplementation across villages is differential participation in the measurement of outcomes. This could happen, for instance, if better off and worse off participants to Atole and Fresco, respectively, came to be measured. This has been investigated and no evidence of this kind of bias has been found, but this must be kept in mind and verified in each analysis.

Another way that other interventions could have been associated with ingestion of the supplements is

if attendance rates were different between Atole and Fresco villages. This is, indeed, the case. Attendance rates were much higher from birth to 3 y in Atole than in Fresco villages (Schroeder et al. 1992). As noted above, this presents no problem if this was because of self-selection within a village. However, differential attendance can result in differential exposure to programmatic influences other than the supplements. For instance, it could have been that those who came to the feeding centers also received better medical care because the clinic and the feeding centers were in the same building. Or, maybe, the socialization experienced in the feeding centers fostered better scores in the behavioral tests. Fortunately, these influences due to differential attendance rates can be taken into account because, in this data set, it is possible to differentiate between nutritional ingestion from the supplements and differential attendance rates. All analyses carried out to date on various outcomes indicate ingestion remains significant after controlling for attendance (see below).

Summary. The randomized design permits a strong inference ($P < 0.005$) that the intervention caused improvements in the outcomes. The component of the intervention that caused the impact must be elucidated by other analyses.

Dose response to supplementation

The analysis. One component of the intervention is, for instance, the knowledge gained by the measurers about the kind of supplement each village received. As discussed above, this could have biased the measurers. A dose-response analysis can be used to exclude the possibility because even though the measurers knew which villages received which supplement they did not know how much supplement each villager ingested. Therefore, any relationship between the amount of supplement (dose) and growth (response) could not be due to measuring biases. All the behavioral and most of the biomedical outcomes published to date have been examined by this method.

One reason for this approach is that sometimes village level analyses cannot identify a statistically significant impact of supplementation, even though it is present. This may be because of lack of statistical power because there were only two villages in each treatment. Had there been four villages per treatment, the *t*-test values would in all likelihood have been doubled with a large improvement in significance probabilities. There is no single appropriate number of replicates (villages) within a treatment. Instead, the number of replicates depends upon power analyses (Cohen 1988) of two-stage sampling designs (Snedecor and Cochran 1980) for each outcome of concern. Nevertheless, it is obvious that analyses of some outcomes will have less power than has growth because they are less reliably measured and because they are affected

to a greater extent by nonnutritional factors than is the case for growth.

Power also may be reduced if baseline values differed by supplement type, as was the case for length in the small village (Table 1), and these data were not available or taken into account. In that case, the impact of Atole relative to Fresco would be obscured. If one does not correct length after supplementation (Table 1) by the baseline data before supplementation, the difference between Atole and Fresco villages at the end of the study would be 2.15 ± 0.78 cm with a *t*-test of only 3.91 with 2 deg of freedom and a two-tailed probability of only $P < 0.10$. This compares to the much greater statistical significance of $P < 0.005$ when the baseline data are taken into account.

Another and unexpected obstacle to using the randomized design analyses occurred when it was found that mothers in Atole and Fresco villages consumed similar mean amounts of energy despite the much greater energy density of the Atole. The greater volume consumption of supplement in Fresco villages was unexpected. At the time the study was designed, birth weight was expected to respond to maternal protein supplementation. But what if birth weight responded to energy supplementation instead? Mean energy intakes in Atole and Fresco villages were so similar that differences in birth weight between supplement types would not be expected and indeed, none was found. A further complication is that baseline data on birth weight were not collected precluding analyses according to the randomized design. However, dose-response analyses revealed a 29-g increment in birth weight for every 10,000 kcal (41,840 kJ) of supplement ingested during pregnancy (Lechtig et al. 1975a). This birth weight increment was similar for women living in Atole (23 g/10,000 kcal or 41,840 kJ) and Fresco (30 g/10,000 kcal or 41,840 kJ) villages even though there was no protein in the Fresco. This corresponded to the lower end of the range (25–84 g of birth weight per 10,000 kcal or 41,840 kJ) expected from factorial calculations of the anticipated response (Lechtig et al. 1975b). As discussed below, the actual dose response is higher after controlling for data reliability. The congruence between the finding and its theoretical expectation is important in deciding the credibility of the inference that energy from the supplement affected birth weight.

Potential biases. Dose-response analyses are amenable to statistical significance testing. However, in contrast to analyses appropriate for the randomized treatment design, the statistical significance does not relate to the causality of the association, only to the evidence for an association between supplement and the outcome of interest. Unmeasured or poorly measured confounding still remains a possible explanation for the findings. However, if due attention is paid to critical issues, it is still possible to credibly infer causality.

For example, mothers with short durations of pregnancy had less time to partake of the supplements and consumed less than other mothers. They also bore smaller children because of shorter gestational ages. Therefore, the association between gestational age and supplement intake could explain the dose response between birth weight and supplementation. Also, attendance to the feeding centers was voluntary, as was consumption. Those who came to the centers differed in many characteristics from those who did not and these differences may account for the variation in birth weight (Johnson 1988). However, when these confounding measures and gestational age were included in multivariate analysis, the dose response between birth weight and supplementation was not reduced (Lechtig et al. 1975a). The dose response actually tended to increase indicating that those who would otherwise have borne smaller babies tended to ingest more energy from the supplement and subsequently had larger babies. Thus, the inference that the increase in birth weight was due to the supplementation was strengthened. Multivariate analyses of this type are the conventional means of demonstrating that a relationship is not due to confounding factors (Snedecor and Cochran 1980). Almost all evidence of economic impact (Judge et al. 1980), and most of the evidence about public health impact, come from these kinds of analyses.

However, it may have been that factors associated with both supplementation and growth were inaccurately and unreliably measured (Habicht et al. 1979) or not measured at all. Therefore, controlling statistically for confounding is always open to question (Kupper 1984), unless the factors controlled for are perfectly measured and are a complete proxy for all confounding variables.

It is most likely that attendance to the supplementation centers, or that amount of supplement ingested, is more directly related to potential confounders than is nutrient ingestion from the supplement, and, equally important, that all confounding associated with nutrient ingestion is mediated through attendance or amount ingested. If this is so, attendance and amount ingested are complete proxies for confounding related to nutrient ingestion. Both of these variables are almost perfectly measured. If these proxies are statistically controlled for in the analyses, and the nutrient ingestion continues to be associated with the outcome, one can be reasonably sure that this is not because of some confounding associated with self-selection for supplementation. Fortunately, this is the case for supplementation during pregnancy and birth weight. The association between birth weight and energy remains statistically significant even when attendance or amount ingested is controlled for, while the converse is not true; the association of amount ingested or of attendance is not statistically significant when energy is controlled for.

There is one situation in which the above analysis is not completely convincing: if the correlates of ingestion differ between Atole and Fresco villages. This is because there is no power to differentiate between attendance or amount ingested from nutrients ingested within a treatment group; there is too much colinearity. The power comes from comparing women with identical nutrient intake, but different volumetric intakes (and vice versa), across the treatments. In pregnant women, no such difference in correlates is seen (Johnson 1988).

The above differentiation between nutritional impact and confounding associated with amount ingested is not possible for the micronutrients added to the supplements. They were added in equal concentrations to both Atole and Fresco. Therefore, it is more difficult to exclude the possibility that a dose response to the micronutrients is due to the effect of socialization in the supplementation centers or to self-selection rather than to any nutritional effect. For birth weight and growth, this is not an issue because no dose response with these micronutrients was observed once energy was taken into account.

There is another, quite different way to control for confounding that does not rely on statistical control of differences among women. Constant differences among mothers were controlled for by relating differences in birth weight to differences in supplement intake across consecutive pregnancies (Lechtig et al. 1975a). This excludes all unvarying characteristics of the mother as a source of confounding (e.g., early childhood nutritional history, genetics, etc.), although it does not exclude factors that may change across pregnancies.

Summary. The above dose-response analyses depend upon demonstration of a statistical association between supplement intake and the outcome of interest. This association could be causal, but it also could be because of other factors, maybe unknown or poorly measured.

Credibility that the association is causal depends, therefore, on a constellation of findings and other evidence from the literature supporting the causal inference, and rejecting other explanations. Credibility is ultimately a qualitative judgement call in contrast to the quantitative assignment of probability that the randomized design permits.

Combining randomized and dose-response analyses

Rationale. Dose-response analyses are useful for revealing the underlying patterns relating supplementation to the outcomes. They also permit analyses that test alternative hypotheses about the cause of the relationship and that take these confounding factors into account. Finally, as discussed below, they permit estimates of the dose response and corrections for those

estimates. The statistical significance tests, however, only relate to the associations and not to the causality of the relationships. Inferences must therefore depend on how effectively one has dealt with confounding factors.

The importance of the statistical significance of a causal relationship between supplementation and outcome was emphasized earlier in analyses based on the randomized design. Analyses that combine both dose response and randomized approaches are more persuasive than either alone. Of course, this can only be done for outcomes that can be analyzed on the basis of the randomized design. Finally, credibility is further increased if the pattern of dose response across different kinds of children is as expected. Such analyses have been done (Rivera et al. 1991) to assess the impact of the Atole on the recovery from moderate wasting (<90% weight-for-length according to WHO's reference data (1983)) in children 6-24 mo old. Recovery is defined as recuperation of weight to >90% weight-for-length after 3 mo.

Analysis according to randomized design. The apparent recovery rate from moderate wasting was 50% in the Atole villages and 38% in the Fresco villages (Table 2). According to the randomized design analysis the statistical significance of the difference was $P = 0.07$ (two-tailed test) or $P = 0.035$ (one-tailed test). This statistical significance relates to the causal relationship between the intervention and the recovery rates. It fully takes into account the effects due to nonintervention factors that may cluster within a village.

Dose-response analysis. For these analyses high Atole ingestors were compared with high Fresco and low Atole and Fresco ingestors (Rivera 1988; Rivera et al. 1991). High Atole ingestion was defined as an ingestion of $\geq 10\%$ of the recommended dietary intake of energy (RDI) from the supplement. Forty-five percent of the wasted children in Atole villages were high ingestors, and 55% were low ingestors. High and low Fresco ingestors were defined as being, respectively, above and below the 55th percentile of Fresco volume ingested. For these analyses, those with high Atole

TABLE 2		
Recovery rates from moderate wasting (<90% weight-for-length) 3 mo after the diagnosis in children 6-24 mo in age		
Supplement type	Recovery rates	
	By village	Means of two villages
Atole	0.49, 0.52	0.50
Fresco	0.42, 0.35	0.38
Supplement effect (Atole-Fresco)		0.12 ¹

¹ $P = 0.035$, one-tailed t-test; $P = 0.07$, two-tailed t-test.

ingestion were compared with those with high Fresco and with low Atole and Fresco ingestion. Of those high ingestors in the Atole villages, 59% recovered while only 44% of low ingestors recovered. Among those in the Fresco villages with high and low ingestion, only 41% and 36% recovered, respectively. The difference between 59% and each of the three other rates of recovery was statistically significant ($P < 0.05$; two-tailed test). None of the three other rates were statistically different one from the other, as expected because none of these three groups received enough nutritional supplementation for one to expect an impact.

There is evidence of true supplementation. The high Atole ingestors consumed larger amounts of supplement (17.3% of RDI) than did the high Fresco ingestors (2.5% of RDI) whereas the home dietary intakes did not differ between the two groups (61.4 and 63.5% RDI, respectively).

Examination for potential confounding factors (parents' education and height, maternal modernity and parity, household size and sanitation, child birth weight, breastfeeding, home diet and illnesses) revealed that high ingestors in the Atole and Fresco villages were more similar to each other than to low ingestors.

Only three measures were different between high ingestors in Atole and Fresco villages. These were proportion of time ill with respiratory symptoms, birth-weight and duration of breast feeding. The first two were higher in the Atole villages and the third was lower. When these were taken into account in the above comparison between high Atole and high Fresco ingestors, the difference in recovery rates rose from 18 to 20% with a corresponding rise in statistical significance. Thus, these potential confounders were not the cause of the association between Atole and recovery from wasting.

As discussed previously, another potential bias is the measurer's knowledge about which village received which supplement. For example, an anthropometrist might unconsciously increase all the measurements in the Atole villages. The result would be a difference between Atole and Fresco villages that is "due" to the intervention, but that is not nutritional. However, the anthropometrist did not know levels of ingestion within a village and biases on the part of the measurer cannot explain the dose response. Furthermore, the dose response within the Atole and Fresco groups fully explains the difference between the Atole and Fresco villages found according to the randomized design. Thus, the difference in recovery from wasting between Atole and Fresco villages revealed by the analysis according to the randomized design was certainly not due to measurer bias. Only adding the analysis for dose response to that of the analyses based on the randomized design can exclude this kind of bias.

Credibility for a nutritional impact of the Atole is further improved if one has evidence that malnutrition

is prevalent. In fact, some children with low weight-for-height are thin but not malnourished. When the proportion of these thin nonmalnourished children in the study villages is subtracted (i.e., that proportion found below the criterion in the reference population used) from the denominators of the recovery rates from wasting, all the malnourished children who were good attenders in the Atole villages recovered. This represented a range in the absolute increase of recovery rate of 29–52% in the Atole villages, more than in the Fresco villages, rather than the 18% found in the uncorrected analyses.

Credibility for a nutritional impact is also improved if those children respond most whom one would expect to do so on the basis of knowledge about malnutrition. The sample sizes are too small to use the randomized design for subgroups of children to investigate this issue. However, introducing the appropriate interactions with supplementation into dose-response equations permitted the demonstration that more wasted children, younger children, children who had diarrhea and those who were supplemented for longer periods responded most to the Atole compared with those who received the Fresco (Rivera 1988). These results all correspond to long held expectations, except for the finding in children with diarrhea. The diarrheal findings correspond to new knowledge that the better the nutrition the less diarrhea affects growth (Lutter et al. 1992). In conclusion, all the results from the dose-response analyses give credible evidence of an effect of supplementation on recovery from malnutrition.

Combined analyses. Combining the dose-response analyses with the results of the randomized design eliminates the possibility that the results about growth could have been due to measurer bias or to self-selection for the supplement. The dose-response analyses also contribute importantly to the inference that the impact of the intervention is not only causal but is also nutritional.

Differentiating among the contributions of energy and other nutrients

Some analysts have attempted to separate the specific contributions of energy from those of protein and other nutrients in the supplements.

Pregnant women. This has been done most successfully in pregnant women who ingested similar quantities of energy from the Atole and the Fresco by ingesting much greater quantities of the latter. The dose-response on energy was the same in both the Atole and Fresco groups. This would not have been the case if protein or other nutrients present in Atole but not in Fresco (e.g., calcium, phosphorous) had been limiting in which case the dose response of birth weight to energy would have been higher (steeper slope) in the Atole group. Nor would it have been the

case if micronutrients added to the supplements had been limiting. These measured micronutrients (Table 2 in Martorell et al. 1995) were added in equivalent concentrations per volume to both supplements in 1971. Therefore, the dose-response would have been higher in the Fresco group if the measured micronutrients had been limiting because the micronutrient to energy concentration was higher in the Fresco group. Furthermore, there would have been no dose response in the Fresco group previous to 1971. This indicates that neither protein nor the other nutrients were limiting factors for fetal growth in the home diets of these mothers, but total energy was.

As discussed previously, the same conclusion results from step-wise multiple-regression analysis that reveals that energy is still statistically associated with birth weight, even when protein and the other measured nutrients are taken into account, whereas the converse is not true. Protein and the other nutrients are not statistically associated with birth weight when energy is taken into account. The above inference that energy is more deficient than protein in these diets is borne out by direct examination of the home diets. The mean of the home diet intakes for utilizable protein was slightly above the recommended daily allowances (7.4%) in contrast to the mean energy intake which was 39% less than the recommended daily intake (Lechtig et al. 1975b). Although the reliability of the individual measures of home diet is poor, that of the means is good (Habicht and Martorell 1992). Given this fact, plus some assumptions about the distribution of intakes around these means and about the dietary requirements, and above all, given the results about the relative effects of supplemental energy and protein, we conclude that energy was likely to be much more limiting than protein for these women.

Children. In children there was much less overlap between Atole and Fresco groups in energy supplement consumption (Schroeder et al. 1992). Atole children consumed much greater amounts of energy in the first three years of life. Analyses in Fresco children found similar or larger growth responses to energy intake than in Atole children, resulting in the conclusion that energy and not protein was limiting (Yarbrough et al. 1978). However, the results do not exclude the possibility that energy is limiting at lower levels of supplement intake as seen in the Fresco villages, but that protein may be limiting at higher levels of intake as seen in the Atole group.

The difference in slopes could not be because of a protein effect at lower levels of supplementation. They might be due to the flattening of the dose-response curve (Yarbrough et al. 1978) as energy intake approached adequacy. The higher response in Fresco also could be interpreted as evidence that another nutrient was limiting because their concentrations in relation to energy were much greater in the Fresco compared with the Atole.

In conclusion, none of these competing hypotheses have found resolution to date, possibly because resolution may not be possible in this data set. This lack of clarity about the exact nutrients that were responsible for the supplements' impact does not affect in any way the inference about a causal effect of the supplementation program on child growth to 3 y of age.

One claim that the effect of the supplement on growth was solely due to protein in the supplement (Balderston et al. 1981) is based on incorrect interpretations, as discussed below, from analyses using the home diet.

Taking home diet into account in dose-response analyses. The objective of the INCAP longitudinal study was to improve nutrition. It is therefore important to know whether the energy and nutrients from the supplements were supplemental or simply displaced home diet consumption. A direct approach would appear to be the use of total dietary intake (home diet + supplement) in the analyses. Alternatively, home diet and supplementation may be used as separate variables in a multivariate equation explaining the outcome. Both approaches would capture the net improvement in energy or nutrient intakes when comparing children from Atole and Fresco villages. Unfortunately, the low reliability of the home diet data usually precludes finding associations between the home diet and the outcomes. For example, Schroeder et al. (1995) found home diet energy to be much less related to growth increments than supplement energy. Rivera et al. (1995) increased the reliability of home diet energy by combining as many as eight separate surveys per subject and by creating a dummy variable, above or below the median; even though the diet variable was statistically significant and in the expected direction, analyses showed its inclusion in the model did not affect the regression coefficient for supplement. This indicates that the range in energy and protein intake from the home diets is too small to be important in the analyses. For this reason, home diet is also usually a de facto constant when included in analyses involving supplement because of the imprecision of its regression coefficients with the outcomes. Nonetheless, some analyses (Rivera et al. 1995, Schroeder et al. 1995) include home diet to increase persuasiveness as many readers would be troubled by the omission of this variable.

Inappropriate inclusion of home diet in the analyses and poor interpretation of the results can lead to false inferences as exemplified by the analyses that led to the claim that it was protein and not energy supplementation that produced better growth in children consuming Atole (Balderston et al. 1981). This claim was made on the basis of two findings: The first was: "The large effect [on growth] of increments of Atole supplementation for children eating the same home diet—contrasted with the small effect of increment in home diet for children consuming the same amount

of supplement—is not consistent with the hypothesis that the total energy value of the supplement is what accounts for the gains of children in Atole villages" (1981:59). In fact, this contrast was even stronger for protein (not reported) and was solely because of the poorer reliability of home diet compared with that of supplement intake. When the independent variable is very poorly measured, the estimates of magnitude (regression coefficients) are biased so they approximate 0 (Habicht et al. 1979). The correlation coefficients are also small, but this does not depend upon whether or not home diet is an independent or dependent variable.

The second finding in Balderston et al. (1981) was that in multiple regression analysis of the total diet (i.e., sum of home diet and supplement) there was a strong association between total protein ingestion and growth and none between total calorie ingestion and growth. This was entirely because of the fact that the Atole had a greater impact on total protein intake (on the average 35%) than on total energy intake (on the average 17%; estimates from Martorell et al. 1982 and from WHO 1985). As noted above, the variability in supplement ingestion is much better correlated with growth than is variability in home diet because of home diet's poor reliability. Therefore, for equal variability of supplement intake, supplemental nutrients that contribute a greater proportion of total dietary intakes will be more highly correlated with growth. This is particularly the case in children under 3 y of age because energy and protein intakes from the supplements are highly correlated with each other; energy ingestion from the Fresco was very low in relationship to Atole.

Home diet data are important, however, in investigating the degree to which the supplement replaced rather than supplemented the home diet. One cannot estimate the amount of supplement used to replace the diet at different levels of diet because this requires using the total diet as the independent variable. On the other hand, the estimates of amount of home diet replaced by the supplement at different levels of supplementation can be estimated without bias because in this analysis supplementation is the independent variable and it is measured almost perfectly. The estimated level of replacement for supplemental energy was 22% for pregnant mothers (calculation from Table V in Lechtig et al. 1975a), but negligible for children (Martorell et al. 1982).

These figures give the apparent proportion of supplement that substitutes for the diet. It might be that high ingestors of supplement would not have had enough food at home to bring their home diets to the same levels of intake as low ingestors, even if they had not consumed the supplement. This is indeed likely because high-energy ingestion from the supplement is related to lower socioeconomic status (Johnson 1988, Schroeder et al. 1992).

These magnitudes still will be somewhat underestimated because the outcomes are affected by variations in nutrition, not because of the supplement. The most important variation is home diet. This variation reduces the power to find associations between measures of supplementation and outcomes so that both the statistical significance and the regression coefficients relating a nutritional component of the supplement to the outcome understate this nutritional relationship (Habicht et al. 1979). In principle, this effect of home diet is no different from the influence of other factors that affect an outcome and that are randomly distributed across different levels of supplementation. The omission of home diet from the analysis of the impact of supplement on outcome measures is no different from the omission of other variables that affect growth but that are not related to the supplement, if one has taken the confounding because of home diet into account. Such omissions are inevitable. The confounding because of home diet is dealt with by correcting for the apparent substitution of supplement for home diet. For instance, the magnitude of the effect of supplement on birth weight is ~20% higher than the figures published when this correction is made. This means that the response of birth weight to actual energy supplementation is ~35 g per 10,000 kcal (41,840 kJ).

Conclusions

This paper presents the statistical significance of causality for an effect of the supplementation of the growth of 3-y-old children. It presents credible evidence for an effect of the supplement during pregnancy on the birth weight of infants. Finally, it presents evidence and the significance of causality for the effect of the supplement on the recuperation of malnourished children. The combination of the probability tests for causality with tests of association for credibility is necessary to make the most persuasive argument that the supplement had a nutritional effect on the outcome of concern. Credibility analyses are always possible and should always be done. Where the probability analysis for causality cannot be done, more analyses for credibility are necessary.

ACKNOWLEDGMENTS

We thank Catherine Geissler for comments on an early draft of this paper.

LITERATURE CITED

- Balderston, J. B., Wilson, A. B., Freire, M. E. & Simonen, M. S. (1981) *Malnourished Children of the Rural Poor*. Auburn House, Boston, MA.

- Cohen, J. (1988) *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Habicht, J.-P. & Martorell, R. (1992) Objectives, design and implementation of the INCAP longitudinal study. *Food Nutr. Bull.* 14: 176-190.
- Habicht, J.-P., Yarbrough, C. & Martorell, R. (1979) Anthropometric field methods: criteria for selection. In: *Human Nutrition, Nutrition and Growth* (Jelliffe, D. G. & Jelliffe, E. F. P., eds.), vol. 2, pp. 365-387, Plenum Publ. Corp., New York.
- Johnson, C. S. (1988) The role of participation with nutritional supplementation during pregnancy: a comparison of data from Indonesia and Guatemala. Masters thesis, Cornell University, Ithaca, NY.
- Judge, G. G., Griffiths, W. E., Hill, R. C. & Lee, T.-C. (1980) *The Theory and Practice of Econometrics*. John Wiley and Sons, New York.
- Kupper, L. L. (1984) Effects of the use of unreliable surrogate variables on the validity of epidemiologic research studies. *Am. J. Epidemiol.* 120: 643-648.
- Lechtig, A., Habicht J.-P., Delgado, H., Klein, R. E., Yarbrough, C. & Martorell, R. (1975a) Effect of food supplementation during pregnancy on birth weight. *Pediatrics* 56: 508-520.
- Lechtig, A., Yarbrough, C., Delgado, H., Habicht J.-P., Martorell, R. & Klein, R. E. (1975b) Influence of maternal nutrition on birthweight. *Am. J. Clin. Nutr.* 28: 1223-1233.
- Lutter, C. K., Habicht, J.-P., Rivera, J. A. & Martorell, R. (1992) The relationship between energy intake and diarrheal disease in their effects on child growth: biological model, evidence and implications for public health policy. *Food Nutr. Bull.* 14: 36-42.
- Martorell, R., Habicht, J.-P. & Klein, R. E. (1982) Anthropometric indicators of changes in nutritional status in malnourished populations. In: *Proceedings Methodologies for Human Population Studies in Nutrition Related to Health* (Underwood, B., ed.), NIH Publication #82-2462, pp. 96-110, U.S. Government Printing Office, Washington, DC.
- Martorell, R., Habicht, J.-P. & Rivera J. A. (1995) History and design of the INCAP longitudinal study (1969-77) and its follow-up (1988-89). *J. Nutr.* 125: 1027S-1041S.
- Rivera, J. A. (1988) Effect of supplementary feeding upon the recovery from mild-to-moderate wasting in children. Doctoral thesis, Cornell University, Ithaca, NY.
- Rivera, J. A., Habicht, J.-P. & Robson, D. S. (1991) Effect of supplementary feeding on recovery from mild-to-moderate wasting in preschool children. *Am. J. Clin. Nutr.* 54: 62-68.
- Rivera, J. A., Martorell, R., Ruel, M. T., Habicht, J.-P. & Haas, J. (1995) Nutritional supplementation during preschool years influences body size and composition of Guatemalan adolescents. *J. Nutr.* 125: 1078S-1089S.
- Schroeder, D. G., Kaplowitz, H. & Martorell, R. (1992) Patterns and predictors of participation and consumption of supplement in an intervention study. *Food Nutr. Bull.* 14: 191200.
- Schroeder, D., Martorell, R., Rivera, J. A., Ruel, M. T. & Habicht, J.-P. (1995) Age differences in the impact of nutritional supplementation on growth. *J. Nutr.* 125: 1060S-1067S.
- Snedecor, G. W. & Cochran, W. G. (1980) *Statistical Methods*, 7th ed. Iowa State University Press, Ames, Iowa.
- World Health Organization (1983) *Measuring change in nutritional status*. WHO, Geneva, Switzerland.
- World Health Organization (1985) *Energy and protein requirements*. WHO Tech. Rep. Ser. 724.
- Yarbrough, C., Habicht, J.-P., Klein, R. E., Martorell, R., Lechtig, A. & Guzman, C. (1978) Response of indicators of nutritional status to nutritional interventions in populations and individuals. In: *Evaluation of Child Health Services: The Interface Between Research and Medical Practice* (Bosch, S. & Arias, J., eds.). DHEW Publ. No. (NIH) 78-1966, pp. 195-207, U.S. Government Printing Office, Washington, DC.