

November 18, 1996

**RECONCILING MISSION AND PVO REPORTING REQUIREMENTS FOR
THE MONITORING AND EVALUATION OF TITLE II FOOD AID PROGRAMS
IN ETHIOPIA**

Frank Riely¹

Food Security and Nutrition Monitoring (IMPACT) Project

I. Summary and Recommendations

While the Mission has made admirable progress and is working in close collaboration with the relevant PVOs, there is still a number of necessary steps done before an integrated information system is fully operational at the Mission-level. It is important to note that the obstacles faced by the Ethiopia Mission in integrating information from various PVO M&E systems are primarily a result of the diversity of the Title II program in Ethiopia, the novelty and complexity of the information management problem, and some limitations in Agency guidelines that have further tended to complicate Mission-level reporting.

In general, the objectives and indicators identified by the Mission are well-defined and measurable. The Mission has done a skillful job of defining its own strategic framework in a way that maximizes overlap with the objectives of the various PVO programs. In fact, the only real concern with the current information system as defined by the Mission is that it is perhaps too ambitious in some of its reporting requirements.

The somewhat generalizable conclusions can be summarized as follows:

(1) **Need for Integration of Food Aid into the Mission's Portfolio:** Because the PVO Title II programs are not fully integrated into the overall development portfolio, and because the Mission has been allocated only a small number of Strategic Objectives (SOs) for its entire program, the Title II activities have been relegated into their own special objective which defines food security in very general terms and identifies only a small number of general indicators at the SPO-level. It is very difficult to squeeze a large number of small and diverse programs targeted to address numerous dimensions of food security into a single SPO that necessarily defines food security in more general terms. It

¹ The Food Security and Nutrition Monitoring Project (IMPACT Project) is funded by the USAID, Bureau for Global Programs, Field Support and Research, Office of Health and Nutrition, Contract Number. DAN-5110-Q-00-0014, Delivery Order 16, Task Order 246-16-01. Funds for the technical assistance to the Ethiopia Mission was provided by the Bureau for Humanitarian Assistance, Office of Food for Peace, Development Program. This report is based on several IMPACT Project inputs to the Ethiopia Mission and most recently the October, 1996 visit of Frank Riely. Opinions expressed as those of the author and not meant to reflect USAID.

is more difficult to expect each of these diverse programs to contribute to impact measured by the small number of SPO-level indicators defined. This problem really highlights the difficulties arising from the lack of food aid integration and a number of more general shortcomings of current practice in food-assisted programs.

(2) Need for a clearer link between PVO and Missions Result's Frameworks: Because the Agency's strategic framework guidelines do not take into account the hierarchical relationship between Missions and PVOs, there are gaps between the indicators that address the PVOs' own information needs for their diverse programs and the information required by the Missions to show impact on food security in general. The SO-level indicators in the PVOs' DAPs overlap with the Mission's IR-level indicators. However, the DAPs and M&E strategies for the individual PVOs do not include the indicators that the Mission has defined at the SPO-level. Because the PVO programs are so small, secondary data, even where available, may not be able to adequately capture the impact of those programs. The implication is that PVOs are being required to collect additional information that may or may not be directly related to their own program needs.

(3) Need to ensure more Standardized M&E and Common Indicators: It appears that the development of PVO strategic frameworks and DAPs was prior to, or coincident with, the development of the Mission's strategic framework. Some of the PVOs completed their various baselines before the Mission had even defined its indicators. More effort is required to ensure that common indicators are defined in a consistent fashion with the framework and with other PVOs. Even though the PVOs participated fully in the Mission's strategic framework development, greater coordination is still required at the detailed technical level, on indicator definition, questionnaire development, sample design, etc., sufficient to ensure consistency across all operational units. In Ethiopia, it is likely that some PVOs will have to repeat at least part of their baselines to ensure that they collect information that is consistent with the ultimate Mission-level definitions.

In the future, it might be more appropriate to organize the entire process as a joint Title II program evaluation, with PVOs pooling their resources to come up with a common evaluation design. To the extent there is overlap in PVO program objectives, common modules of a questionnaire ought to be developed jointly. Individual PVOs could then be responsible for additional modules that address unique aspects of their programs. Common sampling methods could also be developed to ensure consistency in the interpretation of the data. This approach would likely lead to significant cost savings. A single contractor could be identified locally to conduct the joint Mission-level baseline for all common PVO program components. PVOs could share the cost of consultants to address other issues as well. At a minimum, however, PVOs probably need to coordinate, again, at a fairly detailed technical level within countries.

If there is a problem within countries of consistency between PVOs, there will likely be a corresponding problem of consistency across Missions when they begin to report results to USAID/Washington. This problem suggests the need for the development of "canned", but flexible, questionnaire modules for various program types

and indicators which can be used by PVOs worldwide, as well as a set of best practices in evaluation design. The IMPACT Project has begun this effort with the development of a set of indicator guides that will be available in the Spring of 1997.

(4) **More work is needed in Baseline and M&E Design:** In general, there is a need to focus more on the quality of the baseline and the rationale for the overall evaluation design. In Ethiopia, some PVOs clearly over-sampled in their program areas, either including substantially more households in the sample necessary to make reasonable inferences regarding impact, or by including households and communities which do not directly benefit from the program, without being able to distinguish them from those that do receive benefits. Other PVOs clearly under-sampled. Similarly, little thought was given to the possible need for the explicit inclusion of control groups or other aspects of evaluation design. An important problem that does not seem to have been addressed is that of including in a baseline households that have already been participating in the program for a number of years (where the bulk of the program's impact was likely to have been obtained prior to the baseline).

Baselines seem to be viewed and conducted along the lines of diagnostic "baseline surveys" traditionally used for program planning purposes, rather than as "baseline surveys" as understood in the context evaluation. The implication of this will be reduced evaluation rigor that will further reduce the likelihood that PVOs will actually be able to measure the impact of their programs, even for programs that actually have significant impacts. This is very much a strategic problem that may have important implications when these programs come up for renewal of funding and allocation decisions need to be made by USAID/Washington and the Mission.

(5) **Need to Streamline Reporting needed by the Mission:** Again, some of the reporting requirements outlined in the Ethiopia Mission's strategic framework seem overly ambitious and burdensome. These concerns highlight the need for more thinking about how to improve existing information from routine monitoring and better integrate it into the food aid decision-making process. PVOs and the Mission need to develop a rationale for the collection of this information, rather than reporting simply to meet the requirements to show annual progress toward impact, as mandated by Agency guidelines. Without an explicit rationale for how to use the information, there may be a tendency to collect too much, leading to an unnecessary diversion of resources from operations and a possible erosion of overall data quality. Clarification of what it means to "*show progress toward impact*"-- and whether that means annual reporting on trends in the impact indicators themselves -- would also be useful.

(6) **Using Food Security as a Objective or a Goal?:** Because the Agency has decided to define food security in such general terms, the Ethiopian conundrum raises the question whether food security in itself is an appropriate, measurable objective for food aid or other programs. It seems more appropriate to define food security as a program goal, rather than a specific objective. Given the rigor imposed by the evaluation

framework, it is critical to focus on specific dimensions of the food security problem, rather than make statements about food security in some general sense.

(7) Need for better use of Emergency Information for Development

Programming: Given the increasing use of Title II resources in emergency programs, the Mission's focus on the evaluation of emergency response capacity is very important. However, the evaluation of response capacity is inextricably linked to the evaluation of the population-level impacts of the emergency programs themselves and both problems need to be addressed together. The key to better-informed management of emergency food aid programs is in providing detailed information on the timing and location of commodity distributions and linking that to information on food security conditions in the PVO program areas, including information from PVO-based early warning systems. By making that linkage, early warning information can become an important tool for the on-going management of emergency programs, providing a stronger base for ex post impact evaluations. Improvements in the capacity to evaluate the impact of these programs will likely have important implications for developing more accurate estimates of emergency needs in future crises as well.

Next Steps: In terms of continued support by IMPACT to the Ethiopian Mission, there is ample scope, particularly in the review of baselines and the rigor of their sampling methodologies (Recommendation 3 in the report) and the finalization of the evaluation and reporting plan (Recommendation 4). There is likely to be more than adequate expertise within-country to address the issues in Recommendation 1. The REDSO office in Nairobi ought to be able to address the issues in Recommendations 2 and 6 (and, in fact, is in the process of doing so). IMPACT might also consider support in the further development of an M&E plan for Ethiopia's emergency programs (Recommendation 5), as well as the development of a capacity to conduct cost-benefit analysis (Recommendation 7).

II. Description of the Mission SPO Framework

The Ethiopia Mission faces a fairly daunting task in the definition of a Strategic Framework for its Title II food aid programs and the development of a consistent, efficient and effective information system to support food aid decision-making. The large number of PVO programs in Ethiopia and their diversity, in terms of both their program content and geographic distribution, implies a very complex information management problem that will undoubtedly exceed that required for the Mission's other Strategic Objectives. Given the varying level of development of individual PVO information systems to date, reconciliation of those individual sources of information into an effective information system at the Mission level will necessarily take time and will require close collaboration across all the institutions involved at a fairly detailed technical level. Over the long term, the management of that information system will require considerable on-going attention if it is to be used effectively as a decision-making tool.

The Ethiopia Mission has already undertaken the important first steps in this process with the joint development of a Strategic Framework for its Food Security Special Objective in collaboration with its PVO Title II Cooperating Sponsors. This SPO Framework was developed in a series of workshops with the relevant PVOs, utilizing key technical inputs provided by USAID/REDSO, the Greater Horn of Africa Initiative and the Food Security and Nutrition Monitoring (IMPACT) Project. This initial collaboration provides an important structure for the continued development and refinement of an integrated information system to support decisions at the Mission-level.

As indicated in the current Title II SPO Framework, there are two levels of reporting:

At the objective level, all PVOs, regardless of program focus, will be held accountable to report annually on:

- the nutritional status of children
- household food self-sufficiency
- number of households in "need"

This reporting requirement implies that every PVO will have to include these measures in their baseline and final evaluation surveys. The baseline information will have to be representative of the entire PVO program area and not just of segments of the program which are more or less directly related to one or the other indicator. These indicators will also be required for annual reporting, suggesting the need for annual surveys to obtain this information.

At the intermediate result level, PVOs will be required to report on the following indicators, also on an annual basis, according to the specific emphasis of their programs:

- | | |
|-----|--|
| IR1 | percent increase in crop yield
percent increase in total value of agricultural production |
| IR2 | percentage change in income levels
percentage change in asset holdings |
| IR3 | number of clients using selected health services |
| IR4 | land area protected or reclaimed |
| IR5 | improved estimation of needs
improved timing of response
improved targeting of relief |

It is important to emphasize that not every PVO has program activities which correspond to each of the IRs listed in the Mission framework. While all PVOs will be

required to report on each of the three indicators at the SPO-level, they will only be required to report on IR-level indicators which are consistent with their own program content. Therefore, a PVO which focuses only on activities in the agricultural sector, will only be required to report on the indicators listed under IR1, and will not be required to report on other IR-level indicators. Again, IR-level indicators will be included in annual reporting, which may imply annual surveys in some cases.

III. Discussion of Objective-level Indicators and Reporting

Because of the breadth of its other programs, the Mission has adopted a single special objective for its Title II food aid programs, which has been defined broadly as a food security objective. Given the Agency's strategic framework guidelines, the Mission has identified a set of indicators at the SPO-level thought to represent general food security conditions and a set of indicators at the IR-level which are more directly related to the PVO program activities which address specific dimensions of food security. In fact, to ensure consistency, the indicators listed at the IR-level in the Mission's strategic framework actually represent SO-level indicators in each of the PVOs' program frameworks.

This structure is not without complications, however. The PVOs' own individual strategic frameworks and DAPs provide for indicators only up to their SO-level, which, again, overlaps with the Mission's IR-level indicators. Therefore, within the context of the individual PVO DAPs, there is no specific intent or need to collect additional information that would correspond to the Mission's SPO-level indicators. Therefore, there is no obvious source of data to meet the Mission's own reporting requirements at the SPO-level.

The requirement for PVOs to collect and report on indicators at the Mission's SPO-level arguably goes beyond the direct information needs of the individual PVO programs. In the narrow context of their DAPs and the PVOs' own management needs, the point of PVO M&E systems is to collect information on impact that is as directly related to specific program outputs as possible. By doing so, PVOs would be better able to minimize potential distortion of impact measures from factors that are beyond the control of the program (such as climate) and will be better able to actually attribute changes in food security conditions to their specific program activities.

At the same time, because the Mission cannot have more than one objective for its Title II program activities—individual objectives might otherwise correspond to each of its defined intermediate results, for example—it is somewhat constrained to the use of a small set of more general indicators at the SPO-level. It is not immediately clear how to reconcile, on the one hand, the constraints placed on the Mission by the designation of a separate SPO for its Title II programs and its resulting information requirements, and, on the other hand, the need for streamlined information systems on the part of the PVOs.

This complication arises due to the following combination of factors:

- the very great difficulty in reconciling a large number of relatively small and diverse PVO programs into a clear operational focus at the Mission-level;
- Agency strategic framework guidelines which are clear for individual operational units, but which haven't adequately taken into account the hierarchical relationship between the Missions and Title II PVOs;
- the limited integration of food-assisted activities into the Mission's development portfolio, which might otherwise obviate the need for a separate Title II special objective; and
- the very broad definition given by the Agency to the term "food security", with its multiple and largely separable dimensions, which perhaps makes it a more appropriate focus at the goal- or sub-goal-level within a strategic framework, rather than as a clearly measurable objective.

Another rationale for the current indicator selection and reporting requirements at the SPO-level is that, by evaluating program performance on these combined criteria, conclusions may be drawn regarding whether the PVO programs are actually focused on the important elements that determine overall food security. This use of the SPO-level indicators is intended to support an orderly re-orientation of PVO program content over time, where necessary, rather than a more abrupt and disruptive revision of program activities.

In fact, as determined by the Mission, rather than being held strictly accountable to show impact on these indicators, all PVOs will be responsible only to report on these dimensions of food security. Therefore, in a hypothetical example, a PVO which is involved only in agricultural development projects will be responsible for reporting on nutritional status, although not accountable to demonstrate impact by the project completion date. If, upon final evaluation, improvements in nutritional status have not been observed along with improvements in agricultural productivity, the PVO may be required to address nutritional issues more directly in its follow-on DAP.

The rationale for this SPO Framework is understandable. There is an important concern regarding not only whether PVOs are able to meet their stated objectives, but whether the stated objectives are themselves appropriate to the actual food security conditions in their program areas. However, while this rationale is understandable, there are a number of possible concerns with this approach:

First, the purpose of defining indicators at the SPO-level is to evaluate whether impact has been achieved according to those measurable criteria. Given the indicators defined at the SPO-level, even though they are reasonable indicators of general food

security status, it is unlikely that the current set of PVO programs (or, therefore, the Mission) will be able to show impact across each indicator for their entire program areas.

Looking at the objectives and impacts defined by World Vision for its program, for example, the emphasis on crop production, incomes and family planning is unlikely to have a significant impact on child nutritional status within a 3-5 year period. There are simply too many possible confounding factors that might limit the impact of even substantially increased crop production and incomes on nutritional status. Although responsible to report on nutritional status, World Vision has apparently no program resources directed specifically to nutrition-related issues.

Second, the requirement to hold PVOs responsible for reporting on indicators for which they will not be held accountable to demonstrate impact seems to contradict other Agency guidelines on strategic framework development and the principles of 'managing by objective.' It is not clear that the Mission and the PVOs can avoid being held accountable to show impact on these indicators. Even where PVOs agree to these reporting requirements, there would seem to be some risk on the part of the Mission and the PVOs to list indicators in the SPO Framework for which they do not intend to demonstrate impact and for which they are not likely to be able to demonstrate impact, even if the intent was there.

Given the fact that the PVOs have defined a range of impact indicators at the intermediate result level which are more directly related to their program activities and are more likely to show impact, there seems to be little reason for the Mission and the PVOs to expose themselves to the unnecessary cost of collecting information on the defined objective-level indicators, or to the potential risk of being held accountable to them by USAID/Washington, according to the strategic framework guidelines. It is most likely that the Mission will show impact at the IR-level and the emphasis of PVO information systems should be focused on those indicators.

If PVOs are not to be held accountable to show impact at the SPO-level, these additional reporting requirements—where they are not directly related to individual PVO programs—represent an unnecessary burden on PVO information systems. This additional information burden could easily erode the quality of information across the entire M&E system and may reduce the likelihood that PVOs will be able to demonstrate impact at the IR-level as well (regardless of whether or not they actually achieved some impact) simply due to flaws in an over-burdened information system.

Third, there is also a particular concern regarding the definition of the indicator of population in need of food assistance based on Government estimates. Government criteria in making those estimates are primarily related to agro-climatic factors and are likely to vary significantly with rainfall. Only limited and uncertain account is made of individual household socioeconomic characteristics (household capacity to purchase food from other income sources is not fully considered in every part of the country) and no direct consideration is made of the nutritional status of children.

Therefore, to the extent that a PVO program has significant impact on incomes and nutritional status, that good performance would not necessarily influence the indicator of population in need. On the basis of that indicator, the impact of PVO programs could be severely under-estimated.² To the extent agricultural development programs would influence that measure, it might then be redundant to the measure of household food self-sufficiency (depending on how the self-sufficiency indicator is itself ultimately defined).

Fourth, the issue of whether the PVO program objectives are appropriate can and should be understood in other ways and well-before the final evaluation phase of the program. The design phase of the PVO programs should have included significant information-gathering regarding critical food security constraints in selected program areas, including the review of existing literature on food security in Ethiopia. PVO objectives should also be linked to Mission and Government priorities, as described in relevant national food security policy documents.

While there may be an argument for the explicit testing of whether PVO, Mission and Government objectives are appropriate, this is clearly a matter that should be fully addressed earlier in the program, not in the final evaluation phase. Especially when the information could and should be obtained in advance, there would seem to be little point in risking resources over a 3 to 5-year period only to discover *ex post* that the initial objectives of the program were inappropriate. That understanding and the process of re-designing the program, if necessary, will take a significant period of time, certainly more time than would be available between the analysis of final impact and the development of a follow-on DAP. Some suggestions on addressing this issue are provided below.

Finally, the interpretation of these indicators for their intended purpose is simply not clear without a range of other information that would have to be obtained through additional information-gathering activities. If nutritional status is found to be inadequate in an area where agricultural productivity is low, it is not clear that additional investments in the agricultural sector would still be appropriate. It may be that agricultural potential is also low in that area, providing limited returns to investments in the agricultural sector. In other cases it may be appropriate to address inadequate nutrition through additional investments in the agricultural sector, rather than an increased emphasis on nutrition programs.

It is, therefore, recommended that the use of the Mission's SPO-level indicators be reviewed and strong consideration be given to eliminating them, except as they might apply to individual programs at the IR-level. If, however, it is determined that all PVOs will be held responsible for reporting on these indicators, immediate attention must still be given to the following:

²The choice of this indicator does suggest the need (and an interest on the part of the PVOs and the Mission) to look more closely at disaster prevention and mitigation interventions. However, other than various PVO early warning systems currently in place, it is not clear that this set of issues are clearly addressed in the design of the current PVO programs.

- **Ensuring consistency in indicator definitions:** (a) is nutritional status to be measured according to child anthropometry and, if so, by what measure and for what age group; (b) will food self-sufficiency be defined in terms of the adequacy of household food production, some comprehensive measure of income, food consumption or expenditure, some measure of monthly food frequencies, or a simple subjective estimate derived from respondent recall; and (c) how will PVOs manipulate government needs estimates to conform to the set of communities participating in their individual food aid program areas, which are likely to only subsets of the regions for which government data are available?
- **Defining the frequency of reporting:** while it may be determined that SPO-level indicators should be included in PVO baselines, will these indicators have to be reported on an annual basis, even if annual sample surveys are required?
- **The “representatives” of SPO-level indicators:** although baseline data are more likely to be based on a sample that is representative of a PVO program area, to what extent will SPO-level indicators reported on an annual basis have to reflect conditions specific to the program area? For example, it may be relatively easy for PVOs to report on nutritional status in the communities where nutritional surveillance programs are underway, but will annual nutritional status data be required from other parts of the program area where only agricultural activities are being implemented? On the other hand, it may be possible for PVOs to report information through existing early warning systems, but which often include data from communities where the PVO is not otherwise operational. If these indicators are to be used to draw conclusions across all elements of the PVO programs, then it seems necessary to require that they be representative of the entire program area, and not just isolated pieces.

IV. Discussion of Intermediate Result-level Indicators and Reporting

There are two specific concerns regarding the intermediate result-level indicators: whether they are appropriate to the stated intermediate results and whether reporting requirements as currently stated are realistic.

First, the indicators defined for IR1 and IR2 do seem to be appropriate, depending on how they are ultimately defined. Again, as with the SPO-level indicators, some effort to ensure consistency in their definition and use across PVOs is necessary. Although methods to obtain the information may vary, the definition of crop yield is straightforward. In terms of how to define the indicator of the value of agricultural production, some decisions need to be made regarding which agricultural commodities should be included in the definition, particularly whether the value of livestock holdings (or off-take) should also be included. Similarly regarding the definition of the value total asset holdings for IR2, the exact composition of the asset base included in the defined indicator should be clarified and made consistent. For the measure of income, a decision must be made whether

income is defined as cash income only, or cash income plus the value of home-produced commodities.

For IR3, the stated indicator is an output measure, not an impact measure, and is, in fact, not a very good measure of output. The number of clients presenting at clinics and other service delivery points has little to do with program coverage, the quality of services provided by field staff, the availability of rations, vaccines, ORS and other materials necessary for the provision of quality services. In terms of performance monitoring, this indicator would actually say very little regarding progress toward impact on beneficiary well-being.

After reviewing the impact indicators already reported by the PVOs, the following indicators are proposed for IR3. Note, given existing PVO objectives and indicators, the following would require no additional information-gathering beyond that already planned:

- percentage increase in the population (mothers) demonstrating improved practices regarding targeted health and nutrition problems
- percentage decrease in the prevalence of targeted health and nutrition problems
- percentage increase in program coverage of services for targeted health and nutrition problems

Specific indicators reported by PVOs would depend on the particular emphasis of their individual health and nutrition programs. For example, if a PVO health program is not targeted toward diarrheal disease, it would not be expected to report on the use of ORS by mothers in the program area, nor on the prevalence of diarrhea directly. It may be a subject of further discussion between the Mission and the PVOs whether morbidity and malnutrition prevalences should be measured directly, or whether improved practice and program coverage is sufficient to show impact.

IR4 also lists indicators of outputs, rather than impacts. In fact, there is very little similarity in the impact indicators defined by the individual PVOs (CRS and FHI) reporting NRM activities. This issue requires some additional thought and the technical assistance currently planned by the Mission on this IR is certainly needed. There is a question whether the timing of the planned TA will be too late to influence the design of PVO NRM program baselines.

IR5 will be covered in detail below.

V. Implications for Annual Reporting

Some impact-related indicators defined in the Mission's strategic framework, such as that from growth monitoring activities, may be available on an annual or more frequent basis, to the extent that such information is already a regular part of program monitoring. Similarly, some information on savings and asset accumulation patterns may be obtainable through routine monitoring of rural savings and credit programs. A number of PVOs also currently conduct annual crop assessments in their program areas which might provide annual information on impact indicators for IR1. Of the PVOs interviewed, none were currently collecting annual information on incomes. And all thought that annual reporting on income and assets would be unduly burdensome, since it would likely involve annual surveys.

However, for any given impact indicator at the SPO-level or IR-level, there will be a number of PVOs not currently collecting information on an annual basis. Also, for a number of PVOs that are collecting information relevant to these indicators, that information is likely to be based only on a small part of the relevant program area (for example, CRS will be able to report annually on nutritional status based on their planned rolling KPC baseline design, but each year will represent a different segment of the program area), or on a population that goes beyond the program area (CARE's annual crop assessment includes a large number of communities which do not actually participate in its agricultural program). Therefore, in both cases, annual reporting may not be in any way reflective of progress towards impact. It is not clear whether these annual surveys could or should be amended to ensure that they are more representative of the program area.

With the exception of reporting on initial baseline survey results (preferably by Spring 1997), it is recommended that annual reporting be limited to information which can be obtained directly from existing or slight modifications to planned program monitoring. If impact indicators are available on an annual basis, they should reflect the entire program area (for SPO-level indicators) or the set of households participating in a particular program activity, say health programs (for indicators at the IR-level). A requirement to do annual crop, income, nutrition and other surveys, where they are not already an integral part of the PVO's management information system, is, again, likely to be unduly burdensome and could undermine the quality of other key information for program management.

Rather than assess trends in direct impact measures, annual or semi-annual reporting should be on the basis of key outputs which directly indicate progress toward impact. For example, rather than report annually on changes in the prevalence of diarrheal disease (where that information is not readily available from clinic records), reporting might focus on the coverage of the program (percentage of mothers trained in proper treatment of the disease, or the percent who presented with diarrhea who received ORS packets) and changes in the quality of health services (in terms of the training of field staff, measures of whether or not staff actually provide appropriate advice, and whether they

have on hand and manage properly adequate supplies related to the treatment of the disease).

Similarly for agricultural sector activities, annual information on coverage (percent adopting suggested practices and technologies, or numbers receiving extension visits), as well as measures of the quality of services (training of extension staff, whether or not they actually provide good advice, and the availability and delivery of key inputs at the appropriate time) might also be sufficient to indicate progress toward impact.

So, in summary, annual reporting might be based on the following indicators:

- impact-related indicators collected through routine program monitoring (growth monitoring, investment of loan funds, seedling survival rates, etc.)
- indicators of program coverage (percent of targeted population trained, receiving goods and services, other)
- indicators of program quality (training, knowledge and practice of field staff, availability of material inputs on a timely basis)

Some additional effort would be required to obtain a list of these indicators from individual PVOs and to ensure maximum compatibility for reporting up to the Mission level. This additional effort would be well worth the time in view of the huge information burden represented by annual reporting on impact measures.

While the list of annual reporting indicators above might be seen as minimalist, it is again important not to place too large an information burden on PVO programs. A more appropriate approach would be to focus on improving the quality and use of output indicators that are currently being reported through routine program monitoring. The focus on impact and the implementation of well-designed baselines and final evaluation surveys themselves represent a significant step forward, not only in directing activities to the population-level and in understanding impact, but also, potentially, in improved program management. Additional, excessive annual reporting requirements—by requiring annual surveys to report on impact indicators—risk not only undermining the quality of information from every source, but also diverting resources and management attention from the actual implementation of program activities.

The key to a successful M&E system is not only the quantity and quality of information obtained, but also the ability of PVOs, Missions and the Agency to integrate that information into the decision-making process. Rather than burden PVOs and the Mission with significantly greater quantities of information with very ambitious annual reporting requirements, more significant improvements in program management could be obtained through improving the quality of information obtained through routine program monitoring and, importantly, developing better strategies to use existing information from current annual reporting.

For each piece of additional information required in the SPO framework, a very strong rationale should be provided for its use in program decision-making, both by the PVO and by the Mission. Without that strong rationale and, over time, demonstrated feedback from information to action, the willingness to collect additional information and the quality of that information are likely to erode.

VI. Monitoring and Evaluation of Emergency Programs Under IR5

The issue of monitoring and evaluation of emergency programs should be broken down into separate questions:

- How to monitor and evaluate efforts to improve the institutional response capacity of PVOs and their counterparts?
- How to monitor and evaluate the impact of any eventual response on the populations affected by an emergency?
- How can information be developed to assist PVOs in the on-going management of emergency programs, particularly in the transitions into and out of the emergency response mode?

Institutional Capacity. Although the Mission has established a separate IR for emergency programs, only two PVOs (CARE and FHI) currently have listed improved emergency preparedness and response as an explicit strategic objective for their Title II food-assisted programs. While CARE's indicators suggest a focus on improved response capacity, the FHI indicator seems to have little to do with response capacity. Neither program activities seem to have much to do with the indicators as currently defined in the Mission's SPO framework.

At the same time, a number of PVOs have undertaken early warning systems and other related emergency preparedness and response initiatives. It is not entirely clear whether these other activities are being funded under the Title II program and whether, therefore, some review of the PVOs' DAPs is in order to determine the need to highlight these activities more explicitly in their individual strategic frameworks. In fact, part of the justification for PVO development programs in Ethiopia is to maintain their presence in the case of an emergency. Therefore, it might be appropriate for more of the PVOs in Ethiopia to list emergency response capacity as an explicit program objective. In any case, the evaluation of enhanced emergency response capacity should only be considered if significant Title II program resources are actually being directed to capacity-building activities.

The indicators defined in the current SPO Framework address the question of institutional response capacity. As such, however, they are somewhat problematic and

incomplete. In terms of the indicator related to improved needs assessment:: First, it is not clear how that would be measured as an indicator per se. There is very seldom any attempt to validate the results of needs assessments, which would be required to assess whether capacity has improved and which could only be determined through the evaluation of the population-level impact of an actual emergency operation. Second, given the outline of PVO DAPs currently, it is unclear to what extent any of the PVOs are currently involved in explicit efforts to improve their needs assessment capabilities. Therefore, it is not clear that the PVOs could show impact for this result.

In terms of the indicator of improved timing of response: Again, it would be difficult to actually measure this indicator (without a corresponding evaluation of population-level impacts of an actual emergency operation). Also, while PVOs do have early warning systems that can improve the timing of their requests for emergency resources, whether or not they actually receive those resources and the timing of their receipt of those resources is largely out of their immediate control. Because the timing of PVO response is also dependent on donor and government actions—essentially external confounding factors—it is questionable whether the individual PVOs should be held accountable for this indicator. At the same time, this would seem appropriate as a Mission-level indicator, to the extent that the Mission is actively working with the government to develop more efficient emergency response mechanisms at a more general level.

The indicator of improved targeting is very closely related to the issue of needs assessment, since any evaluation of targeting effectiveness necessarily implies being able to identify those in need. Again, it is not clear how this indicator could ever be effectively measured in any rigorous fashion, since there is unlikely to be any baseline measure of targeting effectiveness on which to base comparisons. Again, it will only be possible to develop some sense of targeting effectiveness in the context of an *ex post* evaluation of the population-level impact of an emergency program.

While there are problems with the operationalization of specific indicators defined in the SPO framework for IR5, the issues they represent are still clearly important dimensions of emergency response capacity. At the same time, the list of issues represented by those indicators is not complete enough to ensure adequate response capacity. A number of other institutional measures are also required to ensure adequate emergency response, including efficient commodity monitoring/accountability systems, explicit contingency plans and training of staff and counterparts in disaster prevention, mitigation, and relief strategies and interventions. Evidence of real linkages between disaster responses and development programs, as well as post-disaster transition strategies might also be important components of institutional response capacity. Some assessment of resource availability relative to likely size of emergency response needs under various scenarios might also be a component in the evaluation of response capacity. Finally, some assessment of the sustainability of the PVO efforts in improved response capacity might also be included in the overall evaluation of IR5.

In terms of the monitoring and evaluation of emergency response capacity, it may be more appropriate to use more qualitative methods, rather than quantitative indicators. PVOs might collaborate in the development of a more complete checklist of capacities required to ensure adequate emergency response. PVOs might also collaborate on the development of a joint work plan and time line for the attainment of each of those capacities. Annual monitoring of progress toward improved emergency response capacity might then depend on the extent to which each PVO has adequately addressed each of the issues on the checklist on a timely basis within the context of the benchmarks defined in the joint PVO time line. The formal evaluation of response capacity (which could be an on-going process) might depend not only on the existence of a needs assessment methodology or contingency plan, or on the numbers of staff trained on particular topics, but also on some independent assessment of whether those methodologies, strategies, contingency plans and training sessions were sufficiently rigorous to ensure an improved response.

Population-level Impacts. The ultimate measure of improved response capacity would be based on the evaluation of the population-level impacts of emergency food aid programs. It is important to develop information on population-level impact indicators during emergencies, not only to understand response capacity, but also to improve the on-going management of the emergency programs and to improve the information base for making future needs assessments. With a proper understanding of the population-level impact of an emergency response of a certain magnitude, it should be possible to develop a more accurate assessment of the magnitude of need in future emergencies.

For a variety of practical reasons, statistically rigorous conclusions regarding emergency program impacts are not likely to be easily obtained from traditional baseline and final evaluation surveys. In addition to the problem of developing a well-designed baseline in a very short time period, and the concern for any diversion of resources away from the actual response, a variety of factors arising during an emergency tends to bias the results of these types of surveys. For example, mortality and population migration have been shown to severely bias nutrition survey results during disasters. At the same time, the use of on-going surveys (monthly or quarterly) to monitor short-term trends in food security conditions during an emergency (rather than absolute levels) is fairly common, particularly surveys to collect anthropometric data. This information can be useful for general program management. If external confounding factors can be controlled for, some weak conclusions regarding program impact might be also obtained from the analysis of trends in this type of survey data, as well as on program monitoring data.

Additional information for the improved management of emergency programs may also be available from routine program monitoring activities. Information from growth monitoring in health programs, loan repayment rates in rural credit programs, school participation rates from education programs, demand for employment in food-for-work programs and other data easily obtained from on-going PVO monitoring activities can provide important insights into on-going changes in local food security conditions. Rather than collecting additional information, some effort might be required to determine how

best to analyze this data in the context of emergency program management. Where necessary, additional monitoring might still be implemented on an ad hoc basis during each crisis, such as regular rapid nutrition surveys.

A key, underutilized source of information for the management (and evaluation) of emergency programs is in PVO early warning systems. For example, movements in prices and coping behavior—information which is commonly developing in Early Warning (EW) activities—can be quite useful important to understand on-going emergency conditions. To optimize the use of this information, however, it may be necessary to undertake some slight revisions to current EW systems, asking questions in a slightly different way so that information is clearer, can be more easily entered into a computer data base, and can be analyzed in an historical context. As with other indicators in an M&E system, it is important also to ensure consistency in the definition of these indicators across PVOs as well.

In order to use this type of information in the monitoring and evaluation of emergency programs, it is critical that information on nutritional status, coping behavior and others can be effectively linked to actual food aid distributions. When linked to commodity tracking data, impact monitoring can be a particularly important management tool. Understanding where the food aid was delivered, when and how much, is essential in developing any kind of reasonable understanding of impact. By linking commodity tracking to other information on food security conditions, PVOs and the Mission could have a potentially sensitive tool for making on-going decisions about commodity distributions, particularly regarding the appropriate timing of transitions out of an emergency mode of operation.

In addition to their use in the on-going management of emergency programs, accurate commodity tracking information linked to other program information would be important inputs into *ex post* evaluations of the emergency programs. That data, combined with qualitative assessments by external experts, would provide perhaps the best possible basis for an evaluation of emergency programs. This final process of assessing the impact of the actual response would be critical in understanding future capacity-building and preparedness needs. Also, understanding, not only how much food was actually distributed in an emergency, but also whether that response was adequate to meet prevailing conditions, would provide a critical base on which to make more refined decisions on food aid needs in future crises.

In sum, IR5 information needs would seem to include not only the development of a method to monitor and evaluation institutional capacity, but also to monitor and evaluate the population-level impacts of actual emergency programs. Whether or not PVOs are better able to respond to emergencies would depend on whether they have developed a pre-determined set of capacities which have been evaluated for their rigor by external experts. The process of defining, developing and evaluating requisite capacities should be undertaken as a joint venture by the various PVOs in Ethiopia. The impact of PVO emergency programs at the population-level could be determined by linking commodity

tracking data to food security-related information developed from routine program monitoring, PVO early warning systems or, where necessary, ad hoc surveys. An *ex post* evaluation of the emergency programs should be a requirement for all PVOs and could be based in part on the monitoring information described above. Again, that information could also support improved, on-going management of emergency operations as well.

VII. Next Steps

The most important step necessary for the continued development of Title II information systems in Ethiopia is the establishment of some formal process of collaboration between the PVOs and the Mission in Ethiopia. While the development of the Mission's strategic framework in Ethiopia was based on an admirable collaborative effort, that process must carry through to specific technical details in the development of PVO information systems if the Mission is to make best use of the data available to it. With some modifications, the model of collaboration provided by the recent vulnerability mapping exercise in Ethiopia may provide a useful guide in this regard. PVOs should identify problems common to each, pooling their resources to address those problems and calling on outside expertise where necessary to develop joint solutions. Indeed, to the extent there is overlap in PVO program objectives and indicators, it may be useful view this process as the joint evaluation of all Title II programs, rather than as a series of independent evaluations of various PVO activities.

Although the Ethiopia Mission has made major strides in defining a common strategic framework and in initiating a joint evaluation design process, it is important to recognize the complexity and novelty of the problems faced in operationalizing this integrated information system at the Mission level. It is unlikely that this process can be finalized prior to the up-coming round of R4 reporting in early Spring 1997. The following activities and time references are presented in terms of what might be realistic, even with a very concentrated effort over the next 6-9 months.

Within the context of a joint evaluation design process, the following steps would seem to be particular priorities:

1. PVOs should initiate a process immediately to confirm whether or not their program objectives are appropriate to existing food security conditions in their program areas. The steps included in this process might include:
 - Identify other food security-related activities (government, other donors, other PVO/NGOs) currently in PVO program areas and establish their complementarity to existing PVO programs. (end-December 1996).
 - Implement a joint PVO food security situation analysis to identify and prioritize the most pressing constraints to improved food security in PVO program areas (this might be supplemented by data from PVO baselines and annual reporting). (Completed by end June 1997)

- Develop a mid-term program revision strategy, if indicated by the food security situation analysis, to ensure that program re-design process can begin before the completion of the current DAP and analysis of final evaluation data. (Strategy developed by end-December 1997; strategy implementation and program re-design through 1998).
 - Make necessary adjustments to program revisions based on final evaluation results for follow-on DAP.
2. Take immediate steps to finalize, clarify and make explicit the SPO Framework and PVO reporting requirements (to be completed by end-December 1996). It is especially important for the Mission to clarify the reporting frequency expected for both SPO-level and IR-level indicators, as well as the extent to which annually reported indicators should be representative of the entire PVO program areas. Requirements to do so include:
- For strategic-level indicators: (a) clarify and confirm their intended use in program decision-making at the PVO and Mission levels, (b) define required frequency of reporting, (c) clarify required level of program area to be covered by the indicator data, and (d) provide consistent definition for all PVOs.
 - For IR-level indicators: (a) clarify and confirm their intended use in program decision-making at the PVO and the Mission levels, (b) define required frequency of reporting, (c) provide consistent definitions, (d) clarify required level of program area to be covered by the indicator data, (e) review according to the following:
 - For income-related programs (IR2), review feasibility of collecting income and asset information on an annual basis.
 - For health and nutrition programs (IR3), upgrade reporting requirements to include impact-level indicators appropriate to the specific objectives of each PVO program, including: (a) percentage increase in the population (mothers) demonstrating improved practices, (b) percentage decrease in disease and malnutrition prevalence, and (c) percentage increase in program coverage of services.
 - For natural resource management programs (IR4), review current indicators, revise in consultation with proposed consultant, provide a consistent definition for all PVOs.
 - Define output-level indicators that show "progress toward impact", including:

- impact-related indicators collected through routine program monitoring (growth monitoring, investment of loan funds, seedling survival rates, etc.).
 - indicators of program coverage (percent of targeted population trained, receiving goods and services, other).
 - indicators of program quality (training, knowledge and practice of field staff, availability of material inputs on a timely basis).
3. Undertake a detailed review of both completed and proposed baselines (by end-February 1997) to assess whether or not they meet the reporting requirements of the SPO Framework and develop a technical assistance program, as warranted. Questions to be asked in the baseline review (with baseline reports and questionnaires in-hand) and addressed in training include:
- Do completed and proposed baselines include information that would enable the construction of both SPO-level and IR-level indicators in a fashion consistent with agreed-upon definitions?
 - Will sampling methods used or proposed allow for adequately rigorous conclusions regarding impact according to defined SPO-level and IR-level indicators?
 - Have the samples been drawn in a fashion consistent with the actual program area and which can distinguish program participants from non-participants?
 - Have control groups been specified in the course of developing a sample design for indicators, where necessary?
 - Have PVOs developed detailed analysis plans (report outlines which include proposed table shells) for both baseline and final evaluation survey results to ensure complete and rigorous data collection, data analysis and reporting in a way that is likely to optimize final evaluation results so that they can actually be used effectively by decision-makers?
4. Finalize evaluation strategies and reporting plans (by end-March 1997). This would include formal completion of the Mission's PMP, in close collaboration with relevant PVOs. If possible, it may be useful for PVOs to compile their agreed upon annual reporting data a few weeks prior to the R4 reporting deadline and review in a joint workshop to ensure consistency and define remaining steps necessary for the full integration of information at the Mission-level.
5. Develop a methodology for the monitoring and evaluation of emergency response capacity and the population-level impacts of emergency programs (end-September 1997). This would entail the following:

- development of a list of key capacities necessary to ensure adequate emergency response;
 - development of a joint PVO work plan to address any gaps in these capacities;
 - development of a common evaluation strategy, including the assessment of the rigor of training, contingency plans, analytical methodologies and others by external experts;
 - assessment of the current status of commodity tracking information and other program monitoring and their usefulness in emergency program management and evaluation;
 - revision of PVO early warning systems to promote greater relevance to program management and greater consistency across PVOs;
 - assessment of the need for additional, ad hoc data collection during an on-going emergency program and development of contingency plans for implementation; and
 - on the basis of likely available data, development of a detailed analysis plan for the monitoring and evaluation of PVO emergency programs in Ethiopia.
6. Assess requirements for efficient information management at the Mission-level (perhaps by end-June 1997). Given the complexity of the Title II program and the diversity of information sources, the Mission would be highly justified in assigning a technical expert to the management of this information system on a full-time basis. In addition, other steps necessary to support the Mission's management of this information include:
- the design of a database structure to accommodate data inflows from an integrated PVO information system,
 - the development of appropriate analytical methods to synthesize data from diverse PVOs, in order to report on results at the Mission level (i.e. would Mission-level results be reported as simple averages across all PVOs on common indicators, as weighted averages according to the quantities of food aid distributed by individual PVOs or some other method?),
 - the development of detailed analytical plans to support R4/PMP reporting requirements.
7. Finally, by end 1997 or mid-1998, when monitoring systems have been refined and have been operating efficiently for some time, it will be necessary to review the steps

required to enable effective program cost-benefit analysis, as the final step in the program evaluation. The steps toward cost-benefit analysis will include:

- development of financial accounting systems that enable full costing of PVO program outputs,
- review of necessary steps to link program outputs to impacts in a way that enables cost-benefit analysis, and
- review of necessary steps to link cost-benefit analysis to the development of follow-on DAP benchmarks.

BCD:\IMPACT\BHR\ETHIOPIA\FRNOV18.RPT(WP51)Nov.18'96