THE WORLD BANK

BANGLADESH Development Series

South Asia Human Development Unit

The Bangladesh Integrated Nutrition Project
Effectiveness and Lessons

The Bangladesh Integrated Nutrition Project
Effectiveness and Lessons

Bangladesh Development Series – paper no.8

The World Bank Office, Dhaka

December 2005

www.worldbank.org.bd/bds

Document of the World Bank

Abbreviations and Acronyms

$M	Million U.S. dollars
ANC	antenatal care
BCC	behavioral change communication
BFD	breast-feeding
BINP	Bangladesh Integrated Nutrition Project
BMI	body mass index
BRAC	Bangladesh Rural Advancement Committee
CBNC	community-based nutrition component
CEA	cost-effectiveness analysis
CIDA	Canadian International Development Agency
CNC	community nutrition center
CNO	community nutrition organizer
CNP	community nutrition promoter
DFID	Department for International Development (UK)
DHS	Demographic and Health Survey
EBFD	exclusive breast-feeding
ESDP	essential services delivery package
EU	European Union,
GM	growth monitoring
GMP	growth monitoring and promotion
GOB	government of Bangladesh
H&N	health and nutrition
HAZ	height-for-age Z-score
HH	household
HKI	Helen Keller International
HNP	Health, Nutrition, and Population
HNPSP	Health, Nutrition and Population Sector Program
Ht	height
ICCDR, B	International Centre for Diarrhoeal Disease Research, Bangladesh
IDA	International Development Association
IEC	information, education, and communication
IMED	Implementation Monitoring and Evaluation Division, Ministry of Planning
IUGR	intrauterine growth retardation
K&P	knowledge and practice
LW	lactating women
MCH	maternal and child health
MDG	Millennium Development Goal
MUAC	mid-upper-arm circumference
NGO	nongovernmental organization
NNP	National Nutrition Program
NS	statistically insignificant
NSP	Nutritional Surveillance Project (by HKI)
NSS	Nutritional Surveillance Survey (HKI)
NWW	newlywed women
OED	Operations Evaluation Department, the World Bank
OR	operations research
PLW	pregnant and lactating women
PRSP	Poverty Reduction Strategy Paper

PSM	propensity score matching
PSU	primary sampling unit
PW	pregnant women
SCF	Save the Children Fund, UK
SD	standard deviation
SE	standard error
SES	socioeconomic status
SF	supplementary feeding
SIDA	Swedish International Development Cooperation Agency,
SWAP	sector wide approach
TPO	Thana (Upazila) project officer
U2	under 2 (children younger than two years of age)
U5	under 5 (children younger than five years of age)
UNMC	union nutrition management committee
VA	vitamin A
VAC	vitamin A capsule
VAD	vitamin A deficiency
VNMC	village nutrition management committee
WA	weight-for-age
WAZ	weight-for-age Z-score
WDR	World Development Report
WHZ	weight-for-height Z-score
WHO	World Health Organization
Wt	weight

Note*:* Thana = Upazila

The authors of this paper are: David Pelletier[a], Meera Shekar[b], Lidan Du[a] and Kees Kostermans[b].

[a] Cornell University
[b] The World Bank

Contents

Tables

Figures

Boxes

Acknowledgments

We sincerely acknowledge the dedication and steadfastness of the Ministry of Health officials and the staff of the National Nutrition Program, in cooperation with numerous nongovernmental organizations and other institutions, for their commitment to the fight against malnutrition in Bangladesh.

We thank Christine Wallich, Country Director for Bangladesh, for making the funds available for this study; Julian Schweitzer, Sector Director; and Anabela Abreu, Sector Manager in the South Asia Human Development Department, for their interest and support for these and other nutrition activities within the Bank.

The peer reviewers Saul Morris (U.K. Department for International Development), Tahmeed Ahmed (International Centre for Diarrhoeal Disease Research, Bangladesh), James Levinson (Tufts University), John Mason (Tulane University), and several World Bank colleagues provided very useful and frank comments on earlier drafts of the paper. Finally, I would like to thank the members of the HNP team in the Bangladesh World Bank Country Office in Bangladesh for their cooperation in this study.

Kees Kostermans
Lead Public Health Specialist, SASHD-HNP
Task Manager

FOREWORD

Overall Bangladesh is making good progress toward the Millennium Development Goals (MDGs), faster than any other country in South Asia. Especially in regard to the first goal, "Halve, between 1990 and 2015, the proportion of people who suffer from hunger," the country may well be successful. However, although the prevalence of malnutrition among children under five years of age is decreasing, it is still very high—about 40 percent of children under five are stunted and 12 percent are wasted, according to the 2004 Bangladesh Demographic and Health Survey. It is therefore extremely important that the country's commitment and the momentum for improving nutrition be sustained.

The Bangladesh Integrated Nutrition Programme (BINP) represented the first large-scale government intervention in nutrition. The program began in 1995; it was followed in 2002 by the National Nutrition Programme (NNP) and is being followed up now with a sector wide health approach. Nutrition also features significantly in the draft poverty reduction strategy. All of this shows the commitment of the government of Bangladesh in continuing to address malnutrition, which continues to constrain economic growth in the country.

There has been considerable debate in the press and the scientific community and among development partners about the impact of BINP. This study tries to lay that debate to rest by critically reviewing the various evaluations and trying to explain the variation across the results that were obtained.

The team of authors for the present report was chosen very carefully to meet two complementary requirements: (1) good working knowledge of BINP, the country context, and Bank operations and (2) an excellent understanding of the nutrition policy process and absence of vested interests in BINP or other programs in Bangladesh. Of course, all authors also needed to have sound technical skills in nutrition.

The World Bank remains highly committed to supporting Bangladesh in reaching the MDGs. It sees achievements toward the first MDG as a key element in the achievement of other development goals such as the fight against child and maternal mortality and poverty. Improved nutrition is also a driver of stronger economic growth in Bangladesh.

Jacques F. Baudouy
Sector Director
Health, Nutrition and Population,
Human Development Network

Julian F. Schweitzer
Sector Director
South Asia Human Development

Christine I. Wallich
Country Director
World Bank Office, Dhaka

Executive Summary

Background and Purpose

1. The Bangladesh Integrated Nutrition Project (BINP) was implemented in response to the high rates of low birth weight and malnutrition among children and women of childbearing age. The project's aim was to reduce malnutrition in Bangladesh through three intermediate objectives: (1) improve the capacity of national-level institutions in Bangladesh in the areas of advocacy, analysis, policy advice, operations research, and program support; (2) improve the capacity of communities, households, and individuals in the project area to understand their nutritional problems and to take appropriate action; and (3) improve the nutritional status of the population in the project area, with particular emphasis on children and pregnant and lactating women.

2. The project had three primary components: (1) the national-level nutrition activities, (2) the multisectoral nutrition program (garden and poultry raising), and (3) the community-based nutrition component. The allocated funding for each of these three components was $20.6M (32%), $7.6M (12%), and $39.1M (56%), respectively. They were financed jointly by the government of Bangladesh (GOB) and a credit of $59.8M from the International Development Association (IDA); of this, a total of $51.6M was disbursed. BINP covered 59 of a total of 464 Thanas (administrative units) in Bangladesh. The approach is being scaled up through the follow-on National Nutrition Program (NNP) to cover 105 of the country's 464 Thanas; further scaling-up of nutrition action is planned through the recently approved Health, Nutrition, and Population Sector Program (HNPSP).

3. BINP has been the subject of several evaluations aimed at assessing its effectiveness. These evaluations have differed in their design, analytical methods, rigor, findings, and conclusions. Thus, although some reports have claimed success for the nutrition interventions, others have suggested there was little effect. This lack of consensus has fueled considerable controversy and detracted from the important task of building a nutrition vision and a platform for action in a country that has malnutrition rates that are among the worlds highest.

4. The controversies surrounding BINP illustrate an inherent and common tension between the technical/administrative dimension and the political dimension of large-scale programs: the technical/administrative dimension should acknowledge the inherent complexity of large-scale programs, acknowledge that design and implementation problems always exist and, therefore, employ effective mechanisms to identify programmatic strengths and weaknesses so that improvements can be made over time. Typically, the time required for the evolution of such complex realities extends well beyond the five-year intervals of most development projects. In contrast, political support for improving national nutrition typically requires consensus, positive results, and shorter time horizons and can easily become disrupted by technical disagreements. The first step in managing these inherent tensions is to acknowledge their existence and the distinctive requirements in the political and technical domain. The second step is to demonstrate a commitment to not only *initiate* large-scale programs, but also to sustain the commitment over time and to ensure that their design, implementation, and management *can be improved in appropriate time scales*.

5. Notwithstanding the controversies, BINP was implemented between May 1995 and December 2002 and was followed by the NNP, whose implementation was just begun in 2004. Since then, nutrition issues are being incorporated into HNPSP, a sector-wide program; nutrition is also recognized as a key pillar of the Government's Poverty Reduction Strategy Paper (PRSP).

These actions reveal a high degree of long-term commitment by the government of Bangladesh to the cause of improving nutrition. The development partners likewise need to strike an appropriate balance between the technical/administrative dimension and the political dimension of this cause by providing consistent messages, constructive inputs, and a coordinated agenda that will improve nutrition and nutrition programs in Bangladesh. To support that process, this study examines the totality of evidence available on the effect of BINP (from existing reports), providing constructive suggestions on ways to improve program effectiveness over time, and placing the importance of BINP in the larger context of evolving commitment to nutrition in Bangladesh.

More specifically, the key objectives of this study are as follows:

- Document the nutrition trends in Bangladesh during the past decade (from available data sources such as the Demographic and Health Survey (DHS)
- Summarize the evidence for the efficacy of key nutrition interventions from a few key publications
- Review the evidence for the effect of Bank-supported nutrition interventions in Bangladesh and draw conclusions, to the extent possible, on their contributions toward nutrition outcomes in the country
- Comment, to the extent feasible from available data/analyses, on the contributions of the many components of BINP nutrition intervention package with a view to incorporating lessons into the HNPSP and the PRSP
- Propose an indicative package of nutrition interventions for the GOB to consider as part of the essential services delivery package in Bangladesh (ESDP). This includes the implementation of targeting schemes/techniques as well as the role of communities, if any.

6. This study does not reanalyze any available data or collect any additional data, nor does it review all of the literature in the public/private domain about nutrition in Bangladesh or globally.

7 It is hoped that this document will be useful for those who are involved in programming in health and nutrition in Bangladesh and elsewhere in the developing world. Those interested in these programmatic issues may wish to focus specifically on sections 1 and 3, whereas those interested in delving deeper into the technical issues may wish to also examine section 2 and the detailed appendixes.

The Bangladesh Integrated Nutrition Project

8 As noted earlier BINP consists of three primary components; however, part of the first component, the information, education, and communication (IEC) national activities, were never fully implemented because the development of mass media activities and IEC materials was considerably delayed. As a result many NGOs used their own IEC materials in their project areas. In addition, though several national-level activities such as vitamin A supplementation and salt iodization were implemented under this component, the capacity-building activities with national-level institutions did not materialize to the extent or in the manner originally envisaged. Similarly, the garden and poultry component was implemented at a very minimal scale. Therefore, of the three components this study will focus mainly on the community-based nutrition component (CBNC) with a much more limited review of experience with the multisectoral garden and poultry project. Some additional information on the role of this project on nutrition policy development in Bangladesh is also included, as well as some information on a few “unplanned” outputs from BINP.

The main activities in the CBNC include the following:

- Monthly growth monitoring and promotion (GMP) for children under two years of age and pregnant and lactating women (PLW)
- Supplementary feeding (SF) of malnourished PLW and malnourished and growth-faltered children under 2 years of age
- Nutrition education for pregnant women, mothers of children under two, and adolescent girls

9. The program strategy was designed to improve birth weight and maternal and child nutritional status and to produce sustainable changes in feeding, eating, and health care–seeking behaviors of children and pregnant and lactating women. In design, at least, the proposed approach appears to have been appropriate to the problems and the context and consistent with global best-practice knowledge on intervention strategies available at the time.

Sources and Methods

10 This study is based on a desk review of several earlier reports and analyses aimed at assessing the implementation and effectiveness of BINP. These differ in the source of data, evaluation design, analytical methods, and sponsoring organizations. The present review critically examined the methods and results from each of these reports in an effort to draw conclusions based on the weight of the accumulated evidence. The main reports and analyses considered here are as follows:

- *BINP Endline Evaluation, Final Report,* September 2003 (Karim et al. 2003)
- *Thin on the Ground,* 2003 (Save the Children Fund)
- *Baseline Survey of NNP: Initial Results,* January 2005 (ICDDR,B)
- *Impact Evaluation of BINP,* June 2004 (IMED, GOB)
- *Maintaining Momentum to 2015?* February 2005 (OED, World Bank)
  - Reanalysis of data from the endline evaluation report
  - Reanalysis of SCF data
  - Analysis of *2000 Demographic and Health Survey* (DHS) data
- *Process and Impact Study of BINP Garden and Poultry Activities (Karim et al. 2002)*

11 For the main BINP evaluation design, as used in the endline report, there are four categories of methodological concerns: (1) the criteria applied in selecting project and nonproject *Thanas*, (2) noncomparability in some of the measurements across surveys, (3) use of only two *Thanas* and small sample sizes for a control group, and (4) the likelihood that even the control group was exposed to health and nutrition interventions as a result of spillover from BINP itself, separate initiatives, or both.

12 The accumulated evidence on BINP has three important features. First, all available studies have some design weaknesses. Second, the nature of these weaknesses varies across the studies. Third, the techniques available and applied to detect and/or to correct for potential biases vary across the studies. The net effect of these features is that no one study can be considered unambiguously “more reliable” than the others. Instead, the findings from each study must be interpreted in light of their particular strengths and weaknesses and the process of drawing overall conclusions concerning the effect of BINP ultimately depends more heavily on qualitative judgments about study designs than on probability values from statistical tests.

Findings: The Community-Based Nutrition Component

13 Taken as a whole, the information reviewed here presents clear evidence for improvement in nutrition-related knowledge and practices, maternal and child nutritional status, and birth weight *in the project and nonproject areas alike*. The most consistent body of evidence concerning a positive effect of BINP, *beyond that seen in nonproject areas*, relates to knowledge and practices during pregnancy and to the use of available services such as vitamin A and iron supplementation. The evidence also suggests an added effect of BINP on knowledge and practices concerning infant feeding. The evidence presents a mixed picture concerning any additional effect of BINP on child nutritional status, weight gain during pregnancy, or birth weight beyond those seen in the nonproject areas. Whether that is due to a true lack of impact on outcomes per se or weaknesses in the evaluation designs and the data remains unclear. There is some evidence that BINP may have improved birth weight among important subgroups, such as among women who reported eating more during pregnancy (which was one of BINP's key messages) and, especially, among destitute women (presumably with poorer nutritional status) who reported eating more during pregnancy. These results are summarized below:

Young Children

14 Findings on the effects the community-based nutrition component had on young children are described below.

15 **Coverage and targeting of key child inputs:** The evidence suggests a *high level of participation in growth-monitoring sessions (75%–95%) and a moderate coverage of nutrition education sessions (66%).* The evidence on coverage and targeting of SF for children is limited to a prospective study in one *Thana* (where only 21% of eligible children received SF) and the Save the Children Fund, U.K. (SCF) cross-sectional survey in three *Thanas* (where 22% of severely malnourished children were receiving SF). Although the limited geographic coverage of these findings does not provide a firm basis for drawing conclusions about BINP as a whole, they do suggest that the strategies for increasing participation in GMP and nutrition education worked fairly well (with room for further improvement), whereas the coverage and targeting of SF may be problematic and deserving of more extensive examination.

16 **Knowledge and practices concerning child feeding:** The weight of the available evidence suggests that *BINP improved two key aspects of knowledge by about 10 to 20 percentage points beyond that seen in nonproject areas: feeding of colostrum—the first milk that protects infants against infections—and the benefits of exclusive breast-feeding.* The evidence further suggests *a large effect of the project on the practice of initiating exclusive breast-feeding* (30 percentage points higher in project areas) but a mixed picture concerning the practice of initiating complementary feeding, in that there was *no impact on the prevalence of exclusive breast-feeding at five to* six *months* of age. It appears from this that many more mothers in project areas do initiate exclusive breast-feeding at birth, which will improve the health and nutrition of their babies, but most do not continue the practice until the recommended six months of age. Future program efforts may therefore need to focus on continuing exclusive breast-feeding until six months of age and appropriate breast-feeding plus complementary feeding thereafter.

17 **Child nutritional status:** The available analyses provide inconsistent evidence for an effect of BINP on child anthropometry. Some of the analyses with stronger designs suggest a modest effect, but even these have design problems and differing results, which prevent firm conclusions. Specifically, on the basis of the weakest of study designs for assessing project impact (endline cross-sectional surveys), three analyses suggest a positive effect of BINP and five

analyses suggest no effect beyond that seen in nonproject areas. Two of these analyses employed somewhat stronger designs by controlling for potential confounders between project and nonproject areas: one suggests a positive effect of BINP and the other does not (unless the analysis is restricted to those children who participate more regularly in growth-monitoring sessions). *The weight of the evidence on BINP's effect on child nutritional status can be read in either direction* (though there is some evidence that those children who participated more regularly in growth monitoring did improve their nutritional status.) Whether that is due to a true lack of impact on outcomes per se or to the weaknesses in the evaluation designs and the data remains unclear. The one clear message from this is the need for designing and implementing stronger, technically sound evaluations right at the start when programs are being developed.

Pregnant Women and Birth Weight

18 Findings on the effects the community-based nutrition component had on pregnant women and birth weight are described below.

19 **Supplementary feeding during pregnancy:** Roughly *60 percent of malnourished women (with body mass index [BMI] < 18.5) received supplementary feeding, with a clear pattern of higher coverage among those most malnourished (with the lowest BMIs).* About 40 percent of ineligible women (BMI > 18.5) also received supplementary feeding. On the basis of available evidence, we conclude that receiving *supplementary feeding was not associated with improved birth weight in the project areas as a whole*, and it was not possible from the available studies to directly test whether supplementary feeding had an effect on weight gain. That has implications for the design of the supplementary feeding component for pregnant women in the future.

20 **Knowledge and practices related to pregnancy:** Available studies provide reasonably consistent evidence that BINP *improved selected knowledge and practices related to pregnancy by 20 to 40 percentage points beyond that seen in nonproject areas.* There is some additional evidence from the study with the strongest design that one of these practices (eating more during pregnancy) is associated with an 88 gram increment in birth weight. This suggests that BINP strategies for improving maternal knowledge and practices related to pregnancy worked well.

21. **Pregnancy weight gain and birth weight:** *The import of the evidence suggests little or no additional effect of BINP on pregnancy weight gain or birth weight for the population as a whole,* beyond that seen in nonproject areas. However, it is noteworthy that *subgroup analysis suggests sizable effects on birth weight among those who eat more during pregnancy (+88 g) and destitute women (presumed to have poorer nutritional status) who eat more during pregnancy (+270 g). Such large effects among the poorest have not been demonstrated in any large-scale projects to date.* The lower-than-expected effect of BINP on maternal weight gain and birth weight of the population as a whole can be explained by an evolving body of knowledge about the efficiency of various screening indicators (maternal BMI vs. weight) and the differential partitioning of energy between the mother and the fetus among severely versus moderately malnourished women. This evolving body of knowledge also suggests that the absolute weight of women may be a better predictor than BMI of intrauterine growth retardation (IUGR), which is the major cause of low birth weight in Bangladesh. However, the existing research is not adequate for revealing the best predictor of which women may benefit from supplementary feeding. That is a priority for future research; the MiniMat Study conducted by the International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR, B) should be capable of informing the choice of indicator and cutoff points as soon as the data are analyzed and the report is written.

Micronutrients, Health Care, and Garden/Poultry Project

22. Findings on the effects the community-based nutrition component had on the use of health care and micronutrients are described below.

23 **Micronutrients and health care:** *Available information provides fairly convincing evidence that BINP improved the delivery and use of micronutrients (iron/folate, vitamin A, and iodized salt) to a greater extent than that seen in nonproject areas.* Data from most studies showed a 10 to 30 percentage point higher uptake of vitamin A capsules among children under 2 years old and mothers of young children and a 10 to 30 percentage point higher uptake of iron/folate tablets among pregnant women in the project area. *Similarly, the integration of nutrition and health activities at the community level in BINP significantly improved the delivery and use of key health services* such as antenatal care and tetanus toxoid vaccinations. The key lesson from this is that BINP was successful in encouraging use of some available health services such as vitamin A supplementation, iron supplementation, and antenatal care. This has implications for the need for continuing to strengthen the programmatic link between health service delivery for basic health services and nutrition programming.

24. **Garden and poultry project:** *Although well intended, the garden and poultry project experienced a series of implementation problems* and was therefore highly unlikely to have any systematic nutritional effect. These problems included low coverage, poor targeting, high dropout rate, unmet demand for inputs, inadequate services provided by field workers, and managerial arrangements that failed to generate ownership and responsibility. Therefore, the evaluation of the effect of this component is not followed through. The key lesson learned from these problems is the need for greater attention to implementation issues and better mechanisms for quickly detecting and correcting implementation problems.

Overall Lessons and Implications

25. BINP is one of the largest multicomponent nutrition programs ever implemented in a developing country, and it took place in a country with one of the highest rates of poverty and maternal and child malnutrition. In contrast to the World Bank's health, nutrition, and population (HNP) efforts in many other countries and in recognition of the government's capacity limitations and role for community-level service delivery, it was implemented largely through nongovernmental organizations (NGOs), augmented as appropriate through the addition of essential cadres of workers and volunteers at the community level. In the process, it has demonstrated the feasibility of establishing or adapting large-scale structures to support community-based HNP programs in developing countries in which public sector capacities may be limited and innovative solutions are being sought. This experience is already being taken to scale through the NNP, with further mainstreaming of alternative health-service delivery mechanisms envisioned in forthcoming HNP sector wide approach (SWAP) and PRSP efforts (see box 1 page 14). These are major accomplishments that deserve full recognition.

26. Experience from BINP provides several important lessons for future nutrition programs in Bangladesh and elsewhere:

- One of the key lessons learned from BINP is that people can set themselves up for failure if they are overambitious in their objectives. The project aimed to reduce underweight and stunting rates by 50 percent over a course of 5 years (1995–2000). By comparison, the first Millennium Development Goal (MDG) aims to cut underweight rates (non-income poverty)

in half between 1990 and 2015—a period of 25 years; yet there is continued debate about whether the MDGs are achievable. Viewed in that light, BINP targets were very ambitious—they aimed to achieve in 5 years what the MDGs are aiming for during a 25-year period. In addition, the objective for low birth weight reduction aimed to "reduce the incidence of low birth weights by 50 percent" across all groups of pregnant women.

- In addition to setting an ambitious target, the less-than-expected impact of BINP in increasing birth weight may be explained by the evolving understanding that (1) BMI may not be a particularly useful screening indicator for intrauterine growth retardation (IUGR) and (2) food-related interventions may not affect birth weight through a straightforward input-output response as assumed in the past. Instead, for the most severely malnourished women, energy appears to be preferentially partitioned to the mother's own body stores, whereas for moderately malnourished women energy is partitioned to both her stores and the fetus. Although the cutoff points for defining these different partitioning categories are not yet known with precision, *it is likely that a significant proportion of women eligible for supplementary feeding in BINP were not biologically capable of responding with higher birth weights,* though they could still benefit in other ways. That probability suggests a possible need to change the screening indicator, the expectations about the way different women may respond to interventions, and the indicators used to evaluate programs in the future.

- *There is scope for improving the design and effectiveness of each major component of BINP.* Ten redesign suggestions are offered for supplementary feeding related to the nature, size, targeting, delivery, and monitoring of the food supplements and associated incentives; seven suggestions are offered for the behavior change interventions related to the content of nutrition education sessions, expanded target audiences, local monitoring, problem solving, and incentives.

- Experience with the two main interventions in BINP, supplementary feeding and nutrition education, underscores that effectiveness (and efficacy) is not simply an intrinsic characteristic of interventions. Instead, *the effectiveness of an intervention also depends vitally on context and implementation.* Because the context and the implementation process both are multilevel, complex, and dynamic, it is highly unlikely that the *initial* design of a program will "get it right" on all the important parameters. Therefore, *future programs must be designed with a strong capacity to make adjustments during implementation, including decentralized capacity.* Although this did happen to some extent in the transition from BINP to the NNP and is currently taking place in the move toward the sector-wide approach (HNP SWAP), the scope for adaptation and refinement of program design must be further enhanced. The challenge is to make feasible such a flexible program approach in the complex programming environment and limited managerial capacities for nutrition in developing countries such as Bangladesh, even while those capacities are being strengthened in this area. Recent experiences in other Bank operations have demonstrated the potential to build program-relevant capacity *simultaneously* with program implementation, and may be a promising approach as BINP approach expands (Matta et al. 2000).

- The foregoing may appear to be stating the obvious. But the experience with BINP and well beyond reveals that it may not be common practice to follow these principles, especially when dealing with state bureaucracies and procedures. The reality is that *large-scale* programs often lack the capacity to detect and correct problems with design and implementation in a timely manner, such that most lesson-drawing and program adjustments tend to happen on a 5- to 10-year time scale in response to "endline" evaluations." *Typically,*

the findings of summary evaluations lack the fine-grained detail required for making program adjustments, are not timely, and are resisted by most stakeholders; much greater emphasis should therefore be placed on building the capacity and the incentives for making timely and appropriate adjustments during implementation, based on operations research. This implies much greater emphasis on timely process monitoring and on mechanisms to ensure that information is fed into programmatic changes in a timely manner.

- BINP offers many lessons on the way evaluations can/should be done. Most notable is the need to design and analyze evaluation data with *differential responsiveness* in mind, to adopt *more explicit and sophisticated program theories to guide analysis and to address the common problems in translating evaluation findings into program and policy decisions.*

- Now that BINP infrastructure has been established, especially at the community level, and the NNP is being expanded, Bangladesh offers an excellent opportunity to address many of the issues outlined above. Doing so will require many changes, including a broad understanding that a "culture of inquiry" is fundamental to program success, clear delineation of the scope for discretionary decisions at each administrative level, significant strengthening of management and inquiry capacity at all administrative levels especially as the institutional arrangements for implementing nutrition programs are being defined in the coming year, and realignment of incentive systems to support a performance-based approach to development as is being done for the HNP sector program as a whole.

INTRODUCTION

1. The Bangladesh Integrated Nutrition Project (BINP) was designed in response to the high rates of low birth weight and of malnutrition among children and women of childbearing age. The project had three primary components: (1) the national-level nutrition activities, (2) the multisectoral nutrition program (garden and poultry raising), (3) and the community-based nutrition component (CBNC). The allocated funding for each of these three components was $20.6M (32%), $7.6M (12%), and $39.1M (56%), respectively. They were financed jointly by the government of Bangladesh (GOB) and a credit of $59.8M from the International Development Association (IDA); of this, a total of $51.6M was disbursed from the World Bank by December 2002. BINP covered 59 of the 464 Thanas (administrative units) in Bangladesh. The funding appropriated for each of these three components was $20.6M (32%), $39.1M (56%), and $7.6M (12%), respectively. BINP coverage was 55 of the 464 Thanas in Bangladesh. More recently, the approach is being scaled up in the National Nutrition Program to cover 105 of the country's 464 Thanas.

2. The project has been the subject of several evaluations aimed at assessing its effectiveness in reaching key objectives. These evaluations have differed in their design, analytical methods, findings, and conclusions. Thus, while some reports have claimed success for the nutrition interventions, others have suggested little effect. This lack of consensus has fueled considerable controversy and detracted from the important task of building a nutrition vision and a platform for action in a country whose malnutrition rate is among the worlds highest.

> ***"…the key to reconciling the differing requirements for technical and political decisions is to demonstrate a commitment to not only initiate large-scale programs, but also to ensure that the commitment is sustained, and their design, implementation, and management can be improved in appropriate time scales. "***

3. The controversies surrounding BINP are linked to an inherent and common tension between the technical/administrative dimension and the political dimension of large-scale programs: the technical/administrative dimension should acknowledge the inherent complexity of large-scale programs, acknowledge that design and implementation problems always exist and, therefore, employ effective mechanisms to identify programmatic strengths and weaknesses so that improvements can be made over time. Typically, the time required for realizing such complex programming realities extends well beyond the 5-year intervals of most development projects. In contrast, political support for improving national nutrition typically requires consensus, positive results, and shorter time horizons and easily can become disrupted by technical disagreements. The first step in managing those inherent tensions is to acknowledge their existence and the distinctive requirements in the two domains. Thereafter, the key to reconciling the differing requirements for technical and political decisions is to demonstrate a commitment to not only *initiate* large-scale programs, but also to ensure that commitment is sustained and the design, implementation, and management of programs *can be improved in appropriate time scales*.

4. Notwithstanding the controversies, BINP was implemented between 1995 and 2002 and was followed by the National Nutrition Project (NNP), whose implementation was just begun in 2004. Since then, nutrition issues have been incorporated into the sector wide approach through the Health Nutrition Population Sector Program (HNPSP); nutrition is also recognized as a key pillar of the Government's Poverty Reduction Strategy Paper (PRSP).

These actions reveal a high degree of long-term commitment by the government of Bangladesh to the cause of improving nutrition. The development partners likewise need to strike an appropriate balance between the technical/administrative dimension and the political dimension of this cause by providing consistent messages, constructive inputs, and a coordinated agenda that will improve nutrition and nutrition programs in Bangladesh. This study hopes to start to move in that direction by examining the totality of evidence available on BINP's effect, providing constructive suggestions on ways to improve program effectiveness over time, and underscoring the importance of BINP in the larger context of evolving political commitment to nutrition in Bangladesh.

More specifically, the objectives of this study are as follows:

- Document the nutrition trends in Bangladesh during the past decade.
- Summarize the evidence for the efficacy of key nutrition interventions from a few key publications (Note: The study was not expected to do, nor does it attempt, a full review of the literature).
- Review the evidence for the effect of Bank-supported nutrition interventions in Bangladesh and draw conclusions, to the extent possible, on their contributions toward nutrition outcomes in the country.
- Comment, to the extent feasible, on the relative contributions of the many components of BINP nutrition intervention package with a view to incorporating lessons into the health, nutrition, and population (HNP) sector wide approach (SWAP).
- Propose an indicative package of nutrition interventions for the GOB to consider as part of the essential services delivery package (ESDP) in Bangladesh. That includes the implementation of targeting schemes/techniques as well as the role of communities, if any.

5. It is hoped that this document will be useful for those who are involved in programming in health and nutrition in Bangladesh and elsewhere in the developing world. Those interested in these programmatic issues may wish to focus specifically on sections 1 and 3, whereas those interested in delving deeper into the technical issues may wish to also examine section 2 and the detailed appendices.

1. NUTRITION ISSUES IN BANGLADESH AND EFFICACY OF KEY INTERVENTIONS

A. TRENDS IN UNDERNUTRITION OF CHILDREN

1.1 There is generic agreement on the fact that the rate of malnutrition in children in Bangladesh, although one of the highest in the world, has decreased substantially during the past two decades. Figures 1.1 through 1.3 (data for each of which is drawn from the World Health Organization [WHO] Global Database on Child Growth and Malnutrition; for details see endnote 1 and notes under figure 1.3) show the trend of moderate and severe (combined) childhood malnutrition in Bangladesh between 1982 and 2004 in comparison with the corresponding Millennium Development Goal (MDG) of halving malnutrition (underweight) between 1990 and 2015.[1] If current trends continue, Bangladesh may well be able to achieve the nutrition MDG, but not the child mortality or maternal mortality MDG. Furthermore, though Bangladesh in the1980s started with much higher malnutrition rates than its neighbor India, today these levels are almost the same as those in India, which has done better than Bangladesh on all counts of economic development and social sector indicators.

1.2 The factors responsible for these improvements in nutritional status are suggested by three sets of analyses performed by the Operations Evaluation Department of the World Bank (OED) (2005). The first analysis applies Smith and Haddad's (2000) cross-country regression equations to Bangladesh. It suggests that income growth accounted for only 23 percent of the decline in stunting between 1980 and 2000. Association analysis of underlying factors suggest that the stunting decline is further accounted for by improvements in safe water (25.5%), female literacy (13.8%), improvements of gender equality (11.3%), and daily electricity supply (10.4%). Combined, these factors accounted for an estimated 82 percent of the decline in stunting, on the basis of regression equations from cross-country analysis. However, these associations cannot be assumed to be causal.

1.3 The second analysis is based on an updated cross-country study conducted by the OED, using more recent DHS data and a set of independent variables different from that of Smith and Haddad (2000). The application of the resulting equations from this analysis suggests that the decline in stunting in Bangladesh from 1980 to 2000 is accounted for by improvements in wealth, proxied by an assets index (8.3% share); health service index (19% share); secondary education (27% share), and access to media, proxied by watching television (46% share).

[1] The reasons for the rebound in child malnutrition rates after 2000 to 2001 are not well understood but it is unlikely that it is due to drastically worsened nutritional status. One possible reason could be the fact that the two surveys from different sources (2000 DHS survey and 2001 Helen Keller International [HKI] Nutritional Surveillance Project [NSP]) adopted different sampling schemes. DHS used a two-stage sample from the master sample maintained by GOB Bureau of Statistics, and the primary sampling units (PSU) are the lowest level of administrative units in urban and rural areas. HKI NSP used a multistage clustered sampling design in rural areas and used, up to this 2001 survey, a convenience sampling procedure for selection of samples in urban slums. The sample size of the 2001 survey was also nearly 10 times larger compared with all previous surveys.

Figure 1.1 Trends in Prevalence Rates of Underweight (weight-for-age) among Under-Five-Year-Old Children in Bangladesh

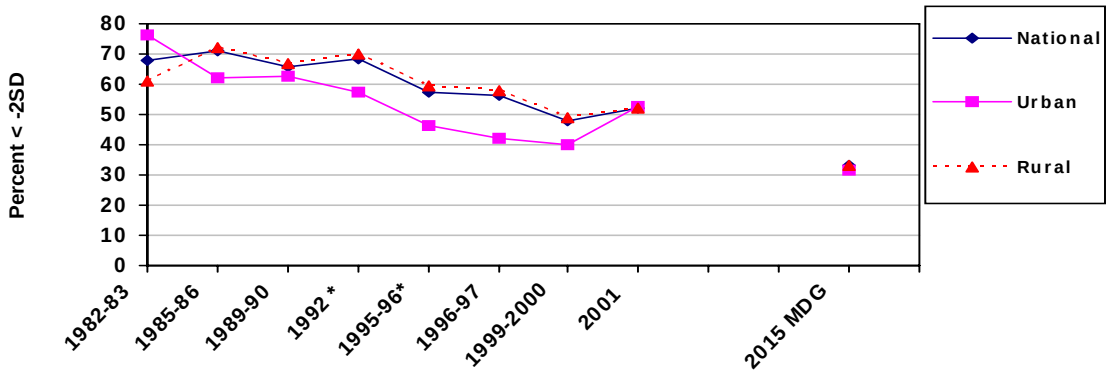


Figure 1.2 Trends in Prevalence of Stunting (height-for-age) Among Under-Five-Year-Old Children in Bangladesh

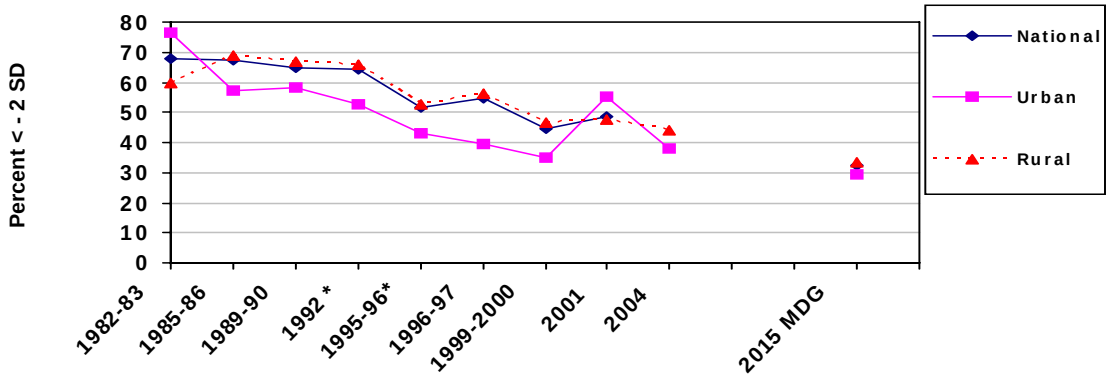
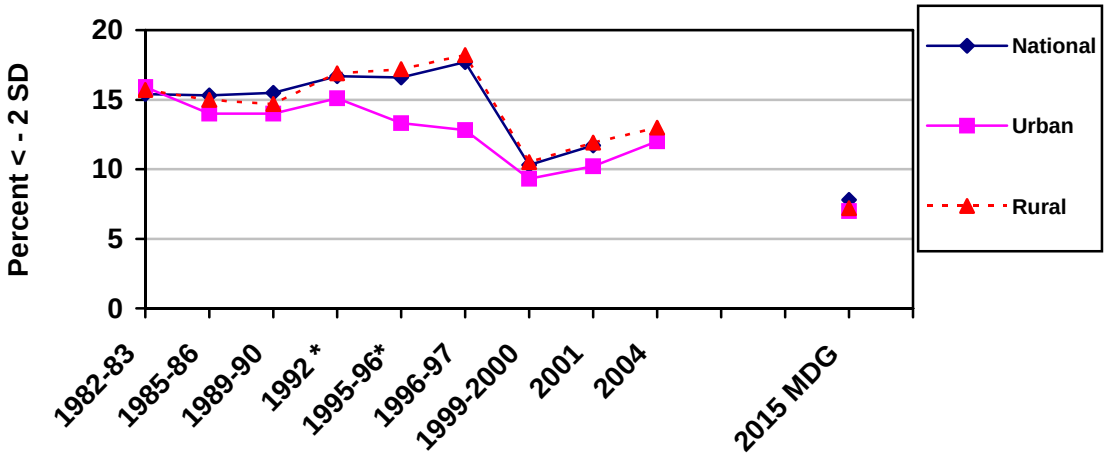


Figure 1.3 Child Wasting Prevalence Changed Very Little Among Under-Five Year Old Children in Bangladesh



Data source: WHO Global Database on Child Growth and Malnutrition, which draws on the following:

1982–3	HKI and Institute of Public Health. Bangladesh Nutritional Blindness Study
1985–6	Government of Bangladesh. Report of the child nutrition status module, Bangladesh Household Expenditure Survey 1985–6
1989–90	Government of Bangladesh. Report of Child Nutrition Status Survey 1989–90
1992	Ministry of Planning. Child Nutrition Survey of Bangladesh 1992
1995–6	Ministry of Planning. Child Nutrition Survey of Bangladesh 1995–6
1996–97	Bangladesh Demographic and Health Survey (DHS) 1996–7
1999–2000	Bangladesh Demographic and Health Survey (DHS) 1999–2000
2001	HKI. Annual Report of the Nutritional Surveillance Project (NSP) in 2001
2004	Bangladesh Demographic and Health Survey (DHS) 2004

Note*:* * In the survey during 1982–3, children under six were included in the urban samples whereas children under five were included in the rural samples. Both urban and rural samples in the survey during 1992 and during 1995–1996 included children under six.

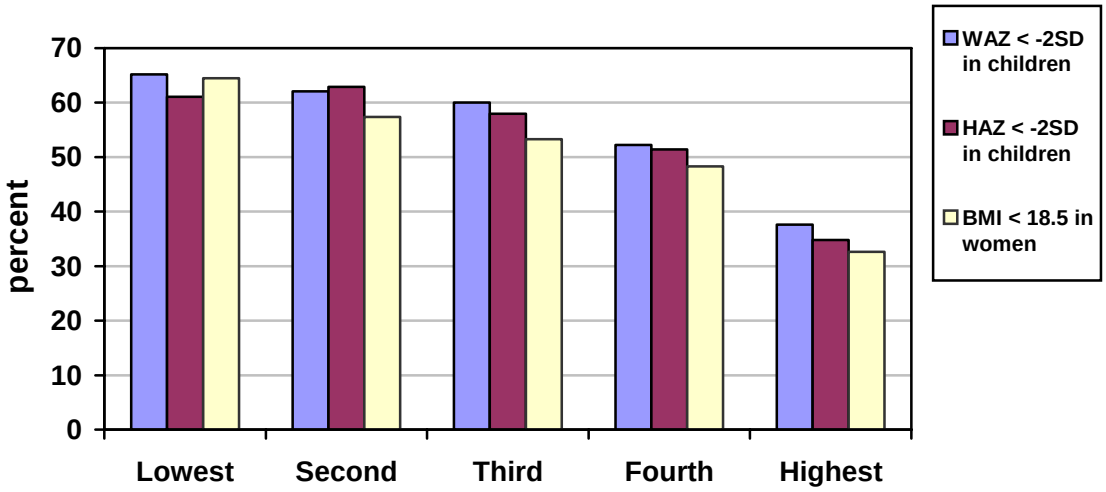
1.4 Although the analyses above have a number of limitations related to variable definition and the use of cross-country regression equations, the third set of estimates is based on regression equations from household-level analysis of DHS data collected in 1996 and 1999. This analysis, over a much shorter time frame, suggests that the decline in stunting rates during this period are accounted for by improvements in wealth (10% share), mother's secondary education (6% share), electricity in the household (4% share), and birth order effects related to declines in fertility (6% share). In this case a much larger share of the stunting decline remains unaccounted for (74%). As shown in figure 1.4, at the time BINP was being planned and initiated there was a clear gradient in malnutrition among women and children across income quintiles, but a significant fraction of women and children were malnourished even in the higher quintiles.

"Marked improvements in child nutritional status in Bangladesh from 1980 to 2000 are consistent with and partially accounted for by significant changes in physical infrastructure and economic and social development during this period."

1.5 Together, these analyses suggest that the marked improvements in child nutritional status in Bangladesh from 1980 to 2000 are consistent with and partially accounted for by significant changes in physical infrastructure and economic and social development during this period. These trends, which continued throughout the period of BINP's implementation, represent the backdrop against which BINP was operating and suggest that (1) the evaluation of BINP's effectiveness must take into account these marked changes already under way, through the use of appropriate comparison communities, and (2) to be considered effective, BINP must accelerate the decline in malnutrition beyond that provided by this background of development.

1.6 Figure 1.4 shows that although the poorest in Bangladesh have the highest levels of malnutrition (nearly 65% of them are underweight or stunted), even the richest quintile has very high malnutrition rates (35%–40%), thereby suggesting that malnutrition in Bangladesh is not necessarily eliminated through expansion of income alone (as proxied by assets) and that more direct efforts may be required to improve knowledge and the access to and use of health and nutrition services.

Figure 1.4 Malnutrition Rates Among Children Under Five Years of Age and Women Varied by Income Quintiles



Data source: Gwatkin et al. (2003) based on DHS 1996/7
Note*:* WAZ = weight-for-age z-score; HAZ = height-for-age Z-score.

B. MICRONUTRIENT DEFICIENCIES (PRESCHOOL CHILDREN AND WOMEN)

1.7 **Vitamin A deficiency:** According to the cutoffs used by the World Health Organization (WHO), vitamin A deficiency (VAD) among preschool children is no longer a public health problem in Bangladesh. The prevalence of night blindness in this age group in 1997 was 0.6 percent compared with 3.6 percent in 1983, and the prevalence reduced further to 0.3 percent (based on mothers' report) in 1999. That could be due to improved coverage/use of vitamin A supplements (capsules) among children under five, which increased steadily from 48.4 percent (among children under three), to 66.8 percent in 1996/7, and to 73.3 percent in 1999/2000 (DHS reports 1993/4, 1996/7, 1999/2000). However, roughly 25 percent of children under five still had subclinical vitamin A deficiency (serum retinol <20 µg/dL). One survey also found that night blindness, a sign of Vitamin A deficiency, among pregnant and lactating women in rural areas was still of public health significance, with a prevalence higher than 2 percent (HKI 1999a). This may be a manifestation of the low levels of vitamin A supplementation coverage post partum among women.

1.8 Factors associated with vitamin A deficiency, including subclinical deficiencies, among preschool children are insufficient breast-feeding, low vitamin A content in breast milk, and delayed and inappropriate complementary feeding. Deficiencies occur more often in children at younger ages. Several maternal factors, such as maternal malnutrition, lack of formal education, and high fertility as well as household characteristics, such as high vulnerability, also plays a role. Night blindness and subclinical deficiency among pregnant and lactating women appear to be associated mostly with household socioeconomic status (SES) factors, such as landlessness, high vulnerability, and poor sanitation/latrine, in addition to younger age, diarrhea, reproductive illness, and the women's own under nutrition.

Iron deficiency anemia: The National Vitamin A Survey by HKI in 1997 studied mother-child dyads and found that the age-adjusted anemia prevalence in preschool children was 52.7 percent with very few severe anemia cases. The prevalence was 78 percent in children younger than one year of age and decreased steadily with increased age (HKI 1999b).

1.9 Nearly 50 percent of pregnant and lactating women and close to 40 percent of other women of reproductive age are anemic. Iron deficiency is considered the main reason for anemia (Ahmed 2000). Malaria is endemic only in the mountain area (hill tracts). Iron intake appears to be low because of a lack of food diversity and a mainly plant-based diet. Intestinal helminths (ascaris and hookworm) are reported to be prevalent in Bangladesh in all age groups, but data concerning helminthic infection in relation to anemia are limited. Very high anemia prevalence in children younger than one year could be associated with a high prevalence (35%–50%) of low birth weight (LBW). Women's anemia appears also to be related to household SES, for instance, household vulnerability, taking a loan to buy foods in the past month, poor sanitation (open latrine), and low intake of fats and oils. Maternal characteristics such as increased fertility, under nutrition, low body mass index (BMI), level of education, and diarrhea are also positively associated with anemia (HKI 1999b). In another study of the hemoglobin concentration of Bangladeshi women, bioavailable iron in the diet, women's height and mid-upper-arm circumference (MUAC), and intake of iron tablets were identified as significant predictors of hemoglobin concentration (Bhargava et al. 2001)

C. EFFICACY OF KEY NUTRITION INTERVENTIONS

1.10 The previous section on trends in under nutrition suggests that some of the decline in malnutrition since the early 1980s could be associated with general improvements in physical infrastructure and economic and social development. This section briefly reviews some evidence concerning the efficacy of key interventions employed by BINP in an effort to accelerate the decline in malnutrition beyond that provided by these broader changes taking place in the country. This section draws on several recent reviews to highlight key features of these interventions, but is not intended as a comprehensive review of the literature.

Birth Weight and Women's Nutritional Status

1.11 The major cause of low birth weight (LBW) in developing countries such as Bangladesh is maternal under nutrition, that is, small maternal size (low weight and short stature) at conception and low weight gain during pregnancy (Allen and Gillespie 2001). The underlying rationale in the design of BINP's maternal food supplementation intervention is that the prevalence of low birth weight will be reduced through improvement in the mother's pregnancy weight gain (OED 2005). As described below, there are reasons to question that program theory in light of the evolving understanding of maternal and fetal nutrition.

1.12 Allen and Gillespie reported that, on the basis of a systematic review of intervention trials selected into the Cochrane databases (14 trials); maternal food supplements that provided balanced protein and energy were found to be the only intervention that improved birth weight. These supplements were defined as to contain less than 25 percent of the energy in protein form, and the maternal weight gain was only modest. The best predictors of low birth weight that is due to intrauterine growth retardation (IUGR), rather than prematurity, are preconception weight of less than 40 kg, short stature, and low attained weight during pregnancy (at week 20, 28, or 36).

1.13 The review by Allen and Gillespie on the interactions of prepregnancy BMI, weight gain, supplementation, and birth weight suggests a complicated picture with important implications for the interpretation of BINP and redesign of the NNP with regard to maternal supplementation. The review suggests, first, that pre-pregnancy BMI (which was used as a screening indicator for supplementation in BINP) is a poor predictor of response to food supplements in regard to improved birth weight. That is because, in general, low BMI women tend to gain more weight

during pregnancy (due to greater metabolic efficiency and their tendency to increase intake spontaneously) thereby partially offsetting the effects of low BMI itself. This has been seen in studies in Indonesia, Pakistan, China, and rural Mexico, where the highest weight gains occurred among the initially thinnest women. Thus, Allen and Gillespie conclude: "Based on these considerations it does not seem useful to choose low BMI as the indicator of IUGR risk, compared to low pre-pregnancy weight or attained weight [at a given stage of pregnancy]. It appears to be low lean tissue mass, rather than low fat mass, that predicts IUGR." A WHO Collaborative Study suggested that a preconception weight of 40 kg may be a useful cutoff for predicting IUGR risk in developing countries (assuming an average height of 150 cm) and noted that the predictive accuracy is even greater when using attained weights during pregnancy (WHO 1995).

1.14 The second insight from this review is that added energy from maternal supplementation appears to be preferentially partitioned to the *woman's* own body stores when women are *severely* malnourished at the onset of pregnancy, thereby producing little or no increment in birth weight. By contrast, supplementing *moderately* malnourished women leads to energy partitioning to both the mother *and* the fetus, thereby producing a more significant increment in birth weight and a more modest increment in maternal weight gain (Fellows 1985; Winkvist et al. 1992; Olson 1994). Although the exact cutoff points defining severe and moderate malnutrition for the purpose of predicting differential increments in weight gain and birth weight are not known, roughly half of the women in BINP had BMI < 18.5 by early/midpregnancy. This makes it likely that a significant fraction of the targeted and supplemented women may have been biologically incapable of responding to the food supplements with higher birth weights, although they may have benefited in regard to their own weight gain.

1.15 In regard to the best timing of food supplementation to increase birth weight, the evidence is inconclusive. Some observational evidence favors supplementation in the third trimester of pregnancy, whereas other observations (including intervention studies) support earlier supplementation (second trimester). Some evidence also shows that supplementation during one pregnancy that extends into lactation or longer could produce better birth weight in the next pregnancy, even if it does not improve birth weight from the first pregnancy. Together with the partitioning theory described above, these observations suggest a need to consider the well-being of the maternal-fetal dyad at several stages in the reproductive cycle as a focus for intervention and as a focus for measuring results, rather than a more narrow focus on critical periods during a single pregnancy with birth weight as the sole or primary focus.

"A significant fraction of the targeted and supplemented women may have been biologically incapable of responding to the food supplements with higher birth weights, although they may have benefited in regard to their own weight gain."

"These observations suggest a need to consider the well-being of the maternal-fetal dyad at several stages in the reproductive cycle as a focus for intervention, and as a focus for measuring results, rather than a more narrow focus on critical periods during a single pregnancy with birth weight as the sole or primary focus."

1.16 There has been very little research on the effect of non-nutritional interventions on birth weight. Some authors suggested that prolonged standing and strenuous work could have an adverse effect on birth weight especially among women who were underweight at conception and gained little weight during pregnancy.

1.17 Intervention for adolescent girls is also not straightforward. Food supplementation after menarche did not appear to generate catch-up growth in height, and it is unclear how much of the increased weight will eventually benefit the fetus. A more important and effective intervention could be to delay the age of marriage and time of first childbirth for girls. The linkage between poverty and early marriage/childbirth should be more explicitly addressed in women's empowerment and income-generation programs as companions to nutrition interventions.

Child Growth

1.18 The most common immediate causes of poor child growth are poor maternal health and under nutrition in utero (because catch-up growth for babies born with low birth weight is difficult); inappropriate breast-feeding; early or delayed introduction of complementary foods of inadequate quantity and quality; impaired absorption due to parasitic and other infections; and often a combination of those factors. The most important issue is that intervening after 18-24 months is less effective than during the first year of life. Some key lessons learned from interventions addressing the issues above include that energy intake is the strongest predictor of both height and weight, whereas sufficient protein intake alone has no added benefit on growth if children's energy intake is deficient. However, homemade complementary foods are generally low in energy, unless mothers are specifically taught how to make them energy-dense, and protein intake is ensured only when children are still being breast-fed after the first 6 months.

1.19 Fortified complementary foods as well as multiple micronutrient supplements (such as foodLET[1]) are critical in providing enough micronutrient intake to children. The results from efficacy trials are mixed. Results from a multicountry study on foodLET among infants between 6 and 11 months reported that only daily multiple micronutrient supplement is efficacious in reducing all micronutrient deficiencies of concern in the study as opposed to weekly supplementation and daily supplement of iron only (Smuts et al. 2005). These measures, however, might not always be feasible options in developing countries because the actual implementation is subject to many resource, belief, and behavioral constraints in those settings. This last concern could also affect food-based interventions aiming to improve children's (and women's) dietary intake and quality.

1.20 Home gardens and animal raising interventions usually focus on alleviating micronutrient deficiency, especially for vitamin A and iron. To have an effect on nutritional status, the assumptions have to be met that increased production and income from sales will lead to increased consumption, through improved knowledge and practices. An extensive review of project experiences and other evidence found that home garden interventions together with strong nutrition education/behavioral change components and promotional activities are more likely to

[1] FoodLET is "a crumbly hybrid between a food and a tablet, which is a chewable, flavored, reasonably soluble, water dispersible, and affordable product, with a good shelf life" (Gross et al. 2005). It is a concept developed during the first International Research on Infant Supplementation (IRIS) meeting in Rio de Janeiro in 1999. Trials were performed in four countries: Indonesia, Peru, South Africa, and Vietnam. Four different formulations of foodLETs were tested in an attempt to develop a credible evidence base to guide policy decision making in this area. Additional work on impact, costs, and delivery systems for a scale-up is needed.

result in increased vegetable intake (Ruel 2001). As for animal-raising interventions, evidence from household studies showed that the increased income from sales of animals may not significantly improve dietary quality if the nutrition education component is absent or weak. Because of the nature of such interventions, *efficacy* trials are rare, and hence the impact of home gardens on nutrition outcomes has not been effectively demonstrated (Schipani et al. 2002). The design and evaluation of *effectiveness* studies are weakened by participation and selection biases in project versus control households/villages, making it difficult to provide convincing evidence of the effects of home gardening interventions in preventing VAD or iron deficiency anemia. However, apart from possible effects on nutritional status, some authors call attention to the additional potential benefits of this intervention for increasing household food security and empowerment of women in having more control on income and influence in household decision making (Bushamuka et al. 2005).

Nutrition Education, Growth Monitoring, and Behavior Change

1.21 Abundant evidence exists that women's literacy is one of the most consistent determinants of the health and nutritional status of children (e.g., Caldwell 1993; Hobcraft 1993; Thomas et al. 1991). There also is convincing evidence that well-designed behavior change programs can have a positive effect on mothers' knowledge and practices, children's food intake, and child nutritional status (Caulfield et al. 1999). In the few programs that have assessed dietary intakes, an improvement of 70 to 165 kcal/d was found; the improvement was due to large shifts in maternal knowledge/attitudes and infant feeding practices. The additional caloric intake indicated weight-for-age and height-for-age gains of +0.24 to +0.87 standard deviation (SD) at 12 months (Caulfield et al. 2005). However, the effectiveness of these programs depends vitally on the quality of their design and implementation. Moreover, their effects can be modified or negated by a number of conditions, which can be broadly grouped under (1) power, including factors such as time, income, food, decision autonomy, social support, and access to services and the social structures that shape them and (2) beliefs, including cultural norms, values, and expectations, and psychological features such as self-efficacy, attitudes, and mental health (Pelto and Backstrand 2003). These power and belief factors interact with each other in many ways and at all levels of society, to ultimately shape child-feeding practices and women's practices during pregnancy. "New knowledge" from nutrition education may become incorporated in this complex, but by itself may be inadequate to influence practices. If the goal is to change practices, rather than simply to inform and educate, the findings from research suggest that well-designed efforts must be made to address how new information is incorporated in the power and belief complex (Contento 1995) and/or how other aspects of that complex must change (Pelto and Backstrand 2003).

1.22 Because of the complex and contingent relationships in this area, the literature seldom discusses the *efficacy* of nutrition education, growth promotion, or behavior change communications. Rather, the primary focus is on articulating the preconditions or requirements for achieving *effectiveness* in specific contexts and on evaluating the effectiveness of specific projects in specific contexts. There also are numerous manuals, principles, and best practices (cf. Allen and Gillespie, pg. 69.) that can serve as guides for design, implementation, and monitoring. This is an appropriate approach in light of the complexities and contingencies involved. However, this does pose a dilemma for those parties wanting to know "what the effective interventions for infant and child feeding are" because the answer depends heavily on the capacity for designing the intervention to fit the local context as well as the capacity to implement the intervention as planned. Thus, once-and-for-all judgments cannot be drawn about the effectiveness of such approaches, as though they apply to all contexts uniformly. Instead, each design team must decide

whether and how these approaches seem warranted in the given context and whether the requirements for successful implementation can eventually be met through quality improvement approaches and adaptive management.

1.23 In summary, major forms of malnutrition in young children and pregnant and lactating women (low birth weight and energy and micronutrient deficiencies) were included as project objectives in BINP. Many of the established intervention strategies and approaches (such as growth monitoring and promotion [GMP], community participation, nutrition education, food supplementation, adolescent education) were adopted in the project to address the identified causes of malnutrition in this context (such as knowledge and practices concerning pregnancy and child feeding). Literature available at the time, and confirmed since then, supports that addressing malnutrition could reduce infant/child mortality (one of the ultimate goals of BINP) by reducing low birth weight and child malnutrition (Jones et al. 2003). In addition, BINP addressed several key opportunity areas (low birth weight, child feeding, particularly exclusive breast-feeding, and iron deficiency as included in BINP) that will effectively lead to significant economic gains through the saving of health care resources and higher productivity from improved physical and cognitive capacities in a population (Behrman et al. 2004). Finally, there were other national-level (capacity-building and information, education, communication/behavioral change communication [IEC/BCC]) and intersectoral activities in BINP (food security through vegetable garden and poultry raising). In design, at least, the proposed project approaches appear to have been appropriate to the problems in the context of the time and considering the then-available information on potential impact of interventions.

2. EFFECTIVENESS OF BINP

A. BACKGROUND OF BINP

2.1 As mentioned previously, BINP was carried out from 1995 to 2002 in response to the high malnutrition rate, particularly among women of childbearing age and children (Karim et al. 2003; IMED 2004). The project was designed with three components: the national-level nutrition activities (capacity building, IEC/BCC, project management, and monitoring and evaluation), the multisectoral nutrition program (garden and poultry), and the community-based nutrition component (CBNC). The IEC activities were never fully implemented because of inadequate financing of mass media activities and material development. As a result, some NGOs used their own IEC materials in their project areas (IMED 2004, Kabir and Levinson unpublished). In addition, it appeared that many of the capacity-building activities on national-level institutions did not take place as planned (IMED 2004). Therefore, the component on national-level nutrition activities will not be reviewed in this study; however, box 1 below lays out how BINP has contributed to nutrition policy development in Bangladesh.

2.2 This study will focus mainly on the community-based nutrition component (CBNC) with a brief review of garden and poultry activities. The central activities of the CBNC included the following:

- Monthly growth monitoring and promotion (GMP) of children and pregnant and lactating women (PLW)
- Supplementary feeding (SF) of malnourished PLW and malnourished and growth-faltered children under 2 years of age (U2) plus referral to health services
- Nutrition education for mothers and adolescent girls

The strategy was designed to produce sustainable changes in feeding and eating behaviors of children and PLW.

2.3 The CBNC activities were implemented by contracted NGOs and by some field-level government staff. The purpose of GMP and SF was to serve as tools to generate community participation and for demonstration and nutrition education[2]. In addition, there were two primary strategies adopted in the provision of project activities— one targeted the "newly married couples" and the other targeted the vulnerable groups, which was implemented in most project *Thanas*.

The following are specific objectives of the CBNC:

- Reduction of severe protein-energy malnutrition (PEM) in U2 children by 50 percent
- Reduction of moderate PEM among U2 children by one-third
- Reduction of low birth weight (LBW) incidence by 50 percent
- Improvement in weight gain to 7.0 kg in at least 50 percent of pregnant women
- Virtual elimination of iodine deficiency disorders
- Virtual elimination of vitamin A deficiency disorders (VAD) and their consequences, including blindness
- Reduction of iron deficiency anemia among children and PLW by one-third

[2] This appears to recognize a distinction that GMP in itself is not an intervention, but an effective approach to offer opportunity for communities to understand and take action to prevent poor growth (Griffiths et al. 1996). Moreover, "growth monitoring in isolation from activities that *improve* children's growth is not going to achieve anything" (Mason et al. 2005).

Box 2.1 Mainstreaming Nutrition in Bangladesh Sustaining Commitment in the Face of Controversy

In the early 1990s the levels of undernutrition in Bangladesh (stunting and underweight as well as micronutrient deficiencies) were among the highest in the world. The Bank's first nutrition investment in Bangladesh in 1995, the Bangladesh Integrated Nutrition project (BINP), was a sector investment loan of $59M. On the basis of a thorough sector review of nutrition issues, the project focused on improving the poor maternal knowledge and inadequate caring practices for young children, which were identified as a key cause of undernutrition in the country. Implementation of the project got off to a slow start—nutrition actions did not fit neatly into one sector, the capacity for implementing nutrition actions was minimal, and many people still believed that poor nutrition was primarily an outcome of a lack of food. Furthermore, the process of contracting NGOs for implementing the project at community levels was cumbersome, bureaucratic, and slow. By 1999 the program was on the ground, and following a visit to the country by then World Bank president Mr. Wolfensohn, a second investment ($92M) was agreed on for a scale-up of the project (from the original 59 *Upazilas* to *139 Upazilas)* covering approximately 30 percent of the rural population in Bangladesh. In 2000 a new project, the National Nutrition Project (NNP), was prepared. However, at that time, the evaluation data for the BINP were poorly analyzed and implemented. The initial BINP evaluation reports were strongly contested by international NGOs (SCF, UK, in particular), and the subsequent discord and discussions (as well as the continued lack of implementation capacity and bureaucratic processes for contracting NGOs) seriously detracted from the implementation of the NNP, which was considerably delayed. In 2003 a stronger evaluation effort was put in place for the BINP, followed by an even stronger baseline evaluation for the NNP. Both of these efforts suggested that the BINP had indeed had a strong impact on changing maternal knowledge and caring behaviors and that this had translated into improved outcomes.

The stage was now set for incorporating nutrition into the Health and Population Sector Program. In December 2004, under the strong leadership of the government of Bangladesh, development partners in Bangladesh endorsed and appraised a large sectorwide operation (the HNP SWAp) based on the government's own Sector Investment Plan. In the process of developing this project, groups such as SCF made a strong plea for why nutrition must be included in the HNP SWAp. A total of nearly $750M grant and credit monies have been committed for the HNPSP for 2005 to 2010 by the pooling partners (DFID Department for International Development, EU-European Union, SIDA-Swedish International Development Cooperation Agency, CIDA-Canadian International Development Agency, the Dutch government, WB-World Bank and others), with a potential for additional funds if progress is good), and an additional $400M has been committed by other development partners through parallel arrangements. Nutrition is firmly included in this SWAp operation, and an monitoring and evaluation component has been built into it; disbursements under the HNPSP will be closely linked to performance, and good performance will be rewarded with supplemental grant funding from bilateral funds, thus making the M&E component even more critical. Finalizing the longer-term institutional arrangements for nutrition is a key component of the new HNPSP. Nutrition issues are also being strongly reflected in the country's new draft PRSP—nutrition being one of the first pillars of the seven-pillar PRSP. This remarkable progress in Bangladesh offers many lessons for nutrition investments in other countries. Among the key lessons are the crucial importance of professional evaluation in confirming what works (and what does not), building coalitions with detractors, the perils of a too early scale-up, the need for stronger investments in capacity building and institutional arrangements for nutrition, and the potential role of nonpublic providers for nutrition services. The role of political commitment, both within the Bank and in countries, is key to ensuring sustained investments as against a one-off short-term investment in nutrition. Recognizing and managing the inherent tensions that arise in sustaining political commitment, while also allowing scope for fine-tuning technical/administrative dimensions of intervention strategies, are critical for all those working in development.

B. Evaluation of the Community-Based Nutrition Component (CBNC)

Evidence Available for the Review

2.4 BINP has been the subject of numerous analyses and reports aimed at assessing the effectiveness with which it has met its stated objectives. These analyses have differed in the source of data, analytical methods, and sponsoring organizations, with corresponding differences in interpretations and conclusions. The primary reports on this topic have been reviewed in detail for the present report. This section restricts itself to the overall findings emerging from this review. Supporting details are provided as appendix material (appendices B–G) and are referenced below.

This study draws primarily on the following reports and analyses:

a) BINP Endline Evaluation, Final Report, September 2003 (Karim et al. 2003)
b) *Thin on the Ground,* 2003 (Save the Children Fund)
c) *Baseline Survey of NNP: Initial Results,* January 2005 (ICDDR,B)
d) Impact Evaluation of BINP, June 2004 (IMED, GOB)
e) *Maintaining Momentum to 2015,* February 2005 (OED, World Bank)
   - Reanalysis of data from the endline evaluation report
   - Reanalysis of SCF data
   - Analysis of the *2000 Demographic and Health Survey* (DHS) data

2.5 The key design features and findings in each of these sources are summarized in appendix table B1. In broad terms, BINP endline contains results from the baseline (1995), midterm (1998), and endline (2003) surveys, in project and nonproject ("control") areas. The next three reports (b, c and d) are based primarily on cross-sectional endline surveys in project and nonproject areas. The unique contribution of the OED report (e) is an effort to overcome some of the methodological limitations of all the preceding reports through the application of multivariate modeling to data from the endline evaluation report (a), the SCF survey (b), and the 2000 DHS data.

> ***"No one study can be considered unambiguously 'more reliable' than the others. Instead, the findings from each study must be interpreted in light of their particular strengths and weaknesses, and the process of drawing overall conclusions concerning BINP's impact ultimately depends more heavily on qualitative judgments concerning study designs than on probability values from statistical tests."***

2.6 In the main BINP evaluation design, as used in the endline report, there are four categories of methodological concerns: (1) the criteria applied in selecting project and nonproject *Thanas*, (2) noncomparability across surveys in some of the measurements, (3) the use of only two *Thanas* and small sample sizes for a control group, and (4) the likelihood that even the control group was exposed to health and nutrition interventions as a result of spillover from BINP itself and/or separate initiatives. The first two of these concerns are described in box 2 and appendix D, the third is examined in box 3 and the fourth is discussed in many project reports as documented in table B3 of appendix B.

Box 2.2 The BINP Intervention Group

The stated selection criteria for BINP *Thanas* included representation of various levels of food insecurity, avoidance of disaster-prone areas, presence of basic health infrastructure, and at least 80 percent full immunization coverage (SCF report, citing Staff Appraisal Report of the World Bank). There are anecdotal and informal reports suggesting that these criteria were not strictly implemented and that other issues of accessibility, geographical balance, and priorities were important. For present purposes the main question is whether or how the selected *Thana* may have differed in ways that could affect their response to the program and thereby bias the estimates of the project's effect. That remains an open question, but some of the sociodemographic characteristics of households with U2 children at baseline are shown below (based on Karim et al. 2003 table IV-2):

Indicator	BINP (6 *Thanas*)	Control (*2 Thanas*)
Water source (tube, tap, or pipe) %	95.7	99.1
Owns homestead land %	8.9	0.5
Grows vegetables at home %	33.3	18.3
Is landless %	47.8	38.8

Box 2.3 The BINP Control Group

In addition to the questions concerning the criteria for selecting the project and nonproject *Thanas* (appendix C), the use of only two control *Thanas* in the six-cell evaluation design of the endline report imposes enormous uncertainty on the comparisons with BINP *Thanas*. The data in table A26 of the endline report can be used to quantify this uncertainty, based on estimates of the variation in mean HAZ changes (i.e., the pooled variance in the secular trend) across the eight *Thanas*, using *Thana* as the unit of analysis. The estimated SD for this change is 0.28 for baseline to endline and 0.13 for midterm to endline. These parameters can be used to estimate the standard error (SE) and 95% confidence interval, within which the true mean change in HAZ is located for project and nonproject *Thanas* (see table below). These confidence intervals in the control areas are very wide relative to the difference in mean HAZ changes. In other words, this implies that the true change in HAZ from midterm to endline in the control areas is somewhere between -0.36 and 0.00 (with 95% confidence), and even wider from baseline to endline. This underscores the extreme sensitivity of the double difference in this design to the choice and characteristics of the two control *Thanas*. The use of large sample sizes at the child level does nothing to alter these potential biases; to the contrary, it only increases the statistical power to erroneously attribute the observed double difference to the effects of the BINP.

	Base-End		**Mid-End**	
	Control	Project	Control	Project
No. *Thana*s	2	6	2	6
Mean HAZ change (OED)	+0.42	+0.46	–0.18	+0.08
Double difference (OED)	+.04		+.26	

SD of *Thana* mean changes	0.28		0.13
SE of *Thana* mean changes	0.20		0.09
95% confidence interval for mean HAZ change in control areas	0.02 to 0.82	-0.36 to +0.00	

Analogous calculations were performed on the cross-sectional variation in mean HAZ at endline, which is an especially important comparison in light of the extensive reliance on endline comparisons for many of the variables in the BINP endline report and the OED's reanalysis of it. In the case of HAZ the calculations show that, although the observed endline mean HAZ for the control areas is -2.07, if two control *Thanas* were to be chosen at random in 100 successive draws, the observed means would range from -1.92 to -2.22 95% of the trials.

2.7 Table B3 (appendix B) reveals three important features of the accumulated evidence concerning BINP. First, all of the available studies have some design weaknesses. Second, the nature of the design weaknesses varies across the studies. Third, the techniques available and applied to detect and/or correct for potential biases vary across the studies. The net effect of these features is that no one study can be considered unambiguously "more reliable" than the others. Instead, the findings from each study must be interpreted in light of their particular strengths and weaknesses, and the process of drawing overall conclusions concerning BINP's effect ultimately depends more heavily on qualitative judgments about study designs than on probability values from statistical tests.

Child Nutritional Status

2.8 The available studies provide inconsistent evidence for an effect of BINP on child anthropometry. Studies with the stronger designs suggest a modest impact, but even these studies have some design problems. The most relevant findings are shown in table 2.1 and highlighted below:

(a). The most common study design involves cross-sectional comparisons of project and nonproject areas, most of which are at endline. Positive effects of BINP are suggested in analyses 1A, 1B, 3B, 3C, 3D, and 5B. Analyses suggesting no positive impact are 1A, 2, 3A, 4, 5A. (Note that analyses 1 and 2 are based on the same data set, as are analyses 3A, 3B, 3C, 3D, the difference among them being the age group examined and/or the use of mean Z scores versus prevalence estimates). A further cross-sectional comparison was conducted at midterm in analyses 1A and 1B, and these suggest a positive impact on mean WAZ (weight-for-age Z-score), mean WHZ (weight-for-height Z-score), and the prevalence of low WHZ (see table B1-3 for details). Analysis 5A suggests positive effects on all three anthropometric indicators at midterm but no effects at endline. Analysis 5B also suggests positive effects, based on DHS data in 2000, which falls halfway between the midterm (1998) and endline (2003) surveys.

(b). Among these cross-sectional comparisons the stronger designs are found in analyses 5A and 5B, but even these have important limitations (see the footnotes for table 2.1 in appendix A), and it is relevant to note that the results for analysis 5A shown in the table are based on a comparison of all children 6 to 23 months of age in project versus nonproject areas; when the analysis is restricted to those receiving more regular program inputs (as indicated by participation in growth monitoring), small but statistically significant effects are observed on all anthropometric measurements at endline as well as midterm. Although analyses 5A and 5B appear to be a stronger basis for estimating project effects than the other cross-sectional

comparisons, they still have some troubling methodological limitations as described in comments j and k in the table.

Table 2.1 Summary of Child Anthropometry Findings

Report/ Analysis	Indicator	Group	Baseline	Midterm	Endline	Midterm-Baseline	Endline-Midterm	Endline-Baseline	Technical Footnotes
1A. Endline	HAZ	Project	-2.36	-1.98	**-1.90**	+0.38	+0.08	+.46	A
(6–23)		Control	-2.49	-1.89	**-2.07**	+0.60	-0.18	+.42	
		P-C	+.13ns	-0.08	**0.17 ***	-0.22	**+0.26***	+.04	
	WAZ	Project	-2.30	**-2.04**	-1.87	+0.26	+0.17	+.43	
		Control	-2.24	**-2.14**	-1.94	+0.10	+0.20	+.30	
		P-C	-0.06 ns	**+0.07***	+0.07 ns	+0.16	-0.03	+.13	
	WHZ	Project	-0.77	**-0.98**	**-0.84**	-0.20	+0.14	-.07	
		Control	-0.73	**-1.19**	**-0.76**	-0.46	+0.43	-.03	
		P-C	-0.04 ns	**+0.21***	**-0.08***	+0.26	-0.29	**-.04 ***	
1B. Endline	%HAZ<-3	Project	34.0	20.4	**15.8**	-13.6	-4.6	-18.2	b, c
(6–23)		Control	31.4	18.5	**22.3**	-12.9	+3.8	-9.1	
		P-C	2.6	1.9	**-6.5***	-0.7	**-8.4***	**-9.1***	
	%WAZ<-3	Project	30.6	17.7	11.8	-12.9	5.9	-18.8	
		Control	27.3	18.9	13.3	-8.4	6.6	-14.0	
		P-C	3.3	-1.2	-1.5	-4.5	+0.7	-4.8	
	%WHZ<-2	Project	22.1	**15.3**	**11.5**	-6.8	-3.8	-10.6	
		Control	21.3	**18.9**	**8.7**	-2.4	-10.2	-12.6	
		P-C	0.8	**-3.6 ***	**2.8 ***	-4.4	+6.4	**+2.0 ***	
						HAZ	***WAZ***	***WHZ***	D
2. SCF	Z-score	Project				-1.73	-1.87	-1.04	
(0–23)	Z-score	Control				-1.78	-1.92	-1.06	
	Z-score	P-C				+.05 ns	+.05 ns	+.03 ns	
3A. NNP	Z-score	BINP				-1.5 ns	-1.6 ns	-0.8 ns	E
(0–23)	Z-score	NNP				-1.6 ns	-1.7 ns	-0.8 ns	
	Z-score	Control				-1.5 ns	-1.6 ns	-0.8 ns	
3B. NNP	Z-score	BINP				**-1.9 ***	-2.1 ns	**-1.2 ***	F
(24–60)	Z-score	NNP				**-2.1**	-2.1 ns	**-1.1**	
	Z-score	Control				**-2.0**	-2.1 ns	**-1.2**	
						HA<-3	***WA<-3***	***WH<–3***	G
3C. NNP	%	BINP				**17.3**	12.9	0.8	
(24–60)	%	NNP				**22.1**	14.2	0.7	
	%	Control				**18.9**	14.9	0.5	
3D. NNP	%	BINP				**15.4**	**9.6**	0.5	H
(36–47)	%	NNP				**24.9 ***	**13.9 ****	1.3	
	%	Control				**20.5 ***	**13.2**	0.7	
4. IMED	%	Project				7.7	14.0	12.2	I
(0–59)	%	Control				7.4	14.0	12.4	
	%	P-C				0.3?	0	-0.2 ?	
5A. OED-Endline	HAZ	P-C		**+0.08***	+.01 ns				J
(6–23)	WAZ	P-C		**+0.14***	+.01 ns				
	WHZ	P-C		**+0.16***	+.01 ns				
5B. OED-DHS	HAZ	P-C			**+.27***				K
(6–59)	WAZ	P-C			**+.24***				
	WHZ	P-C			**+.15***				

Note*:* "ns" - Not Significant; "?" - Significance unknown. Shaded cells represent a double difference (P-C at $t = 2$ vs. P-C at $t = 1$); bolded numbers indicate a significant difference between project and control area, or a significant double-difference. For technical footnotes to this table see appendix A.

(c). In principle, analyses 1A and 1B should have the strongest designs for testing project effects by virtue of data being available at baseline, midterm, and endline for project and nonproject areas alike. In practice, interpretation of these results is highly problematic because the control data are based on only two *Thanas*, both of which have health and nutrition interventions under way that are broadly similar, and the endline survey was conducted in a different season (Feb/Mar) than the baseline and midterm surveys (Sep/Oct). With these complications in mind, the double-difference results suggest greater reduction in severe stunting in the project area (HAZ<-3), but a greater reduction in moderate and severe wasting in the nonproject area (WHZ<-2). Boxes 2 and 3 address the high level of uncertainty associated with these findings.

Knowledge and Practices Concerning Child Care and Feeding

2.9 The community-based component of BINP sought to improve knowledge and practices through monthly growth monitoring (GM) sessions and nutrition counseling and discussion sessions. The endline report indicates that at midterm 85 percent of women participated regularly in child GM sessions and another 9 percent participated occasionally. At endline the corresponding figures were 76 percent and 18 percent, respectively. Participation in group nutrition discussions outside the GM sessions was reported by 66 percent of mothers. BINP endline report found little or no difference in GM participation patterns across seven sociodemographic variables tested and no differences by child nutritional status, consistent with the intention to cover all children with GM. The OED reanalysis of these data using multivariate models reveals that participation in GM at midterm is somewhat less likely among younger mothers, daughters-in-law (in two *Thanas* only), those with lower incomes, and those few in the very highest education category. However, with the exception of the daughter-in-law effect in two *Thanas* (which reduces the probability of participation from .88 to .65), the practical importance of this is not clear, given the generally modest coefficients and the very high levels of participation observed overall.

2.10 Against this background of participation, data on knowledge and/or practices are available from all reports reviewed above and, thus, are subject to the same design limitations. When all sources of data are considered, the evidence concerning changes in knowledge and practices does not suggest a consistent set of results (table 2.2). The Implementation Monitoring and Evaluation Division, Ministry of Planning (IMED), and SCF surveys report knowledge levels in the project areas that are 10 to 20 points higher than in nonproject areas. However, the NNP baseline survey, based on much broader sampling design than the IMED and SCF surveys, suggests these differences are of the order of 1 to 8 points and some of them may not reach statistical significance. The endline report did not include any analysis of child care/feeding knowledge. The OED report contains a reanalysis of the SCF data on colostrum and breast-feeding, controlling for potential confounders. In those analyses there are statistically significant effects of living in BINP areas, even after controlling for confounders, and significant effects related to participation in nutrition education activities. The order of magnitude remains in the range of 10 to 20 points between project and nonproject areas. That is the strongest evidence for an effect of the project on knowledge, but it is based on the survey with the most restricted geographic coverage (three project and three control *Thanas*) and an endline cross-sectional design. Of the two surveys with broader geographic coverage, one agrees with this finding (IMED) and the other does not (NNP baseline). It is unclear whether this discrepancy is due to differences in the way questions were asked, the sampling, the timing of the surveys, or other factors, Since the NNP survey was conducted at a time when BINP had ended and NNP had not yet started.

2.11 With regard to practices, the results also vary according to the survey and the specific practice in question. The *feeding of colostrum* is widely practiced in all areas in the endline surveys, but the difference between project and nonproject areas does vary from 4 to 14 points across the surveys. The data on exclusive breast-feeding indicate that *initiation immediately after birth* is about 30 points higher in project areas at midterm and endline (based only on BINP endline report), but *continuation of exclusive breast-feeding (EBFD) for the recommended five to six months* is low in all areas and not significantly different between project and nonproject areas (based only on the SCF and NNP baseline reports).[4] The *timing of complementary feeding* is not different in project and nonproject areas in BINP baseline, BINP endline, and NNP baseline surveys, but the SCF survey suggests proper timing is about 8 points higher in project than in nonproject areas. The IMED survey suggests a higher prevalence of eating more food during lactation, but that is not confirmed in the NNP baseline.

2.12 In summary, the most significant improvements in young child nutrition would be expected under conditions of immediate initiation of breast-feeding at birth (including use of colostrum); continuation of exclusive breast feeding (EBFD) for 5 to 6 months; appropriate timing and quality of complementary feeding before, during, and after illness; and continued breast-feeding for at least 24 months. The available evidence provides only a few snapshots in this larger picture and does not offer a consistent body of evidence on differences between project and nonproject areas. The most impressive results relate to practices immediately after birth: nearly universal feeding of colostrum in project and nonproject areas alike and a 30 point greater prevalence in immediate initiation of breast-feeding in the project areas. However, continuation of EBFD for 5 to 6 months is low in all areas (4%–13%) and proper timing of complementary feeding is reported by only about 50 percent of mothers, with little or no difference between project and nonproject areas. The feeding of colostrum and immediate initiation of EBFD is supported by evidence that knowledge of the benefits of these practices is 10 to 20 points higher in project areas.

[4] Note that the WHO recommendation for exclusive breast-feeding (EBFD) for five to six months was not clarified until the BINP was completed (2002). The message promoted in the BINP may have been three to four months, but project documents available for this review are not clear on that point. Data from the NNP baseline survey suggest that roughly 40 % of mothers introduced some food other than breast milk by three months (with no differences among NNP, BINP, and control areas). Corresponding figures at four and five months of age are 50 % and 60 % to 70 %, respectively. Thus, it does not appear from these data that BINP influenced the duration of EBFD, even if a shorter duration is considered.

Table 2.2 Knowledge and Practices Concerning Child Feeding

Indicator	BINP Endline					OED (0/6–23mo)	IMED (U5)	SCF (6–23 mo)	NNP baseline (0–23mo)	
	BINP-control baseline	BINP-control midterm	BINP-control endline	Baseline -endline	Midterm -endline	BINP-control (SCF data)	BINP-control	BINP-control	BINP-NNP	BINP-control
1. Was aware of the benefits of colostrum						*	88/69 *	63/53 *	93/87 *	93/91
2. Fed child colostrum	66/78 *	98/86 *	96/92*	*	(C) *	ns	91/77 *	77/73 *	96/91 *	96/95
3. Was aware of EBFD						*	87/71 *	78/69 *		
4. Initiated BFD immediately after birth	15/12 * ^^	61/31* ^^	84/56* ^^	*	ns				34/28 *	34/33
5. Practiced EBFD to 5 to 6 months						ns		4/5 ns	8/12 *	8/13 *
6. Aware of complementary foods at 5–6 months							81/60 *	64/64 ns	95/87 *	95/94?
7. Gave complementary foods at 5 to 6 mo	27/22 ns		55/53 ns	ns				57/49 *	43/38 *	43/42
8. Ate more food during lactation							61/43 *		26/27	26/26

Note*:* Shaded rows contain knowledge variables; unshaded rows contain self-reported practices

* Difference is statistically significant

ns Difference is not significant

? Significance of difference not tested

(C) Difference favored control area

^^ Baseline asked an open-ended question; midterm and endline asked close-ended (immediately, within 24 hrs, after 24 hrs, never)

Supplementary Feeding for Children

2.13 Eligibility for supplementary feeding (SF) in BINP was determined through community-based growth monitoring. As noted, reported participation in GM sessions was very high, with 94 percent participating occasionally or regularly. There are no data available for estimating the coverage and targeting of SF from a population perspective, in part because a child may be eligible because of a severely low weight-for-age (WA) (<60% of reference) or low weight gain over a given period.[5] In addition, there are no untreated comparisons for estimating the impact of the SF program in BINP areas.

2.14 In an effort to partially assess the targeting and subsequent progress of children receiving SF, the SCF report examined the administrative records for 998 children (6–23 months) who had reportedly received this input. The OED reanalysis of those data revealed that 5.4 percent of children receiving SF had WAZ < -4 on entry and another 59.5 percent had WAZ -4 to -2 on entry. Insofar as these are similar to the prevalence's in the project areas as a whole, this implies either poor targeting or that most children are being enrolled based on the low weight gain criterion. The SCF analysis further revealed that, of those severely malnourished on entry, 58 percent had improved their WA status by the end of the SF period. Of those moderately malnourished on entry, 28 percent had improved by the end of the SF period. Improvements were seen in 12- to 23-month-old children but not the 6- to 11-month-olds. These differential improvements in older children and more malnourished children were confirmed in a multivariate reanalysis in the OED report. However, both sets of analyses are limited by the lack of a control group and the threat of regression to the mean, especially if most children are screened into the SF program based on changes in weight for age rather than absolute levels.

2.15 Another source of information for examining coverage and targeting is the study by Nahar (2003) in one BINP *Thana*. In this study, which linked prospective measurements of infants with community nutrition center (CNC) records, it was estimated that 70 percent of infants should have been eligible for SF (based on absolute WA and/or the weight gain criteria) but only 30 percent of these actually received SF. Of the 30 percent not eligible, 13 percent did receive SF. Of the supplemented infants, 27 percent received SF for the full 90 days, 29 percent received SF for 61 to 89 days, and 44 percent received SF for 60 days or less. In addition to the Nahar study in one *Thana*, the SCF cross-sectional survey in three *Thanas* reported that only 22 percent of severely malnourished children were receiving SF. It is unclear how representative these findings from limited geographic areas may be in regard to the overall BINP areas. However, it is useful to mention that even in some of the best-designed and implemented programs such as PROGRESA in Mexico; similar issues of targeting and inclusion/exclusion continue to arise (Skoufias 2005). These imprecision's in targeting are partly a reflection of the sensitivity and specificity of the indicators used for selection of children, an partly that of the realities and imperfections of implementation in large scale programs.

[5] Normal weight gain in children is defined as <600 grams in 2 months for 6 to 12-months-olds or <300 grams in 2 months for 12- to 23-month-olds; and for children with I or II degree malnutrition, <600 grams in 3 months for 6- to 12-month-olds or <300 grams in 3 months for 12 to 23-month-olds.

Maternal Weight, Pregnancy Weight Gain, and Birth Weight

2.17 BINP sought to reduce the chronic energy deficiency of women, improve their weight gain during pregnancy, and lower the incidence of low birth weight. There are three sources of data related to these outcomes: BINP endline report, the OED reanalysis of that report, and the NNP baseline survey, with summary results shown in table 2.3.

2.18 These three sources provide little or no evidence for improved outcomes in BINP areas. Mean weights and BMIs improved significantly from midterm to endline, as did the prevalence of low weights and low BMIs, but that improvement occurred equally in project and nonproject areas. Mean monthly weight gains appear higher in project areas at endline, but a derived estimate of total weight gain and the prevalence of low weight gains suggest a difference in the opposite direction, probably the result of the indirect calculation methods applied. A multivariate analysis of these same data in the OED report indicates, at most, a 1 kg greater total weight gain during pregnancy in the project areas, which is small but more consistent with the monthly weight gain estimates.[6] There are no significant differences in mean birth weight or the incidence of low birth weight at midterm, endline, or in the "double difference." Finally, the NNP baseline survey, with much broader geographic coverage in both BINP and the non-BINP areas, suggests that the outcomes in BINP areas are very similar to those in NNP and control areas; the one suggestive difference is in the incidence of low birth weight, which is lower in BINP than in new NNP baseline areas, but lowest of all in the control areas.

2.19 Although the findings above suggest that the project has had no overall effect on maternal weight gain or birth weight, it is important to note that the ability to produce such an effect depends on the ratio of women physiologically capable of responding to the interventions versus those not capable, in the particular population in question (as described in section 1.3.1). When the intervention is food (whether via supplements or changed eating practices) and the outcome is birth weight, the biological partitioning of energy between the mother and the fetus is a major determinant of responsiveness in birth weight. Results from animal studies and evaluations of supplementary feeding interventions suggest that energy is preferentially partitioned to *the mother* when she is severely malnourished and greater proportions are partitioned to the fetus when the mother is moderately malnourished or well-nourished (Fellows 1985; Winkvist et al. 1992; Olson 1994). This suggests that food-related interventions in populations with moderate maternal malnutrition may improve birth weight without necessarily being mediated via maternal

[6] The significance of a 1 kilogram increment in weight gain depends on the perspective taken. In numerical terms it represents a 12 % increase over the usual gain of about 8 kilograms, which is not inconsequential. In regard to its potential impact on birth weight, as discussed in the OED report, it would translate into an increment of only 20 grams in birth weight, which is only a 0.7 % increase over the usual birth weight of about 2,700 grams. However, when the implications of maternal partitioning are taken into account, the 270 gram birth weight increment detected in the OED multivariate analysis (for undernourished women) represents a 10 % increase in usual birth weight and a significant reduction in the incidence of low birth weight. (See next paragraph in text concerning maternal partitioning).

weight gain during pregnancy, contrary to the theory guiding the design and evaluation of BINP. This is partially indicated in the results of the OED multivariate analysis (table 2.3), which show that certain subgroups do have higher birth weights, despite there being no significant improvement in weight gain during pregnancy in the sample as a whole.

Specifically, there are impressive and statistically significant effects for those women reported eating more during pregnancy (+88 g), especially if the latter are destitute and presumed to have poorer nutritional status (+270 g). [7] This has implications for both the design of future nutrition programs and their evaluation, as discussed in a later section.

Table 2.3. Maternal Weight, Weight Gain, and Birth Weight

	BINP Endline Data			
	Midterm		Endline	
Mothers of U2 children	BINP	Non-BINP	BINP	Non-BINP
Height, mean, cm	149.0	149.1	151.2	150.5
Height <145 cm [b, c]	22.3%	21.7%	11.4%	13.7%
Weight, mean, kg [b, c]	39.6	39.7	43.1	42.7
BMI, mean	17.8	17.9	18.8	18.9
BMI < 18.5 [a]	68.8%	66.7%	48.5%	46.8%
Pregnant Women				
Mean monthly weight gain, kg	1.2	1.1	1.2	1.0
Mean total weight gain, kg (sig. not tested)			8.2	9.1
Proportion of pregnant women with Wt gain < 6 kg (sig not tested)			43.0%	19.5%
Wt gain < 7 kg (sig not tested)			54.7%	36.6%
Birth weight OED				
Bwt < 2.5 kg	25.6%	29.9%	15.9%	17.6%
Bwt, mean, g (OED)	2,677	2,712	2,770	2,800
Difference, g (OED)	34.3		30.9	
Double diff, g (OED)	3.5			
	Significant subgroup effects			
Impact on birth weight for women "eating more"	44 g		88 g	
"Eating less"	-		-44 g	
"Eating more" among destitute women (presumed w/ poorer nutritional status)	-		270 g	
	NNP Baseline Survey (significance not tested)			
	NNP Areas	BINP Areas	Control Areas	
Mean weight at 5 mos gestation, kg	47.3	47.8	49.1	
Wt gain from mos 5–9, kg	3.59	3.55	3.98	
Birth weight < 2.5 kg	22.3%	16.2%	12.9%	
Birth weight, mean, g	2,780	2,750	2,780	

* OED = results are taken from the OED multivariate reanalysis of data in BINP Endline Report.
a. BINP-control diff sigt at midterm; b. BINP-control diff sigt at endline; c. Double-diff sig;

[7] These findings are presented only in the text of the OED report (with no corresponding tables) as follows: "In addition, the effect is larger for women of poorer nutritional status. The interaction of the 'eating more' dummy with the income level of the household of residence (result not shown) produces a much larger coefficient (+270 grams) for the 'destitute' women." This result deserves more detailed analysis and presentation, informed by the partitioning theory described in section 1.3.1.

Knowledge and Practices Related to Pregnancy

2.20 Information about pregnancy-related knowledge and practices is available from four separate surveys described in the endline report, IMED, SCF, and NNP baseline. Results are summarized in tables 2.4 and 2.5.

2.21 This domain represents the most consistent body of evidence for a positive effect of BINP. All four surveys show that most indicators of knowledge and practices are higher at endline in the project than nonproject areas and that these differences are statistically significant (when tested). For many knowledge and practice measures the differences are substantial, on the order of 20 to 40 percentage points, although these differences are smaller in the NNP baseline survey. Multivariate reanalysis of the SCF data and BINP endline data in the OED report indicates that the knowledge differences related to food intake, rest, and avoidance of hard work during pregnancy remain highly statistically significant after controlling for potential confounders.

Table 2.4. Knowledge and Practices Related to Pregnancy: BINP Endline Report

	Baseline		Midterm		Endline	
	Project	**Control**	**Project**	**Control**	**Project**	**Control**
Eating more is beneficial (%)						
Mothers of U2 children	64	65			92	68
Currently pregnant	-	-			95	65
Recent pregnancy	-	-			-	-
Ate more in last pregnancy (%)						
Mothers of U2 children	57	80	56	22	59	29
Currently pregnant	-	-	57	84	-	25
Recent pregnancy	-	-	-	-	-	-
More rest is beneficial (%)						
Mothers of U2 children					89	58
Currently pregnant					90	64
Recent pregnancy					-	-
Rested more last pregnancy (%)						
Mothers of U2 children	14	24	65	35	67	46
Currently pregnant	-	-	23	10	71	45
Recent pregnancy	-	-	26	10	50	39
Hard work is harmful (%)						
Mothers of U2 children					94	83
Currently pregnant					95	92
Recent pregnancy					-	-
Never did hard work last preg. (%)						
Mothers of U2 children			61	82	53	44
Currently pregnant			60	81	61	51
Recent pregnancy			-	-	49	61

Notes: Shaded areas represent knowledge; unshaded areas represent practices

Table 2.5. Maternal Knowledge and Practices: NNP Baseline, IMED, and SCF Surveys

	NNP Area	BINP Area	Control Area
Attended at least one antenatal care (ANC) (%)	46	**60**	48
NNP mothers of U2 children	40	**48**	39
NNP pregnant women	-	**44**	27*
IMED total sample	-	**41**	20*
IMED poor sample			
Attended 3+ ANC (%)			
NNP mothers of U2 children	41	**48**	38
NNP pregnant women	29	**34**	25
SCF Mothers of U5 children	-	**42**	10 *
Iron tablets taken during last preg. (%)			
NNP mothers of U2 children	36	**50**	38
NNP pregnant women	27	**36**	26
IMED total sample	-	**26**	13*
IMED poor sample	-	**24**	10*
SCF Mothers of U5 children	-	**58**	24*
Tetanus toxoid received during last pregnancy (%)			
NNP mothers of U2 children	82	86	88
NNP Pregnant women	67	66	67
Vit A capsule taken after last preg. (%)			
NNP mothers of U2 children	8.3	**7.9**	5.2
NNP pregnant women	-	**-**	-
Eating more is beneficial (%)			
NNP mothers of U2 children	69	**80**	72
NNP pregnant women	74	**82**	77
Ate more last pregnancy (%)			
NNP mothers of U2 children	15	20	20
NNP pregnant women	22	**25**	21
IMED total sample	-	**44**	23*
IMED poor sample	-	**42**	19*
SCF mothers of U5 children	-	34	30*
Rest is beneficial (%)			
NNP mothers of U2 children	53	**69**	58
NNP pregnant women	56	**68**	60
SCF Mothers of U5 children		**78**	70*
Rested more last pregnancy (%)			
NNP mothers of U2 children	40	45	45
NNP pregnant women	46	55	55
IMED total sample	-	**42**	34*
IMED poor sample	-	**38**	29*
SCF mothers of U5 children	-	**60**	54*

Note: *Shaded rows represent knowledge; unshaded rows represent practices.*

** Indicates statistically significant difference. Statistical significance of group differences was not tested in the NNP; bolded numbers indicate suggestive differences but statistical tests were not reported.*

Supplementary Feeding of Pregnant Women

2.22 The coverage and targeting of SF for pregnant women can be more reliably assessed than that of children because a single eligibility criterion was employed, that being BMI <18.5 as early as possible in pregnancy. OED's reanalysis of the data from the endline report indicates that about 75 % of women participated in the weighing sessions (intended for all pregnant women) and 49 % (at midterm) and 45 % (at endline) participated in SF. At midterm, 39 % of women with normal BMI (>18.5) received SF, compared with 49 % of those with mild BMI deficits (17–18.5), 56 % of those with moderate BMI deficits (16–17), and 59 % of those with severe deficits (<16). By endline the prevalence of low BMI among mothers of U2 children had decreased significantly in the project areas (48% versus the 69% at midterm) but the coverage of the low BMI women had improved further, with corresponding figures of 34 %, 51 %, 60 %, and 69 %, respectively. Overall, it appears that 50 % to 60 % of women with BMI < 18.5 received SF and about 40 % of ineligible women (BMI > 18.5) also received SF.[8]

2.23 The prospective study in one *Thana* by Nahar (2003) also found that 63 percent of BMI-eligible women did receive SF, of which 53 percent were fully supplemented once enrolled (at least 95% of sachets) and 50 percent were only partially supplemented (less than 95% of sachets). That study reported a much lower rate of errors of inclusion (4%).

2.24 The OED reanalysis of data from BINP Endline Report was not able to demonstrate any effect of maternal SF on birth weight and did not attempt to demonstrate an effect on maternal weight gain because of the strong selection bias inherent in such analysis. The study in one *Thana* by Nahar (2003) found no difference in the incidence of low birth weight between SF and non-SF mothers with BMI < 18.5. As noted above, however, the partitioning of energy between mother and fetus varies according to maternal nutritional status, such that the effects of SF on birth weight may not be evident in the sample as a whole. It is not clear whether or how the OED analysis or the Nahar study tested the effects of SF among different subgroups of women.[9]

Micronutrients

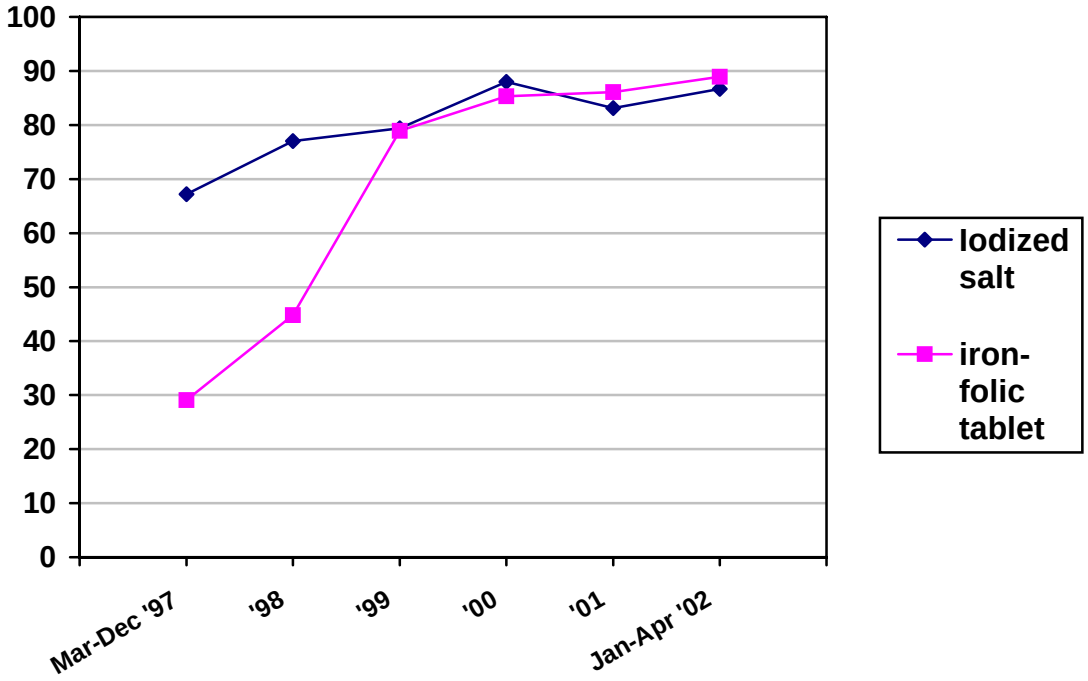
2.25 One of the objectives of BINP is to virtually eliminate micronutrient deficiencies of vitamin A, iron, and iodine. Because biological measurements were not taken at the baseline, various evaluations used coverage of iodized salt, postpartum vitamin A supplements, and iron-folic acid tablets to estimate the project's performance on these micronutrient objectives. Figure 2.1 shows the increasing trends on the coverage of IFTs among pregnant women and percentage of households using iodized salt in BINP areas during the period of the project (from one source). Table 2.6 provides information on children's vitamin A supplementation and that of women's vitamin A and iron intake from other sources. The increasing trends in coverage of iodized salt, vitamin A supplements among both postpartum women and U2 children, and iron-folic acid

[8] Mistargeting can occur through errors of inclusion or exclusion. Several of the reports express great concern for the errors of inclusion (giving SF to ineligible women), which is a legitimate concern from the perspective of overall cost and equity. However, the more relevant concern from an impact perspective is that a high fraction of eligible women be covered (i.e., minimizing errors of exclusion). From a strictly technical perspective, a high degree of mistargeting to the ineligible can be tolerated so long as the coverage of the eligible is high.

[9] The only information from the Nahar study available in time for the present review is a dissertation abstract and PowerPoint presentation.

tablets suggest that BINP did succeed in improving the delivery of these inputs in BINP areas.[10] There are no data available in the reports reviewed here to assess whether this has led to corresponding improvements in iodine, iron, or vitamin A nutritional status. Such data have been included in the NNP baseline survey for iodine and iron and will be available for future evaluations.

Figure 2.1 Coverage of Iron-Folic Acid Tablets and Iodized Salt Use Increased in First-Phase BINP Thanas during Project Years



Data source*:* IMED

Maternal Health Care–Seeking Behaviors

2.26 Similar to the case of micronutrients, there were improvements in health care–seeking behaviors among women of young children in their most recent pregnancy, as demonstrated especially in antenatal care (ANC) and tetanus toxoid (TT) injections (table 2.6). The mean month of pregnancy when the initial ANC check-up occurred is the fifth month (NNP table 22), such that the mean number of check-ups remains low. Most women (60%–70%) received no more than two TT injections based on NNP data. Beyond these particular findings, however, local knowledge about the situation in BINP areas suggests that CNCs were providing not only nutrition services (GMP, counseling, etc.) but also some ANC and distribution of IFTs and VACs. This "one-stop shopping" for maternal and child nutrition and health services is a significant achievement that should be carried over into the design and implementation of future programs.

[10] In each case the differences between project and nonproject areas are greater in the endline evaluation than in the NNP baseline survey, possibly because the latter survey had broader geographic sampling and took place two years after the end of the BINP. However, the trends reflected in the endline evaluation provide fairly convincing evidence for an effect of the BINP in improving the delivery of these inputs.

Table 2.6. Health Care Practices of Mothers of U2 Children During Pregnancies with That Child and Practices in Relation to Children 12 to 23 Months Old

INDICATOR	BASELINE (1995)		MIDTERM (1998)		ENDLINE (2003)		NNP [1]		
	Project (*N*=535)	**Control** (*N*=212)	**Project** (*N*=4465)	**Control** (*N*=1502)	**Project** (*N*=3729)	**Control** (*N*=1238)	**BINP** (2086)	**NNP** (5026)	**Control** (1707)
1. Mean no. of antenatal checkups during last pregnancy (SD) [b, c, e]	-	-	1.4(1.9)	0.6 (1.3)	**3.6(2.7)**	**1.4 (1.7)**	-	-	-
2. Attended at least one antenatal checkup during last pregnancy [b, c, d]	11.8%	13.7%	48.6%	24.4%	**81.0%**	**55.1%**	60%*	45.6%	48.3%
3. **Iron tablets** taken during last pregnancy [b, c, d, e]	16.6%	21.7%	57.3%	28.9%	**83.9%**	**41.8%**	50.3%*	36.5%	38.3%
4. Mean no. of **iron tablets** taken per month during last pregnancy (SD) [c]	-	-	-	-	**23.3(16.5)**	**3.5 (7.4)**	-	-	-
5. TT vaccinations received during last pregnancy [a, b, c]	81.8%	74.8%	91.4%	84.6%	**93.5%**	**87.7%**	86.3%	81.8%	88.3%
6. **Postpartum VAC** taken after last pregnancy [b, c]	-	-	56.3%	10.0%	**65.0%**	**13.0%**	7.9%	8.3%	*5.2% vs. NNP
7. VAC given to children 12 to 23 months old, in the past 6 months—%/(number of children)	39% (82)	42% (48)	84.1% (2,121)	72.3% (675)	**90.6% (1,732)**	**61% (590)**	65.2%	64.1%	72.8%

Note*:* Shaded row contains information for children 11 to 23 months of age.

1. Significance in differences not reported in the original report but provided by Dr. Tahmeed Ahmed during the peer review of the first draft of this study

a. Indicates a statistically significant difference/association ($P < .05$) between project and control areas at the time of the baseline study according to a Pearson's chi square test or an independent sample *t*-test as appropriate

b. Indicates a statistically significant difference/association ($P < .05$) between project and control areas at the time of the midterm evaluation (without controlling for baseline differences) according to a Pearson's chi square test or an independent sample *t*-test as appropriate

c. Indicates a statistically significant difference/association ($P < .05$) between project and control areas at the time of the endline evaluation (without controlling for baseline differences) according to a Pearson's chi square test or an independent sample *t*-test as appropriate

d. Indicates a statistically significant difference ($P < .05$) in odds ratios or means from baseline to endline between project and control areas according to a Breslow-Day test or a two-by-two analysis of variance (ANOVA) as appropriate

e. Indicates a statistically significant difference ($P < .05$) in odds ratios or means from midterm to endline between project and control areas according to a Breslow-Day test or a two-by-two ANOVA as appropriate

C. GARDEN AND POULTRY PROJECT COMPONENT

2.27 The main objectives of the garden and poultry project under BINP project are as follows:

- Improve household food security by increasing homestead food production
- Increase per capita consumption and nutritional status of children and women
- Increase income of the economically vulnerable households through women's participation in the project.

The project was designed to provide the following services to households and communities:

- Garden project: seeds and seedlings and privately run community nursery to meet the demands of more households that seek to benefit from participating in the project
- Poultry project: hybrid birds that have high egg yield and demand timely inputs. A system of eight workers was developed, primarily to help with the provision of microcredit
- as well as poultry immunization and marketing.

Evidence Available for the Review

2.28 No final evaluation of this component was available except for an operations research (OR) on its process and impact (Karim et al. 2002) as almost the sole reference. A few evaluations and documents that discussed BINP as a whole provided some additional information (IMED, Kabir and Levinson).

2.29 The OR study was carried out to compare three types of households: (1) BINP households receiving garden and poultry interventions (in addition to CBNC), (2) BINP households without garden and poultry interventions (with only CBNC), and (3) households outside BINP areas receiving garden and poultry interventions. Five phase I through III BINP *Thanas* and seven non-BINP *Thanas* were selected for the study. The study intended to assess the (1) targeting and coverage of the intervention, (2) implementation of project by the recipients and project staff, (3) production of eggs or vegetables and fruits, (4) change in household income or food consumption in children and women of reproductive age, and (5) change in nutritional status of these two population groups.

Observed Results

2.30 Because of several design and sampling weaknesses in data collection and processing, the results presented in the OR study had only very limited meaning and utility with regard to answering most of its stated study objectives (box 4). Thus a decision was made not to summarize the results in this study. However, this study and other reports (IMED) researched the implementation issues of the garden and poultry project, and they are presented here.

- *Coverage and targeting*. The goal of covering 10 beneficiaries in each CNC was not achieved. And a large proportion of participants were not eligible (36% and 44% for garden and poultry, respectively).

- *High dropout rate*. Dropout rates among BINP participants were high within a year of implementation at a level of 40 % in the garden project and 50 % to 60% in the poultry project (Karim et al. 2002). The IMED evaluation reported a nearly complete dropout from the garden project and found about half of their respondents were still raising poultry at the time of their study (IMED 2004).

Box 2.4 Design Flaws with the Garden and Poultry Study

1. Study samples were not selected following a random sampling procedure. Instead, they were sought to meet predetermined sample size requirement and are not truly representative of either the BINP or non-BINP area. In fact, the original selected Upazilas had to be changed because of the "great difficulty finding project participants."

2. Throughout the study, the presentation of data was not stratified by those people who were still keeping garden or poultry and those who had dropped out. The study reported that 74 % of the BINP and 96 % of non-BINP samples in the study were not running a garden at the time this study was conducted. And the %ages for samples that had dropped out from poultry raising were 44 % and 70 % in BINP and non-BINP areas, respectively.

3. Of those who were counted as "still participating," no data were given on how many of them were already running the gardens or raising poultry BEFORE the BINP intervention began.

- *Inputs.* There were considerable inputs that need to be actually purchased by the participants in the poultry project—feed, vaccine, and medicines—because the hybrid bird selected for the project needed much special care. As a result of these problems, bird morbidity and mortality were high and resulted in sales of the birds to recover losses.

- *Services*. It was reported that the eight-cadre system provided only limited support to participants. More attention of the staff was focused on loan repayment.

- *Implementation*. In BINP *Thanas* where the same contracted NGO was responsible for both CBNC and garden and poultry, the performance of NGOs and participants was reported to be better. The intersectoral projects did not appear to complement the CBNC, as expected in the original design, and did not benefit from the intensive support and monitoring associated with the CBNC.

- *Managerial arrangements.* Field-level staff of government departments in charge of agriculture and livestock performed little monitoring work because of a lack of ownership and primary responsibility for the project.

Lessons from the Garden and Poultry Project

2.31 No conclusions can be drawn about the impact of this component on the nutritional status of women and children owing to the weaknesses in the OR study. Although this component was well intentioned, the series of implementation inefficiencies reported in the OR study and a few other evaluations suggest that it is unlikely for garden and poultry to have any systematic nutritional effect. Special considerations should be devoted to the selection and provision of inputs (seeds, fertilizer, bird species, feed, etc.) and services (vaccination, training, sales channels, microfinancing, etc.) and how to make these available or affordable to potential beneficiaries if they cannot be supplied free of charge. All these factors as well as the resources needed (personnel and materials) for the delivery and supervision activities need careful attention. The level of complexity and sophistication involved in the implementation raises questions about whether it is feasible for such activities to be carried out on a large scale in the Bangladesh context, especially if there is an expectation of achieving a nutritional effect. As reviewed in section 1.4.2, strong education and IEC/BCC components must be in place to increase the chance of having an effective garden and poultry intervention to improve nutrition. This might have been achieved if there had been much greater coordination between the garden and poultry component and the CBNC, but this was not achieved. It appears that some of these issues had been corrected in the design of the NNP, for example, now these two components are managed by the same NGO in the area, instead of by different NGOs (appendix H), but the actual performance during implementation remains to be seen, and many implementation and scaling-up challenges remain.

D. FINDINGS: THE COMMUNITY-BASED NUTRITION COMPONENT

2.32 Taken as a whole, the information reviewed here presents clear evidence for improvement in nutrition-related knowledge, practices, maternal and child nutritional status, and birth weight *in the project and nonproject areas alike*. The most consistent body of evidence on a positive effect of BINP, *beyond that seen in nonproject areas,* relates to knowledge and practices during pregnancy and to the use of available services such as vitamin A and iron supplementation. The evidence also suggests an added effect of BINP on knowledge and practices concerning infant feeding. The evidence presents a mixed picture on any additional effect of BINP on child nutritional status, weight gain during pregnancy, or birth weight, beyond those seen in the nonproject areas. Whether this is due to a true lack of impact on outcomes per se or due to weaknesses in the evaluation designs and data remains unclear. There is some evidence that BINP may have improved birth weight among important subgroups, such as among those who reported eating more during pregnancy (which was one of the key messages of BINP) and, especially, among destitute women (presumably with poorer nutritional status) who reported eating more during pregnancy.

Young Children

2.33 **Coverage and targeting of key child inputs.** The evidence suggests a *high level of participation in growth–monitoring sessions (75%–95%) and a moderate coverage of nutrition education sessions (66%).* The evidence on coverage and targeting of supplementary feeding for children is limited to a prospective study in one *Thana* (where only 21% of eligible children received SF) and the SCF cross-sectional survey in three *Thanas* (where 22% of severely malnourished children were receiving SF). Although the limited geographic coverage of these findings does not provide a firm basis for drawing conclusions about BINP as a whole, they do suggest that the strategies for increasing participation in GMP and nutrition education worked fairly well (with room for further improvement), whereas the coverage and targeting of SF may be problematic and deserving of more extensive examination.

2.34 **Knowledge and practices concerning child feeding.** The weight of *the available evidence suggests that BINP improved two key aspects of knowledge by about 10 to 20 percentage points beyond that seen in nonproject areas: feeding of colostrum—the first milk that protects infants against infections—and the benefits of exclusive breast-feeding.* The evidence further suggests *a large impact of the project on the practice of initiating exclusive breast-feeding* (30 points higher in project areas) but a mixed picture on the practice of initiating complementary feeding, in that there was *no impact on the prevalence of exclusive breast-feeding at five to six months* of age. It appears from this that many more mothers in project areas do initiate exclusive breast-feeding at birth, which will improve the health and nutrition of their babies, but most do not continue this practice until the recommended six months of age. Future program efforts may therefore need to focus on continuing exclusive breast-feeding until six months of age and appropriate breastfeeding plus complementary feeding thereafter.

2.35 **Child nutritional status.** The available analyses provide inconsistent evidence for an effect of BINP on child anthropometry. Some of the analyses with stronger designs suggest a modest impact, but even these have design problems and differing results that prevent firm conclusions. Specifically, on the basis of the weakest of study designs for assessing project impact (endline cross-sectional surveys), three analyses suggest a positive impact of BINP and five analyses suggest no impact beyond that seen in nonproject areas. Two of these analyses employed somewhat stronger designs by controlling for potential confounders between project

and nonproject areas: one suggests a positive impact of BINP and the other does not (unless the analysis is restricted to those children participating more regularly in growth-monitoring sessions). *The weight of the evidence on the effect of BINP on child nutritional status can be read in either direction* though children who participated regularly had better nutritional status. It remains unclear whether this is due to a true lack of impact on outcomes per se or due to the weaknesses in the evaluation designs and the data. The one clear message from this is the need for designing and implementing stronger, technically sound evaluations right at the start when programs are being designed.

Pregnant Women and Birth Weight

2.36 **Supplementary feeding during pregnancy.** Roughly 60 % of malnourished women (with BMI < 18.5) received supplementary feeding, with a clear pattern of higher coverage among those most malnourished (with the lowest BMIs). About 40 % of ineligible women (BMI > 18.5) also received supplementary feeding. On the basis of available evidence, we conclude that receiving supplementary feeding was not associated with improved birth weight in the project areas as a whole and it was not possible from the available studies to directly test whether supplementary feeding had an effect on weight gain. This evolving body of knowledge suggests that the absolute weight of women may be a better predictor than BMI of intrauterine growth retardation (IUGR), which is the major cause of low birth weight in Bangladesh. However, the existing research is not adequate for revealing the best predictor of which women may benefit from supplementary feeding. This is a priority for future research; the MiniMat Study conducted by the International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR,B), should be capable of informing the choice of indicator and cut off points as soon as the data are analyzed and the report written.

2.37 **Knowledge and practices related to pregnancy.** The available studies provide reasonably consistent evidence that BINP improved selected knowledge and practices related to pregnancy by 20 to 40 percentage points beyond that seen in project areas. There is some additional evidence from the study with the strongest design that one of these practices (eating more during pregnancy) is associated with an 88 gram increment in birth weight. This suggests that BINP strategies for improving maternal knowledge and practices related to pregnancy worked well.

2.38 **Pregnancy weight gain and birth weight.** The import of the evidence suggests little or no additional impact of BINP on pregnancy weight gain or birth weight for the population as a whole, beyond that seen in nonproject areas. However, it is noteworthy that subgroup analysis suggests sizable impacts on birth weight among those who eat more during pregnancy (+88 g) and destitute women (presumed to have poorer nutritional status) who eat more during pregnancy (+270 g). The lower-than-expected impact of BINP on maternal weight gain and birth weight may be explained by an evolving body of knowledge on the efficiency of various screening indicators (maternal BMI vs. weight) and the differential partitioning of energy between the mother and the fetus among severely versus moderately malnourished women.

Micronutrients, Health Care, and Garden/Poultry Project

2.39 **Micronutrients and health care.** The available evidence is fairly convincing that BINP improved the delivery and utilization of micronutrients (iron/folate, vitamin A, and iodized salt) to a greater extent than that seen in nonproject areas. Data from most studies showed a 10 to 30 point higher percentage uptake of vitamin A capsules among children under 2 years old and

mothers of young children as well as uptake of available iron/folate supplements among pregnant women. Similarly, the integration of nutrition and health activities at community level in BINP *significantly improved the delivery and utilization of key health services such as antenatal care and tetanus toxoid vaccinations. The key lesson from this is that BINP was successful in encouraging the use of some available health services such as vitamin A supplementation, iron supplementation, and antenatal care*.

2.30 **The garden and poultry project.** *The garden and poultry project experienced a series of implementation problems* and was therefore highly unlikely to have any systematic nutritional impact. These problems included low coverage, poor targeting, high dropout rate, unmet demand for inputs, inadequate services provided by field workers, and managerial arrangements that failed to generate ownership and responsibility. Therefore, the evaluation of the effect of this component is neither relevant nor useful. The key lesson learned from this is the need for greater attention to implementation and better mechanisms for quickly detecting and correcting implementation problems.

The implications of these findings for the expansion and strengthening of the community-based nutrition program in Bangladesh are examined in the final section.

3. OVERALL LESSONS AND IMPLICATIONS

This section draws out lessons and suggestions in three areas: (1) possibilities for intervention redesign; (2) improvement in program impact; and (3) lessons for monitoring and evaluation.

A. POSSIBILITIES FOR INTERVENTION REDESIGN

3.1 Although the present review was not intended to undertake a detailed study of how to redesign BINP's interventions for use in future programs such as the NNP, there is interest in having suggestions from a wide range of sources. Thus, the following suggestions are offered based on an overall consideration of BINP experience as reviewed here. This is done with great reservation because the feasibility of these suggestions and their fit with the local context has not been properly examined. The information required for such a detailed examination is not available for this review and goes well beyond the current terms of reference. Thus, it is expected and hoped that these suggestions might be considered, debated, accepted, modified, or rejected by decision makers in Bangladesh, as appropriate, based on their understanding of the feasibility, costs and contextual issues.

> ***"Was food supplementation the wrong intervention for this context? Or was it a correct intervention but poorly implemented? Or was it a correct implementation but not sufficiently tailored to this context?"***

Supplementary Feeding

3.2 The review of interventions in section 2 of this study reveals that SF as an intervention has been shown to be effective in improving child growth and pregnancy outcomes in numerous studies, yet the available evidence from BINP does not show that it improved growth in BINP children, weight gain in pregnant women, or birth weight in infants. There are numerous potential explanations for this, including mistargeting (especially errors of exclusion); sharing of supplementary food; substitution for other foods; inadequate ration size or incomplete participation; palatability issues; biological partitioning; inadequate feeding practices before, during, and after illness; infection itself; restrictions on women's mobility; and high work and time demands placed on women.[11] All of these explanations, and others, are plausible and have been offered in various evaluation reports. Moreover, there are many layers of factors underlying these immediate explanations related to training, supervision, incentives, cultural norms, and so on. Many of these issues were, or could have been, anticipated when BINP was designed, and corrective or complementary measures could have been designed into the program. Some were, and others were not.

3.3 This complexity raises several important questions. Was supplementation the wrong intervention for this context? Or was it a correct intervention but poorly implemented? Or was it a correct implementation but not sufficiently tailored to this context? These are very different questions that lead to very different policy and program recommendations, which are addressed in this final section. Below are some design changes to be considered in light of these potential explanations:

[11] There also is the nontrivial possibility that the BINP did have an effect on these primary outcomes but the evaluations were not of sufficient quality to detect them. This is discussed in section 3.2.

The following are some supplementary feeding suggestions/options for change:

1. Increase the ration size so that more food will be available to the targeted child or woman even after some of it is shared with others.
2. Improve the palatability of the supplement to increase its consumption.
3. Design highly focused social marketing and IEC strategies to discourage food sharing.
4. Change the eligibility criteria for children to Gomez-based status indicators (i.e., eliminate the weight change criteria) to simplify screening, improve targeting efficiency, and make it possible to verify the eligibility status of individual children. [12]
5. Introduce incentives for early registration and weighing of pregnant women; begin SF as soon as they are registered and deemed eligible; and for very underweight women postpartum, extend the period of SF to meet body weight targets in preparation for the next pregnancy.
6. Introduce special incentives for women to attend postpartum clinics.
7. Related to suggestion number 4, introduce local verification procedures (e.g., review of log books by community nutrition organizers [CNOs] and a village committee) and occasional random remeasurements and rechecks on log books by the CNOs and district supervisors; post these results in a public location; and introduce recognition and other incentives for good performance.
8. Clarify exit criteria, eliminate incentives for retaining women and children on SF beyond these limits, and introduce alternative forms of assistance for those who exit but still show poor weight status.
9. Introduce incentives for individual women to meet pregnancy weight gain and child weight targets, for example, by linking them to eligibility to receive inputs for the garden and poultry project, microcredit, or other programs.
10. Introduce incentives for individual villages to meet weight gain and child weight targets, for example, by linking this to eligibility for community development grants for specified types of projects.
11. Introduce carefully designed recognition and other rewards for the CNPs, CNOs, and higher-level supervisors/managers for meeting targets in their jurisdictions.

3.4 In light of the suggestions above and keeping in mind the cost and human resource implications imposed by the supplementary feeding component, decision makers in Bangladesh may wish to review whether this component provides the benefits commensurate with the costs, and/or how this can be modified to minimize costs and maximize impact.

3.5 In addition to the above, the questions related to screening indicators (BMI vs. weight) and the differential responsiveness among severely versus moderately malnourished women deserve careful consideration. The evidence in this area may not yet be sufficiently firm for making programmatic changes, but studies under way in Bangladesh and elsewhere may help inform decisions on these matters in the near future. These considerations must also be front and center in reviewing and setting goals and targets.

[12] Note that the primary utility of weight-change indicators is among older children (>24 months) because use of static measures would result in SF being given to a large number of stunted children who are unlikely to benefit. Because the main focus of the NNP is younger children (<24 months) the advantages of using a static indicator may outweigh the disadvantages.

Behavior Change Strategies

3.6 Abundant evidence exists that a woman's literacy and nutrition-related knowledge can have a positive impact on the health and nutritional status of her children when certain other conditions are met (e.g., Caldwell 1993; Hobcraft 1993; Thomas et al. 1991). These conditions can be grouped under (1) power, including factors such as time, income, food, decision autonomy, social support, access to services, and the social structures that shape them and (2) beliefs, including cultural norms, values, and expectations and psychological features such as self-efficacy, attitudes, and mental health (Pelto and Backstrand 2003). These factors interact with each other in many ways and at all levels of society to ultimately shape things such as child-feeding practices. "New knowledge" from nutrition education may become incorporated in this complex, but by itself may or may not be adequate to influence practices. If the goal is to change practices, rather than simply to inform and educate, the findings from research suggest that well-designed efforts must be made to address how new information is incorporated in the power and belief complex (Contento 1995) or other aspects of that complex also need to change (Pelto and Backstrand 2003).

> ***"…was nutrition education a well-considered and correct choice for this context, despite going against the prevailing wisdom from theory and research? Was it a wrong choice that happened to benefit from some unknown contextual features that made it successful? Why did it appear to change some practices but not others?"***

3.7 The considerations above would appear to imply that it was futile for BINP to include a rather simple nutrition education component (based on the highly simplified and largely discredited Health Beliefs Model it employed), yet this review finds fairly consistent evidence that BINP improved exclusive breast-feeding and some key pregnancy-related practices by a sizable degree. So, was nutrition education a well-considered and correct choice for this context, despite going against the prevailing wisdom from theory and research? Was it a wrong choice that happened to benefit from some unknown contextual features that made it successful? Why did it appear to change some practices but not others? Again, these are very different questions, which have profound implications for the design of future programs.

3.8 Although BINP appears to have succeeded in improving selected practices (e.g., EBFD initiation and some pregnancy practices), there remains a substantial proportion of the population still not demonstrating these practices. Moreover, much work remains to accomplish some of the more difficult behavior changes, such as extending the duration of EBFD, providing timely and appropriate complementary feeding, reducing work and increasing intake during pregnancy, and attending antenatal care early and frequently. Although specific strategies for accomplishing this cannot be specified here, some of the generic strategies or options related to the behavior change component are as follows:

> ***"…much work remains to accomplish some of the more difficult behavior changes, such as extending the duration of exclusive breast-feeding, providing timely and appropriate complementary feeding, reducing work and increasing intake during pregnancy, and attending antenatal care early and frequently. Some of this may be accomplished by strengthening the link between the increased demands with the supply-side of access to basic health services"***

- Strengthen the counseling skills of community nutrition workers—this should be a specific objective in strengthening the quality of training for workers.
- Ensure a focus in nutrition education sessions on the practicalities of doing or changing specific behaviors, drawing extensively on the experience of women in these groups that have succeeded or have attempted to do so.
- Develop or strengthen strategies to involve "significant others" in the woman's life, including mothers-in-law, husbands, older children, employers, and others with whom she can share experiences and behavior change goals.
- Develop or strengthen strategies for involving other community leaders and organizations as facilitators of behavior change and reinforcers of social norms.
- Develop or strengthen methods for documenting significant obstacles to behavior change (as revealed in number 2 above) and for initiating discussion in appropriate programmatic or community venues for the way these obstacles are to be addressed. This could link to suggestion numbers 8 and 9 under the SF section above.
- Develop or strengthen methods for monitoring progress with behavior change, displaying results in pubic places, and recognizing progress. This also could link to some of the suggestions for SF above.
- Support and reinforce these activities through a variety of IEC strategies.
- Support this demand creation with a supply of adequate-quality HNP services. For example, if women are being encouraged to take iron pills in pregnancy, then they must be available at the health centers; if they are being encouraged to access antenatal care more regularly, then these services should be made more available. A greater integration of health services with nutrition interventions is warranted.

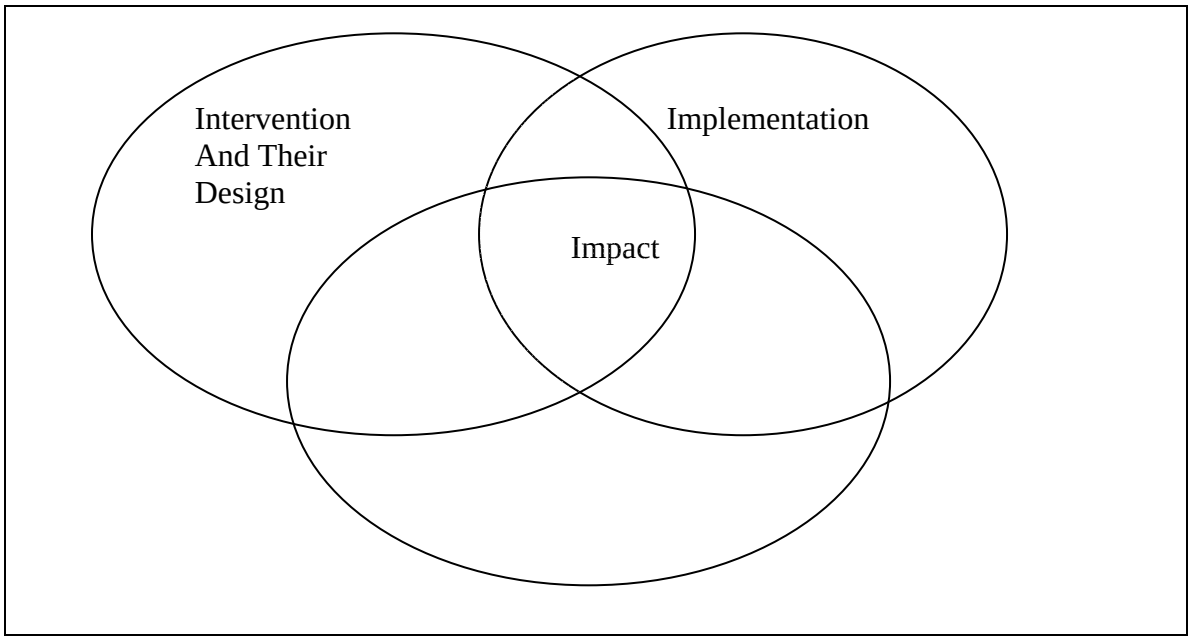
B. IMPROVING PROGRAM IMPACT

3.9 The conventional view of programs as being a set of vertically planned and implemented interventions leads to a basic set of questions during an evaluation:

- Are the chosen interventions the right ones?
- Were they implemented as planned?
- Did they produce the expected changes?
- Which interventions or adaptations are appropriate for the future?

3.10 However, these seemingly simple questions become complex very quickly and reveal the limitations of this conventional view of programs. That is illustrated with the two primary interventions in BINP. The evident lack of impact of SF may be due to faulty implementation, the wrong intervention for the context, the wrong choice of indicators/goals (as was the case for maternal supplementation), or a failure to adapt the intervention and the implementation to the context. However, the apparent positive impact of nutrition education on practices suggests it may have been a well-considered and correct choice (despite the many theoretical reasons to predict otherwise), or it was a wrong choice that happened to benefit from some unknown contextual features that made it successful, or it may have improved the "simpler" practices (initiation of exclusive breast-feeding and feeding of colostrum) but may not suffice for the more difficult ones (e.g., duration of exclusive breast-feeding and quality of complementary feeding).

Figure 3.1 Determinants of Impact



3.11 These examples representing the major BINP interventions reveal that answers to the four evaluation questions posed above depend very much on one's assumptions about the context, the intervention, and how the two are combined in the design and implementation of the program (figure 3.1). One of the overall lessons BINP appears to be offering is that an intervention that has been "shown to be effective" in one setting may not be effective in a new setting (and that applies especially to evidence coming from highly controlled efficacy trials) and, perhaps more surprisingly, interventions that have been declared "ineffective" in some settings may be effective in certain other settings (even if some of the reasons and mechanisms remain unclear).

3.12 The importance of context is illustrated in table 3.1 in relation to the 11 supplementary feeding suggestions offered above. These hypothetical examples demonstrate the varied ways in which context can affect program design and redesign, the fundamental importance of incorporating local knowledge into design and redesign decisions, and the need for continuous assessment and improvement procedures to adjust the program to local contexts over time. In the face of such contextual complexity, efforts to follow the conventional view of programs shown in figure 3.1 should be considered highly suspect.

3.13 If that is so and if the choice and effect of interventions are so highly dependent on context and implementation, what are the implications for future nutrition programs in Bangladesh and beyond?

3.14 First, the fundamental importance of harmonizing interventions, implementation, and context is now widely understood in most development communities. Indeed, it is fair to say that the dominant view in development circles has shifted away from promoting a highly structured and centralized approach, in favor of adaptive management based on more decentralized decisions, incentives, and accountability systems. This is evident in numerous World Bank

publications, among others, most recently World Development Report in 2001 (Attacking Poverty) and 2004 (Making Services Work for Poor People). However, there still appears to be a wide disconnect between the espoused approaches and the approaches adopted or promoted in practice.

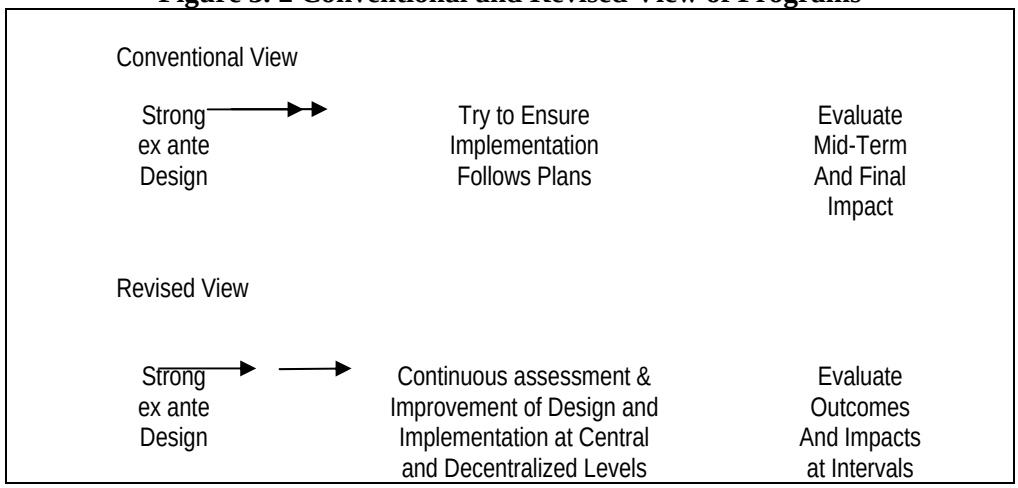
> ***"…the dominant view in development circles has shifted away from promoting a highly structured and centralized approach, in favor of adaptive management based on more decentralized decisions, incentives, and accountability systems."***

3.15 In simple or highly predictable systems harmonization can be accomplished by accurately diagnosing the relevant contextual features ex ante and fashioning the product accordingly (e.g., an assembly line). The conventional view of programs vainly tries to adapt this approach to the complex and dynamic behavior of households, communities, and organizations, with only partial feedback on performance and outcomes gathered by central authorities through evaluations every few years (figure 3.2). The result is a large number of mismatches between interventions and context (e.g., mistargeted services, low participation rates, nonperforming staff, etc.), long feedback loops (e.g., years) between the time mismatches occur and information on these becomes available, and transfer of feedback information to the administrative apex rather than to more decentralized levels where context-relevant corrective actions could be more readily identified, more quickly implemented, and more easily monitored.

Table 3.1 Contextual Factors That May Invalidate the 10 Supplementary Feeding Suggestions

Suggestion	Illustration of Possible Countervailing Contextual Factors (hypothetical)
1. Ration size	Amount fed to children is influenced by their perceived needs, not by the total size of the ration; costs for supplementary food are high
2. Palatability	Reports of poor palatability arise from a minority of mothers and communities; food preparers routinely improve palatability based on local norms
3. IEC to discourage food sharing	Food sharing is fulfilling important social needs especially among the most destitute households, such as reciprocity with relatives or neighbors that help out in lean times
4. Eliminate weight change criterion	This criterion has significant symbolic importance among professionals, senior administrators, and policy makers (because it is believed to be an important feature of the program), and the image and political support of BINP may be severely damaged by eliminating it
5. Supplementary feeding (SF) for women	Controversy over leakage of women's SF is such that efforts to extend the SF to earlier in pregnancy and into the lactation period would further complicate negotiations; costs for supplementary feeding (financial costs and human resource and political costs) are high
6. Attendance	Introduce special incentives for women to attend postpartum clinics.
7. Local verification procedures	CNPs are young and vulnerable members of the community, and efforts to increase accountability and accuracy must be handled with great care
8. Exit criteria	Local power relations are such that stricter criteria will be disproportionately applied to poorer women and children rather than the less needy
9. Women's incentives	Most incentives will go to better-off women unless special assistance is provided to the poorest to help meet targets
10. Village incentives	Incentives will lead to greater falsification of monthly reports and further politicization of the program
11. Recognition and rewards	These may be well received with no unintended consequences

Figure 3. 2 Conventional and Revised View of Programs



3.16 To the credit of BINP designers, BINP expressed a desire to develop a "culture of inquiry" when it was first designed, precisely in recognition of the need to adjust design and implementation to context. Although the present review was not charged with examining this aspect of the program, the information reviewed suggests that this culture of inquiry was itself only partially implemented. For instance, a positive sign is that a large number of changes have been made in the design of the NNP based on BINP experience (see appendix H). In addition, a large number of operations research projects (18 according to one report) reportedly have been undertaken. However, there are concerns related to the quality of some of these research projects and the nonutilization of many of these findings. Moreover, a substantial body of local knowledge, anecdotal information, and conventional wisdom has accumulated on specific implementation problems, but this did not appear to be always translated into programmatic changes during BINP implementation (appendix J). Some reports indicate that there are significant administrative obstacles to utilizing findings from OR, but as yet there does not appear to be a strategy to correct that. As a result, the primary focus of the "culture of inquiry" in BINP appears to have been on the use of midterm and endline evaluations to assess project inputs, outputs, and outcomes at an overall project level. The existence of more decentralized decision making, and especially the authorizing of such decision making, is not apparent from any of the documents reviewed.

"To the BINP designers' credit, the BINP expressed a desire to develop a "culture of inquiry" when it was first designed, precisely in recognition of the need to adjust design and implementation to context."

3.17 Thus, although BINP recognized the need to adjust interventions and implementation to context, there appears to be "a gap between knowledge and practice" on that issue, on the part of the government and its partners. Now that BINP infrastructure has been established and is being expanded nationally, Bangladesh offers an excellent opportunity to bridge this gap and, more fundamentally, to learn what is involved in bridging the gap on a national scale. This will require many changes. The following changes are suggested, with the recognition that each raises a host of complications, concerns, obstacles, and conflicts, all of which are themselves part of the "Bangladesh context." Ultimately, therefore, these require greater elaboration and analysis by the relevant parties in Bangladesh who possess greater contextual knowledge on these matters. However, as these discussions proceed, it is important that the parties resist the temptation to revert to the conventional model shown in figure 3.2, whose limitations as a model for development are by now well understood.

The following are suggestions for improving the effect of programs:

- *Foster a broader understanding* that a culture of inquiry (or adaptive management, quality improvement or other named approaches) is fundamental to program success. This should begin by identifying the key national decision makers and other stakeholders implicated in this notion and developing a systematic set of strategies to explore the issue with them, learning from previous experiences and field visits in Bangladesh and elsewhere; challenging some core assumptions (e.g., local groups are not capable or cannot be trusted; adequate local accountability mechanisms cannot be established); and evaluating various options for proceeding. This should not be approached as a set of once-off advocacy efforts, but a systematic plan for engaging the same set of individuals and organizations in exploring this issue over a sustained period (e.g., one year) with the intention of reaching decisions at the end of that period. Some of the more experienced NGOs in Bangladesh may be given this mandate and the opportunity to innovate linked to the delivery of results. We recognize that such recommendations are very hard to implement in the context of a country such as Bangladesh where capacities are very limited, wherein rules and procedures are held sacrosanct, and where divisive forces have struck many notes of discord during the past few years. In this context, we realize that it would be naïve and unrealistic to expect a sea change; however, we do hope that it may be possible to follow through on some of this agenda in the interests of development and with the full support and encouragement of the development community.
- *Anticipate experimenting with alternative approaches* on an appropriate scale as the design and implementation of the NNP and HNPSP proceed. The decision making discussed in point 1 above will be greatly facilitated if it is only a set of pilot strategies being contemplated in the next five years.
- *Identify some key administrative changes needed in each pilot area*, including but not limited to clearly delineating the scope for discretionary decisions by implementing NGOs as well as at different administrative levels in government, significantly strengthening the adaptive management and inquiry capacity at all levels, and realigning incentive systems and accountability mechanisms to support the approach. The design of these changes should be informed by the best possible experience and expertise with such issues in Bangladesh and beyond Recent experiences in other Bank operations have demonstrated the potential to build program-relevant capacity *simultaneously* with program implementation, which may be a promising approach as BINP expands (Matta et al., 2000).
- *Consider the initiation of such innovative efforts in other areas*, beyond nutrition, for example in innovations for demand-side financing for health to broaden the base of experience and learning and to expand this into a larger theme in national development discourse.

C. LESSONS AND SUGGESTIONS FOR MONITORING, OPERATIONS RESEARCH, AND EVALUATION

3.18 The materials made available for this review represent only a small fraction of the project documents that exist on BINP and the terms of reference did not include reviewing BINP's monitoring and evaluation system. Nonetheless, the available documents are sufficient to demonstrate the central importance of M&E to the success of future nutrition programs and, thus, a few key lessons and suggestions are offered here.

Operations Research and Monitoring

3.19 BINP was successful in contracting for at least 18 operations research projects. It is unclear whether this is an appropriate number or not, relative to the full range of implementation issues needing attention, nor is the quality of these projects clear. It appears that some changes were made at the time of the midterm evaluation. However, *numerous sources refer to difficulties in translating the findings from these projects and from program monitoring into programmatic changes.* This is a common finding in many settings, so BINP and Bangladesh are not unique in this regard. The causes of this problem in other settings also are well known in general terms, such as poor timing, quality, or inappropriate focus in the research; lack of interest, investment, involvement or incentives on the part of program staff; rigid and over centralized decision making; an overall administrative culture based on rules, routines, and procedures rather than continuous inquiry, learning, innovation, and adaptive management; and the lack of human and organizational capacity to implement adaptive management.

3.20 In light of these problems, a systematic plan for improving the design and use of OR and monitoring systems should be developed for the HNPSP. This plan should be developed and implemented by and with the relevant program decision makers and functionaries at various levels, beginning with a participatory or inclusive assessment of views, issues, and constraints on the supply side (research organizations, M&E units, etc.) and the user side (program decision makers at all levels).

Lessons for Evaluation

3.21 The methodological problems associated with the evaluation systems in BINP are well documented in other sources and summarized in appendices B and C, so no attempt is made to reiterate them here. Rather, this section highlights four fundamental issues relevant to the design of future evaluation and the analysis of data from them: differential response, realism of objectives, program theory, and use of findings.

3.22 Differential response. One of the basic insights that emerge from experiences in effectiveness evaluation and survey analysis more generally is that not all individuals or households respond to interventions or other exogenous influences in the same way. Instead, differential responsiveness or effect modification is a common phenomenon. In practical terms, this means there is a need to systematically undertake subgroup analysis and test for statistical interactions between program variables (e.g., living in a program vs. nonprogram area, or having different levels of exposure to program inputs) and other variables that may modify their effects (e.g., child's age, number of siblings with whom supplementary food might be shared, woman's initial nutritional status, household resources that may affect responsiveness to nutrition education, etc.).

> ***"…not all individuals or households respond to interventions or other exogenous influences in the same way."***

3.23 In policy terms, the exploration of differential responsiveness is important for two reasons: (1) programs deemed *ineffective* (based on main effects analysis) may be deemed effective (based on subgroup or interaction analysis) and (2) programs deemed *effective* (based on main effects analysis) may actually be failing to improve outcomes for important subgroups (such as the poor), which would not be revealed through main effects analysis alone. In both cases, the overall judgment about a program is at stake and, more important, insights on program redesign can arise from analysis of differential response.

3.24 The present review of BINP evaluations bears out those principles. Of the four primary evaluations or sources of data (endline, IMED, NNP, and SCF) none of them undertook a systematic analysis of differential responsiveness.[13] The study that did explore differential responsiveness to some extent is the OED report, and that report documented positive effects of BINP on important subgroups that were overlooked in earlier reports:

- At endline, child nutritional status was better in project areas among those children that participated most in GMP.
- At endline, the ability to translate nutrition knowledge into improved practices was differentially distributed across women varying in wealth and mobility.
- At endline, birth weights were 88 grams higher among those that reported eating more during pregnancy and 270 grams higher among the destitute (with presumably poorer nutritional status) that reported eating more during pregnancy

3.25 Although it is unclear whether or how the OED analysis was guided by a systematic strategy for exploring differential responsiveness, the examples above are sufficient to demonstrate the importance of undertaking such an analysis. Clearly, such findings can substantially alter judgments about BINP's effectiveness and can provide clues about program redesign. These findings underscore the need for future evaluations to anticipate the possibility of differential responsiveness when they are designed and to systematically explore differential responsiveness when the data are analyzed.

3.26 One of the dangers inherent in that suggestion is that analysts may undertake a large number of subgroup and interaction analyses, identify a smaller number that are statistically significant, and report only on the significant findings ("data dredging"). Such an approach is prone to making type I errors (declaring positive findings when none actually exists). One safeguard against that problem is to report the total number of such analyses undertaken so that it can be compared with the number found to be statistically significant. The other complementary method is to adopt a disciplined approach to the analysis based on an explicit program theory, as described below.

*3.27 Realism of objectives***.** Many programs and BINP is not an exception, have been designed to "reduce malnutrition levels by half." Although these ambitions are well intentioned, there is little evidence in the literature that such ambition levels are feasible. Most countries with successful nutrition interventions have reduced their malnutrition rates by 1 to 2 percentage points per year (Mason et al 2005). Given that context, reducing underweight rates from nearly 70 percent to 35 percent in a five year period was not only ambitious; it was a completely unrealistic

[13] The operative word here is *systematic*. Several reports did present results stratified by subgroups, such as the IMED (poor vs. whole sample), the NNP (child's age), and the SCF (child's age and nutritional status at entry to supplementary feeding). However, it is not clear from the reports whether the interactions implied by these analyses were statistically significant, and these examples are but a small fraction of the number and type of statistical interactions that could be relevant.

goal. In comparison, the global MDG is to reduce underweight rates by half between 1990 and 2015—a 25 year horizon. The question that remains is whether any program, however successful, could have met the unrealistic targets that were set. When setting such ambitious targets, does the development community set itself up for failure? Should future programs set more realistic goals and targets? This issue also links to the previous issue of differential response. Future programs may need to be much more nuanced in setting goals and objectives, not only in regard to ambition levels, but also in regard to the specific population subgroups that are expected to benefit from such programs. Evaluations must then be designed carefully to address those questions and to measure program outcomes against realistic and relevant targets. Failing that, many successful programs will continue to be evaluated as "unsuccessful" in meeting their objectives.

3.28 *Program theory.* The endline evaluation and OED report were clearly organized according to an explicit "program theory" showing how the coverage and targeting of program inputs were intended to produce specific outputs and outcomes. That framework served the basic purpose of helping to organize and present a large volume of analyses and findings in a coherent manner. However, the framework employed for that purpose was still not sufficiently detailed to guide the multivariate analysis, including and especially those of differential response.

3.29 Although OED's detailed framework did include attention to some of the anticipated effect modifiers (e.g., household resources and women's autonomy may affect the size of the knowledge-practice gap), it was lacking in certain other critical respects. For instance, the complexity of the knowledge-practice relationship was underspecified and the implications of maternal partitioning were not incorporated. That had important effects on the types of analyses undertaken and the interpretations provided. For instance, the determinants of the knowledge-practice gap were overanalyzed, whereas the determinants of the apparently large effects of BINP on practices were under analyzed; and it was assumed that program effects on birth weight would be mediated through weight gain during pregnancy. To be fair, both of these are issues in which scientific knowledge is evolving, which is a situation facing many analysts or program evaluators.

3.30 These observations underscore the importance of developing a robust and explicit program theory to guide (1) the design of programs, (2) the design of evaluations, and (3) the analysis of evaluation data. Moreover, in light of the close connections among intervention design, implementation details, and context (described in section 3.2 above), this program theory should seek to integrate knowledge from each of these domains (implying inputs from diverse parties, each with specialized knowledge from one or more domains) to specify not only an explicit set of main effects, but also the likely subgroup effects and differential responses that should be anticipated (in program design and evaluation design) and explored (in the analysis of evaluation data).

3.31 *Use of evaluation findings.* The use of evaluation findings faces some common problems, analogous to those described for operations research and monitoring. Some general observations and principles are as follows:

3.32 A common belief or aspiration is that evaluations are efforts to generate objective, empirical evidence about program performance to assist rational policy decisions. In reality, evaluations and their findings are often used by various stakeholders to support their own views, interests, and objectives. These may include improving nutrition; using resources efficiently; protecting, expanding, redesigning, or terminating a program, its inputs and resources; allocating the authority (power), responsibility (control), accountability (blame), liability (risk), or visibility (credit) associated with the program; and promoting or discrediting various interventions (e.g., supplementary feeding, nutrition education, capsules, home gardening), approaches (e.g., GMP,

social marketing, community-based), or organizations (e.g., government ministries, NGOs, international agencies). It is common for a given stakeholder to simultaneously hold multiple interests and objectives and for multiple stakeholders to steadfastly claim the same goal (e.g., improving nutrition) while differing in their related views, objectives, and interests. Because these views, interests, and objectives may vary among stakeholders, the ultimate use of evaluations and their findings can become a function of (political) power relations among stakeholders.

"A common belief or aspiration is that evaluations are efforts to generate objective, empirical evidence about program performance to assist rational policy decisions. In reality, evaluations and their findings are used by various stakeholders to support their own views, interests, and objectives."

3.33 As a consequence, the strength and technical quality of evaluation findings come under careful scrutiny as they become available and enter policy debates. However, the technical quality demanded at that stage usually exceeds what key stakeholders were willing to pay for at the beginning and what the evaluation organizations were able to achieve on the ground. The net result is ambiguity in evaluation findings, which further enlarges the space for interpretations and policy debates to be based on stakeholders' interests, as opposed to empirical evidence. Typically, as seen in BINP, the effect of the program is more "in the eye of the beholder" than in the evaluation findings themselves.

3.34 It once was possible for certain organizations (e.g., international agencies, specialized government units, research institutes, and universities) to maintain a virtual monopoly on evaluation expertise and thus have a disproportionate influence in policy debates. However, the ability to conduct, interpret, and critique evaluations has become more widely distributed over the past few decades (e.g., across NGOs), and this trend is likely to continue. Together with advances in communications and openness and the rising influence of civil society organizations, these trends suggest that the technical scrutiny and political character of evaluations will only intensify in the future.[14]

3.35 Assuming it is neither possible nor desirable for a small number of organizations to reestablish or maintain a monopoly on evaluation expertise, the considerations above lead to three suggestions. First, governments and their partners should make greater investments in the technical quality of evaluations (including the monitoring, operations research, and process evaluation components). That implies capacity strengthening as well as the direct costs of specific evaluations. Second, stakeholders associated with a given program should come to some common understanding about the types of program and policy decisions for which evidence is needed and agree on some reasonable types and standards of evidence for each. That is distinct from setting realistic goals and targets for the program (which also is important), and some lessons and guidance from the Multi-Country Evaluation of IMCI on how this may be done have recently become available (Bryce et al. 2004). Third, governments and their partners should anticipate that sound evaluations sometimes can settle certain empirical questions (e.g., when they produce convincing evidence for impact) but contentious issues always will remain (e.g., the debate shifts to further fine-tuning of targeting strategies). Information on these and other issues always will be

[14] The "mature form" of this system can be seen in the present-day United States, where a wide spectrum of think tanks, policy analysis centers, advocacy organizations, professional and scientific societies, and public relations and political organizations routinely use technical expertise to advance organizational, policy, or political agendas. Although developing countries are moving in that direction but still far from that state, the underlying principles are the same in regard to the difficulty of simultaneously addressing the technical and political dimensions of evaluations.

inadequate, as illustrated in box 5 for the cost-effectiveness analysis on BINP. Thus, governments and their partners should create, borrow, or adapt evolving procedures for making decisions under conditions of uncertainty in multistakeholder contexts (National Research Council 1996). Fourth, information from evaluations must not be used as a "policing tool". Instead, this information should be used constructively to strengthen and fine-tune the technical dimensions of the intervention strategies, without detracting from the political dimension of building commitment for sustained investments in improving nutrition. This delicate balance between the technical and the political is critical.

Box 3.1 Implications for Cost-Effectiveness Analysis

Cost-effectiveness analysis (CEA) is a method for quantifying the unit cost of achieving a given outcome. It can be used to (1) inform judgments about continuing or terminating a program (if there is consensus on the social value of saving a life, lifting a child from malnutrition, preventing a case of low birth weight, etc.); (2) compare one program against others to see whether there are more efficient means to achieve similar ends; and (3) identify potential changes in a program that may improve the effects, lower the costs, or both to improve its efficiency. It is important to distinguish these purposes when undertaking a CEA because the information requirements and comparisons differ considerably. Ideally, the purposes and assumptions of a CEA should be specified in advance (similar to the program targets), so that appropriate data can be collected and decision parameters can be agreed on in advance.

To fulfill its functions well, a CEA should be based on reasonably firm estimates of program effects and costs, should account for all relevant outcomes (or be able to separate the costs associated with each outcome of interest) and, as noted, should distinguish which of the three objectives above are of interest.

The BINP illustrates a case in which some of the key requirements for the CEA were not met. Specifically, (1) despite the many evaluations of the program, as the present report reveals, it has not been possible to develop firm estimates of the effects on key outcomes (child malnutrition, maternal weight gain, and low birth weight); (2) the CEA excluded some of these outcomes from consideration (maternal weight gain and birth weight) while including one that was not an objective of the program (child mortality), and it did not examine a range of related or intermediate outcomes (e.g., improvements in micronutrient and health service delivery, maternal knowledge and practices); (3) the CEA estimates for reducing child malnutrition and mortality appear to attribute all program start-up and recurrent costs to these child objectives (by simply excluding the cost of food directed to mothers)[1], and the CEA provides no estimates of the marginal C/E for the child component; (4) the analysis did not specify which of the three purposes it was designed to fulfill, but implicitly adopted the second purpose by comparing the BINP to a (hypothetical) rice ration program; (5) there was no analysis of the potential improvements in the CEA that might be obtained if the program were to be redesigned based on lessons learned during the BINP. Finally, although the CEA did make use of the subgroup analysis showing that regular attendance at growth-monitoring sessions is associated with somewhat greater improvements in nutritional status, it did not make use of the subgroup analysis showing substantially higher birth weights among the thinnest women and those who eat more during pregnancy.

This illustration reveals that a CEA is an extension of outcome evaluation and is affected by the same key issues discussed in this section. Those issues include the need to consider differential responsiveness, apply a systematic program theory, use a revised ("continuous improvement") view of programs, and give attention to the evaluation-utilization issues before the design of evaluation systems. These are some of the higher-order lessons offered by the BINP experience.

[1] This is inferred from a brief description of the cost analysis provided in annex G of the OED report.

D. SUMMARY: BUILDING ON BINP TO IMPROVE IMPACT AND COST-EFFECTIVENESS

3.36 The MDGs have brought into sharp relief one of the major policy issues in international health in recent decades: the relative emphasis to place on building/strengthening health systems, capacities, and infrastructures versus high-impact but possibly less sustainable strategies for reaching targets at-scale. Although BINP was not designed with the MDGs in mind, it is worthwhile reflecting on BINP in this context because the question is central to the choices facing Bangladesh today as well as the MDGs more generally.

3.37 BINP is one of the largest multicomponent nutrition programs ever implemented in a developing country, and it took place in a country whose rates of poverty and maternal and child malnutrition are among the world's highest. In contrast to HNP efforts in many other countries, it was managed largely through established administrative structures and procedures, augmented as appropriate through the addition of essential cadres of workers and volunteers at the community level implemented primarily through NGOs. In the process it has demonstrated the feasibility of establishing or adapting large-scale structures to support community-based HNP programs. These already are being taken to a national scale through the NNP, with further mainstreaming envisioned in forthcoming HNP SWAP and PRSP efforts. These are major accomplishments that deserve full recognition. And they should not be discounted by an overemphasis on impact at outcome level. That said, impact, mainstreaming, and sustainability clearly must be major foci for the future.

3.38 The issues raised in section 3 suggest that continued progress in Bangladesh will require a stronger focus on achieving a good fit between the design of interventions and the implementation and contextual factors. Insofar as contextual features and implementation issues both can vary across locales, this is likely to require greater decentralization of decision making as well as the capacity for timely monitoring, operations research, and program adjustments. Finally, the demand for high-quality process and impact evaluation is likely to intensify globally and is fundamental for ensuring effective as well as politically sustainable programs in the future.

> ***"This is likely to require greater decentralization of decision making as well as the capacity for timely monitoring, operations research, and program adjustments."***

3.39 In short, Bangladesh and its partners have an opportunity to pursue a combination of the systems-strengthening route and the targeted route of direct interventions toward the MDGs and to boldly and systematically confront the challenges of the approaches recommended in the World Development Report 2001 and 2004. The usual criticism that Bangladesh is not suitable or ready for a decentralized or self-correcting approach like that suggested here may be a well-taken description of the *current* situation. The activities suggested above are intended to augment the process of moving Bangladesh in a direction that ultimately will be required.

APPENDIX A. TECHNICAL FOOTNOTES TO TABLE 2.1

Note that the double difference in HAZ (baseline vs. endline) is not significant using means, but *is* significant using prevalence of severe stunting (<-3 Z) or moderate + severe stunting (<-2 Z); the significant difference at midterm for WAZ shown here is *not* confirmed using prevalence estimates; the findings for WHZ based on means are confirmed using prevalence estimates.

Note that HAZ and WAZ prevalence estimates here are based on <-3 Z but the estimates for WHZ are based on <-2 Z. There were no significant differences in WHZ using severe wasting (<-3 Z).

The seemingly inconsistent results involving stunting and wasting at midterm and endline have puzzled many analysts of these data. A possible explanation may involve the lagged nature of changes in these two indicators. Earlier work in Bangladesh has shown that HAZ is at its annual peak when WHZ is at its annual low, and vice-versa (Brown 1982; Briend 1989). Similarly, in Nepal, children with higher WH at the beginning of the harvest season experienced greater increments in height (than those with low WH) during the postharvest period. This is broadly similar to the patterns in BINP data: the double difference (base-mid) is -0.22 for HAZ but +0.26 for WHZ, whereas the double difference (mid-end) is +0.26 for HAZ and -0.29 for WHZ.

Statistical tests were not applied, but results shown are unlikely to be significant. Results also were stratified by severe malnutrition (<-3 Z) and moderate malnutrition (-2 to -3 Z) for each indicator. Again, there were no meaningful or suggestive differences. The patterns across the three indicators strongly suggest that length/height measurements were systematically overestimated, thereby producing much lower stunting estimates and much higher wasting estimates than other surveys.

Results are based on samples from three project *Thanas* and three nearby nonproject *Thanas* with roughly similar demographic characteristics.

Results for the 0 to 23-month-old age group also were examined using prevalence estimates (total, severe, and moderate separately). There were no significant differences. Results for this age group also were stratified by gender, three age groups (0–5, 6–11, and 12–23), and moderate or severe categories for each of the three indicators, with potential contrasts between BINP-NNP and BINP-control. Of the 72 cells thus defined there were four statistically significant findings, involving severe or moderate stunting or underweight.

HAZ is significantly lower, and WHZ significantly higher, in BINP versus NNP areas.

The original report further suggests a significantly lower rate of moderate stunting (-2 to -3 Z) in BINP versus NNP areas, but the difference is only 1 percentage point and is likely a typographical error meant to apply to moderate underweight (43.3 in BINP areas vs 40.9 in NNP areas). That would be consistent with the significantly higher levels of moderate wasting in BINP versus NNP areas (12.4 vs. 9.6).

Results for the 24 to 60-month age group were stratified by gender, three 12-month-old age groups and moderate or severe categories for each of the three indicators, with potential contrasts between BINP-NNP and BINP-control. Of the 72 cells thus defined, there were eight statistically significant findings, seven of which are in the 36 to 47-month-old age group involving lower levels of severe stunting or underweight among BINP children. The results shown here for 36 to 47-month-olds are for sexes combined, but boys and girls have broadly similar results in this age group.

These results represent a reanalysis of data from BINP midterm and endline surveys, using two-stage multivariate analysis to control for potential confounders between project and control areas and a matching procedure (using Helen Keller International [HKI] survey data) to control for propensity to participate in BINP. Results shown here are for the nearest-neighbor matching method. Notable is that these results confirm higher WAZ and WHZ in project areas at midline (shown in the same column above) and also bring the results for HAZ in line with those results. Although no differences are shown here at endline, further analysis of those children receiving project inputs reveals slightly higher values for HAZ (+0.03), WAZ (+0.06), and WHZ (+0.04), with the WAZ finding being statistically significant ($p < .05$). In this latter case distribution analysis reveals the primary difference in the WAZ distribution is a smaller number of severe cases (< -4 Z) and mild cases (-2 to -3 Z).

Although the OED analysis applied state-of-the-art methods (indeed, extended on those methods) to overcome the weaknesses inherent in a cross-sectional endline evaluation design, the results remain ambiguous. Key methodological threats in such a design include the inability to control for secular trend (due to lack of access to the baseline data) and possible area- and individual-level selection bias involved in the choice of BINP areas. The methods employed by OED cannot address the secular trend, and they appear likely to reduce but not eliminate the potential bias from selection for the following reasons:

- Technical appendix 4 (table 1) of the OED report contains a test of the area-selection bias and the extent to which it is corrected by the propensity score matching (PSM) methods. The results show that such a bias does exist (with mean HAZ and WAZ being significantly higher in BINP areas) and that the PSM methods reduce these differences to nonstatistically significant levels. Although the OED authors conclude from these tests that the matching procedure eliminates the geographic placement bias, the numbers in their table actually demonstrate that this bias is not fully corrected by matching. Indeed, the residual bias in HAZ (0.06) is nearly as large as the project effect suggested in their analyses (0.08 to 0.10 at midterm).
- Technical appendix 4 (table 3) of the OED report contains a test of the participant-selection bias, specifically; the possibility that participants and nonparticipants may differ in unmeasured variables that are correlated with nutritional status, which would yield biased estimates of project effects. As with geographic bias, the result confirms that participants have statistically significantly higher mean HAZ, WAZ, and WHZ than nonparticipants and that the PSM methods reduce that difference to nonsignificance in most cases. However, here again, a residual bias exists which is a meaningful fraction of the estimated project effect: the residual bias is 0.01 to 0.04 for HAZ (compared with an estimated project effect of 0.08 to 0.10 at midterm); the bias is 0.04 to 0.08 for WAZ (compared with an estimated project effect of 0.14 at midterm); and the bias is 0.06 to 0.09 for WHZ (compared with an estimated project effect of 0.13 to 0.16 at midterm).

These results represent analysis of national DHS data from 2000, using two-stage multivariate analysis to reduce selection bias, endogeneity, and right-censoring due to mortality. Of all the

results shown in this table, this is the only set based on such methodological controls, which compares children in BINP areas to others in a representative national sample and uses survey instruments common to BINP households and non-BINP households. Remaining methodological threats are the same as those identified in point j above, but were not assessed to the same extent. These threats to validity could be reduced by testing for BINP area effects in the 1997 DHS and by pooling the 1997 and 2000 DHS data to test the significance of the double difference (BINP x year interaction).

APPENDIX B. SUMMARIZED TABLES FOR SECTION 2

Table B.1. Existing Evaluations of BINP (CBNC) Adopted Different Methodological Approaches and Had Different Results

Evaluation	Objectives	Study design and sampling	Data collection and indicators	Results
Endline evaluation	To compare pre-post and BINP and non-BINP differences in : 1. nutritional status in U2 children and women of reproductive age 2. pregnancy weight gain 3. birth weight 4. iron tablet use in PLW and adolescent girls 5. postpartum vit A coverage	Longitudinal study (6 first-phase *Thanas* and 2 control *Thanas*) Three-stage stratified systematic sampling 1,438 pregnant women (PW) 1,362 mothers and their newborns 4,967 U2 children 4,967 mothers of U2 children 525 adolescent girls	3 Questionnaires (U2 children, pregnancy weight gain, and birth weight) covered the five objectives of concern and basic HH sociodemographic information 1.Anthropometry: in children, in women, and adolescent girls 2.Participation 3. Knowledge and practice (K&P) 4. Weight measurements for PW and newborns	1. No difference in child Wt and Ht and WAZ; women and adolescents' BMI between BINP and non-BINP areas, but differences in HAZ and WHZ (favors control group though) are significant. 2. Improved overall K&P of mothers on pregnancy and childcare behaviors. 3. Women in BINP had higher one-month and per-trimester weight gain but lower overall gain (insignificant); no difference in birth weight.
OED	To examine the evidence of the effectiveness of BINP	Secondary data analysis using data from the endline and SCF reports as well as HKI NSS and DHS data Same design and sampling as for endline and SCF studies; DHS and HKI data have probability sampling and national coverage.	Same as for endline and SCF; DHS has growth of U5 children plus questionnaires 1. Coverage and targeting in GMP and SF 2. K&Ps 3. Anthropometry of children (U2& U5) 4. Weight measurements in mothers and newborns	1. Partial but inconsistent evidence for impact on child growth 2. Likely impact on K&Ps about child feeding 3. Little evidence of effect on weight gain or birth weight for overall sample, but significant effects on vulnerable subgroups 4. Consistent evidence of effects on K&Ps concerning nutrition in pregnancy
IMED	To review the target and achievements of project; to enquire into implementation and functions of IEC activities; to assess the appropriateness of CBNC; to measure maternal and child health goals; to measure demographic	Cross-sectional study (12 *Thanas* of all phases and 12 control *Thanas*) Three-stage stratified systematic sampling 240 newlywed women (NWW), 1,485 PW 1,487 lactating women (LW), 576 mothers of U5 children, 3,024 U5 children 36 CNPs, 12 CNOs, 9 TPOs 36 village nutrition management committee (VNMC) groups	Surveys (NWW, PW, LW, mothers of U5 children) Interviews (community nutrition promoters [CNPs], community nutrition organizers [CNOs], *Thana* project officers [TPOs]) Focus Group Discussions (VNMC, UNMC, and adolescents)	1. Better results shown in more health and nutrition practice variables than awareness variables among PW and LW in BINP area than in non-BINP samples. 2. No differences between BINP and non-BINP areas in severe and moderate child malnutrition. U2 children had lower malnutrition rates than U5 children (data not shown)

	and nutritional health situations	12 union nutrition management committee (UNMC) groups 2 groups of adolescents	Anthropometric measurements of PLW, NNW, and U5 children 1. K&P 2. Anthropometric measurements	
SCF report	To assess rates and reduction of U2 child malnutrition in BINP against BINP objectives and as opposed to rates and reduction in non-BINP areas	Cross-sectional (3 first-phase BINP *Thanas* and 3 control *Thanas* with matched household (HH) socioeconomic status (SES) SF Register data Results from baseline and midterm evaluations Two-stage stratified sampling procedure	Basic HH information Anthropometry in U5 children, their mothers, and PW Secondary register data (weight and age) of children in SF 1. WAZ, HAZ, and WHZ 2. K&P 3. Children's nutritional status pre- and post-SF	1. No differences in severe and moderate underweight, wasting, and stunting in children U2/5 between areas 2. Mothers in BINP area had better knowledge and practice. Minimal practical difference with few exceptions (iron and antenatal care) 3. Young children (<12 mo) did not improve WAZ after graduation from SF. Inadequate and incorrect enrollment and food sharing in SF
NNP baseline (initial results)	To assess various nutritional and socioeconomic indicators in the intervention and control *Thanas*: 1. To obtain statistically valid samples form 44 "new" NNP *Thanas* 2. To obtain valid data from appropriate samples from 53 "old" BINP *Thanas* 3. To obtain valid data from 16 control *Thanas*	Cross-sectional survey (not explicitly discussed in text). Developed primary sampling unit (PSU) across all six divisions 360 PSUs for 44 NNP *Thanas*; 120 PSUs for 16 control *Thanas* 228 PSUs for 53 BINP *Thanas*	Questionnaire surveys: - HH - U5 and mothers for U2 group: 5,026 NNP, 2,086 BINP, 1,707 control PW: 2,723 NNP, 2,193 BINP, 1,063 control - Adolescent girls Primary: Anthropometry of women & children Pregnancy weight gain and birth weight K&P	1. No difference in any nutritional status indicators among 6- to 23-month-old children by gender or across areas. 2. No indications of whether significant tests were performed on women's anthropometric measurements, mean pregnancy weight gain, or mean birth weight (partial data only) across three areas. 3. Several K&P variables were reported with significant differences among areas.

Table B.2. K&P and Nutritional Status Outcome and Impact Variables for Women (during pregnancy and about child feeding) and Children That Have Been Found Significant in Evaluations [1,2]

Evaluation	BINP Endline (6–23mo)					OED (0–23mo)	IMED (U5)	SCF (6–23 mo)	NNP baseline (0–23mo)	
Indicator	BINP-Control Baseline	BINP-Control Midterm	BINP-Control Endline	Baseline -Endline	Midterm-Endline	BINP-control	BINP-control	BINP-Control	BINP -NNP	BINP-Control [3,4]
Children										
Height (cm)	70/70 ns	71/71 ns	71.5/71 ns	ns	ns				69.2/68.9 ns	69.2/69.4 ns
Weight (kg)	7.7/7.8 ns	7.8/7.7*	8.1/8.1 ns	ns	ns				7.5 /7.4 ns	7.5/7.5 ns
HAZ mean	-2.36/-2.49 ns	-1.98/-1.89 ns	-1.9/-2.07*	ns	*	*			-1.5/-1.6 ns	-1.5 /-1.5ns
HAZ < -3 SD (severe stunting %)	34/31.4 ns	20.4/18.5 ns	16/22*	*	*		7/8?	11.6/12.4 ns	10.4/11.5 ns	10.4/11.3 ns
HAZ < -2 (Severe and moderate stunting %)	62/66 ns	39/48.4 ns	45.7/51*	ns	*		7.7/7.4?	27.5/27.6 ns	23.5/24.2 ns (-3 <Z < -2)	23.5/23.2 ns (-3 <Z < -2)
WAZ mean	-2.3/-2.24 ns	-2.04/-2.14*	-1.87/-1.94 ns	ns	ns	*			-1.6/-1.7 ns	-1.6/-1.6 ns
WAZ < -3 SD (Severe underweight %)	30.6/27.3 ns	17.7/18.9 ns	11.8/13.3 ns	ns	ns		14/14?	11.4 /12.2 ns	10.5/11.7 ns	10.5/11.4 ns
WAZ < -2 (Severe and moderate underweight %)	59.5/56.5 ns	55.5/57.5 ns	46.8/49.4 ns	ns	ns		13/12.9?	35.2/36.3 ns	30.2/29 ns (-3 <Z < -2)	30.2/29.3 ns (-3 <Z < -2)
WHZ mean	-.77/-.73 ns	-.98/-1.19*	-.84/-.76 (C)*	ns	(C)*	*	**(5)**		-.8/-.8 ns	-.8/-.8 ns
WHZ < -3 (Severe wasting %)	7.4/5.3 ns	2.2/2.1 ns	.6/.7 ns	ns	ns		12.2/12.4?	1/1.2 ns	1.6/1.7 ns	1.6/1.1 ns
WHZ < -2 (Severe and moderate wasting %)	22.1/21.3 ns	15.3/18.9 (C)*	11.5/8.7 (C)*	ns	(C)*		28.5/29.6	13.3/14.1 ns		
WHZ < -2 but > -3 (moderate wasting only %)	14.8/16 ns	13.1/16.8 (C)*	10.9/8 (C)*	ns	(C)*				11.6/11.8 ns	11.6/12.2 ns
Birth weight (kg)		2.7/2.7?	2.8/2.8?		ns	ns			2.78/2.75?	2.78/2.84?
% of birth weight < 2.5 kg		25.6/29.9?	15.9/17.6?		ns	ns			16.2/22.3?	16.2/12.9?

Table B-2 Continued

Evaluation / Indicator	BINP Endline (6-23mo)					OED (0-23mo)	IMED (U5)	SCF (6-23 mo)	NNP baseline (0-23mo)	
	BINP-Control Baseline	BINP-Control Midterm	BINP-Control Endline	Baseline -Endline	Midterm-Endline	BINP-control	BINP-control	BINP-Control	BINP -NNP	BINP-Control [3,4]
PW										
Mean one-month weight gain by all 3 trimesters (kg)		.5/.5 ns 1.1/1.1 ns 1.6/1.3*	1/.8 * 1.3/1.2 * 1.4/1.1 *	ns	ns	ns			3.55/3.59? (3rd	3.55/3.98? trimester only)
Total pregnancy weight gain (kg)			8.2/9.1			ns				
Mother of U2/5										
(during pregnancy)										
Knowledge of *eating* more %	63.6/65.1 ns		92/67.5*	*		*	86/71*		46/38 * (benefit to	46/35* mother & fetus)
Ate more than usual %	56.7/80*	56.1/22.1*	58.6/29.4*	*	*		44/23*	34.3/30.2*	19.8/15?	19.8/20?
Knowledge of *resting* more %			89/58*			*		77.8/70.5*	37/28.5* (benefit to	37/25.6* mother & fetus)
Rested more %	14/24(C)*	65/35*	67/46*	*	*		60/51*	60.2/54*	45.4/40.4?	45.4/45.3?
Knowledge of less *lifting and heavy work* %			94/83*			*	87/78*			
Never *lifted* heavy objects %			53/44*				42/34* **(6)**			
Awareness of dangerous sign %							24/16*			
(child care)										
Awareness of the benefits of olostrums						*	88/69 *	63.3/53* (should give)	93.5/87?	93.5/91?
Fed child colostrum	66/79©*	98/86*	96/92*	*	*		91/77*	77.5/72.6*	95.5/91?	95.5/94.6?
Initiated breast-feeding (BFD) immediately after birth	14.5/12 ns	62/31*	84/56*	*	ns				34/28?	34/33?

Table B-2 Continued

Evaluation / Indicator	BINP Endline (6–23mo)					OED (0–23mo)	IMED (U5)	SCF (6–23 mo)	NNP baseline (0–23mo)	
	BINP-Control Baseline	BINP-Control Midterm	BINP-Control Endline	Baseline -Endline	Midterm-Endline	BINP-control	BINP-control	BINP-Control	BINP -NNP	BINP-Control [3,4]
Awareness of benefits or length (in **SCF**) of exclusive breast-feeding (EBFD)						*	87/71*	78/69*		
EBFD to 6 months								4/5 ns	8/12 ns	8/13 ns
Awareness of comp. feeding at 5 to 6mo							81/60*	63.6/64.4 ns (when to give)	98/93?	98/97?
Gave complementary food at 5 to 6 mo	27/22 ns		53.4/50.5 ns	ns				56.6/48.6*	43/38.4?	43/42?
Eating more food during lactation							61/43*		25.5/27?	25.5/26.5?
(Anthropometry)									**(PW**	**Only)**
Mean height (cm)	150/150 ns	149/149 ns	151/150*	ns	*				149.9/150.1?	149.9/150.4?
Mean weight (kg)	42.5/42 ns	39.6/39.7 ns	43.1/42.7*	ns	*				48.6/48.6?	48.6/48?
Mean BMI (kg/m^2)	18.8/18.7 ns	17.8/17.9 ns	18.8/18.9 ns	ns	ns		**(7)**		21.6/21.5?	21.6/21.2?
% BMI < 18.5	46/50.5 ns	68.8/66.7 ns	48.5/46.8 ns	ns	ns					

Note*:* ns = no differences, ? = significance unknown

* data are significantly different, test results in NNP columns were provided by Dr.Tahmeed Ahmed during the peer review of the first draft of this study

C. difference favored control area

(1). Comparisons of results are always between the project and control areas at the time when the evaluation was conducted or when the endline evaluation was conducted. Under BINP endline, the last three columns showed whether there are differences from baseline to endline, from midterm to endline, and from baseline to midterm between BINP and control areas.

(2). Age of children sample in endline evaluation, OED, and SCF are U2 and 6 to 23 months specifically for the endline evaluation. U5 were included in the samples of IMED and NNP baseline (data of U2 were presented here) instead.

(3). BINP *Thanas* here include second- to fourth-phase *Thanas* only.

(4). Children under 6 mo of age were included in the sample, and there were significantly fewer children who were severely stunted (HAZ < **-** 3 SD) in BINP area than in NNP area.

(5). The evaluation reported that "in Weight for Health index, normal children in the intervention *Thanas* are 84.8 percent as opposed to 73.8 percent in comparison *Thanas* (significant at .00002)." However, data were not reported and what "normal" means was not defined.

(6). "Exposure to heavy work during pregnancy" was stated in the table of the evaluation report; it was unclear whether the percentage represented the proportion of women who did heavy work or that of those who did not.

(7). The evaluation reported BMI of mothers "are almost same because of many comparable features of the women: in both areas.

Table B.3. Presence of Potential Threats to Internal Validity in Available Evaluations and How They Were Addressed

Evaluations / Threat	Endline	OED	IMED	SCF	NNP
Use of control group	Yes, only 2 *Thanas*	Yes, as in endline and SCF studies; plus some child anthropometry analyses using (1) the HKI super-control with household-level matching and (2) DHS data	Yes, 12 *Thanas*	Yes, 3 *Thanas*	Yes, 16 NNP comparison *Thanas*
Use of baseline info to control for secular trend	Yes	Midterm vs. endline were compared because of problems with baseline	No	No	NA
Noncomparable control group	Yes. Thin description is provided on the search of the candidate control group with least differences on a few predetermined SES variables. Acknowledged significant health and nutrition (H&N) activities present in control areas	Same as endline and SCF. HKI super-control and DHS analyses employ better controls but still suffer from nonrandom selection of BINP *Thanas.*	Yes. Provided 15 to 20 basic variables (number of village women/men and children, schools, health facilities, etc.). No significant tests but more health and education units in BINP *Thanas*	Uncertain. Claimed that matching was done based on comparable SES (showing some variables) and that there were no major H&N activities (no data shown)	Uncertain. No specific discussion on the comparability of NNP, BINP, and control *Thanas*. No significant tests on SES variables. % of ever educated women in BINP area reported being lower (0.7%–1.2% lower than other areas).
Cross-contamination of control group	Control *Thanas* had received considerable development assistance. Intended to use DHS 2000 and HKI data as "super" control, but not reflected in data analysis	As with endline and SCF analyses; HKI super-controls and DHS analyses also affected by national secular trends but not contaminated by BINP per se.	Control *Thanas* had considerable exposure to health programs (19 ongoing programs vs. 26)	Not discussed	Not discussed

Table B.3 Continued

Evaluations Threats	Endline	OED	IMED	SCF	NNP
Approaches to control for confounders	None	Ordinary Least Square and 2-step regressions; household-level matching with HKI super-control data	None	None	None
Incompatible questions in rounds of surveys	Yes	Yes	NA	NA	NA
Poor choice of indicators	Yes. K&P indicators tied too closely to key education messages, thereby increasing reporting bias; no triangulation.	Yes. K&P indicators tied too closely to key education messages, thereby increasing reporting bias; no triangulation.	Yes. K&P indicators tied too closely to key education messages thereby increasing reporting bias; no triangulation.	Yes. K&P indicators tied too closely to key education messages thereby increasing reporting bias; no triangulation.	Yes. K&P indicators tied too closely to key education messages thereby increasing reporting bias; no triangulation.
Poor measurement of indicators	– High level of errors in age and height (<6 mo) – Problems in weight gain measurements for PLW – Heaping with birth weight data – No information provided on reliability of K&P measurements	Same as for endline and SCF. DHS data provide highest-quality and most uniform measurements	Not discussed, and because little data were presented it was not entirely clear what had been measured and how it was done	Not discussed	Not discussed
Stratified analyses by age and/or gender	No	Tested and controlled for age and gender effects in regression models	Stratified nutritional status by age (data not shown though)	No	Data presented by age and gender (e.g., child nutritional status)
Tests for interactions	No	Inconsistently and with little documentation of which ones were tested but were not significant	No	No	No

Table B.3 Continued

Evaluations Threats	Endline	OED	IMED	SCF	NNP
Statistical significance reported	Yes, in most variables with few exceptions	Yes.	Some were reported but none on anthropometry	Yes	Sometimes (anthropometry)
Subgroup analysis	No, and acknowledged	Through interactions in models, though inconsistently and with incomplete documentation	Tested practices in poor subsamples between areas	No	Compared some data by *Thanas*
Individual-level sample size adequate for power	Yes, for most analyses as conducted, but not in baseline control	For most analyses as conducted, but not in baseline control; sample sizes and power for interactions unclear	Yes, for most analyses as conducted	Yes, for most analyses as conducted	Yes, for most analyses as conducted
Seasonality of measurements	Acknowledged and indicated there is no conclusive data on trend; endline different season from base and midterm	As for endline and SCF. HKI super-controls matched on season; DHS not affected	Not addressed	Not addressed	Not addressed
Age of children	Focused on 6- to 23-month-olds only due to high error in Ht measurements of < 6 mo	As with endline and SCF. DHS analyses 6 to 60 mos.	Focused on Under 5 with the belief that mothers of older children (>24 mo) had more exposure to BINP interventions	Under 2 (6–23 mo) for change in nutritional status and SF register data	Under 5 (data presented separately, Under 2, 24-60 months, and by one-year intervals

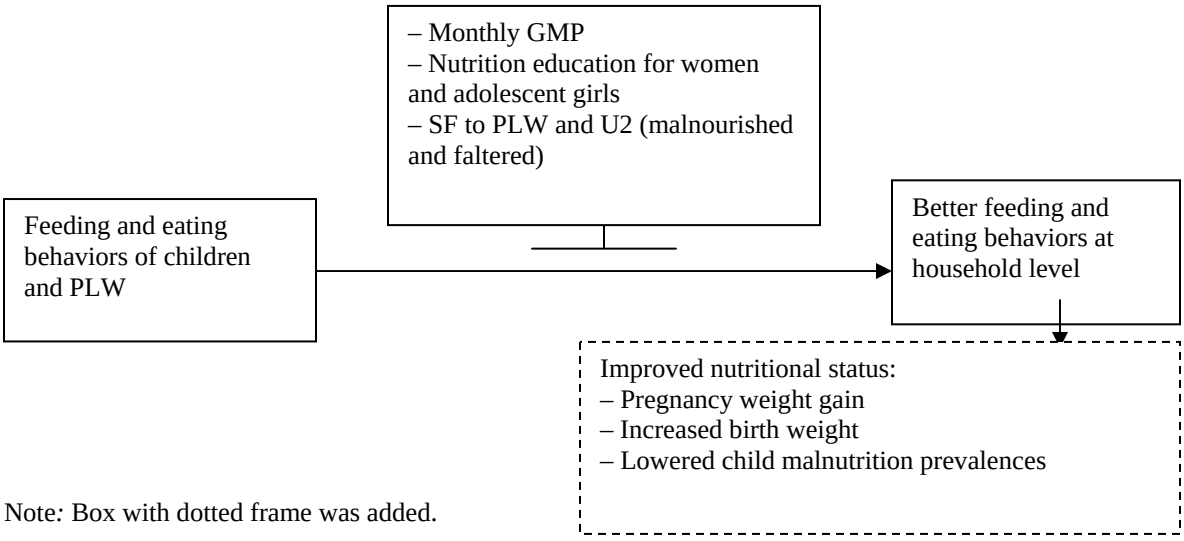
Note*:* NA = not applicable

APPENDIX C. DESCRIPTION OF ENDLINE EVALUATION OF BINP CBNC

The endline evaluation described most clearly the implementation strategy of the CBNC (figure C.1), and it was the only evaluation that had a longitudinal design. In theory, comparisons of many indicators could be made in three formats—(1) between BINP and control *Thanas* at different time points of baseline, midterm, and endline surveys, (2) changes between any of these three time points within each area, and (3) changes during a period between two areas.

However, the strength of this design was compromised in several ways, including incompatible questions, varied ways of asking supposedly the same questions, and missing questions in some earlier surveys (baseline and/or midterm), which make certain comparisons less valid. In addition, there was very little qualitative research to complement the results of the data analysis.

Figure C.1. Program Strategy of Intervention in CBNC



Note*:* Box with dotted frame was added.

Issues related to the selection of BINP and control *Thanas* were covered earlier (box 1 and 2). The compositions of BINP and control *Thanas* were further complicated by a few issues. (1) Every division in Bangladesh contributed one of the first six BINP *Thanas* to achieve a balanced geographical presence. (2) One of these six intervention *Thanas* was experimenting with a "newly married couples" (NWC) strategy, whereas the other five were following a general "vulnerable group" strategy. It should be noted that this one *Thana* had several significantly different features compared with the other five at the baseline (mean age) and in some changes (higher improvement in nutritional status). (3) Activities were run by government in three *Thanas* and by contracted NGOs in the other three. The difference between these two implementing models lies primarily in the source of payment to frontline workers (CNPs/CNOs). Yet it appeared to be more complicated than that because NGOs were in charge of project supervision and they were in an uneasy situation in government-run *Thanas*. (4) The only brief qualitative research reported in this evaluation identified the extensive presence of NGOs (BINP implementing NGO for all six *Thanas* was Bangladesh Rural Advancement Committee and others) with a lot of nutrition and health activities including GMP, but details were not provided or available (table C.1).

Table C.1. Presence of Other NGOs in BINP and Control Thanas

District	BINP Thanas	Implementing NGO	Control Thanas
Magura	Mohammadpur	BRAC (+ 6–8 others)	
Faridpur	Faridpur Sadar	BRAC (+ 13 others)	
Bogra	Gabtoli	BRAC	Shonatala (BRAC + 6 others)
Barisal	Banaripara	BRAC	
Chandpur	Shahrasti	BRAC	Hajiganj (6–8 others)
Maulavi Bazaar	Rajnagar	BRAC	

Despite these limitations, this evaluation provided much information on various outcome and impact variables as covered in earlier sections. However, critical in-depth analyses appeared to be missing in several aspects. The following are a few examples:

1. Potential factors associated with stunting (reported), underweight, and wasting were not examined.
2. To enhance plausibility of overall BINP effects, the effects of the intervention activities (nutrition education and SF) and practices on nutritional status were not examined.
3. The association between pregnancy weight gain and birth weight, two closely related issues, were not examined.
4. Although the evaluators laid out a nice conceptual framework for the evaluation (figure C.1), little research was done to understand the implementation process pathway, such as the inputs category (development and dissemination of IEC, identification and training of CNPs, establishment of CNCs, formation of Women's Group and Adolescent Girls' Forum.

Many of these issues were taken up in the OED report described next.

APPENDIX D. DESCRIPTION OF OED

The OED Report (from the Operations Evaluation Division of the World Bank) contains results from three sets of analyses: reanalysis of data from the endline evaluation (appendix C), reanalysis of data from the SCF survey and register data (appendix F), and analysis of the 1997 and 2000 DHS data. The analyses are structured in relation to the logical framework of BINP, examining the coverage and targeting of activities and inputs (GMP, nutrition education, and supplementary feeding); improvements in knowledge, practices, and nutritional outcomes; and the statistical association among these factors. In addition, the multivariate analyses provide stronger tests of the differences (and double differences, over time) between BINP and non-BINP areas by controlling some of the potential confounding factors between these areas. In general, this study makes far better use of the available data for evaluating BINP, applies stronger analytical techniques, and attempts to overcome some of the limitations associated with earlier analyses and the data themselves. Some key aspects of each analysis are noted below.

Endline evaluation: the OED analysis systematically examined the coverage and targeting of GMP, nutrition education, and SF of children and pregnant women, including examination of leakage in the delivery of SF. It examined the changes in maternal knowledge and practices between BINP and non-BINP areas and applied two-stage multivariate models in an attempt to attribute some of these changes directly to the project itself and/or degree of participation/exposure among individual women. Further, it applied two-stage models to the knowledge-practice gap in an effort to examine the factors constraining mothers from translating nutrition knowledge into improved practices.[15] With respect to nutritional outcomes, the OED report applies multivariate models to test for project effects on child nutritional status, maternal weight gain, and birth weight. A major contribution of the OED analysis is the use of HKI national nutritional surveillance data as a "super-control" for evaluating the impact of BINP on child nutritional status. This is important because of the problems associated with the baseline data in the endline evaluation (see appendix C and boxes 1 and 2 in section 2).

SCF survey: the OED applied multivariate, two-stage models to reanalyze data from the SCF cross-sectional endline survey with a particular interest in testing the effects of the project on knowledge and practices and exploring some of the resource, mobility, and other factors associated with the knowledge-practice gap. In addition, the report reanalyzes the SCF register data to explore the factors associated with the changes in weight among children receiving supplementary feeding.

[15] The gap analyses are motivated by a concern that the BINP may have improved knowledge but was not successful, or as successful, in improving practices. The OED analyses are restricted to the subset of women responding correctly to a given knowledge question, and they seek to identify the characteristics of women who do versus do not also report the corresponding practice. Those analyses are not reviewed here because (1) they ignore all women that report the practice but did not respond appropriately to the knowledge question; (2) they assume all behavior change, or the only relevant behavior change, must be mediated via knowledge acquisition; (3) they assume women can accurately represent their knowledge and practices concerning complex issues in response to close-ended questions; (4) they assume that the other important influences on behavior change are reflected in the exogenous control variables; (5) they ignore the influence of ceiling effects and differential rates of change in knowledge versus practice; and (6) for all these reasons, they systematically underestimate the potential effect of the project as reflected in tables 3.4 and 3.5. The weaknesses in this working theory (that knowledge leads to behavior change), which has guided both the design of the nutrition education component and the evaluation, are discussed in section 3.1 of the present study.

DHS data: OED applied multivariate models (OLS and two-stage) to the 1997 and 2000 DHS data in an effort to test for differences in nutritional status among children living in BINP and non-BINP areas of the country. Although they are only cross-sectional data, these analyses represent one of the strongest tests for BINP effects because of the nationally uniform and generally high quality of DHS data, the wider range of variables available for use in the models, and the use of a national control group, which has the potential to better account for national secular trends. In addition, the analysis not only tested for project main effects, it also explored interactions in an effort to identify variation in results among subgroups and some of the factors potentially responsible for differential project effects.

APPENDIX E. DESCRIPTION OF THE IMED

The IMED report covered the implementation and operational process aspects of BINP more than the other evaluations reviewed here. This evaluation applied both quantitative (surveys) and qualitative (interviews) methods in data collection with the aim of assessing both the process and outcome variables of BINP. But the timing and cross-sectional design of this evaluation (data collection February to June 2004) appeared to have prevented it from assessing the process variables, which would demand operations research during the years when the project was being implemented.

In regard to sampling, the evaluation selected its sample BINP *Thanas* from all four phases of BINP. Sampling was done more from later-phase *Thanas*, and it was unclear whether a one-on-one matching was done in the selection of the control *Thanas* (table 1in IMED). Although the evaluation claimed high comparability between BINP and control *Thanas*, the number of government-run satellite clinics, schools, and colleges were found to be much higher in BINP area despite the comparable size of the population in the control area (table 3 in IMED). The data also show that a similar total number of H&N programs were going on in both BINP and control *Thanas* (26 vs. 19), indicating that while BINP was in process, there were still other active H&N programs in BINP area and outside. In addition, some NGOs (BRAC and SOPIRET) operated in both areas. There appeared to be more NGOs active in BINP *Thanas* than in the controls.

Analysis and presentation of quantitative findings in the evaluations appeared to be very limited. Except for awareness and practices, other variables (including nutritional status) were presented only in simple descriptive statistics without significant tests of the differences between BINP and control areas.

Qualitative data, on the other hand, provided a great deal of information on the process aspects of the project. The interviews with CNPs, CNOs, and TPOs revealed obstacles that CNPs faced in cooperating with other health facilities, running the CNCs, ensuring attendance, and delivering the many services they were responsible for. CNOs expressed interest in requiring a list of tools and equipment. TPOs emphasized that CNPs/CNOs should be better paid and the food supplements be replaced by more education and training to women. These results could be made more useful by exploring what the perceived reasons are for these obstacles and the validity of the suggestions and requirements, which could inform future intervention design and implementation. Focus group discussions (FGD) by the group of diverse participants from the VNMC and UNMC levels revealed that these committee members had perceived many positive health and nutrition changes as a result of BINP. And the problems identified by them appeared to agree with those identified in interviews with CNPs/CNOs and TPOs on attendance, decreased quality of food supplements, payment issues, and so on. Some recommendations were also generated from the FGDs on community mobilization, cooperation, and supervision, although working mechanisms were not specified.

The report suffered from not backing up some of its statements with evidence, and sometimes the statements contradict their own and others' data. The following are a few examples:

1. "Effectiveness of the community-based infrastructure—all CNCs, VNMCs, and UNMCs were found operating."
2. "Efficiencies of the manpower (CNPs and CNOs) at the grassroots level, whose development through training was fully satisfactory."
3. "Findings suggest that the major focus of BINP interventions was on the poor."

4. The better practice was due to that "the women trained from CNCs might have acted as Change Agents in their community influencing directly the behavior of other women."
5. "… Awareness is comparable on many variables in both areas" while "the level of practices in BINP area compared to control area is very high."

Surprisingly little data analysis on the changes in nutritional status was presented in this evaluation except for a table on moderate and severe underweight, wasting, and stunting as measured by the corresponding Z-scores in children (table 11 IMED). The data appeared fairly incomparable with the results from the endline report by Karim et al. (table AE.1) and raised the question of possible mistakes in reporting. Because all children under five were included in this evaluation, a higher malnutrition rate should be expected here as opposed to the data in Karim et al., which included only U2 children (6–23 mo) in data analyses. The very low stunting rate and very high wasting rate, compared with all other surveys in BINP, strongly suggest systematic overestimation of length measurements in the field.

Women's BMI was not reported except a statement saying the value was not different between the two areas. LBW data were not collected because of the time and resource constraints this evaluation faced. Provision of supplementary foods was not addressed in the evaluation although it was reported that all women's groups were found making the food supplements. Most important, the linkages between awareness/practice and/or receipt of supplementary foods and women's and children's nutritional status were not examined.

Table E.1. Inconsistency in Child Nutritional Status by IMED and Endline Evaluation

Indicator	Endline evaluation (1st phase Thanas) 6–23mo		IMED evaluation (all-phase Thanas) U5 *		SCF evaluation 1st phase Thanas (6–23 mo)	
	BINP	Control	BINP	Control	BINP	Control
Severe Stunting (<-3 SD)	15.8	22.3	7.0	8.0	11.6	12.4
Moderate Stunting (<-3, <-2 SD)	29.9	28.7	7.7	7.4	27.5	27.6
Severe Underwt (<-3 SD)	11.8	13.3	14.0	14.0	11.4	12.2
Moderate Underwt (<-3, <-2 SD)	35.0	36.1	13.0	12.9	35.2	36.3
Severe Wasting (<-3 SD)	0.6	0.7	12.2	12.4	1.0	1.2
Moderate Wasting (<-3, <-2 SD)	10.9	8.0	28.5	29.6	13.3	14.1

Sources: IMED table 11; Karim et al. tables VII-5, 7, and 8; and SCF table 5.

Implementation Issues

This evaluation showed an explicit interest/objective in the process aspects of the project. First, the report documented that many of the field activities were carried out with the NGOs' own IEC materials as a result of the delay in receiving centrally produced materials, though the quality of these existing materials is unclear. Second, in regard to the training component, more details on the contents and methods of the various kinds of training given and some measurements of the change in trainees' competence in their functions would be more useful information than just a percentage of how many CNPs and CNOs were trained. Third, no informative data on capacity

building of national-level institutions were recorded in the evaluation. This observation raised the question of how to justify the spending on this aspect of BINP. Fourth, the evaluation found that no specific strategy to use the operations research studies had been developed in support of BINP activities. Fifth, contradictory to the overall statement the report made that the field infrastructure—all CNCs, VNMCs, and UNMCs were found operating, the report's chapter on process outputs of the CBNC stated that only 80 % of the VNMCs, CNC were active. This inconsistency and the definition of "operating" were not explained. Last, it would be desirable to see some comments on the lessons learned from the major delay (16 months) in the commencement of BINP as well as the institutional structural issues of BINP at the central level.

APPENDIX F. DESCRIPTION OF SCF REPORT

The Save the Children's evaluation of several World Bank–funded community-based nutrition interventions included BINP (SCF 2003). This study was not set out to be a comprehensive evaluation of BINP, but rather with a focus on the results related to child nutrition, the supplementary feeding (SF), and mothers' health and nutrition (H&N) knowledge and practices (K&P).

The design of the evaluation is a combination of critiquing the results of baseline and midterm reports, analyzing the register data from BINP (for SF), and an SCF cross-sectional survey in 2001/2. The critique on the midterm survey included the lack of a significance test and possible sampling bias.

The selection of BINP and control *Thanas* for the cross-sectional survey was briefly described. SCF randomly selected three *Thanas* of the six first-phase BINP *Thanas*. The controls were matched from a nearby area based on similar socioeconomic status (SES) and the condition of no major ongoing nutrition projects, in the hope of minimizing unequal baseline malnutrition prevalences. Yet in the report, data on only six household-level indicators were presented as evidence of their matching efforts.

SCF did not find any significant differences between BINP and non-BINP *Thanas* on severe and moderate malnutrition among young children (6–23 mo) (table 5 SCF). SCF results and comparison with endline and IMED evaluations are shown in appendix table E.1.

Concerning mothers' K&P, the evaluation found that BINP samples reported significantly higher positive response in more practice variables than knowledge variables (11 vs. 6). Yet the number of variables the evaluation had screened in total was unreported. No further analysis was made to test the association between the K&P variables and nutritional status.

The SCF evaluation was unique in that it collected register data for all children (n = 998) who had ever enrolled in SF from their survey samples. The average pre- and post-SF WAZs for younger (under 12 mo) children were not different, suggesting that SF had no effect in improving their WAZ. However, the WAZ in older (12–23 mo) children increased significantly. In addition, the nutritional status of 42 % severely malnourished and 70 % moderately malnourished children remained unchanged after they graduated from SF.

The fact that on the day of the SCF survey, only 22 % of the severely malnourished children (6–23 mo) were enrolled in SF suggests that some SF implementation issues perhaps need to be reconsidered. It would be useful to report whether those 78 % of children had ever been enrolled in SF, (if so) their entry and exit WAZs, dropout, and duration on SF, especially the issue of replacing home foods with the supplements. Also, a discussion on the potential effect of the graduation criteria (if any) of SF and the policy of a maximum feeding period of four months (three months with a possible extension of one more month) would be useful, along with a discussion of releasing children who had not achieved expected change.

Concerning project implementation, SCF reported that its evaluation did not identify major problems with the capacity of the community workers without providing details on what capacities were studied. The evaluation reported that the interviews with CNPs revealed that the CNPs spent a "disproportionate amount of time on supplementary feeding," but no further data were found in the annex of the report, as indicated in the main text.

APPENDIX G. DESCRIPTION OF NNP BASELINE EVALUATION

The baseline evaluation of the National Nutrition Program (NNP) was conducted to obtain valid data for new NNP *Thanas*, old BINP *Thanas*, and control *Thana*s. It was not an evaluation of BINP outcome and effect but it aimed to supply comparisons between the areas as a baseline for future evaluations. Nonetheless, these data provide yet another source of data for assessing the effectiveness of BINP, the expectation being that nutritional status, knowledge, and practices should be better in BINP areas than in the NNP or control areas.

The number of *Thanas* selected to represent these three areas were 44, 53 (all *Thanas* except those in the first phase), and 16, respectively. The selection of NNP and control *Thanas* was not described in the evaluation report, but these *Thanas* appeared to be selected from all divisions in Bangladesh.

Data collection was conducted in five phases and followed mainly a cross-sectional design with an exception of longitudinal data collection on weight gain of women in the third trimester.

The evaluation report is composed of nearly 200 tables comparing a large number of variables among the three areas. For unknown reasons, results of significant tests were not always reported in most of the tables, except in anthropometric measurements. It should be noted here that only the initial results were available to this review. Data were stratified by age and gender in child anthropometry tables. Results were not consistently superior in BINP area, in fact, very few differences were found between BINP versus NNP or control area. Some of that may be due to the gap in time between BINP completion and NNP inception, during which program operations were interrupted or ceased altogether.

APPENDIX H. DESIGN AND IMPLEMENTATION CHANGES BETWEEN BINP AND NNP

	BINP	NNP
1	Run by project office headed by project director (Deputy Secretary level)	Run by project mgt unit headed by executive director (Joint Secretary /Addlsecy level)
2	Executive director has financial authority/control	Executive director has financial authority/control
3	Three diff components implemented by 3 diff NGOs in the same Upazila	Three diff components implemented by same NGOs in one Upazila
4	The project directors of the subcomponents had separate offices	Project directors of the subprojects have office in the same building
5	SF for PLW given for 6 months postpartum	SF for PLW given for 2 months postpartum
6	Vulnerable Group Development -NNP linkages absent	Vulnerable Group Development -NNP implemented in 53 upazilas for 2 yrs, the rest included in 2005
7	Implemented in phases: first 6 Upazilas, then 17, then 17, and fourth phase 19	All 59 (+2) BINP Upazilas and 44 new upz| implemented at one time
8	Training component implemented directly by project office	Diff parts of training component contracted out to IPHN, BRAC, and NIPORT by project office
9	BCC component implemented by Project Director office with a Deputy Project Director being in charge. Did not succeed well	BCC component contracted out to UNICEF. Quality is still a question.
10	Newlywed couple intervention in only one Upazila	Newlywed couple intervention in all Upazilas
11	Adolescent girls forum only in a few selected Upazilas	Adolescent girls forum in all Upazilas
12	Implementation began soon after credit effectiveness	Implementation began 3 yrs after credit effectiveness
13	Upazila selection was scattered	Upazila selection was clustered and awarded to NGOs in clusters
14	Feeding on site	Feeding on site
15	Food composition unchanged in BINP and NNP	Food composition unchanged in BINP and NNP
16	LBW prevention strategy	LBW prevention strategy
17	Growth card developed after extensive consultation, and it followed the Gomez classification	Growth card changed to Z score. As a result, the number of severe malnutrition cases increased. So for SF, the Gomez criteria are followed
18	Referral of malnourished children and mothers with other complications not formalized. Some NGOs on their own initiative setting up referral system through a cooperative set up in each Upazila	Referral is still not working.
19	First three phases used salter scale and bathroom scale. Fourth phase used the *Uniscale*	Using uniscales only. However, target on use was inadequate according to reports from field.

APPENDIX I. BINP UPAZILAS BY IMPLEMENTING NGOS

Division	BINP	Implementing NGO(s)
Dhaka	20	
	Bhedargonj/(Vhadergonj?) (phase II)	TMSS
	Gopalgonj Sadar (phase II)	Proshika
	Bhanga (phase II)	Proshika
	Madaripur Sadar (phase II)	BRAC
	Tarail (phase II)	SARD
	Shibpur (phase II)	BRAC
	Narshindi Sadar (phase II)	SHED
	Monohordi (phase III)	BRAC
	Damudia (phase III)	TMSS
	Tungipara (phase III)	Proshika
	Modhukhali (phase III)	BRAC
	Jamalpur Sadar (phase III)	BRAC
	Nikli (phase III)	BRAC
	Bhairab (phase IV)	SARD
	Kuliarchar (phase IV)	SARD
	Bhaluka (phase IV)	Proshika
	Kapasia (phase IV)	BRAC
	Nakla (phase IV)	BRAC
	Tongibari (phase IV)	Proshika
Barisal	4	
	Barisal Sadar (phase II)	Proshika
	Amtali (phase III)	BRAC
	Gournadi (phase III)	SARD
	Jhalkathi Sadar (phase IV)	BRAC
Chittagong	9	
	Chandpur Sadar (phase II)	BRAC
	Fatikchari (phase II)	SHED
	Teknaf (phase II)	SHED
	Chokoria (phase III)	SHED
	Matlab (phase III)	BRAC
	Ramu (phase IV)	SHED
	Chaddagram (phase IV)	BRAC
	Laxmipur Sadar (phase IV)	Proshika
	Kachua (phase IV)	BRAC
Khulna	6	
	Dacobe (phase II)	BRAC
	Sreepur (phase II)	BRAC
	Dumuria (phase II)	BRAC
	Baliaghata (phase III)	BRAC
	Rupsha (phase III)	BRAC

	Fultala (phase IV)	BRAC
Rajshahi	11	
	Adamdighi (phase II)	BRAC
	Tetulia (phase II)	TMSS
	Sherpur (phase III)	TMSS
	Sariakandi (phase III)	BRAC
	Boda (phase III)	TMSS
	Natore Sadar (phase IV)	TMSS
	Sirajgonj Sadar (phase IV)	Proshika
	Rowmari (phase IV)	BRAC
	Mithapukur (phase IV)	TMSS
	Dimla (phase IV)	BRAC
	Lalmonirhat sadar (phase IV)	Proshika
Sylhet	3	
	Kulaura (phase II)	BRAC
	Sreemangal (phase III)	BRAC
	Sunamganj Sadar (phase IV)	BRAC
Sub-Total	53	27 BRAC/26 others
Sub-Total	6 (phase I)	6 BRAC
TOTAL	59 (phase I–V)	33 BRAC/ 26 others

BRAC Upazilas: Phase I – 6
Phase II – 8
Phase III – 10
Phase IV – 9

APPENDIX J. SAMPLE OF MISSED OPPORTUNITIES FOR MIDSTREAM ADJUSTMENTS

While this study was being developed, some information related to the implementation of BINP emerged from various and mostly informal sources that were not documented or only briefly so in many of the formal evaluations. This information, if it had been noticed and addressed in a timely fashion, might have resulted in improved efficiency and effectiveness of the project. The instances noted in this review represent but a small fraction of the "local knowledge" that typically exists about programs, much of which lends itself to problem solving if structures, mechanisms, and capacities for doing so have been built into the program. A few examples follow:

1. Problems with the palatability and acceptance of the food supplements have been noted.
2. The turnover of project officer at the lower level was high, and it affected the stability of management, transfer of information, maintaining of motivation, and problem solving.
3. An operations research (OR) study reported a dramatic increase in energy intake for women receiving food supplements (600–800 kcals higher than pre-SF energy intake) and suggested that this did not replace any of the usual intake by these women. Yet, this did not result in the expected weight gain in the women. This should have initiated a follow-up study, an assessment of the quality of the OR studies in general, and/or a focused discussion about possible design changes, but it did not.
4. Women's groups (that produce supplemental food) did not update the roster of names for food supplement recipients as they should have. This made it difficult to track the distribution of food and assess targeting and leakage.
5. Use of the results from the OR was minimal. The information was not made publicly available, and the quality of some of the OR appears questionable.
6. There were impressions that the CNPs who worked in GOB-run Upazilas (i.e., were paid by the government) felt that they were government employees and did not fully follow the supervision from the NGOs.
7. Some implementing NGOs rounded birth weight to exactly 2,500 grams to meet the LBW objective. This massive rounding resulted in an unusually high frequency of babies born with an exact 2,500-gram birth weight.

REFERENCES

Ahmed, F. 2000. "Anaemia in Bangladesh: A Review of Prevalence and Aetiology." *Public Health Nutrition* 3 (4): 385–93.

Allen, L., and S. Gillespie. 2001. "What works? A Review of the Efficacy and Effectiveness of Nutrition Interventions." United Nations Administrative Committee on Coordination / Sub-Committee on Nutrition, Asia Development Bank, Manila

Behrman, J.R.; Alderman, H.; Hoddinott, J., *Copenhagen Consensus: hunger and malnutrition, Economic reasons for fighting hunger and malnutrition*,

Bhargava, A., H. E. Bouis, N. S. Scrimshaw. 2001. "Dietary Intakes and Socioeconomic Factors Are Associated with the Hemoglobin Concentration Of Bangladeshi Women." *Journal of nutrition* 131 (3): 758–64.

Briend, A., Kh. Z. Hasan, K. M. A. Aziz, B. A. Hoque, and F. J. Henry. 1989. "Measuring Change in Nutritional Status: A Comparison of Different Anthropometric Indices and the Sample Sizes Required." *Eur J Clin Nutr* 43: 769–78.

Brown, K. H., R. E. Black, and S. Becker. 1982. "Seasonal Changes in Nutritional Status and the Prevalence of Malnutrition in a Longitudinal Study of Young Children in Rural Bangladesh." *Am J Clin Nutr* 36 (8): 303–13.

Bryce, J., C. G. Victora, J. P. Habicht, P. Vaughan, and R. E. Black. 2004. "The Multi-Country Evaluation of the Integrated Management of Childhood Illness Strategy: Lessons for the Evaluation of Public Health Interventions." *American Journal of Public Health* 94 (3): 406–15.

Bushamuka, V. N., S. de Pee, A. Talukdar, L. Kiess, D. Panagides, A. Taher, and M. Bloem. 2005. "Impact of a Homestead Gardening Program on Household Food Security and Empowerment of Women in Bangladesh." *Food Nutr Bull* 26 (1): 17–25

Caldwell, J. C. 1993. "Health Transition: The Cultural, Social and Behavioral Determinants of Health in the Third World." *Social Science and Medicine* 36: 125–35.

Caulfield, L. E., S. L. Huffman, and E. G. Piwoz. 1999. *Interventions to Improve Complementary Food Intakes of Six to 12-Month-Old Infants in Developing Countries*. Washington, DC: LINKAGES Project, Academy for Educational Development.

Caulfield, L.E., S.A. Richard, J.A. Rivera, P. Musgrove, and R.E. Black, 2005, Stunting, Wasting, and Micronutrient Deficiency Disorders, Chapter 28 in 2nd edition, *Disease Control Priorities in Developing Countries*, Disease Control Priorities Project (DCPP) Fogarty International Center, National Institutes of Health, Bethesda, MD.

Contento, I. 1995. "The Effectiveness of Nutrition Education and Implications for Nutrition Education Policy, Programs and Research: A Review of Research." *J Nutr Educ* 27 (6): 278–365.

Fellows, W. D. 1985. *Partition of Nutrients Between Dam and Fetuses in Rats Underfed Acutely and Chronically.* MS thesis. Cornell University, Ithaca, NY.

Griffiths, M., K. Dickin, and M. Favin. 1996. *Promoting the Growth of Children: What Works—Rationale and Guidance for Programs*. Washington DC: Human Development Department, World Bank.

Gross, R., Benade S., Lopez G., and the International Research on Infant Supplementation initiative. The International Research on Infant Supplementation initiative. *J Nutr*. 2005 Mar; 135(3):628S-630S.

Gwatkin, D., R, S. Rutstein, K. Johnson, E. A. Suliman, and A. Wagstaff. 2003. *Socio-economic Differences in Health, Nutrition, and Population,* 2nd ed. Washington, DC: World Bank.

Helen Keller International. 1999a. *Vitamin A Status Throughout the Lifecycle in Rural Bangladesh*. Bangladesh: Helen Keller International.

Helen Keller International. 1999b. *Iron Deficiency Anemia Throughout the Lifecycle in Rural Bangladesh*. Bangladesh: Helen Keller International.

Helen Keller International. 2001. *The Nutrition Surveillance Project in Bangladesh in 1999—Toward the Goals of the 1990 World Summit for Children*. Dhaka, Bangladesh: Helen Keller International.

Hobcraft, J. 1993. "Women's Education, Child Welfare and Child Survival: A Review of the Evidence." *Health Transition Review* 3 (2): 159–75.

ICDDR, B: Centre for Health and Population Research. 2005. *Baseline Survey of National Nutrition Programme: Initial Results.* Dhaka, Bangladesh.

IMED (Implementation Monitoring and Evaluation Division), Ministry of Planning. 2004. *Impact Evaluation of Bangladesh Integrated Nutrition Project*. Dhaka, Bangladesh: IMED.

Jones G, Steketee RW, Black RE, Bhutta ZA, Morris SS; Bellagio Child Survival Study Group., How many child deaths can we prevent this year? *Lancet.* 2003 Jul 5; 362(9377): 65-71.

Kabir, A. F. M. I., and F. J. Levinson. Lessons Learned in the Bangladesh Integrated Nutrition Project, unpublished paper.

Karim, R., F. J. Levinson, M. Akhtaruzzman, and S. Lamstein. 2002. *Process and Impact Study of BINP Garden and Poultry Activities.* BINP Operations Research Project No. 18. Institute of Nutrition and Food Science, University of Dhaka and The Friedman School of Nutrition Science and Policy, Tufts University.

Karim, R., S. A. Lamstein, M. Akhtaruzzman, K. A. Rahman, and A. Nural. 2003. *The Bangladesh Integrated Nutrition Project Community-Based Nutrition Component—Endline Evaluation*. The Institute of Nutrition and Food Science, University of Dhaka and The Friedman School of Nutrition Science and Policy, Tufts University.

Mason, J.B., Sanders, D., Musgrove, P., Soekirman, and R.G. Galloway. 2005. Community Health and Nutrition Programs, chapter for in 2nd edition, *Disease Control Priorities in Developing Countries*, Disease Control Priorities Project (DCPP), Fogarty International Center, National Institutes of Health, Bethesda, MD

Matta, N., R. Ashkenas, and J.-F. Rischard. 2000. Building Client Capacity Through Results. Internal Report. Washington, DC: World Bank.

Nahar, S. Impact of Food Supplementation on Pregnancy Weight Gain and Birth Weight in Rural Bangladesh. PhD thesis, Cambridge University, 2003.

National Research Council. *Understanding Risk: Informing Decisions in a Democratic Society*. Washington, DC: National Academies Press, 1996

Olson, R. K. Developing Indicators that Predict Benefit from Prenatal Energy Supplementation. PhD thesis, Cornell University, Ithaca, NY, 1994.

Operations Evaluation Department (OED). 2005. *Maintaining Momentum to 2015? An Impact Evaluation of Interventions to Improve Maternal and Child Health and Nutrition Outcomes in Bangladesh*. Washington, DC: World Bank.

Pelto, G., and J. Backstrand..2003 "Interrelationships Between Power-Related and Beliefs-Related Factors Determine Nutrition in Populations." *Journal of Nutrition* 133: 297A–300S.

Ruel, M. 2001. *Can Food-Based Strategies Help Reduce Vitamin A and Iron Deficiencies? A Review of Recent Evidence*. Washington, DC: International Food Policy Research Institute.

Save the Children (SCF). 2003. *Thin on the Ground—Questioning the Evidence Behind World Bank-Funded Community Nutrition Projects in Bangladesh, Ethiopia and Uganda.* United Kingdom: SCF.

Schipani, S., F. van der Haar, S. Sinawat, and K. Maleevong. 2002. "Dietary Intake and Nutritional Status of Young Children in Families Practicing Mixed Home Gardening in Northeast Thailand." *Food Nutr Bull*. 23 (2): 175–80.

Skoufias, E. 2005. *PROGRESA and Its Impacts on the Welfare of Rural Households in Mexico*. Research Report No. 139. Washington, D.C.: International Food Policy Research Institute.

Smith, L., and L. Haddad. 2000. *Explaining Child Malnutrition in Developing Countries: A Cross-Country Analysis*. Washington, D.C.: International Food Policy Research Institute.

Smuts CM, Lombard CJ, Benade AJ, Dhansay MA, Berger J, Hop le T, Lopez de Romana G, Untoro J, Karyadi E, Erhardt J, Gross R; International Research on Infant Supplementation (IRIS) Study Group, Efficacy of a foodlet-based multiple micronutrient supplement for preventing growth faltering, anemia, and micronutrient deficiency of infants: the four country IRIS trial pooled data analysis, *J Nutr*. 2005 Mar;135(3):631S-638S

Thomas, D., J. Strauss, and M.-H. Henriques. 1991. "How Does Mother's Education Affect Child Height." *The Journal of Human Resources* 26: 183–211.

Winkvist, A., K. Rasmussen, and J. P. Habicht. 1992. "A New Definition of Maternal Depletion Syndrome." *American Journal of Public Health* 82 (5): 691–4.
World Health Organization (WHO). 1995. "Maternal Anthropometry and Pregnancy Outcomes: A WHO Collaborative Study." *Bulletin of the World Health Organization*, Supplement to volume 73.

World Bank *Bangladesh Development Series* publications

1. Bangladesh PRSP Forum Economic Update – Recent Developments and Future Perspectives, paper no 1, November 2005.

2. End of MFA Quotas: Key Issues and Strategic Options for Bangladesh Readymade Garment Industry, paper no 2, December 2005.

3. Bangladesh Country Water Resources Assistance Strategy, paper no 3, December 2005

4. Comparative Advantages of Health Care Provision, paper no 4, December 2005

5. Targeting Resources for the Poor in Bangladesh, paper no 5, December 2005

6. Improving Trade and Transport Efficiency – Understanding the Political Economy of Chittagong Port, paper no 6, December 2005

7. Revitalizing the Agricultural Technology System in Bangladesh, paper no 7, December 2005

8. The Bangladesh Integrated Nutrition Project – Effectiveness and Lessons, paper no 8, December 2005

All *Bangladesh Development Series* papers are available at www.worldbank.org.bd/bds

THE WORLD BANK
1818 H Street, N.W.
Washington,
D.C. 20433, USA
Tel: 1-202-473-1000
Fax: 1-202-477-6391
www.worldbank.org

WORLD BANK OFFICE DHAKA
Plot: E-32, Agargaon,
Sher-e-Bangla Nagar,
Dhaka-1207, Bangladesh
Tel: 880-2-8159001-28
Fax: 880-2-8159029-30
www.worldbank.org.bd