

AGENCY FOR INTERNATIONAL DEVELOPMENT  
 WASHINGTON, D. C. 20523  
**BIBLIOGRAPHIC INPUT SHEET**

FOR AID USE ONLY

1. SUBJECT CLASSIFICATION	A. PRIMARY Public Health
	B. SECONDARY General Public Health

2. TITLE AND SUBTITLE  
 Evaluation Plan for the DEIDS and Related Projects

3. AUTHOR(S)

Klaus, David J.

4. DOCUMENT DATE July 1974	5. NUMBER OF PAGES 87p.	6. ARC NUMBER ARC 614-K63
-------------------------------	----------------------------	---------------------------------

7. REFERENCE ORGANIZATION NAME AND ADDRESS  
 American Institutes for Research  
 Washington, D.C. 20016

8. SUPPLEMENTARY NOTES (Sponsoring Organization, Publishers, Availability)

9. ABSTRACT

The Development and Evaluation of Integrated Delivery Systems (DEIDS) Program has been initiated to explore new methods for providing low-cost, far-reaching health services in the developing countries. These efforts are designed to improve overall health and well-being, particularly for those with very low incomes, by furnishing needed medical, nutritional and family planning services where these are not now available. The basic approach of the DEIDS Projects and similar Integrated Health Delivery Programs is the expansion of indigenous health services through the use of innovative techniques intended to increase their outreach, effectiveness and impact.

This report describes the development of the model, illustrates its application using the Ecuador DEIDS Project as an example, discusses the features and utilization of the model in planning future health service delivery programs, and presents recommendations as to further steps which should be taken in support of the DEIDS effort.

10. CONTROL NUMBER PN-AAB-157	11. PRICE OF DOCUMENT
12. DESCRIPTORS Integrated Delivery Systems, Health Services, Model, Ecuador, Developing Countries	13. PROJECT NUMBER 930-11-995-030
	14. CONTRACT NUMBER AID/CM/otr-C-73-201
	15. TYPE OF DOCUMENT Evaluation Plan

411  
K63

AIR 46000-7/74-FR

EVALUATION PLAN  
for the  
DEIDS and Related Projects

Final Report  
to the  
United States Agency for International Development  
Work Order No. 5 of Contract No. AID/cm/otr-C-73-201  
Project No. 930-11-995-030

David J. Klaus  
American Institutes for Research  
Washington, D. C. 20016

July 1974

## PREFACE

This Evaluation Plan is the product of a concentrated, seven-week effort. Because of obvious time constraints, it often was necessary to choose between including a significant issue even if it could be considered only briefly and omitting all discussion on some topic if it could not be presented thoroughly and systematically. The former approach was taken wherever possible in the hope of at least identifying critical aspects of a comprehensive evaluation scheme so that these could be elaborated upon and incorporated into further planning efforts.

It also should be emphasized at the outset that the description of the DEIDS Project in Ecuador and the enumeration of the targets and interim achievements intended for this effort are quite incomplete relative to what actually is expected. Specifics were drawn from the Ecuador Project for illustrative purposes and neither the scope nor the accuracy of the examples presented in this report should be viewed as unerringly representative of that program. Indeed, as will be more fully examined later in this report, operational planning for the Ecuador Project is far from complete and thus many of the details presented to illustrate the suggested approach to program evaluation are entirely speculative.

The evaluation model described in this report draws heavily on the present state-of-the-art in impact assessment methodology, but it also incorporates a number of new components. In this sense, particularly, work on designing the model has been uniquely productive and gratifying. Many of these innovations emerged directly from interactions with the scores of people who made substantive contributions, suggested alternative points of view, or doggedly pursued their own biases until they were reflected in the evaluation design. Among these individuals, Mark Ward, Walter Furst, James Brown and Lloyd Florio of AID, Aaron Ifekwunigwe and Eugene Boostrom of UCLA, Herbert Dalmat and Donald Rice of APHA, and Paul Schwarz, Robert Krug and John LeSar of AIR all willingly made themselves

available for project discussions on several occasions and did their best to make this conceptualization as meaningful and useful as possible.

A draft version of this report was reviewed by a substantial number of the participants in the DEIDS Program. Some twenty-five individuals representing AID, APHA and various of its subcontractors, and AIR met on 22 August 1974 to consider the implications of this evaluation model in future DEIDS activities. As a result of this discussion, clarifications of some points have been made in this final version of the report, and the order of the chapters has been modified slightly. No substantive changes have been made which alter the content appearing in the draft version or in its conclusions and recommendations.

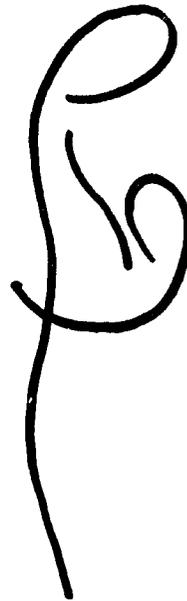
## TABLE OF CONTENTS

PREFACE .....	ii
<b>PART I: <u>OVERVIEW AND RECOMMENDATIONS</u></b>	
1. OVERVIEW .....	1
Introduction .....	1
Features of the Approach .....	2
The Ecuador Project .....	3
Description of the Model .....	4
Focus of the Model .....	8
2. RECOMMENDATIONS .....	11
Aim of DEIDS Evaluation .....	11
Specific Recommendations .....	12
<b>PART II: <u>TECHNICAL DISCUSSION</u></b>	
3. CHARACTERISTICS OF THE MODEL .....	16
Decision-Oriented Assessment .....	16
Types of Decisions .....	18
Hypothesis-Oriented Assessment .....	22
Stages of an Action Program .....	26
Diagnosing Sources of Failures .....	31
4. APPLICATION OF THE MODEL .....	33
DEIDS Project in Ecuador .....	33
Ultimate Outcomes for Ecuador DEIDS .....	35
Design Options for Ecuador DEIDS .....	38
Targeted Achievements for Ecuador DEIDS .....	39
Complete Lists of Targeted Achievements .....	43
Internal Operations for Ecuador DEIDS .....	48
Planned Inputs for Ecuador DEIDS .....	52
Interrelationship of Entries .....	55
Diagnostic Routines .....	57
5. THE DEVELOPMENT OF INDICATORS .....	60
Assessment Indicator Criteria .....	60
Examples of Possible Indicators .....	62

6. DISCUSSION AND ANALYSIS .....	68
Evaluation and Project Design .....	68
Data Gathering Alternatives .....	69
Program Evaluation Alternatives .....	72
Multiprogram Comparisons .....	74
Process Studies .....	75
Next Steps in Evaluation Design .....	76
Costs of Impact Assessment .....	79

# **EVALUATION PLAN FOR DEIDS**

**Part I**



## 1. OVERVIEW

### Introduction

The Development and Evaluation of Integrated Delivery Systems (DEIDS) Program has been initiated to explore new methods for providing low-cost, far-reaching health services in the developing countries. These efforts are designed to improve overall health and well-being, particularly for those with very low incomes, by furnishing needed medical, nutritional and family planning services where these are not now available. The basic approach of the DEIDS Projects and similar Integrated Health Delivery Programs is the expansion of indigenous health services through the use of innovative techniques intended to increase their outreach, effectiveness and impact.

A very important aspect of these demonstration projects is the identification of especially promising low-cost health delivery practices which could be adopted or applied by other developing countries in designing health programs consistent with their own needs and resources. While the design of any program must reflect local priorities and circumstances, they all are likely to share similar problems. And, although the practical experience of responsible program planners obviously can do much to insure that glaring errors are avoided, the availability of reliable information on the probable contribution of alternative program components can enhance the potential success of any future effort.

Guidance for future action programs is only one facet of an effective evaluation design. The expectable interest of program participants in ascertaining overall accomplishments is another. Systematic early warnings to alert program managers to correctable defects and inadequacies is still another. Even the process of designing an evaluation can contribute to program effectiveness by making intended outcomes explicit so that all parties are able to direct their efforts toward common goals and so that the criteria for success are agreed upon before any investment plans are implemented.

The possibility of assessing several projects concurrently adds further to the range of information that can be investigated by an evaluation model. Later in this report, the use of this assessment approach for ascertaining health needs where programs do not now exist will be considered. In the meanwhile, it should be made clear that this model is not limited to any one DEIDS Project. Instead, it represents a beginning in the development of a technological tool which could be used to collect information from all pertinent efforts directed at supplying health services to those who are beyond the reach of conventional delivery programs.

In this sense, this approach is a first step in the design of a comprehensive impact assessment model suitable for monitoring and evaluating a wide variety of health-related delivery systems, regardless of their specific objectives or the circumstances which differentiate the setting of one project from another. Although several iterations will be required before the model can be used in this way, it intentionally has been designed to go beyond the immediate needs of any one DEIDS Project in order to permit the extension of the resulting methodologies to other programs and to allow comparability among the findings of all pertinent efforts.

The remainder of this report describes the development of the model, illustrates its application using the Ecuador DEIDS Project as an example, discusses the features and utilization of the model in planning future health service delivery programs, and presents recommendations as to further steps which should be taken in support of the DEIDS effort. So that each portion of the discussion can be related to the others, it may be helpful at this point to present a brief overview of the model, its application to one representative country project, and the principal conclusions and recommendations.

### Features of the Approach

One fundamental characteristic of this approach to program assessment is that it is decision-oriented. Determining what decisions are to be made is essential for deciding what data are needed, when they should be

collected, and what range of results can be expected to influence the course of these decisions. A failure to take account of decision requirements can result in an unresponsive evaluation design. Even more important, the collection of information which has no conceivable effect on program decisions can be an expensive and unnecessary use of valuable resources. As will be seen, evaluation designs which emphasize the collection of data for its own sake, which attempt to remedy failures to adequately define the problem in the first place, or which explore aspects of the program which are beyond manipulation may be interesting in a theoretical sense but have little practical value.

A second feature of this approach is that it is based on a network of hypotheses, or program rationales, which make explicit the dynamics of the cause-effect relationships being tested. Although expressed, as are most impact evaluations, in the form of states-of-affairs at various stages of program realization, the methodological focus in this model is on the hypotheses that relate events at one stage to those at the next. The certainty with which outcomes can be attributed to inputs under program control is vastly enhanced by this technique and it presents an important breakthrough in the design of action program assessments.

The third important component of the model is its emphasis on impact-referenced indicators of accomplishments. While changes in attitudes, beliefs or opportunities may be a relevant contributor to desirable outcomes, it is difficult to justify the attainment of these changes as evidence of success, to say nothing of the problem of measuring such changes with any degree of confidence. The point of view incorporated in this model is that all program benefits should be easily and openly observable, and expressed in terms of events which are incontrovertibly evidence of improvement. This position is not as much one of precluding disagreement over the way in which an outcome is measured as it is one of precluding disagreement over the impact of what was achieved.

### The Ecuador Project

The development of a workable assessment model cannot be accomplished in the abstract. For this reason, one of the DEIDS Projects was selected

as a starting point. As will be seen, specifics from other projects in the same program successively can be added to the framework to produce a composite and generalizable evaluation model for use over a wide range of health service delivery systems. However, the examples in this report refer to what are, or might be, the goals and constraints of the DEIDS Project in Ecuador.

Very briefly, the aim of the Ecuador Project is the development and implementation of an innovative delivery system for providing integrated health services to at least two-thirds of the young children and women of fertile age in a selected rural highland area of southern Ecuador at a continuing cost affordable by the host country. Both preventive and curative services will be offered, with emphasis on reducing morbidity and mortality through improved child and maternal health, better nutrition and sanitation, and more effective family planning. A successful demonstration of the feasibility of the approach is expected to lead to its adoption throughout rural Ecuador.

#### Description of the Model

Three cycles of decisions are associated with any action program. Evaluation is directed at supplying timely, tangible information to increase the accuracy and certainty of these decisions by supplying the evidence needed to choose among plausible alternatives. Planning decisions are those made at the beginning of a program, those concerned with defining the problem to be solved so that appropriate goals can be established and so that needed inputs can be determined. Operating decisions are made during the course of a program in response to early warnings of program defects so that appropriate remedial actions may be taken. Concluding decisions are made at the end of a program and concern the continuation, expansion and replication of the program on the basis of obtained results.

Obviously, no evaluation can contribute constructively to decisions after they have been made. Thus, some completed or nearly completed programs are assessed only in retrospect with the aim of deciding whether the program or certain of its components are worth replicating.

In other instances, as in the DEIDS Program, evaluation is initiated along with other project activities so that progress toward desired outcomes can be monitored systematically and so that management decisions regarding the need for refinements can be made promptly. And in a few cases, there is the opportunity to use evaluative tools in advance of a project to improve the design of program strategies by establishing the identity of already existing components and relationships which could contribute to program success.

To be responsive in these ways, an evaluation model must be directed as fully at understanding the cause and effect relationships underlying the production of change as it is to the quantitative description of program accomplishments. The latter says only what changes have materialized while the former contributes the how; and it is the how which will make possible the creation of future health delivery systems without repeating the long years of developmental testing planned for the DEIDS Projects. An analysis of the change process, then, has to focus on the chains of causally related events which programmatically bring about the accomplishment of change.

Most action programs are too complex to be considered only as a whole if meaningful information on the enabling change process is being sought. For this reason, programs generally are divided into some arbitrary number of stages so that interim events in various chains can be defined for assessment purposes. These stages, it should be made clear, do not fall along any uniform time schedule. At any particular point in a project, some chains will be complete while others have progressed only through their earlier stages.

Events at two of these stages normally are neglected during program evaluation. Ultimate Outcomes are the long-range goals of a program which typically materialize as the cumulation of the efforts of many discrete programs as well as happenstance events. Their remoteness from a program, both in time and causality, excludes them and their attainment even from decisions typically made even after a program ends. Nevertheless, ultimate outcomes identify the thrust and scope of a program and so determine its overall design. The ultimate outcomes established for the DEIDS Project

in Ecuador, as enumerated later in this report, fall into four areas. These include cost, institutional and side-effect objectives as well as outcomes concerned with health.

The very first stage of a program, encompassing Program Design, also is excluded from evaluation whenever planning decisions already have been made. Further information on the way the program has been conceptualized no longer can influence these decisions, and assessment of them is therefore unnecessary. One advantage of the proposed evaluation model, on the other hand, is that it potentially can be used to assess needs and resources during program design and in this way contribute from the very beginning to program success through better and more systematic planning.

Program events occurring between planning and concluding decisions can be thought of in four stages. The first of these, called Planned Inputs, are the initiatives undertaken by program managers to set each chain of events in motion. A number of examples from this stage of the DEIDS Project in Ecuador appear later in this report. Included among them are these specific objectives:

- *health needs in the test area are agreed to by project and government representatives*
- *appropriate systems have been developed for administration, accounting and quality control*
- *sufficient opportunities for reducing rates of infant and childhood mortality and morbidity to expected levels within the scope of the program have been identified*

The second stage of a program concerns Internal Operations, the levels of readiness achieved as direct consequences of planned inputs. Unlike their antecedents, internal operations are not entirely under program control, and their status may reflect unexpected circumstances or extraneous happenings. Some of the examples of internal operations appropriate to the DEIDS Project in Ecuador, selected from among those listed further along in this report, are:

- *six months of training for Health Workers results in sufficient skills*
- *Health Workers can demonstrate preparing tasty food from now undesirable supplies*

- *adequate supply of birth control pills on hand at Health Posts*

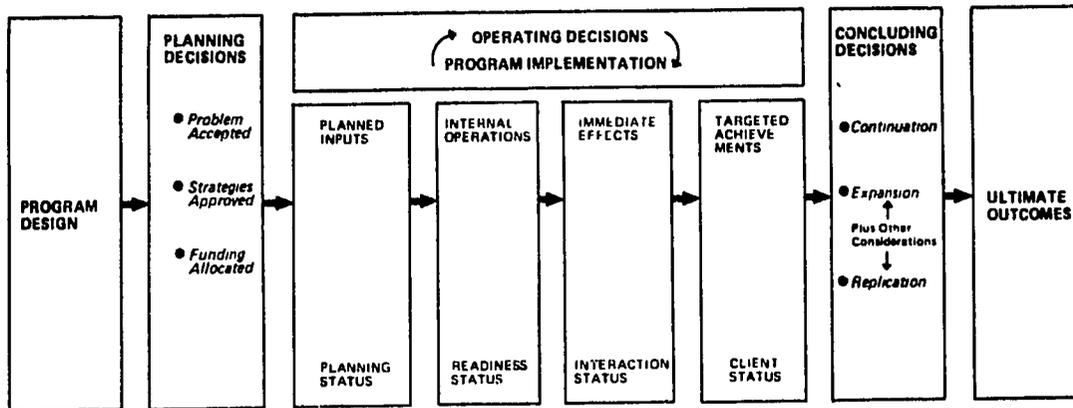
Immediate Effects, comprising the third stage, represent the earliest impact of the program on its beneficiaries, and thus describe the interaction between the program and its clients. In health projects, these immediate effects often are described as "services rendered". Some typical immediate effects expected during the DEIDS Project in Ecuador which are presented later in this report include:

- *a specified percentage of children receive three DPT immunizations by 8 months of age*
- *Health Worker attends a specified percentage of probable normal childbirths*
- *client records maintained by Health Workers*

The last stage of a program considered for impact assessment purposes contains events called Targeted Achievements. Events at this stage emphasize states-of-affairs which are both realistically attributable to the program and clearly evidence of improved well-being. A few examples of targeted achievements from those identified for the DEIDS Project in Ecuador include:

- *breast milk part of infant diet for 90% to 12 mos. and 60% to 24 mos.*
- *no pregnancies for 50% women within 3 yrs. after birth of a presently living child*
- *all training activities required for program continuation have been assumed by the training unit*

These various stages of program events can be seen in perspective in the following figure. As shown in the figure, the area of greatest concern falls within the implementation phase of a program because it is in these stages that evaluation can contribute its most constructive and useful information.



### Focus of the Model

The purpose of this model is not so much to delineate milestones suitable for evaluation efforts as it is to make evident the inter-relationships among events at various stages in a program. The linkages connecting successive events in each stage represent the assumptions about causes and effects that formed the basis of project design. It is hypothesized, for example, that if six months of training is given then proficient Health Workers will result, or that if the Health Worker is present at delivery, then some percentage of deaths due to maternal hemorrhaging will be avoided. (As explained later in the report, one cause may lead to several effects and any effect may be the product of several causes whenever a complex and integrated program is being examined.)

This focus on the hypotheses which connect program events is a key element in the design of this evaluation plan because it permits the convincing attribution of positive outcomes to successively earlier program ingredients. In other words, the use of an hypotheses-oriented assessment approach rather than one which is primarily descriptive in nature establishes not only what changes have occurred, but also makes it possible to conclude, beyond reasonable doubt, that these changes were due to the program itself and not to extraneous causes. And, because each link can be examined separately, those which demonstrably are of value can be identified for later replication even when the program as a whole is less than fully successful.

Dividing a program into component linkages also enhances the efficient use of evaluative resources. The level of uncertainty surrounding any particular hypothesis will vary considerably, depending on the confidence program designers have in that particular prediction. We are more willing to assume that maternal immunization prevents neonatal tetanus, for example, than we are to accept the hypothesis that nutritional advice leads to needed reductions in malnutrition. Although no program event should be totally ignored, greater attention can be given to those linkages of most concern to us because their validity is debatable than to hypotheses which have been repeatedly verified in the past.

In any thoughtfully designed program, some linkages will prove fallible. Otherwise, it is more than likely that the inputs assembled for it far exceeded what was really required to achieve the expected outcomes. In this respect, all programs must be capable of responding rapidly to impending defects so that the underlying deficiencies can be corrected. The report contains some typical diagnostic routines of the kind that could be initiated as soon as information was generated, suggesting that some predicted event was not materializing as expected. As will be explained later, it is much simpler and wiser to postpone the actual development of these diagnostic routines until needs become clear. Preparing all those which might be required in advance would be entertaining but otherwise next to useless.

One final feature of the proposed evaluation plan which deserves mention at this point concerns the development of indicators for each program event to be measured. A number of criteria for selecting indicators to use in determining outcomes are given later in the report. Of these, the most essential is that the indicators used represent objective, overt events which are readily observable and uncontroversial in interpretation. Sufficient strides have been made in impact assessment methodology to preclude any need for subjective sources of, or judgments about, what was achieved. Palpable, direct indicators not only add conviction to ensuing conclusions, but also permit advance agreement among all interested elements as to the precise aims of the program. The development of such indicators often requires considerable time and ingenuity, but the results tend to be well worth whatever effort and

talent was invested in them. Several samples of the kinds of indicators that would be appropriate for events likely to be examined in the DEIDS Project in Ecuador appear later in the report.

## 2. RECOMMENDATIONS

### Aim of DEIDS Evaluation

Workable approaches to evaluation are not created in the abstract. For this reason, the specifics referred to in this report all have been derived from existing plans for the DEIDS Project in Ecuador. The creation of distinctly different evaluation models for each of the intended DEIDS projects would almost preclude aggregating results across projects, however, and one of the most impactful features of the overall program would be lost. The proposed approach has been designed with the aim of providing the needed commonality in evaluation across projects by providing a framework that is fully capable of encompassing all plausible variations among the various projects while at the same time taking advantage of consistencies between projects in their input designs and expected outcomes.

To accomplish this aim, it would first be desirable to complete the detailed matrix for the DEIDS Project in Ecuador. Although the design for this project is still not firm pending some needed field work after the team arrives on site, this step nevertheless is likely to require several months of interaction between project personnel and evaluation experts. Once that was done, work could begin on developing an equivalent matrix for each of the remaining DEIDS Projects. Local needs and circumstances in each separate location are likely to lead to some changes in the particular entries at each stage of any subsequent project, but many other entries will be identical from one project to the next.

The emphasis at this point, as explained more fully later in this report, is not to force all projects into a single mold but, instead, to recognize and identify whatever commonalities do exist. As to Mecca, there are many paths to improved health in the developing countries, and future choices among these alternatives cannot be made unless they all have compared with a consistent evaluation methodology. As each succeeding matrix is assembled, only those entries not present in the

projects analyzed earlier would have to be added. Indicators needed to assess the realization of each critical aspect similarly could be used across projects, with new ones developed only as needed.

This necessarily centralized impact assessment approach neither precludes the active participation of the host countries and the sub-contractors, since they should contribute to the evaluation plan and are essential to its implementation, nor does it in any way limit the gathering of whatever additional information is felt to be useful within the scope of each of the individual projects. On the other hand, this approach insures that the results from the DEIDS Program can be used when decisions among proposed alternatives are made regarding the design of future health delivery systems. This is not the only assessment approach which could be devised for the DEIDS effort. However, any evaluation plan which fails to contribute significantly to the ease and certainty with which subsequent programs are designed is not responsive to the unique opportunities afforded by DEIDS.

### Specific Recommendations

As noted at the beginning of this report, this plan is only a first step in the development of a comprehensive evaluation design which is capable of meeting the expectations surrounding the DEIDS Program. Although other conceptual approaches to the impact assessment of action programs exist, none of these alternatives (as described later in this report) offers a design which is likely to be equally thorough, which carries the same degree of confidence regarding findings, or which could be out at as an appropriate cost. But, as also will be described, a number of steps remain if the proposed design is to be carried forward and implemented. The principal actions which should be taken at this point in order to establish an evaluation program for the DEIDS Projects are summarized below. As will be apparent, several of these recommendations concern the planning of the individual DEIDS Projects but are included here because of the necessary interrelationship between planning and evaluation.

1. It is recommended that some version of the evaluation model

described in this report be adopted for the DEIDS Program. Models which are not decision-oriented, based on hypotheses which link program events, and directed at impact-referenced indicators are not expected to provide the kinds of fundamental programmatic information that should result from the DEIDS Program.

2. It is recommended that a single, overall evaluation design be used for all DEIDS Projects, and that the assessment of the various projects be made a central program responsibility. The systematic accumulation of consistent programmatic information from among the projects is essential if this information is to be used to its fullest extent in planning more effective health delivery programs in the future.

3. It is recommended that expert technical assistance be made available to the DEIDS Program to carry forward the development and implementation of the evaluation plan, and that this assistance continue throughout the full length of the program. Recent methodological advances in the field of impact assessment permit a very constructive role for evaluation in program management and design, and the importance of specific impact assessment skills in the realization of these contributions cannot be minimized.

4. It is recommended that the scheduling of the specific DEIDS Projects be reviewed to provide an appropriate period at the beginning of each project to perfect detailed programmatic plans. While there are tempting reasons to initiate health service activities as soon as possible at each project site, it is evident from even a brief exposure to the DEIDS Project in Ecuador that additional planning efforts are needed if substantial waste among initial inputs is to be avoided. This planning should be done on site because it will entail the collection of specific, problem-defining information.

5. It is recommended that personnel from all planned DEIDS Projects and host countries participate in the development of the matrix of rationales and the indicators for the Ecuador DEIDS Project, and that this coordination continue during the accomplishment of these steps for each successive project. Significant planning benefits will accrue to all projects, and considerable effort will be saved in the long run as each

new project is added to the evaluation scheme.

6. It is recommended that carefully defined standards be established for the DEIDS Program concerning planning and evaluation objectives. As will be noted later in this report, there are pronounced discrepancies among the various interests represented in the DEIDS Program as to the thoroughness and detail that needs to be reflected in project plans, the requirements that are to be met by evaluation efforts, and the nature of the results expected from these demonstration projects with respect to their potential generalizability. These differences are not irreconcilable, but steps must be taken in the very near future to insure that all efforts are directed toward mutually acceptable goals.

7. It is recommended that opportunities inherent in the DEIDS Program for furthering the advancement of impact assessment methodology be recognized, and that provisions be made for utilizing these opportunities to their fullest extent. In the short term, even modest amounts of investment in this area will permit extending the application of the DEIDS evaluation plan to many other health delivery systems and problems. In the longer term, additional benefits from this program would accrue to the planning and assessment of a wide variety of action programs.

8. It is recommended that more careful attention be given to the cost-benefit aspects of the DEIDS evaluation with respect to the probable expense of collecting information that is of questionable value because of its lack of generality, its debatable methodology, or its relative significance. In the case of DEIDS, almost any additional data requirement not essential to the conduct or future replication of the program will represent an added expense, and that additional requirement should be assessed accordingly.

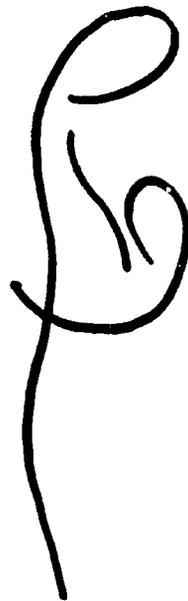
9. It is recommended that the further planning of an evaluation component for the DEIDS Program be consistent with the intended low-cost concept of the Program. On the basis of presently available information, it can be estimated that all centralized evaluation planning, monitoring and reporting for all the DEIDS Projects combined could be accomplished with no more than twenty man-years of expert support over the eight-year life of the program. All additional needs for local data collection and

interpretation could be met by indigenous staff in the host countries who could continue the evaluation function after program support has ended.

10. It is recommended that various project activities not designed to yield generalizable information, such as problem-definition and process studies, be carried out by the staffs of each individual project with consultation, as appropriate, from the central evaluation team. These efforts will be essential to the success of each project and should be considered the responsibility of each project team. Such studies should not be part of the overall evaluation design, on the other hand, since the immediate results of these investigations are likely to be of little use elsewhere until their impact on programmatic accomplishments is demonstrated through the regular evaluation process.

# **EVALUATION PLAN FOR DEIDS**

**Part II**

A handwritten signature in black ink, consisting of a vertical line that curves into a loop at the top, followed by a smaller loop and a final downward stroke.

### 3. CHARACTERISTICS OF THE MODEL

#### Decision-Oriented Assessment

The first step in designing a program evaluation is to determine which decisions are to be made about the program, when these decisions are to be made, and what kinds of results are likely to influence decision outcomes. As noted earlier, an almost infinite variety of data can be collected from even the simplest action program. Yet, most of this information will have little impact on any subsequent decisions. In a few instances, the findings sought will be irrelevant to the purposes and accomplishments of the program. In other cases, the data will be pertinent to decisions which already have been completed, and where the course of action -- rightly or wrongly -- already has been determined (efforts to define the nature or magnitude of a problem after deciding on a program to solve that problem is one example).

A still more significant reason for a failure to utilize evaluative information is that the only plausible outcomes of some data collection effort are not likely to exceed the limits of quite ordinary expectations. We are willing to accept the efficacy of smallpox vaccinations, for example, without further longitudinal studies of the incidence of smallpox among those who were vaccinated. Similarly, we are willing to assume people everywhere want improved health care for themselves and their families without having to survey them for their opinions. Evaluation is appropriate only when there is legitimate doubt as to the outcome. A decision that some particular health delivery system is advantageous will be far more influenced by the program's demonstrated ability to cope with previously unsolved problems than its attainment of routinely accomplished tasks.

Determining what decisions will be made about a program and what information will affect these decisions makes it possible to focus attention specifically on those outcomes which will contribute heavily to the decision-making process. Clarifying decision alternatives in advance

also yields other tangible benefits. First, by obtaining agreement beforehand on which decisions will result from any pattern of results, fundamental program priorities are made clear to all those involved in its planning and implementation. Widely differing opinions as to the goals of the program are avoided by making both operational targets and their measurement standards explicit. Second, accord regarding the decision process concentrates attention on those variables and outcomes presumed to be under program control. In this way, expectations are tempered to the point where they are consistent with the resources made available to the program. And third, by bringing possible outcomes into the open before data collection, much of the understandable reticence toward evaluation by action-oriented program managers can be avoided.

An effective analysis of program decision requirements depends on concurrence as to decision variables among all agencies intending to participate in the decision process. In the case of the DEIDS Program, this analysis would include the points of view of A.I.D., the designers of the programs and, of course, the host countries. In a general sense, decision requirements are identified by the goals established for the program. These goals presumably reflect the conditions under which future commitments to the program would be honored or, put another way, they reflect the expectations which the various participants had when the program was proposed and accepted.

It is unlikely that each goal will be attained exactly at the hoped for level, of course. Some are likely to be missed to a degree while others will be exceeded, and these variations in intended outcomes will be considered at the time the decisions are made. From the point of view of project management, on the other hand, program goals serve from the very beginning as the basis for determining what inputs are required and, later on, for deciding when resources have to be reallocated to improve results in deficient areas. In the absence of decision-oriented goals, any estimate of the cost of a program, the composition of its inputs, or the time required to yield the expected results is exceedingly arbitrary. Similarly, it is naive at best to proceed aimlessly and see what will result from a capricious combination of cost, composition and

time estimates. There are too many real improvements to be made in the world to tolerate solutions in pursuit of problems.

Agreeing upon the goals of a program does not mean that these goals must remain immutable. A properly managed action program is extremely dynamic. Holding the program still for measurement purposes would neither be desirable nor realistic. Changes are expected as the result of unanticipated outside events, ever increasing understanding of the problems being addressed and, more importantly, the successes caused by the program itself. It also is possible that some program goals simply cannot be attained with the available inputs. Any truly imaginative program design should lead to some failures. If it does not, there is every likelihood that potentially innovative components had been deliberately avoided or that the levels of inputs called for were unnecessarily excessive.

### Types of Decisions

Three cycles of decisions generally are associated with any action program. These are planning decisions, operating decisions and concluding decisions or, more simply, those that have to be made at the beginning, during, and at the end of the program. Impact assessment models, including the one presented here, are concerned primarily with outcomes, both those which enable decisions concerning the entire program as a whole and those which guide the program during its course by providing program managers with interim evidence of progress. To make the distinction among cycles of decisions and the information needed at each cycle clearer, they can be considered in greater detail.

Planning decisions are those concerned with defining the problem or problems to be solved so that realistic goals can be established and so that the appropriate inputs then can be assembled. In planning any DEIDS-like project, it probably is essential to determine existing levels and causes of morbidity and mortality, malnutrition, unsanitary conditions and excessive fertility before deciding, for example, on the number of health workers needed, what their training will consist of, or what improvements in health-related conditions can be expected from the program.

Unless the problems are reasonably well understood, any estimate of the cost or character or length of a program to deal with these needs is dubious at best. (One difficulty in preparing this report, in fact, was that the program's designers feel presently available problem information for the test area in Ecuador is inadequate and that both the inputs and outputs stated for the program must remain extremely tentative until further data can be collected. This point will be considered again later in this report.)

The three most usual methods of collecting problem-defining information are an analysis of records, the conduct of sample surveys, and the consensus of experts. All three methods are equivalent in that all three yield only estimates of the "true" state-of-affairs. Thus, the choice among them is appropriately based on the relative credibility and cost of each type of estimate rather than on the superficial appearance of the tabulated data. Wherever adequate records are not available, it also is unlikely that there is sufficient knowledge about the composition or dispersion of the population to develop a sampling technique which is any less biased or more accurate than the opinion of knowledgeable experts. And this last method, perhaps supplemented by some field observations and interviews with locally indigenous leaders, probably will provide as reasonable an estimate in many underdeveloped areas as more detailed methods, and at a far lower cost. The same logic applies to an ever present need to continually review the way the problem has been defined during the course of the program by collecting new data which permits updating the system to new priorities.

Operating decisions are concerned with two aspects of program management. The first of these relates to administrative responsibilities and depends on orderly records of expenditures, training sessions, clinic visits and other data. The primary purpose of this information is to establish that certain events, particularly inputs, did indeed occur. A little later in this report, a distinction will be made between the adequacy of a program activity and the relationship between that activity and its consequences. As will be seen then, an impact assessment model assumes that events occur as planned (e.g., that iron tablets distributed by the program contain iron or that an infant's weight as determined by a

Health Worker meets the program's standards of accuracy) unless the expected outcomes (e.g., reduction of anemia or the appropriate nutritional recommendation regarding the infant) fail to materialize. Should this happen, a diagnostic routine would be required to ascertain what went wrong.

The second aspect of operating decisions concerns interim information on the need to increment or modify the conceptualization of program inputs. Such changes would be required when everything was done as planned and yet the expected outcome was not reached. This would not be an instance of the improper or defective implementation of a planned input but, rather, the inability of that input to generate the required outcome (e.g., the ratio of one Health Worker per thousand inhabitants may be nowhere near sufficient to meet existing demands for essential health services). Failing to achieve an interim target during program implementation, and then failing to demonstrate any plausible deficiency in the way the input was executed suggests that some technological principle adopted for planning purposes was at fault (e.g., the successive swings between breast and bottle feeding as each generation discovers one or the other approach does not produce perfect children). The question in this case is not whether the input occurred, but whether it inherently was sufficient as a cause to yield the desired effect.

Assessments pertinent to these two aspects of program management, the verification of inputs and demonstration of relationships, collectively are referred to as "formative" evaluations even though the underlying reasoning for each is quite different. In either case, however, it is important to first decide what interim outcomes should be reached at any given stage of program development, then determine whether or not these outcomes were achieved, and finally, as necessary, complete a diagnostic routine to determine the source of an identified defect. The resulting information will permit program managers to decide the appropriate remedial action. The timing and fineness of measures contributing to operating decisions depend not on the outcomes themselves but on opportunities program managers have to implement program modifications.

Concluding decisions are those of greatest concern to impact assessment because they affect plans for the continuation, expansion or replication of the program elsewhere. It must be emphasized that evaluations

about a program never can be definitive except in retrospect. At the conclusion of a program, decision makers are called upon to predict the likely success of more or less the same program under more or less similar circumstances at some future time or other place. The accuracy of the judgments they make can be improved and enhanced by providing as much relevant information as possible to help them decide among the options open to them. These kinds of assessments often are described as "summative" evaluations, even though their intent is to aid in planning for the future rather than in judging about the past.

Three related concluding decisions could be expected at the end of the DEIDS Project in Ecuador. The first of these, the one most limited in scope, would be whether to continue the project in the three-province test area. (Actually, this decision most likely would be incorporated into the next one, but it can be considered independently for purposes of this discussion.) If all of the program's goals were achieved, we could presume this decision would be affirmative. This would not be true, on the other hand, for a decision to expand the program (either in area served or scope of services offered) or for one to replicate the program (or some aspects of it) elsewhere. These latter two decisions also would depend on how the problem was defined for the new program, on what its goals were to be, and on what kinds of costs would be deemed acceptable. The generalizability of information derived from impact assessment is limited by the degree to which circumstances and problems are similar between past and future programs, or to the degree these variables have been shown to be irrelevant on the basis of the results of past programs.

One of the dilemmas of impact assessment is that even these concluding decisions are made prior to the time when program goals are likely to be realized. Truly important goals tend to be long-range expectations which normally are a function of many more factors than can be dealt with by any one program. A reduction in birth rate, for example, may take years to be evident and may encompass the contributions of any number of programs and happenstance events. Delaying decisions about the value of a family planning program until such an ultimate outcome has been achieved, or attributing the attainment of that goal solely to the program, would be

unwise and unwarranted. (For the same reasons, the non-attainment of expected ultimate outcomes is not by itself evidence that the program itself was a failure.)

To recapitulate for a moment, three levels of decisions are associated with most action programs. As shown in Figure A, planning concerning the nature of the problem, the strategies appropriate for solving the problem, and the determination of funding requirements are made before a program is initiated. Operating decisions are made during the course of program implementation to correct defects in the design of inputs wherever there are "early warnings" that events are not occurring as planned. Concluding decisions as to the overall effectiveness of the strategies in terms of the probable benefits to be derived from continuing, expanding or replicating the program are made at its conclusion, but long before its ultimate outcomes are likely to be attained.

#### Hypothesis-Oriented Assessment

Although decisions about the program's design and ultimate outcomes are relevant because they globally determine its implementable inputs and direct outputs, most of the emphasis in the remainder of this report will be on what happens between these decisions; that is, during the course of the program. To set the stage for that discussion, it is important to examine why a program ought to be evaluated. As noted earlier, assessment is economically justifiable only when there is doubt about the relationship between causes and effects. Normally, the design of any program represents the best guess of which approach among available options is most likely to yield the desired results within prevailing constraints. The designers are making predictions, or hypotheses, about what will happen if certain actions are taken, and the purpose of program evaluation is to test these predictions.

All action programs can be characterized in this way. If seat belts are made mandatory in cars, then traffic fatalities will be reduced. If time and money is invested in research, then seeds for higher yielding crops will result. If training is given, then the life and productivity of road equipment will be improved. Each successful experience with more

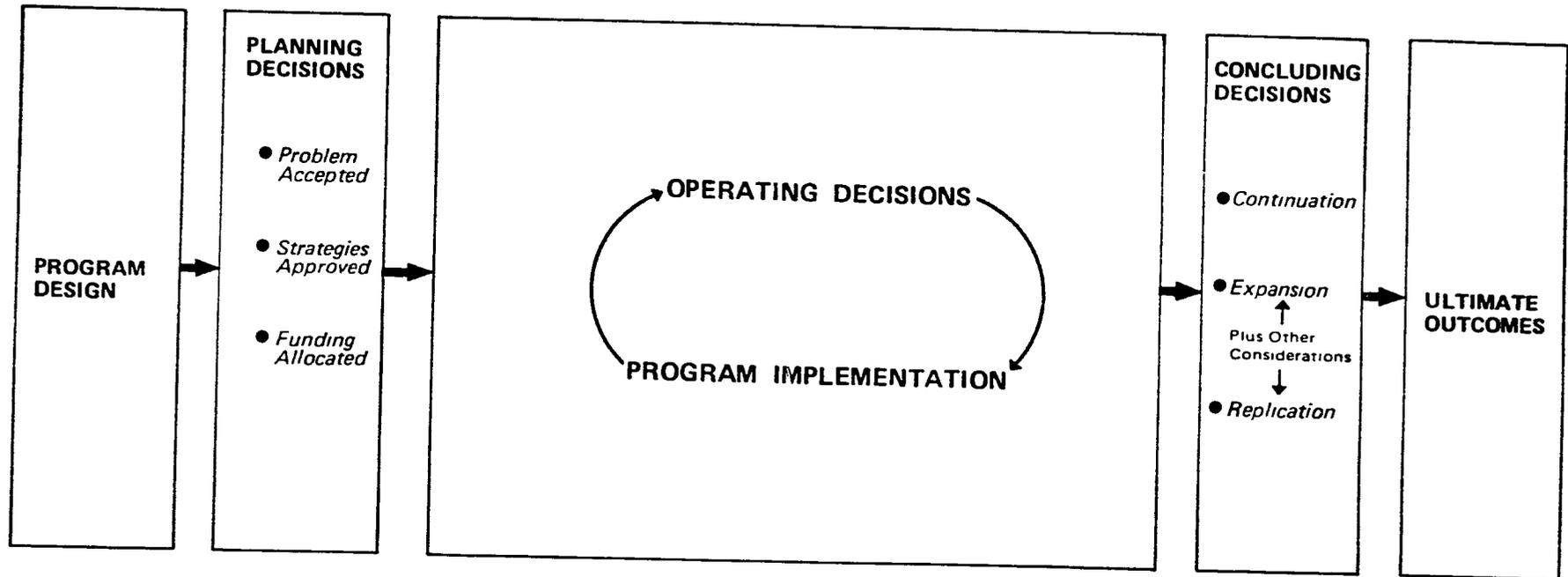


FIGURE A

or less the same variables under more or less the same conditions increases our confidence in the expected outcome. But certainty is never achieved. Some hypotheses about health have been virtually established as fact, such as the relationship between immunization and disease. Others have been found to be not demonstrable, such as the relationship between the availability of sterilization and a significant reduction in birth rates.

For programs of any complexity, it would neither be prudent nor particularly informative only to assess the relationship between events at the very beginning and end of the program as a method of testing programmatic hypotheses. For one thing, the time before testing may correspond to a considerable investment, much of which could be wasted if correctable defects were not discovered as early as possible. For another, the attribution of effects to causes is not an easy task. An adequate evaluation plan not only must reveal changes, but it must also be capable of demonstrating (beyond reasonable doubt) that these changes were due to the program and not to extraneous factors. Because agricultural yields are affected by the weather as well as by fertilizer programs, attributing a rise or fall in production per acre to the use of fertilizer alone would not be convincing evidence as to the efficacy of the program.

Confidence in the conclusions about a program can be enhanced by two methods. First, we can compare the outcomes achieved among individuals affected by the program with what happened among individuals who were not affected by the program. But, the use of an "experimental-control" design depends on assumptions as to the equivalence of the two groups which rarely can be met in real world situations; even the collection of data from the non-affected population is likely to change their behavior (e.g., "Did you brush your teeth today?"). Finally, just the cost of collecting sufficient data to demonstrate the comparability of the two groups can outweigh the value of the comparison.

The second alternative is a design which uses the experimental population itself as its own control. Confidence in attributing an outcome to the program is generated when the outcome occurs and when we are convinced that it could not have occurred in the absence of the program.

(It would be difficult indeed not to attribute an indigenous Health Worker's correct use of a sphygmomanometer to anything but her training.) The smaller the time span between cause and effect, and the less complex the relationship, the greater will be our confidence in the conclusion. Thus, whenever it is possible to break up a long and involved program into a series of discrete hypotheses, or "linkages", confidence as to the attribution of outcomes can be as great or greater with this approach than when control groups are used. And, the cost is generally far lower.

The purpose of formulating linkages between successive program events is to establish a series of expectations, often called program rationales, which are sufficiently narrow in terms of accumulated wisdom to permit us to reject alternative explanations of why some outcome occurred. (Here, the word "reject" is used in the same sense as it is in statistical tests of the null hypothesis; it refers to the probability that some outcome was reached due to unidentified and totally extraneous causes.)

The process is similar to following a route plan. The actions we take, such as making a right turn after 16.3 miles, are accompanied by "tests" of what was accomplished, such as then arriving at a point where some monument will be visible to the left. While it is obvious that we could have arrived at that point over an alternative route, had that been the plan, or even through a series of fortuitously compensating errors, we nonetheless are ready to conclude that the plan did work and we would be willing to recommend it to our friends who later wanted to take the same trip.

A further advantage of this approach is that it potentially permits the future assembly of confirmed links or hypotheses into some other program which may be designed for different outcomes or different circumstances. (Routing plans prepared by auto clubs are, in fact, assembled this way to lead to alternative destinations and to allow for preferences by permitting scenic side-trips.) A method for setting up a training program, a method for recruiting personnel, or a method for distributing supplies under refrigeration throughout remote areas may be segments of a program which, if successful, could be adopted for use elsewhere even if the total program of which they were a part was an over-

all failure. An evaluation scheme which fails to provide this kind of information about a program has not met one of its most important responsibilities, that of pinpointing those aspects of the program which would be generalizable to other settings.

The level of specificity aimed for in formulating chains of successive events which lead from inputs to outcomes is determined by the complexity of the program and the degree of confidence we already have in the cause-effect relationships hypothesized for any segment. If the segments are too large (e.g., from the beginning all the way to the end of the program), we risk the possibility of failing to detect correctable deficiencies while there was still time to do something about them. If the segments are too small, the ultimate cost of evaluation may be an excessive burden, particularly in the sense of the extent to which evaluation interferes with the operation of the program. The size of segments reflected in an evaluation design somewhat arbitrarily depend on collective judgment as to the fallibility of the underlying hypotheses and on the practical convenience of the resulting assessment plan.

#### Stages of an Action Program

Experience suggests that an evaluation design which focuses on four hypothetical stages of a typical action program provides reasonably accurate, meaningful and timely information on what is being accomplished. These stages can be defined both theoretically, in terms of impact evaluation methodology, and concretely, in terms of the requirements of projects in the DEIDS Program. In this description, the two stages which occur outside of the actual program have been omitted. One is the Program Design stage which precedes planning decisions. The other is the Ultimate Outcome stage which is too remote from the program to be considered for evaluation purposes, and typically appears after concluding decisions have been made.

Planned inputs are at the first stage of important events in the chronology of a program, those intended to set the cause-effect sequences into motion. They include all the initiatives undertaken by program managers, and can consist of such different ingredients as upgrading

facilities, preparing training materials, procuring commodities, recruiting personnel or fostering publicity. A simple, operational definition of planned inputs would be the "new things" the program adds to an ongoing state-of-affairs. Perhaps more accurately, on the other hand, planned inputs also take account of decisions as to what changes were not needed in terms of the assumptions and projections that entered into program design.

The focus here is on the status of planned inputs rather than on their quality. For the program to have any hope of success, there must be a reasonable degree of consistency among planned inputs and there must be some provision made for each anticipated requirement. The evaluation process assumes that the nature of every input reflects the best of available knowledge as to how to meet program requirements. It questions only whether any have been omitted or conflict with one another in terms of the entire network of linkages that represents the program. (In the Ecuador DEIDS Project, for example, there is a requirement for each Health Worker to be supervised in the field roughly one full day per month, but no provision has been made to augment the staff of the Health Subcenters to provide both this supervision and a referral capability.) Planned inputs are at the last (and only) stage of the program that is completely under program control and not affected by extraneous events.

Internal operations are the instrumentalities of the program, the doing. This could include the number of Health Posts actually opened, the number of workers actually trained, the number of food packets actually made available for distribution, or the number of handbills actually printed. Included here is not only what is being done by the program, but also the characteristics of the process, or how it is being done. This stage of a program is particularly challenging for evaluation because the nuances of process almost never are discussed in program documentation. Yet, the timing and sequencing of inputs, as well as their realization, are extremely important.

In a practical sense, internal operations refer to the capabilities or readiness of the program to accomplish changes. Thus, this stage represents the status of program readiness. It is a product of various inputs, but it does not include any client component. The status of

internal operations is particularly sensitive to the timing of program events because a number of inputs sometimes must occur serially to produce any desired program state. In order to have Health Workers at their posts, for example, they must be recruited, trained and deployed. We could, and often would, measure the success of each of these inputs as they occurred but, in the perspective of the entire program, we are most interested in whether all of these inputs in combination yielded a readiness to begin serving clients.

Immediate effects are those typically observed during the change process itself. They are concerned with what the program does on behalf of its clients. It is at this stage in a program that its influences are first felt. Here, there is concern for how many clients visit a Health Post, how much food supplement is distributed, or how many mothers receive instruction on sanitation practices. Characteristically, immediate effects are events which are mutually observable to the program and the client. Both or either could be the source of information on whether these events did or did not occur. In the health field, this stage of a program often is characterized in terms of "services rendered".

The intent of evaluation at this stage is to measure the status of interactions between program and client. Once more, it is assumed (at least until shown otherwise) that these interactions are skillfully and completely accomplished. For this reason, we look at the state produced by the interaction rather than the process itself. If both Health Worker and client agree that immunizations were given or that family planning information was presented, we at least temporarily regard these events as equivalent to immunizations having been received or family planning information having been understood.

The concept of "compliance" deserves special comment. Since we only have evidence of compliance because of what results from it (attitudinal statements, even if they are reliable, refer to intention and not compliance itself), it is not possible to assess compliance at this stage of a program. Instead, we can look at events which lead the Health Worker to expect compliance and thus signal an end to her further efforts in that regard. Therefore, an agreement by the client to include vegetables in

the family diet is a signal to the Health Worker to stop lecturing in the same way that the disappearance of a rash is a signal to the Health Worker to discontinue medication.

Targeted achievements are the continuing or lasting benefits derived from the program. Unlike immediate outcomes which include any changes resulting from the program, targeted achievements are events which themselves are uniformly accepted as evidence of improved well-being. The absence of anemia during pregnancy, a child's rapid recovery from an ear infection, or a prevalence of outhouses are all desirable states-of-affairs that are within the scope of a DEIDS Project and are events which could be attributed to program efforts with some reasonable degree of certainty. To make the point clearer, it should be evident that "improved nutrition" is appropriate as a goal or ultimate outcome of a program but not as a targeted achievement. For this stage of a program, it is necessary to specify what constitutes improved nutrition, and which of those many aspects of nutrition can be attained under existing circumstances and within established cost limitations.

Events of interest at this stage of a program concern the status of its clients. Because these events must be ones observable independently of the program itself, they obviously could be the product of exogenous influences. By connecting them to events at successively earlier stages, however, it is possible to conclude with confidence that the results are indeed attributable to the program's inputs. And, by including among the targets all aspects of client status which we think necessarily must be achieved to produce the desired, longer-range goals, targeted achievements are connected to ultimate outcomes. This stage, then, completes the chain of events from program design to ultimate outcomes.

Recapitulating at this point, a program arbitrarily can be divided into stages in order to limit the span of individual programmatic hypotheses. As shown in Figure B, planned inputs describe the status of program planning in terms of the initiatives to be undertaken. These actions are expected to produce a stage of internal operations reflecting the readiness status of the program. Immediate effects are the first changes to occur at the client level and characterize the status of inter-

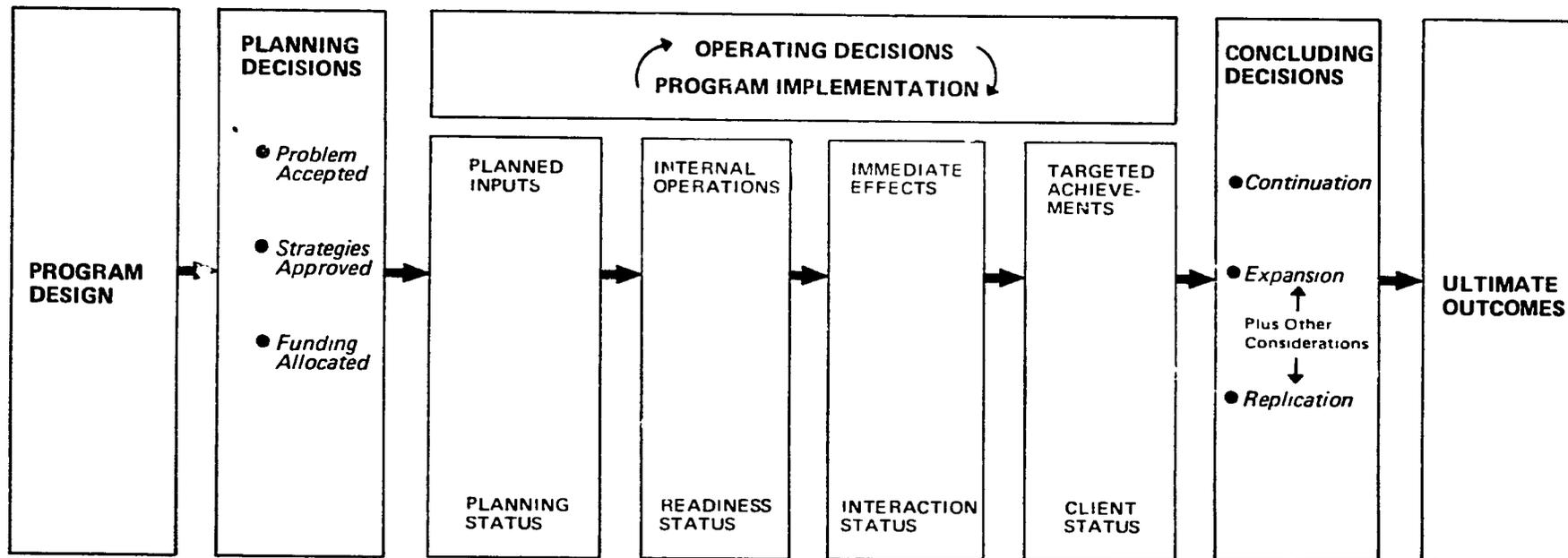


FIGURE B

actions between the client and the program. Targeted achievements are entirely client states that represent desirable end products of direct program influences.

### Diagnosing Sources of Failures

Impact assessment refers to tests of the relationship between events in any two of the program columns. In the best of all worlds, evaluation would be concerned with all of the hypothesized linkages. In a more practical sense, it is neither feasible nor worthwhile to measure every one of these relationships since many are bound to be too trivial or too predictable to warrant the effort. Those which are selected for assessment should be chosen because of their potential impact on the decisions that are to be made upon completion of the program. The needed information is not necessarily restricted to targeted achievements and their antecedents. As described earlier, there is considerable value in knowing which hypotheses were confirmed and which were not even if the desired end states were not achieved.

The same relationship between events used to establish attribution between causes and effects becomes the basis for diagnosing and remedying program defects. Each successive link in the chain represents an "if...then" expectation. After achieving the "if" end of the link, either as an input or the result of an input, we can look to see if the "then" end of the link appears. If the expected result occurs, and the span of the hypothesis is small enough for us to rule out alternative explanations, we conclude that the relationship has been verified. (Similarly, we would not expect the result to appear if its cause was absent. Only when the outcome does not appear as expected is it necessary to initiate a diagnostic effort.)

In order to diagnose a failure in some program component, that link must be further divided into a series of smaller steps. Most women referred for treatment to their nearest Health Subcenter, for example, might not receive that treatment. It then would be necessary to explore some of the possible reasons. Was it because they didn't understand what they were to do? Was it because they had no means of transportation? Was it

because the Subcenter was understaffed, or too crowded, or unable to provide the appropriate treatment? Considerable effort is required even to plan an effective diagnostic analysis, and collecting the needed data can be expensive even when potential weaknesses are explored in their order of probability. For this reason, it is impractical to create in advance all of the thousands of diagnostic routines which might be needed during the course of a program. Each requirement can better be dealt with as it arises. This is particularly true for a highly integrated program, such as DEIDS, where many inputs contribute to many outcomes.

Diagnostic routines, of course, tell us what to remedy but not what to do about it. In some cases, the assumption that "anything done was done well enough" will turn out to be incorrect. This would be a failure in the implementation of an input (e.g., the training given was insufficient for a key skill). In other cases, the problem will have been incorrectly assessed (e.g., hookworm rather than diet was the prime reason for malnutrition). And, in still other cases, the hypothesis itself will have been incorrect in the sense that the input was insufficient or of the wrong kind (e.g., the amount of salary for the Health Workers was not enough to keep them employed). The proper remedial action in each instance would be to alter the program in some way and then again check for the desired outcome.

#### 4. APPLICATION OF THE MODEL

##### DEIDS Project in Ecuador

The DEIDS Project in Ecuador was chosen to illustrate this approach to program impact assessment. As was pointed out earlier, however, the design of this particular project is not yet complete. For this reason, there are a number of gaps and speculative entries in the matrix that will be presented. The detail should be sufficient, on the other hand, to show how the model is used to generate the needed entries for each stage of a project. So that the overall intention of the project can be better understood, its purpose and operations can be very briefly summarized as follows.

The Ecuador DEIDS Project will begin in one province and then gradually expand to two adjacent provinces over a period of about seven years. During this time, some 720 Health Posts will be established, each serving approximately 1000 people. A locally recruited Health Worker will be assigned to the Health Post following a six-month training period. Referral sources and supervisory support for the Health Workers will be provided from Health Subcenters operated by the Ministry of Health. Reorganization of the administrative structure and operational functions of the Subcenters to better meet the needs of the program will be part of the project. Each Subcenter will be expected to provide support to about twenty Health Workers.

Once operational in a given community, the program is expected to provide integrated health services to the residents of that area, with special attention given to women of fertile age and child under five. The Health Workers will be expected to provide the clients of their area with a mix of preventive and curative services such as counseling on child rearing, minor first aid, immunizations, midwifery assistance, the distribution of family planning materials, nutrition advice, and sanitation information.

While it is impossible to present a specific schedule of Health

Worker activities at this stage of the project, a typical day might include the following activities:\*

8:00-10:00 a.m. At the Community Health Post:

- Diagnosis and Treatment of Minor Ailments
- Referral of More Complicated Cases
- Dispensing of Medicines
- Dispensing of Family Planning Information and Materials

10:00 a.m.-12:30 p.m. Home Visits:

- Diagnosis and Treatment of Minor Ailments
- Referral of More Difficult Cases
- Dispensing of Medicines
- Dispensing of Family Planning Information and Materials
- Health Education
- Nutritional Counselling
- Follow-up Care

12:30-2:30 p.m. Lunch Break

2:30-3:30 p.m. At the Community Health Post:

- Diagnosis and Treatment of Minor Ailments
- Referral of More Difficult Cases
- Dispensing of Medicine

3:30-6:30 p.m. At the Community Health Post:  
(Thrice a Week)

- Health Education Sessions
- Nutrition Counselling and Demonstration
- Family Planning Information and Motivation

4:30-6:30 p.m. At Mothers' Clubs in Community Hall, School,  
(Twice a Week) Church or a Private Home:

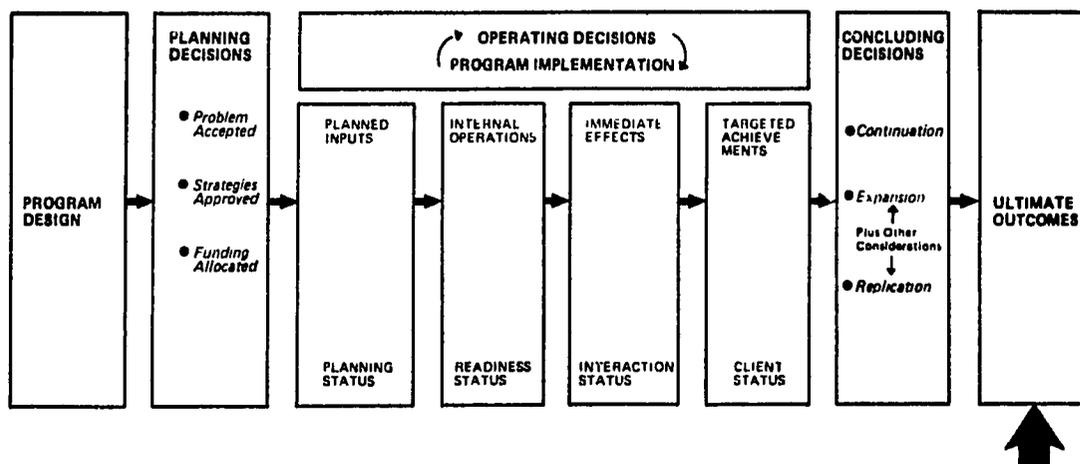
- Health Education
- Nutrition Counselling and Demonstration
- Family Planning Information and Motivation
- Home Economics

---

\* Aaron E. Ifekwunigwe, M.D. Daily Schedule for the Community Health Workers, 18 June 1974.

## Ultimate Outcomes for Ecuador DEIDS

An analysis of the network of events that constitute a program can begin at either end; that is, with the program goals (Ultimate Outcomes)



or with the choice of strategies (Program Design). Let's look first at one possible formulation of the goals of the DEIDS Program being developed for Ecuador, as shown in Figure C. Objectives in four areas are indicated, those related to cost, institutional, health and side-effect outcomes. Since the specific criteria for these objectives are specified in only some instances, it is necessary to estimate them for the remaining areas.

### *ULTIMATE OUTCOMES*

- *At a cost the Government of Ecuador is willing to afford,*
- *Develop an institution that is able to provide health services to two-thirds of the fertile women and children in the test areas,*
- *Which results in improvements with respect to:*
  - *maternal mortality & morbidity (15% reduction in rate)*
  - *infant & young child mortality & morbidity (24% reduction in rate)*
  - *mortality & morbidity due to infections, communicable & other preventable diseases (35% reduction in rate)*
  - *nutrition, fertility rates, and sanitation practices (a demonstrable amount)*
- *Without unacceptable side-effects.*

FIGURE C

Cost objectives can be thought of as whatever per-capita expenditures Ecuador would be willing to provide in order to continue an experimentally established health program indefinitely. Based on cost figures contained in the project plan budgeted for administrative personnel, health worker salaries and office space, this amounts to roughly \$.63 for each of the some 120,000 rural inhabitants in Cunar Province. A number of other costs can be anticipated, however, such as expenses for materials, transportation, supervision and refresher training. A more realistic estimate of what would be an acceptable level of costs, which takes these requirements into account, would be perhaps \$1.50 for each rural inhabitant per year. (Note that this figure refers to all rural inhabitants in the test area. If only two-thirds of the population participates as clients, the amount would be \$2.25 per client. Furthermore, it is intended that most of the available investment will be directed at women of fertile age and children under five, permitting a differentially higher amount to be programmed for services to these groups.)

It is plausible, of course, that a substantially higher amount would be considered acceptable by the time continuation decisions are to be made, but present estimates may be safer as an objective. Furthermore, no provision has been made in this estimate for the possible participation of other donors. It should be clear, too, that not all costs are financial. For this reason, various "opportunity costs" (e.g., recruiting top-notch people for this program would prevent these individuals from being available for other programs) have to be considered as well as "programmed costs" (those already budgeted) and "non-programmed costs" (those not yet budgeted but which may be needed to overcome program deficiencies).

Institutional objectives encompass the various functional, administrative and support services required for the continuing operation of the program. For the DEIDS Project in Ecuador, these include "...the development of administrative and managerial skills in the Provincial Health Office, ... establishment of an effective evaluation unit, ... reorganization of the rural health subcenters, and ... establishment of a network of health posts (staffed by) community health workers." It is intended that these institutional objectives include provisions for the

training and retraining of personnel, the compilation and analysis of vital statistics and health records information, a budgetary control system, a materials distribution system, and the coordination of program efforts with those of traditional health practitioners, among others.

Health objectives for this project proved particularly challenging. The fundamental approach adopted for the project has a broadly based preventive orientation rather than a specifically directed curative one, which causes considerable difficulty when efforts are made to state intended outcomes in objective terms. Part of the problem stems from a hopefully temporary inability to convert rather global ideas into more concrete language. For instance, a desirable end-state of this project is to make people more "sensitive to" and "concerned about" health in general. While it is not too difficult to identify plausible examples of what is meant by these terms (e.g., a growing distaste for shopping at market stalls where the food is covered by flies), it is extremely hard to find ways of describing these outcomes in such a way as to permit agreement on when and if they have been reached.

Another and perhaps more significant part of the problem is a confusion between means and ends. Neonatal tetanus, for example, can be virtually eliminated either by clinical intervention (innoculating the mother during pregnancy) or by improved sanitation (sterilizing possible sources of infection before delivery). The end-state sought is a reduction in neonatal tetanus, and this objective could be achieved by either route. In the same way, the use of community education in the hope of reducing the prevalence of skin infections is an alternative methodology to the use of ointments for the same purpose. Whether the emphasis of a health program is preventive or curative, expected health outcomes will be similar if not identical.

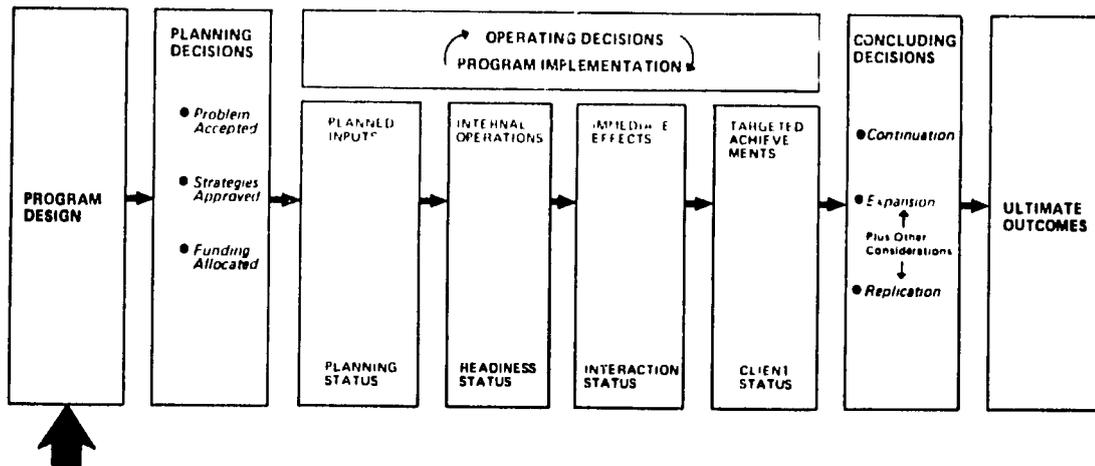
For this project, it is convenient to group health objectives into six categories: maternal mortality and morbidity, infant and childhood mortality and morbidity, mortality and morbidity from infectious diseases, fertility, nutrition and sanitation. As will be seen, there is considerable overlap among these areas as, for example, the influence of fertility on family nutrition and the mortality of women of fertile age, or the

influence of sanitation on almost all other objectives.

Side-effect objectives include those which would be viewed as potentially detrimental, such as a possible surge in paralytic cases of polio following improved sanitation, and those which would be viewed as potentially favorable, such as increased school attendance following better nutrition. Identifying possible side-effects in advance is an interesting way to spend some time, but the resulting list is necessarily endless. Experience has shown that most of the more pronounced side-effects of any program are readily recognizable when they occur and that it is relatively pointless to try to anticipate them in advance except when negative outcomes are so certain as to necessitate plans to deal with them.

#### Design Options for Ecuador DEIDS

Turning now to the choice of strategies, it is possible to consider what guided the selection of inputs. Because planning decisions already have been made, very little emphasis was given to this topic in preparing



the report, and therefore, the list in Figure D is somewhat tentative and incomplete. It should be evident from the entries in the list that other approaches could have been used. This design, as all program designs, represents the best judgment of which available options could best be hypothesized to work out successfully in the test area.

## PROGRAM DESIGN OPTIONS

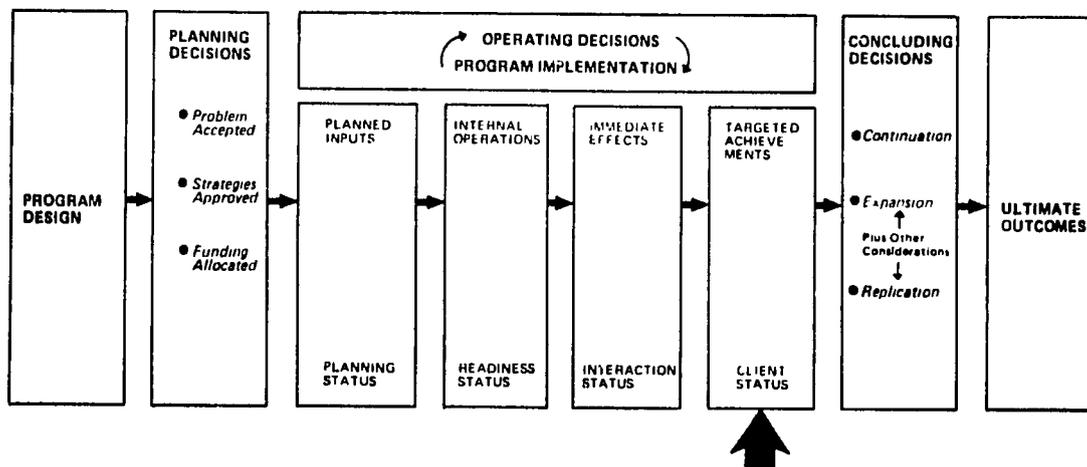
1. *Many auxiliaries vs few professional personnel;*
  2. *Partially community vs entirely central responsibility;*
  3. *Generally integrated vs generally specialized services;*
  4. *Heavily preventive vs heavily curative emphasis;*
  5. *Largely indigenous vs largely external resources;*
  6. *Primarily technical vs primarily commodity assistance;*
  7. *MCH, FP, N & S vs other possible mixes;*
  8. *Public sector vs private sector implementation;*
  9. *Health service vs health research objectives;*
  10. *More clients vs more quality;*
- ... and others*

FIGURE D

Now that boundaries of the matrix have been identified, it is possible to look at the way entries falling within the scope of the program are determined. Once again, while the analysis might begin from either end (or perhaps in the middle, too), starting with the last events in the chain is convenient. At this point in the discussion, the intent is to provide some examples and describe what they represent; the relationships among entries will be considered later on.

### Targeted Achievements for Ecuador DEIDS

Targeted achievements are direct outcomes of the program which generally are regarded as benefits in their own right and, in the case of DEIDS,



represent desired client states. To be considered as a targeted achieve-

ment, an outcome must be one which is consistent with program inputs. Taken together, all targeted achievements must be sufficient to lead us to expect that the ultimate outcomes, or goals, of the program will result. Figure E contains examples of targeted achievements appropriate to the DEIDS Project in Ecuador grouped under the same headings that were used to describe its ultimate outcomes.\* Entries preceded by an asterisk were felt to be particularly essential. Note that each entry must have a quantitative as well qualitative aspect (the base for calculating percentages under Health Objectives will be considered in a moment).

Even from these examples, several generalizations are clear. First, it is possible to describe desirable client states (or their parallels in the case of cost and institutional outcomes) in objective terms and based on criteria which are beyond the direct control of the program. These are states which could result from program efforts, but they also could be a consequence of non-program influences. Second, while this list is by no means exhaustive in the sense of describing the complete DEIDS Program in Ecuador, there already are numerous instances of repeated entries (such as III-A-4 and D-1, or III-D-4 and E-3; and, in a moment, an example will be given of overlap between Institutional and Health Objectives). This is expected to be true not only within programs, but between programs as well. And third, it should be repeated that the entries in any group (when completed) are those felt to be sufficient to reach the goals of the program. Having more than is required potentially represents either too modest a goal or an over-estimate of the inputs needed for the program's success.

---

\* Targeted achievements for nutrition, fertility and the immunization part of infectious diseases were identified for this report by a group including A. Ifekwunigwe, E. Boostrom, H. Dalmat, D. Rice, J. LeSar and the author.

## *SAMPLE TARGETED ACHIEVEMENTS*

### *I. Cost Objectives*

#### *A. Programmed Costs*

- 1. continuing costs of the program do not exceed \$1.50 per capita per year*
- 2. facilities required for the test area are completed by the end of the project*
- 3.*

#### *B. Non-Programmed Costs*

- 1. no further non-programmed costs are anticipated at the end of the project*
- 2.*

#### *C. Opportunity Costs*

- 1. this program does not preclude the initiation of other health projects in the test area*
- 2.*

### *II. Institutional Objectives*

#### *A. Functional*

- 1. staffed Health Posts exist to serve two-thirds of the rural population in the test area*
- 2.*

#### *B. Administrative*

- 1. all evaluation activities required for program continuation have been assumed by the evaluation unit*
- 2. all training activities required for program continuation have been assumed by the training unit*
- 3.*

#### *C. Support*

- 1. the materials distribution system procures and distributes all equipment and supplies recurrently required at the Health Posts*
- 2.*

### *III. Health Objectives*

#### *A. Maternal*

- \*1. X% mothers recover from hemorrhaging during delivery*
- \*2. all "small pelvis" mothers who want are delivered by C-section*
- \*3. toxemia avoided during pregnancy for X%*
- \*4. maternal anemia avoided for X%*
- 5.*

FIGURE E(a)

*SAMPLE TARGETED ACHIEVEMENTS (cont'd)*

*B. Infant and Child*

- \*1. X% recover from severe respiratory illness*
- 2. fluid balance maintained during diarrhea for X%*
- 3. X% infants eat solid food before weaning complete and by 6 mos.*
- \*4. X% mothers immune to tetanus before delivery*
- 5. X% children recover from non-immunizable infectious diseases*
- 6.*

*C. Infectious Diseases*

- \*1. X% children recover from non-immunizable infectious diseases*
- 2. immunity:*
  - \*a. X% newborn immune to tetanus at birth*
  - \*b. X% children immune to diphtheria by 12 mos. and up to 5 years.*
  - c.*
- 3.*

*D. Nutrition*

- \*1. maternal anemia avoided for X%*
- 2. maternal goiter avoided for X%*
- \*3. maternal calorie insufficiency avoided for X%*
- \*4. breast milk part of infant diet for 90% to 12 mos. and 60% to 24 mos.*
- 5. X% infants eat solid food before weaning complete and by 6 mos.*
- \*6. childhood calorie insufficiency avoided for X%*
- 7. childhood vitamin A deficiency avoided for X%*
- 8.*

*E. Fertility*

- \*1. no pregnancies for 50% women with four or more living children*
- \*2. no pregnancies for 50% women within 3 yrs. after birth of a presently living child*
- 3. breast milk part of infant diet for 90% to 12 mos. and 60% to 24 mos.*
- 4.*

*F. Sanitation*

- \*1. potable water regularly used by X% women and Y% children*
- 2. latrines regularly used by X% women and Y% children*
- 3.*

*IV. Side-Effect Objectives*

*A. Desirable*

- 1.*

*B. Undesirable*

- 1. no unacceptable side effects*
- 2.*

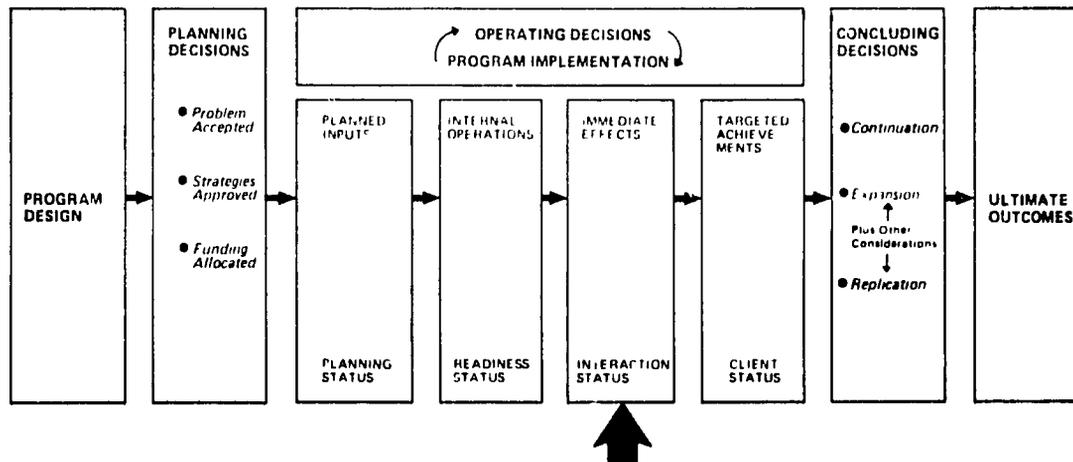
FIGURE E(b)

### Complete Lists of Targeted Achievements

More or less a complete list of targeted achievements for immunizations, nutrition and fertility were prepared for this report by a group of program designers. These targets are presented in Figure F. In considering what base to use where a percentage outcome was appropriate, it was decided that the Health Worker's registry would serve best. This is a list of inhabitants in her area which the Health Worker is expected to compile and maintain through continuous updating. Information on individuals is expected to be gathered for the registry whether or not these individuals in any way participate in the health program. As an Institutional Objective, it is expected that at least 95 percent of all inhabitants in the area will be registered. For Health Objectives, the aim of the program is to produce health improvements for two-thirds of all women and children in the area, or about 70 percent of those who appear on the registry. The same individuals need not be reflected in each targeted achievement, so it is likely that the actual percentage of women and children receiving at least some benefits from the program will be greater than 70 percent. (Note also that direct effects on 70 percent of a target group can have far more impact. Immunization to this level is sufficient for "herd immunity", for example.)

### Immediate Effects for Ecuador DEIDS

The next stage of the project to be considered concerns Immediate Effects which, in the DEIDS Program, refer to the status of interactions



## COMPLETE TARGETED ACHIEVEMENTS

### C. Infectious Diseases (goal = 35% reduction in rate)

- 1.
2. Immunity (note 1)
  - \*a. X% newborn immune to tetanus at birth (note 2)
  - \*b. X% children immune to tetanus by 12 mos. and up to 5 yrs.
  - \*c. X% children immune to diphtheria by 12 mos. and up to 5 yrs.
  - \*d. X% children immune to pertussis by 9 mos. and up to yrs.
  - \*e. X% children immune to measles by 9 mos. and up to 5 yrs. (note 3)
  - \*f. X% children immune to TB by 1 mo. (note 4)
  - g. X% children immune to polio and smallpox (note 5)
  - h. X% children protected against malaria (note 6)

#### Notes:

1. achieving the six principal targets (a-g) is expected to reduce childhood mortality perhaps 5-15%.
2. prevention of neonatal tetanus also will be achieved by sterile deliveries and/or use of cord dressing.
3. may be deleted as target; present cost of the one injection required is \$1.50 for vaccine alone, and continuous refrigeration is necessary; thus, these costs may exceed program resources; alternatives are to wait for lower vaccine costs, use of smaller doses (half U.S. standard), and/or immunization only for high-risk subgroups.
4. the target here is children, but greatest gains will appear after 20 years in the reduction of incidence of TB among women of fertile age.
5. presently low-risk causes of morbidity/mortality and these immunizations not expected to contribute to rate reductions in infectious diseases; will be offered, however, to meet Institutional targeted achievements because of current government policy; note also that polio may increase as a problem as sanitation improves.
6. because of low prevalence in high-altitude test areas and effectiveness of existing malaria eradication efforts in coastal areas, no program input required for this target at present.

### D. Nutrition (goal in a demonstrable improvement) (notes 7, 8, 9)

- \*1. maternal anemia avoided for X% (note 10)
2. maternal goiter avoided for X% (note 11)
- \*3. maternal calorie insufficiency avoided for X%
- \*4. maternal protein insufficiency avoided for X% (note 12)
- \*5. breast milk part of infant diet for 90% to 12 mos. and 60% to 24 mos. (note 13)
- \*6. X% infants eat solid food before weaning is complete and by 6 mos.

FIGURE F(a)

COMPLETE TARGETED ACHIEVEMENTS (cont'd)

- \*7. childhood calorie insufficiency avoided for X%
- \*8. childhood protein insufficiency avoided for X% (note 12)
- 9. childhood iodine insufficiency avoided for X% (note 11)
- \*10. childhood anemia avoided for X% (note 10)
- \*11. childhood vitamin A insufficiency avoided for X%

Notes:

- 7. ultimate outcomes are expected to include increased energy, decreased susceptibility to infection, and improved physical and intellectual growth for children.
  - 8. here, targets are avoidance of malnutrition states for as many individuals as possible; recovery from deficiencies is considered under maternal and childhood diseases.
  - 9. improved childhood nutrition also is expected to benefit the future infants of mothers who themselves had improved nutrition as children; this outcome would not be measurable for at least 20 years.
  - 10. particularly through added iron intake, most significant contributor to this target may be reduction of hookworm considered under preventable diseases and sanitation.
  - 11. iodine supplements to prevent goiter required only for inhabitants of locations within test area easily identified by prevalence of goiter.
  - 12. calorie intake simultaneously must be sufficient for this target to be achieved.
  - 13. also will have some impact on family planning outcomes.
- E. Fertility (goal in a demonstrable reduction in rate of uncontrolled births) (note 14)
- \*1. no pregnancies for 50% women with four or more living children
  - \*2. no pregnancies for 50% women above age 35 (note 15)
  - \*3. no pregnancies for 50% women within 3 yrs. after birth of a presently living child
  - 4. breast milk part of infant diet for 90% to 12 mos. and 60% to 24 mos.
  - 5. 50% recover from tubal infections without loss of fertility (note 16)

Notes:

- 14. it was expressly noted that "These targets have not been presented to the Ecuadorian Government for its approval, and may not be consistent with its family planning goals."
- 15. also considered as a maternal health goal.
- 16. includes treatment of male partner which is considered under infectious diseases.

FIGURE F(b)

between the program and the client. Once again, only examples of entries will be presented to illustrate the accomplishments that can be expected at this stage. All immediate effects taken together must be sufficient to produce the targeted achievements just described. They also must be consistent with intended program inputs in the sense of not requiring skills, facilities, support or other contributions that are neither expected as a result of program efforts nor available from existing conditions. Advocating the inclusion of green leafy vegetables in the diet to prevent anemia and vitamin A deficiencies may be inappropriate, for example, if such vegetables were only seasonally available and the program would have to take responsibility for supplying them at other times. In Figure G, the basis for percentages again is the population represented in the Health Worker's registry.

*SAMPLE IMMEDIATE EFFECTS*

*I. Cost Objectives*

*A. Programmed Costs*

1. *budget covers number of Health Workers required for intended services in test area*
2. *budget covers materials distributed to clients*
- 3.

*B. Non-Programmed Costs*

1. *budget covers cost of transportation for Health Workers*
2. *budget covers cost of processing referrals at Health Subcenters*
- 3.

*C. Opportunity Costs*

1. *demands for curative services do not limit preventive services*
2. *services for nontarget populations or outcomes do not limit intended services*
- 3.

*II. Institutional Objectives*

*A. Functional*

1. *client records maintained by Health Workers*
2. *program supported by traditional health practitioners*
- 3.

*B. Administrative*

1. *records kept of all expenditures*
- 2.

*FIGURE G(a)*

*SAMPLE IMMEDIATE EFFECTS (cont'd)*

*C. Support*

1. *community volunteers aid Health Workers*
- 2.

*III. Health Objectives*

*A. Maternal*

1. *X% pregnant women receive prenatal examination by fourth month*
2. *referral to Health Subcenter accomplished for X% of probable difficult childbirths (small pelvis, etc.)*
3. *Health Worker attends X% probable normal childbirths*
- 4.

*B. Infant and Child*

1. *X% of pregnant women receive two tetanus immunizations*
2. *X% of women with children under five receive instruction on when to contact Health Worker for examination of childhood illness*
3. *X% pregnant women receive demonstration on bathing infants*
4. *X% of children receive antibiotic treatment for severe respiratory illness*
5. *X% of children examined and weight recorded at no more than six-month intervals*
- 6.

*C. Infectious Diseases*

1. *X% of women receive instruction on food storage at least yearly*
2. *X% of children receive three DPT immunizations by 8 mos. of age*
3. *X% of children and fertile women suspected of having hookworm complete treatment*
- 4.

*D. Nutrition*

1. *X% of women instructed on need for green leafy vegetables in diet at least yearly*
2. *X% of women in identified goiter areas receive instruction on need for iodined salt (or substitute)*
3. *protein supplements distributed to X% of children suspected of protein deficiencies*
- 4.

*E. Fertility*

1. *individual family-planning contacts made with X% of all women with four or more living children or who are above age 35*
2. *X% of all women of fertile age informed of at least three alternative methods of contraception*
3. *X% of husbands of women requesting family planning assistance involved in decision process*

*FIGURE G(b)*

*SAMPLE IMMEDIATE EFFECTS (cont'd)*

4. monthly checks made on X% of women receiving birth control pills
- 5.

*F. Sanitation*

1. X% of households surveyed annually for potable water and latrines
2. Health Worker arranges assistance to X% of households wanting a latrine
3. X% of women and children informed of need to wash hands (with soap if available) before eating
- 4.

*IV. Side-Effect Objectives*

*A. Desirable*

1. traditional midwives adopt Health Worker delivery practices
- 2.

*B. Undesirable*

1. Health Worker loses credibility in community after unpreventable deaths occur
- 2.

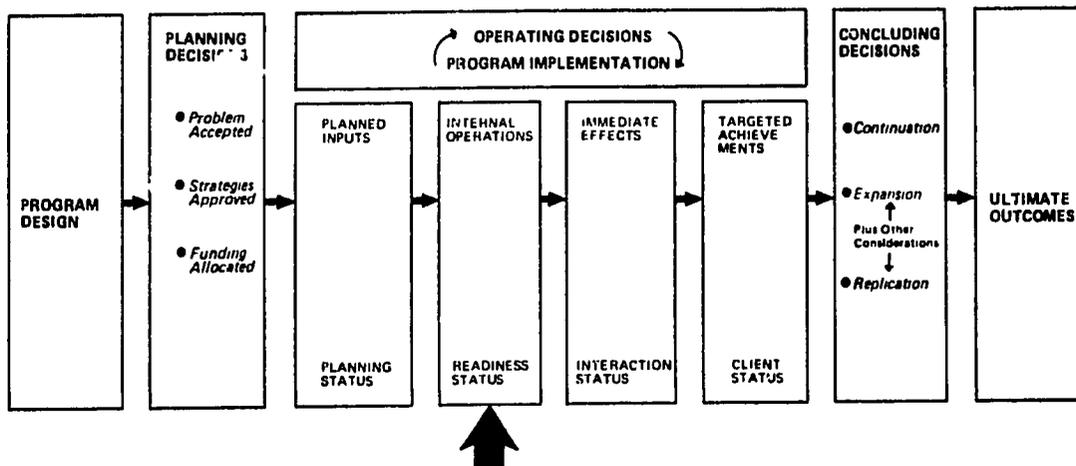
*FIGURE G(c)*

As indicated earlier, entries under immediate effects can be characterized as states-of-affairs which could be observed by either the Health Worker or the client (e.g., regularly receiving a supply of birth control pills is an event that both participate in, but taking the pills daily is not). Also, we assume that what was done was done adequately, in the sense of looking at events as they occur rather than at their later consequences. (When it occurs, we assume a Health Worker's instructions on infant care are adequate if they correspond to the intention of the program; determining whether they resulted in improved infant health is a separate problem.)

Internal Operations for Ecuador DEIDS

The program stage preceding immediate outcomes is Internal Operations. Entries describing the expected state-of-affairs at this stage of a DEIDS Program are concerned with the capacity, or readiness status, of the project. These entries do not involve the clients in terms of providing them with health services; instead, they represent the accomplishment of

various planned inputs. As pointed out earlier, there can be considerable differences in the time at which various entries are completed. For this reason, entries from this stage sooner or later must be accompanied by



a time-line reference to indicate when they appropriately should be accomplished. (As an example, it would be ineffective to establish and equip community Health Posts several months in advance of the deployment of trained Health Workers.) Selected examples of entries needed to specify what must be achieved by internal operations are given in Figure H.

*SAMPLE INTERNAL OPERATIONS*

*I. Cost Objectives*

*A. Programmed Costs*

1. six months' training for Health Workers results in sufficient skills
2. training of replacement Health Workers to overcome attrition accomplished within budget
3. Health Worker salary level sufficient to retain personnel
- 4.

*B. Non-programmed Costs*

1. refrigerators needed at Health Posts for immunization materials
2. materiel inventory losses must be replaced
- 3.

FIGURE H(a)

*SAMPLE INTERNAL OPERATIONS (cont'd)*

*C. Opportunity Costs*

- 1. personnel needs for field supervision diminishes referral capability of Health Subcenters*
- 2. refresher training accomplished without interfering with on-going health services to community*
- 3.*

*II. Institutional Objectives*

*A. Functional*

- 1. proficient Health Workers in place at all Health Posts*
- 2. adequate referral resources available for difficult cases throughout test area*
- 3. all Health Workers have operational manuals*
- 4. procedures established for detecting and combatting epidemics (e.g., cholera)*
- 5.*

*B. Administrative*

- 1. Provincial Health Office reorganized to meet program needs*
- 2. system established for timely collection of information from field*
- 3. Health Workers complete registry of inhabitants in service area*
- 4.*

*C. Support*

- 1. Health Posts equipped with needed equipment and apparatus*
- 2. sources of medical supplies established*
- 3.*

*III. Health Objectives*

*A. Maternal*

- 1. Health Worker has skill to predict difficult childbirths*
- 2. uterine contraction injection materials are at Health Posts*
- 3. potentially pregnant women know of health service program*
- 4.*

*B. Infant and Child*

- 1. Health Worker aware of unfavorable weaning practices in area*
- 2. Health Worker knows alternative ways for transportation of seriously ill children to Health Subcenter*
- 3.*

*C. Infectious Diseases*

- 1. supply of antibiotics at Health Posts*
- 2. sterile syringes and needles at Health Posts*
- 3. Health Worker able to identify potentially fatal conditions*
- 4. Health Worker washes own hands between examinations*
- 5.*

FIGURE H(b)

*SAMPLE INTERNAL OPERATIONS (cont'd)*

*D. Nutrition*

- 1. Health Worker knows which foods to recommend at all seasons*
- 2. Health Worker can demonstrate preparing tasty food from now undesirable supplies*
- 3. Health Worker knows of food supplement supplies available from other donors*
- 4.*

*E. Fertility*

- 1. supply of birth control pills at Health Posts*
- 2. Health Workers have met with traditional midwives in area to solicit cooperation*
- 3. Health Workers have met with women's club leaders*
- 4. Health Workers know of at least three alternative methods of contraception*
- 5.*

*F. Sanitation*

- 1. arrangements made to insure soap available in market*
- 2. Health Worker knows how inexpensive sanitary latrines can be constructed*
- 3. sanitary latrine constructed at Health Post*
- 4.*

*G. Side-Effect Objectives*

*A. Desirable*

- 1. other donors attracted to provide materiel support*
- 2.*

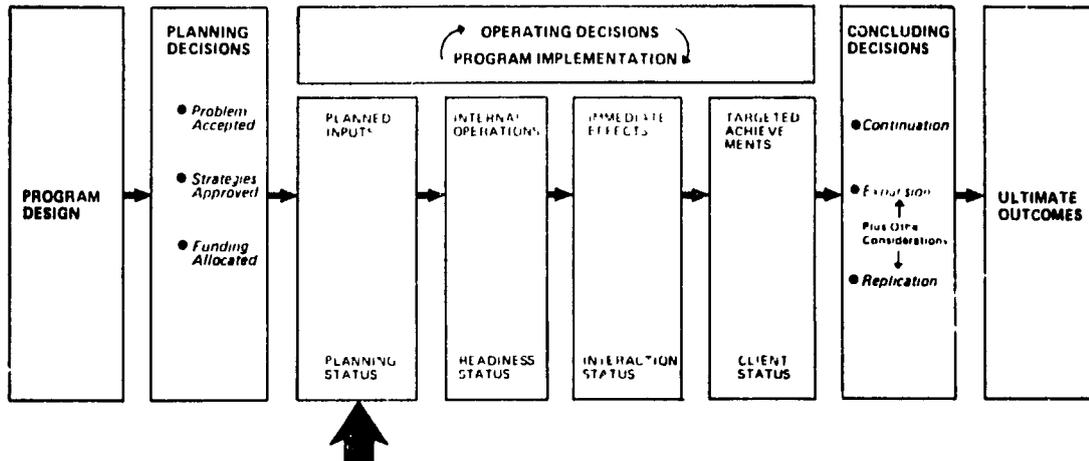
*B. Undesirable*

- 1. residents of areas not supplied with Health Posts in first year initiate competing program*
- 2. medical professionals in area leave because feel efforts no longer required*
- 3.*

*FIGURE H(c)*

## Planned Inputs for Ecuador DEIDS

Planned Inputs describe the first stage of a program but will be the last to be considered here. Entries at this stage are concerned with what



the program designers have accomplished. They do not relate to the success of what will be done but, instead, to the plans and preparations which have been made for institutional development, personnel training, administrative feedback, community participation and so forth. The criteria for planning status entries do not emphasize the quality of the approach (although making the details of the approach explicit facilitates the collection of constructive suggestions from available experts and focuses attention on those courses of action that will need careful watching), but rather their consistency and completeness. Since planned inputs also relate to and should be derivable from the strategies of Program Design, they complete the chains of events represented by this evaluation model. Some examples of planned inputs that might be appropriate for the DEIDS Project in Ecuador are given in Figure I.

## SAMPLE PLANNED INPUTS

### I. Cost Objectives

#### A. Programmed Costs

1. technical assistance budget for preparing and developing the training program is sufficient
2. qualifications for applicants for Health Worker positions are consistent with labor supply in test areas
- 3.

#### B. Non-Programmed Costs

1. too few training instructors available at no cost to the project
2. additional data must be collected before operational procedures can be established
3. students failing to pass proficiency tests must be replaced or given additional training
- 4.

#### C. Opportunity Costs

1. training and deployment of Health Workers must be delayed until Health Subcenters are restaffed
- 2.

### II. Institutional Objectives

#### A. Functional

1. health needs in the test area are agreed to by project and government representatives
2. operational procedures reflect all pertinent health needs
3. decisions made as to health conditions appropriate for referral to Health Subcenter
- 4.

#### B. Administrative

1. Health Subcenters reorganized or established as needed
2. supervisory routines established
3. appropriate systems developed for administration, accounting and quality-control
- 4.

#### C. Support

1. community groups that will provide Health Posts are organized
2. housing for technical assistance personnel made available
- 3.

### III. Health Objectives

#### A. Maternal

1. sufficient opportunities for reducing rates of maternal mortality and morbidity to expected levels within the scope of the program have been identified

FIGURE I(a)

*SAMPLE PLANNED INPUTS (cont'd)*

2. operational procedures, training materials and materiel requirements for each of these opportunities have been prepared
- 3.

*B. Infant and Child*

1. sufficient opportunities for reducing rates of infant and childhood mortality and morbidity to expected levels within the scope of the program have been identified
2. operational procedures, training materials and materiel requirements for each of these opportunities have been prepared
- 3.

*C. Infectious Diseases*

1. sufficient opportunities for reducing rates of mortality and morbidity due to infectious diseases to expected levels within the scope of the program have been identified
2. operational procedures, training materials and materiel requirements for each of these opportunities have been prepared
- 3.

*D. Nutrition*

1. sufficient opportunities for reducing incidence and severity of malnutrition to expected levels within the scope of the program have been identified
2. operational procedures, training materials and materiel requirements for each of these opportunities have been prepared
- 3.

*E. Fertility*

1. sufficient opportunities for reducing unwanted births and lack of concern over family size to expected levels within the scope of the program have been identified
2. operational procedures, training materials and materiel requirements for each of these opportunities have been prepared
- 3.

*F. Sanitation*

1. sufficient opportunities for reducing unsanitary conditions to expected levels within the scope of the program have been identified
2. operational procedures, training materials and materiel requirements for each of these opportunities have been prepared
- 3.

*IV. Side-Effect Objectives*

*A. Desirable*

1. plans made to help field evaluator staff become alert to possible side-effects

FIGURE I(b)

*SAMPLE PLANNED INPUTS (cont'd)*

2.

*B. Undesirable*

1. *project implementation reduces government investment in growth of traditional professional medical services*
- 2.

*FIGURE I(c)*

Interrelationship of Entries

No effort has been made in the above examples to relate entries at various stages of the project to one another as they would be in a final impact evaluation design. These relationships are illustrated by the following example. Here, directly related entries at each stage of the project are shown together. (Some indirect relationships between these and other entries in a completed network of rationales will be discussed following the example.) For this illustration, the avoidance of non-fatal skin and eye infections was selected because of its preventive emphasis; curative efforts leading to recovery from infections which do occur would be represented by a parallel chain of entries elsewhere in the matrix.

As can be seen from the example in Figure J; the two targeted achievements are "skin infections avoided by X% of women and children" and "eye infections avoided by Y% of women and children". One approach to accomplishing these goals would be to prevent intolerable levels of pathogens from building up on the skin and eye, to prevent the occurrence of micro-trauma (small cuts and bruises) to these surfaces, and to maintain the resistance level of the skin and conjunctiva. To attain these states, program-client interactions would be directed at improved cleanliness, increased concern for the skin and eyes, and better nutrition, as described under immediate effects. Note that all three entries at this stage contribute to both outcomes at the next stage, that no qualifications have been made as to whether the three immediate effects must be accomplished simultaneously or whether the advice should be given individually or in groups, and that the nutritional entry obviously is one that will be re-

<u>PLANNED INPUTS</u>	<u>INTERNAL OPERATIONS</u>	<u>IMMEDIATE EFFECTS</u>	<u>TARGETED ACHIEVEMENTS</u>
<p>1. analysis made of:</p> <p>a. prevalence and causes of skin and eye infections.</p> <p>b. current practices regarding skin and eye hygiene and protection against microtraumas.</p> <p>c. availability of good local sources of calories, proteins and vitamins.</p> <p>2. Health Worker training program includes instruction on preventing skin and eye infections.</p> <p>3. Health Worker training program includes instruction on presenting advice.</p> <p>4. teaching aids and reference materials prepared on skin and eyes for Health Worker use.</p>	<p>1. Health Worker has knowledge on:</p> <p>a. how to avoid pathogens, remove them, and prevent their growth on skin and in eyes.</p> <p>b. common causes of and how to reduce occurrence of microtraumas to skin and eyes.</p> <p>c. nutritional needs for calories, proteins and vitamins and how to meet these needs through available food sources</p> <p>2. Health Worker has skill needed to present information to individuals and groups.</p> <p>3. Health Post is equipped with charts and other community instruction materials and with references Health Worker needs to answer questions on skin and eye hygiene and diseases.</p>	<p>1. X% of mothers advised twice yearly on skin and eye hygiene with regard to:</p> <p>a. avoiding pathogens</p> <p>b. washing away pathogens</p> <p>c. washing away growth factors for pathogens.</p> <p>2. X% of mothers advised twice yearly on ways to reduce incidence of microtraumas to skin and eyes.</p> <p>3. X% of mothers advised twice yearly on caloric, protein and vitamin needs in diet.</p>	<p>1. skin infections avoided by X% of women and children</p> <p>2. eye infections avoided by Y% of women and children</p>

FIGURE J

plicated many times at this stage of the program in response to other targets.

Turning next to internal operations, the entries describe what preparations probably would be necessary for the Health Worker to provide this advice. As shown in the example, relevant considerations at this point include technical knowledge, teaching aids and counseling skill. Finally, the entries under planned inputs describe what plans the program would need in order to start this chain of events in motion. At each stage in the sequence, the events listed are those that would have to occur before the results at the next stage could be expected. Not all of the steps that could be taken are included (options chosen would be those felt to be sufficient, low cost and within the scope of the program) nor are all of the outcomes that could result from these efforts indicated (such as the impact of hand washing on gastro-intestinal infections).

#### Diagnostic Routines

Finally, this specific chain can be reexamined to illustrate how diagnostic routines would be created to deal with failures that could be detected during the course of the program. As a simple example, suppose we test Health Workers at the end of their training program to see if they have the "skill needed to present information to individuals and groups" (Internal Operations). Setting aside for a moment the question of how we determine the presence of this skill (until the next section of this report), suppose we find this skill is generally absent. Although the inclusion of instruction on this skill in the training program was expected (Planned Inputs), it is not the only potential source of the difficulty. Let's examine some of the possibilities.

First, we could look at likely errors in the implementation of the training. For instance, it could be that this block of training was inadvertently omitted for this group of trainees, or the instructor was not competent to teach the skill, or the amount of practice the trainees received was inadequate. If we are unable to demonstrate that any of these failures occurred, we next could look at errors in the way the problem was defined. For example, it could be that public speaking is regarded as

culturally inappropriate by the trainees, the Spanish language fluency of the trainees is too low, or there is some other unrecognized problem which was not considered in planning the training. Should none of these explanations turn out to be correct, the last area to investigate would be that the hypothesis linking the two events was mistaken. In other words, it may be our expectation that it is possible to produce speaking skills by training is wrong at least within the limits of present knowledge.

As one more example, suppose skin infections continue to be prevalent (Targeted Achievements) despite the program having accomplished all the states (Immediate Effects) though necessary to produce the desired outcome. This discrepancy could be due to an error in implementation, and that our assumption that whatever was done was done well is incorrect. The specified number of mothers could have attended explanation sessions, for instance, but these could have been presented in such a way as to make understanding the advice impossible. The source of difficulty also could be in the way the problem was defined. Washing with available ( and possibly contaminated) water may do more harm than good, or the recommended use of soap may be unaffordable unless the program supplies it. Finally, the hypothesis may be in error. Instruction to the clients alone, no matter how well it is presented, or how appropriate it is, may be insufficient to achieve this outcome. Perhaps another approach, such as the distribution and use of a medicated skin salve, may be a more effective solution.

As these examples suggest, any attempt to develop a complete set of diagnostic routines for a project as complex as the DEIDS Program in Ecuador would be quite unrealistic. Demonstration programs of this kind are carried out because there are not thoroughly tested technologies for overcoming the problems which need to be solved. Many "best guesses" necessarily are involved, and these will not always be correct. For this reason, it is essential to have an evaluation scheme which can provide an "early warning" of impending failures so that other alternatives may be considered or the need for new approaches identified.

Because interim measurements can be made at each successive stage in any chain of events, information on impending failures can be made

available to program managers as promptly as desired to implement diagnostic routines and corrective actions. In most cases, impact assessments needed to monitor programmatic accomplishments would be conducted serially as each new group of Health Workers are trained or as each new Health Post is established. Experience with the results then can be used to determine the urgency and necessity for subsequent measures of equivalent events (e.g., the adequacy of supervisory assistance to meet ongoing needs in a new section of the test area) or even for repeated measures of a given event (e.g., the availability of family planning materials at certain Health Posts).

## 5. THE DEVELOPMENT OF INDICATORS

Knowing what to measure only partly answers the question of how to carry out an impact assessment. Indicators are the measurement instruments used to determine if an event in the matrix of program rationales indeed occurred. The strategies used in developing usable indicies depend on the importance of the event, the degree of uncertainty surrounding its relationship to subsequent or antecedent events, and the relative value of measurement precision for the cost of the information obtained. To the degree possible, indicators are developed with a number of standards, or criteria, in mind.

### Assessment Indicator Criteria

First, indicators should focus on overt, openly observable events. Covert events simply do not have the veracity needed to be convincing. Our bias is rather uncompromising in this respect because we have no faith whatever in verbal reports of attitude, feeling or satisfaction. It is not that we doubt the existence of these states, but our experience repeatedly has been that such measures are incapable of differentiating between successful and unsuccessful programmatic alternatives. Support for a program by those it affects is better viewed as being evidenced by attendance at clinics, compliance with suggestions and cooperation in community efforts than as the cause of these outcomes.

Second, indicators should be as unobtrusive as possible. Too often, and particularly in pilot demonstration programs, the measurement process itself is a highly significant determinant of what results, and this rarely can be taken into account in considering the accomplishments of the program. The aim of using unobtrusive measures is not secrecy, but rather the avoidance of confounding between the program and the measurement process. For the same reason, the indicators should be as straightforward as possible. It would be much better to observe the performance of a skill than to administer a multiple-choice test of knowledge of that skill.

Third, the indices preferably should be quantifiable in some systematic way to increase the fineness of the inferences that can be drawn. In the course of action programs, the occurrence of an event rarely is "yes or no", but instead is "more or less". The need for quantification is particularly evident when it is recalled that evaluative information should be directed at the management process. "Fine tuning" of the program is not possible unless the measurements themselves are fine enough to provide early, unambiguous evidence of malfunctions.

Fourth, the indicators chosen should be those which generally can be applied continuously to allow, when desirable, the progressive collection of data which yields trends as well as measures of status. The use of indicators which are not bound to specific points in time also simplifies the data collection process by permitting measures to be made successively in a series of locations rather than all at once.

Fifth, multiple measures should be devised for key events, both to insure reliability of measurement in the statistical sense and as a practical precaution with respect to indicators which may not turn out to be feasible or appropriate in the field. An evaluation effort, like a program itself, should have sufficient redundancy at critical points to make failures improbable. Once the various measures have been tried, it would be possible to delete those that were least desirable.

Sixth, the ease and cost of measurement is a heavily weighted consideration in selecting among potential indices. Measures which are time consuming to collect (either for program personnel or special evaluators), those which require a particular expertise or the use of sophisticated apparatus, and those which involve measurement approaches which inherently are so complex as to limit potential reliability are not well suited for use outside of a laboratory setting. The size and representativeness of the sample required to produce useful conclusions also must be considered in this respect. Cost is especially relevant when the program being studied is itself intended to be low cost.

And seventh, it is important that the measures used, where possible, are potentially generalizable to other settings. As will be explained in the following section, a main emphasis of the proposed evaluation approach

is an assessment design that will permit components from various DEIDS Projects to be compared. Although each DEIDS Project will be "custom-fitted" to the particular needs and circumstances of its respective host country, it nonetheless is highly desirable to develop indicators for health delivery systems which could be adapted to other situations, settings and aims.

The design of measurement tools capable of establishing whether a critical event has occurred can be a protracted undertaking. Although the linkages among events which define a program help substantially in delimiting what ought to be measured, considerable ingenuity often is required to come up with acceptable indicators for assessing what has been accomplished. It is not surprising, then, that this step rarely is completed without several cycles of energetic brainstorming and patient field work. Good indicators are more likely to be developed than discovered.

#### Examples of Possible Indicators

In order to illustrate the kinds of indicators that might be appropriate for the DEIDS Project in Ecuador, speculative lists of potential indices were developed for four events that might be significant in the program, one at each stage. Each example includes compilations of both types of data and data sources. Generally, more than one type of data could be collected from any particular data source and, inversely, several sources often are available to provide any particular kind of data. In practice, several combinations of data types and data sources would be considered initially on a trial basis with only those showing reasonable promise retained for more serious examination. The final selection of indicators would be based not only on the confidence we have in them but on the range of data sources that would have to be tapped as well. Limiting their number, even at the loss of some validity or reliability, may be advantageous in terms of cost.

The first example, in Figure K, suggests possible indicators and data sources for an entry that might appear at the "Targeted Achievements" stage of a program under Health Objectives (Infant and Child): "X% newborn

immune to tetanus at birth".

*POSSIBLE INDICATORS: TARGETED ACHIEVEMENT  
(X% NEWBORN IMMUNE TO TETANUS AT BIRTH)*

*Types of Data:*

1. *number of infants confirmed neonatal tetanus*
2. *number of infants confirmed free of neonatal tetanus*
3. *number of mothers with tetanus antibodies at time of delivery*
4. *number of infants with tetanus antibodies at time of delivery*
5. *number of mothers having received tetanus immunizations before or during pregnancy*
6. *number of infants thought to have died from neonatal tetanus*

*Data Sources:*

- A. *interviews with mothers*
- B. *interviews with Health Workers*
- C. *Health Worker records*
- D. *compiled vital statistics*
- E. *observer reports*
- F. *medical (or laboratory) team surveys*
- G. *interviews with village gossip*
- H. *referral records at Health Subcenters*

*FIGURE K*

The next example, in Figure L, illustrates some indicators that might be examined for an entry under "Immediate Effects" for the Cost Objective (Opportunity Costs): "demands for curative services do not limit preventive services".

*POSSIBLE INDICATORS: IMMEDIATE EFFECT  
(DEMANDS FOR CURATIVE SERVICES DO NOT LIMIT PREVENTIVE SERVICES)*

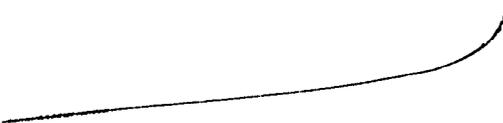
*Types of Data:*

- 1. ratio of Health Worker time spent on preventive contacts to curative contacts as a function of number of requests for curative services*
- 2. ratio of number of recent contacts per individual for preventive vs. curative services*
- 3. number of preventive contacts scheduled but not met*
- 4. existence of projected Health Worker daily schedule(s)*
- 5. number of times preventive efforts made during curative contacts*
- 6. number of times curative services sought from traditional health practitioners*

*Data Sources:*

- A. client interviews*
- B. Health Worker interviews*
- C. Health Worker records*
- D. observer reports*
- E. interviews with traditional health practitioners*

*FIGURE L*



The third example, in Figure M, consists of a list of likely indicators that could be investigated for an entry appearing under the "Internal Operations" stage of the program concerned with Institutional Objectives (Support): "Health Posts equipped with needed equipment and apparatus.

*POSSIBLE INDICATORS: INSTITUTIONAL OBJECTIVES  
(HEALTH POST EQUIPMENT AND APPARATUS)*

*Types of Data:*

- 1. inventory of equipment present at Health Post*
- 2. records of equipment shipped to Health Post*
- 3. Health Worker requests for equipment and apparatus*
- 4. supervisor records of missing equipment*
- 5. inferences from patient records as to equipment used at Health Post to provide health services*

*Data Sources:*

- A. Health Worker interviews*
- B. records at Health Posts*
- C. supervisor reports*
- D. supervisor interviews*
- E. records at Provincial Health Office*
- F. observer reports*
- G. bills of lading or delivery receipts*

*FIGURE M*

The last example, in Figure N, lists indicators which might be appropriate for a Side-Effect Objective (Undesirable) which could appear as an entry in the "Planned Inputs" stage of the program: "project implementation reduces government investment in growth of traditional professional medical services". (Note that attribution of this event to the program would be difficult because it would most likely occur as a consequence of an unforeseen occurrence.)

*POSSIBLE INDICATORS: PLANNED INPUTS  
(PROJECT IMPLEMENTATION REDUCES GOVERNMENT INVESTMENT IN GROWTH  
OF TRADITIONAL PROFESSIONAL MEDICAL SERVICES)*

*Types of Data:*

- 1. allocations in government budgets*
- 2. number of new hospitals constructed*
- 3. number of medical and nursing scholarships awarded*
- 4. number of candidates accepted for nursing training*
- 5. number of physicians sent out of country for advanced training*
- 6. number of alternative rural Health Service facilities established*
- 7. amount of free immunization materials distributed through private physicians*

*Sources of Data:*

- A. published reports*
- B. construction records*
- C. medical society records*
- D. interviews with Provincial Health officials*
- E. interviews with health professions training officials*
- F. surveys of private physicians*

*FIGURE N*

It had been hoped that this report could focus more heavily on the problem of developing than it has. A more thoroughly presented example based on inputs from program designers, for instance, may have helped make clear the process of distinguishing among events in terms of those which would need verification and those which probably do not. Similarly, it would have been helpful to have incorporated the opinions of experts on the practical feasibility of collecting various types of data and of tapping alternative sources of data in Ecuador (e.g., whether certain kinds of records exist, or the opportunity an observer would have to see certain events). This aspect of an evaluation is essential to planning and costing actual data collection, and considerably more would have to be done in this respect before it would be possible to determine how the needed information should be collected in the field and by whom.

## 6. DISCUSSION AND ANALYSIS

### Evaluation and Project Design

Only the bare beginnings of an evaluative plan which is sufficiently detailed and comprehensive to meet the challenge of the DEIDS Program has been presented. A complete evaluation design might take several months to complete even after a project's operational staff had achieved a firm grasp on what had to be done, and had devised at least a tentative program to do it. In the case of the DEIDS Program, the project now at the most advanced stage of planning is agreed to be far indeed from having all of the substantive issues well defined, let alone having their solutions in order. A cogent network of program rationales for the Ecuador Project cannot be compiled until the intended activities of the project have been established. And, in turn, this must wait until there has been an opportunity to collect such basic facts as what are the sanitation practices that should be improved and what are the causes of infant mortality that can be corrected.

On the other hand, it is our belief that these initial steps toward preparing an evaluation plan for the Ecuador Project have been a very valuable contributor to the design process. The project's staff has been well aware that far more thorough, careful and detailed planning will have to be accomplished before the potential benefits of their efforts can be realized. The conceptual framework underlying this evaluation plan provides them with a methodology that so far has been lacking in their initial attempts to relate intended inputs to hoped-for outcomes. It permits them to go beyond consideration of the processes whereby change will be effected and examine the targets of change as well. Only in this way can the prospects of improved health care be reduced to practice.

The absence of a suitable framework has not been the only barrier to more specific planning. The eagerness for positive action shared by almost all those associated with the project has suppressed the realities of such an undertaking. As now described, for example, the first six-

month period of the project is to include the gathering of necessary demographic information on the population to be served and its health problems, a task analysis of what indigenous Health Workers could do to overcome these problems, the design of a comprehensive six-month training program, the preparation of all needed training materials, the development of an operational manual, the training of trainers, the initial reorganization of Health Subcenters in the first test area, and the recruitment of the first group of Health Workers. No one would doubt that all these steps could be accomplished, but how well is a real issue. Earlier in this report, it was recommended that the schedule for each DEIDS Project include a longer preparation phase to permit more attention to be given to essential particulars before the start of the operational phase of a program.

In making these observations, it is recognized that there is a rather fundamental difference between how the Program is viewed by many of its designers and by others who have been participating in its formulation. On the one hand, these projects are seen as focusing on what mix of health services can be provided at a cost affordable to developing nations. On the other, the projects are seen as focusing on the steps that must be taken to ameliorate health deficits in developing countries. These positions are not incompatible, but they do result in discrepant priorities being assigned to various programmatic activities, and particularly those concerned with data gathering.

#### Data Gathering Alternatives

Not all the information which could or should be collected by the intended projects is pertinent to evaluation, and not all of the possible methods for assembling the information that is required are equally suitable in light of the aims of this program. During the development of the model, a number of alternative assessment plans were suggested or reviewed. Several characteristics of these other views deserve individual comment because of their conceptual or methodological contrasts to features in the proposed plan. (One of these, relating to the use of control groups, has been discussed earlier.)

Base-line information is essential both to establishing changes in rates (such as for mortality or fertility) and for determining the types of health services likely to do the most good. The latter, as noted previously in this report, is a necessary contribution to program definition. As such, it is properly the subject of evaluation (as one of the Planned Inputs) rather than part of it. Just as combatting malnutrition is a project activity, defining the problem is part of the project and not part of the impact evaluation procedure. Establishing changes in rate also is beyond the scope of impact assessment because, as noted several times earlier, the attribution of such changes to the accomplishments of any particular action program is virtually impossible. The purpose of trying to identify the uses of base-line data is not to avoid gathering them as part of the responsibility of an evaluation effort (indeed, this may turn out to be an expedient thing to do) but rather to insure that the costly process of data collection will be commensurate with practical needs for that information.

Demographic-analysis information represents a special case of too much emphasis on what data could be collected, and too little on how it might be used. The demographic characteristics of a population are important in assessing the transferability of a successful program to another location and in diagnosing the reasons for any differential impact of a program input on subsegments of the client group. But the more or less random compilation of this kind of information before needs are clear neither insures that the necessary data will have been collected nor that the collection process itself will be at all cost effective. Data collection aimed at determining the differential impact of a program is far more suitable to programs with static, nonrepetitive inputs (such as the construction of a dam) than to programs with ongoing, modifiable inputs (such as the delivery of health services.) Identifying why full impact was not achieved is second best to actually correcting an impact deficiency if this can be done considering the nature of the inputs. The first provides leads as to how problems might be avoided; the second provides answers on what to do to avoid them.

Cost-allocation information is an intended component of the proposed evaluation design (the Cost Objectives at each stage), but the emphasis on cost components is not as large as some might expect it to be. One, perhaps obvious, reason is that present costs are not necessarily good indicators of the costs that would be entailed were the project to be expanded or replicated elsewhere. Any change in the mix or density of services offered, in the opportunities to take advantage of economies of scale, or in the cost of various inputs (e.g., the amount of pay a Health Worker would expect) will have a far-reaching effect on the probable costs of an enlarged or transferred program. Furthermore, in a highly integrated program such as this one, the allocation of cost by program function or service rendered can be both difficult and misleading. Most of the changes which could occur in the program (e.g., altering the Health Worker's activity priorities) will take place with little or no change in budget. It will be important to maintain financial and time records, of course, to establish the total cost of the program and to determine where opportunities for cost savings and new activities exist. But this will not permit any sort of meaningful cost-per-outcome calculations.

Broadly these various data-collection alternatives are aimed at improving the fineness of the evaluation and at increasing the meticulousness with which outcomes are examined. It should be remembered, however, that action programs have neither the stability nor precision of a laboratory experiment. The present state-of-the-art in impact assessment is barely advanced enough to permit gross conclusions regarding the effectiveness of a complex program to be made with a tolerable degree of confidence. The specificity of the analysis can be enhanced somewhat by designing the evaluation so that within-program linkages are clear-cut, as is recommended in this model. But tightening the level of examination to the degree required to establish quantitative inferences between individual inputs and particular outcomes is presumptuous.

One further consideration in the choice of data collection approaches is their relative ability to identify and deal with unanticipated events. Holding progress at bay for the purpose of program evaluation is both intolerable and impractical. The occurrence of favorable as well as unfavorable outside events must be assumed for purposes of assessment

design, and the methodology employed must be capable of accommodating these events in terms of attribution of causality. Breaking up cause-effect relationships into a series of smaller linkages enhances the confidence we have in identifying what produced what. By carrying out the measurement process at or about the time the effect event is expected to occur, every possible opportunity will be created to detect the influence of extraneous happenings. The timing of an event as well as its appearance becomes an important aspect of confidence in attribution.

### Program Evaluation Alternatives

In light of these comments, it is possible to look briefly at a few of the alternative evaluation approaches that have been proposed for use in the DEIDS Program. Perhaps the most elaborate of these is the Special Evaluation Activity Relative to DEIDS/Ecuador (15 March 1974). While many of the aims described in this proposal are of unquestionable interest, the evaluation technology necessary to support the means for establishing answers is absent. There is no valid program assessment methodology which can provide information on the "most cost-effective way to achieve integrated maternal and child health -- family planning programs" or the "best ways to solve program problems." Similarly, the demographic emphasis of the approach, which includes such socio-economic and cultural data as the history and ecology of the area, results in a plan which is quite tangential to the principal needs of decision-makers. A large number of hypotheses are proposed for investigation, but there is no obvious relationship between the probable results of these investigations and the use of resulting information to improve this project or, particularly, to facilitate its replication elsewhere.

It is not so much what this plan includes (which leads to budget and resource requirements approaching those needed to carry out the project itself), however, but what it does not include that is most important. The DEIDS Projects are not being undertaken to examine the innumerable variables likely to affect the availability and utilization of health services in the developing countries, but to do something about them. The primary focus of the DEIDS evaluation plan should be on the effectiveness of optional ways of manipulating programmatic inputs which transcend

the specific requirements of any one project and thus might be generalizable to all future ones.

A second possibility, presented in Report of World Health Organization Informal Consultation on Evaluation of Health Services Research and Development Projects (Fourth Draft of a report on a conference held 8-12 October 1973 by Timothy D. Baker, M.D. et. al.), has a number of features in common with the approach described in this report. It contains, for example, an emphasis on institutional as well as health objectives, on information which is transferable to new applications, and on decision-makers as the principal users of evaluative information. The model also divides a program into stages, although the end-state for measurement purposes (excluding long-term effects which cannot be attributed to the program) consists almost exclusively of services received (comparable to Immediate Effects) rather than attainable and demonstrable improvements in client status (referred to in this report as Targeted Achievements.)

More fundamentally, the approach prepared for WHO does not establish the relationships (or linkages) among events in such a way as to permit concluding, with a reasonable degree of confidence, that outcomes were produced by inputs except through the use of control groups. This distinction in methodology is a critical one in light of the cost of gathering control-group data and the difficulty of establishing equivalence between the two populations. On the other hand, the report goes beyond what has been discussed here in considering the need for consistency in the way data from all health-related projects are compiled to improve the exchange of information among programs, donors and host-country planners.

A third possible source of an evaluation design is inherent in the Logical Framework (the "Log Frame") prepared to accompany the Proposal for the DEIDS Project in Ecuador (12 March 1974). Those familiar with the Log Frame approach are aware that it was intended both as a project planning and design tool and as a preliminary formulation of project evaluation intentions. As it turns out, the one prepared for the Ecuador DEIDS Project does contain many of the same elements as the sample entries listed earlier for events occurring at successive stages of the program.

Again, however, it is not the characterization of events but the delineation of their linkages which is crucial to the establishment of generalizable cause-effect relationships. In the future, it may be possible to refine the analytic procedure represented by the Log Frame approach to include these linkages. At that point, a completed Log Frame would be equivalent to a first approximation of the matrix of rationales needed to initiate an effective evaluation plan. Both planning and evaluation clearly would be improved as a result of this increased precision.

### Multiprogram Comparisons

Program design is an iterative process. With each subsequent success, more of the information needed for increasingly precise control over program events is accumulated, and greater attention can be focused on progressively narrower issues. Perhaps the only reasonable way to compress this process is to accumulate information on a number of roughly equivalent programs simultaneously. This makes it possible to discover not only what works, but which approach works most expeditiously. The specific contrasts that are so difficult to isolate within any integrated program often are clearly evident between programs. Comparisons among different approaches, however, are meaningful only when a consistent evaluation methodology is applied across all relatable programs. For this reason it is essential in the DEIDS undertaking that a common assessment model is adopted for all projects.

There will, of course, be differences among the specific DEIDS Projects depending on host-country resources, local circumstances, and desired goals. In this sense, each project will be "custom fitted" and the evaluation plan for that project will have to be adaptively tailored as well. The specific entires at each stage will vary to the extent that each project varies. Nevertheless, an assessment model such as the one which has been presented permits the aggregation of information on effective and ineffective interim program linkages across programs. Even though some inputs, goals, or within-program methodologies may be different from one program to the next, many others will be shared in common by several of the programs. To the extent these commonalities exist, a single

overall approach to evaluation which focuses on component linkages will permit the maximum possible to be learned from the DEIDS Program about possibly generalizable and replicable features of low-cost health delivery systems.

Attempting to anticipate these common components or linkages in advance can be sorely unproductive. It is far more efficient to begin with one project (such as was done here with the DEIDS Project in Ecuador) and successively incorporate the features of additional projects in step-by-step fashion. Increments in the variety of linkages necessary to accommodate each new project will diminish with every added program. The same general methodology also applies to the development of indicators. Experience from one project can be transferred to the next so long as the events represented by each of the indicators are reasonably identical. This cooperative sharing could be an enormous advantage over the more usual and costly trial-and-error process applied to the development of impact indicators for individual projects.

Finally, use of an assessment model which is directed at the discrete components which constitute an integrated health program allows application of the resulting evaluation technique for assessing the status of health services even in the absence of a program. Individual events which might be selected as Targeted Achievements, for example, could be investigated in advance of a program to determine which improvements were most needed and to estimate what inputs would be required to alleviate these problems. Similarly, it would be possible to inventory pertinent resources or constraining circumstances in a new location to greatly facilitate program planning for that location. In this sense, the evaluation model is used to determine what baseline data are needed for project design purposes.

### Process Studies

Not all of the linkages which could be investigated within the framework of this evaluation design have yet been described. One group of these concerned with outcomes that are peripheral to or remote from the scope of the program. Contentment of the client population with their new

services has not been mentioned, for example, because the intent of the program is to improve health and well-being rather than increase political affinity. Similarly, a number of probable outcomes are not listed among the Targeted Achievements because their attainment, while a direct function of the program, is too distant to have immediate significance. Whatever is done to eliminate infant diarrhea and impetigo, for instance, will be a step forward in the ultimate control of typhoid.

A second group of linkages concerns those relationships where confidence in the inputs is too low to be tolerable for planning purposes. How many interactions between a pregnant woman and her Health Worker are necessary, or what skills should a Health Worker possess to deal with possible injuries, illustrate these kinds of questions. Process Studies is a suitable name for tests of alternative inputs conducted to determine which of several options should be adopted for program use. Because these studies would be project-specific, there would be no reason to orchestrate them across all DEIDS Projects. The results of these investigations will determine what inputs are required (or, more probably, how ineffective inputs can be corrected) and, to the extent they prove satisfactory in light of what outcomes result, they will be identified as successful program components through the regular impact assessment process.

#### Next Steps in Evaluation Design

Although the essential structure of the suggested evaluation model has been presented in this report, considerable effort and skill will be required to complete its application to the DEIDS Project in Ecuador and to determine its utility over the range of other DEIDS Projects now being planned. The time required to construct a reasonably accurate and finished matrix of programmatic events and their linkages should be expected to diminish with each successive project since significant overlaps among projects are anticipated, and a fully generalizable set of components should begin to emerge as this work progresses.

As mentioned several times earlier, additional planning efforts will be required for the DEIDS Project in Ecuador which cannot begin until the specific institutional and health problems to be dealt with are

defined in greater detail. While the matrix could not be completed before the necessary project plans are prepared, simultaneous efforts directed at the assessment activity would likely enhance the planning process itself. Furthermore, it would be senseless to wait until the evaluation matrix was completed before beginning to explore the development of usable indicators. Experience has shown that the process of creating and testing indicators can be long and frustrating. No particular objective would be served by delaying initial work in this area once a reasonable portion of the matrix was in near-to-final form.

Hopefully, both of these two activities, on the matrix and the indicators, would include the active participation of representatives from the remaining DEIDS Projects and the host countries. As suggested earlier in this section, the overall impact assessment model should be as similar from one project to the next as possible to facilitate the generalizability of information across the entire program. Coordination among assessment plans is highly essential, and the completion of the evaluation design should be considered a central program rather than specific project responsibility. At the same time, involvement in the evaluation design process should materially improve the individual plans being made for each project by focusing attention on the importance of conceptualizing these action programs in terms of targetable, meaningful outputs.

Host country involvement during all phases of the development of the corresponding evaluation plan similarly is essential in a number of respects. As repeatedly noted in this report, agreement among all parties in advance of a project as to the targets to be achieved contributes enormously to the effective coordination of the efforts of all participants. Furthermore, the later continuation or expansion of project activities within the host country substantially depends on continued evaluation for monitoring and quality-control purposes as well as to make possible the periodic redesign of the program to meet changing circumstances.

Until the appropriate indicators have been defined more thoroughly than has been possible in this report, the remaining two steps in the evaluation design, the specification of actual data collection procedures

and the determination of data analysis requirements, cannot be considered in any great detail. The features of the proposed model suggest, on the other hand, that neither of these two steps should be particularly complicated or costly. This is particularly appropriate for an intentionally low-cost program because it would be expected, as just described, that the host country would continue to apply these evaluation methods and instruments to gradually improve and expand the program long after external assistance was discontinued.

It will be essential, of course, to prepare plans for data collection and analysis with considerable care. One clear conclusion from past impact evaluation efforts is that the procedures for carrying out an evaluation are normally as critical to its success as its design. Consideration must be given to the problems of how the data are to be collected, from whom, by whom, and when. Some aspects of these problems obviously depend on the circumstances peculiar to that one setting. Nevertheless, it is possible to characterize, in advance, which data collection procedures are likely to work best and what resources are apt to be required for their implementation.

We are confident even at this point, for example, that on-site project staff almost always are the most desirable data collectors. Generally, much of the needed effort can be incorporated into normal program operations so long as the indicators chosen are ones which do not require burdensome amounts of time or highly specialized skills. Using project staff for this purpose also helps institutionalize evaluation itself so that it will be continued systematically throughout the future of the program. This approach also provides the project staff with immediate and meaningful feedback on their own efforts so they can see first-hand the areas requiring further attention. This use of project personnel, whether Health Workers, supervisors or other staff, depends on carefully prepared data collection routines and schedules, however, to prevent the need for evaluative decisions on the part of otherwise fully occupied personnel.

Still another advantage of involving project staff at all levels in the evaluation process is to make the standards for accomplishment as clear as possible. By focusing attention on explicitly enumerated targets,

concerns over evaluation are at least partially alleviated in that there is concurrence beforehand on what ought to be measured and how the measurements will be carried out. A well-designed evaluation emphasizes outcomes which are generally regarded as significant in their own right, and implementing this design as a team effort aids in coordinating the efforts of all participants to achieve these common goals.

### Costs of Impact Assessment

At this point in the development of the DEIDS Projects, accurate estimates of the cost of completing the model are almost impossible. Some fairly rough guesses can be made, on the other hand, based both on the apparent needs of the DEIDS Program and our experience with other impact assessment efforts in developing countries. These costs can be divided among the remaining steps of the evaluation plan.

First, the network of program rationales will have to be completed for the DEIDS Project in Ecuador, a preliminary set of indicators will have to be devised, and work will have to begin on developing the indicators so that needed data can be gathered systematically. Because the completion of the matrix should be closely tied to the intended refinement of project plans, both efforts should proceed simultaneously during the first six months after the on-site staff has arrived in Ecuador. This step would not require the full time support of evaluation experts nor would it occupy an undue amount of the time of field project staff.

Second, it would be highly desirable to involve the planners of the remaining DEIDS Project during the completion of the matrix for the Ecuador Project. Their participation would enhance their own planning efforts and facilitate the use of similar statements of program events among projects. No effort should be made, of course, to force convergence in the approaches used or in the events to be measured across projects. The interchange of the detailed planning information needed to prepare the matrix, on the other hand, would help increase the parallelism among evaluation plans. Work also could begin at this time on the development of rationales for each of the other programs. This step would be conducted concurrently with the previous one, utilizing the same evaluation resources.

Third, work could begin on the development of the needed set of indicators for Ecuador. Since there will be overlap from one project to the next in their indicator requirements, much of this effort could and should be a central program responsibility even if most of the attention initially would be on indicators needed in Ecuador. The appropriate time frame for this effort would be the second six months of the Ecuador Project, roughly corresponding to the period when the first group of Health Workers will be in training. Some work on the indicators would have to continue beyond this point, but at least much of the preliminary effort can be accomplished during this period. Again, evaluation experts representing the program as a whole should be involved in this step, and these efforts should include the participation, of an assistance to, the staffs of the other DEIDS Projects.

In addition, at least one member of the Ecuador Project staff would have to be made available for preliminary field trials of possible indicators. This individual should have the language fluency and site familiarity needed to carry out meaningful field assignments. Presumably, this person then could be chiefly responsible for the continued collection of data during the life of the Ecuador Project. Although periodic monitoring and analytic assistance would have to be made available for the project, especially to the degree that diagnostic routines are required, no on-site evaluation expert would be needed.

Fourth, procedures for collecting and analyzing impact assessment data could be developed, and staffing arrangements made for these ongoing activities. While the design of the evaluation and the development of indicators should be a central program responsibility, data collection and diagnostic analysis of this data should be an activity within each project. Anticipating the magnitude of staff required to meet these requirements is nearly impossible at this time because the amount of data needed will depend on preliminary judgments as to the reliability of the adopted indicators, on the number of linkages within the matrix about which there is sufficient doubt to warrant measurement, and on final agreement as to what information is needed to meet decision requirements. To the degree that service personnel could assume a reasonable amount of this responsibility, it is not likely that more than two or three field evaluators would

be required for the three-province test area in Ecuador, however.

And fifth, continuing supervision and coordination of the entire evaluation effort across all projects will be essential to insure the comparability of the information being obtained and to make certain that as much as possible is learned about the design of low-cost health delivery systems. A single point of reference for all of the projects should be established and maintained throughout the length of the program to facilitate the coordination of procedures and the exchange of findings (even during the projects, some direct benefits may be derived from the aggregation of information on linkages common to more than one project), and to prepare compilations of results across all projects.

This centralized activity also should be responsible for the continuing development of impact assessment methodology as it relates to health delivery systems. This methodology is still emerging, and the DEIDS Program provides the needed opportunities to create both practical and conceptual methodological advances. One of the more significant of these would be the investigation of possible applications of the model, or its adaptations, to other kinds of health programs and to the assessment of health needs in locations where no innovative health programs currently exist.

In summary, the size of staff required to continue the implementation of the evaluation model in the DEIDS framework is not large. Aside from the field staff needed for routine data collection and compilation at each project site, and the limited participation of host country officials and project designers in the evaluation design and indicator development process, the combined assessment needs of all four expected DEIDS Projects probably could be met with the equivalent of no more than three man-years of expert assistance per year for the remainder of the program. A staff of this size would be sufficient to coordinate the development and maintenance of the project matrices, to devise and help refine the needed indicators, to prepare and monitor data collection and compilation routines and to assemble and integrate evaluative information from all four projects.

Actions which should be taken to insure the timely and competent development of evaluation plans for the DEIDS Program have been summarized

earlier in this report. These principally relate to the need for centralizing the evaluation responsibility for all DEIDS Projects, for expert technical assistance in evaluation methodology, and for better and more continuous coordination between planning and evaluation efforts. A concerted attempt to resolve discrepancies in how the goals of the overall DEIDS Program are now being interpreted, and the consequent establishment of standards for project planning and evaluation also were suggested.