A STUDY OF THE PROCESS, EFFECTIVENESS, AND COSTS OF THE EPSDT PROGRAM IN SOUTHEASTERN PENNSYLVANIA

SECOND INTERIM REPORT
Conceptual Modeling

PHILADELPHIA HEALTH MANAGEMENT CORPORATION

PHILADELPHIA, PA.

A STUDY OF THE PROCESS, EFFECTIVENESS, AND COSTS OF THE EPSDT PROGRAM IN SOUTHEASTERN PENNSYLVANIA

SECOND INTERIM REPORT
Conceptual Modeling

PHILADELPHIA HEALTH MANAGEMENT CORPORATION

PHILADELPHIA, PA.

A STUDY OF THE PROCESS, EFFECTIVENESS, AND COSTS OF THE EPSDT PROGRAM IN SOUTHEASTERN PENNSYLVANIA

SECOND INTERIM REPORT
Conceptual Modeling

November, 1979

Patrick H. Irwin, Ph.D., Principal Investigator

Stephen G. Pellegrini, M.S., Project Director

Rosemary Conroy-Hughes, B.A., Research Assistant

TECHNICAL EVALUATION COMMITTEE

Member	Affiliation
Bernard Bloom, Ph.D.	Leonard Davis Institute of Health Economics, University of Pennsylvania
Lillian Chin, M.D.	Philadelphia Model Cities Program
Harold Dickson, Ph.D.	Deputy Director, Health Services Research Institute, University of Texas
Anthony F. DiStefano, O.D., M.Ed., M.P.H.	Associate Director of Public Health, Pennsylvania College of Optometry
Michael Kean, Ph.D.	Executive Director for Research and Evaluation, School District of Philadelphia
Doreen R. Steg, Ph.D.	Department of Human Behavior and Development, Drexel University
E. Peter Wilson, M.D.	Program Director, The Supportive Child/Adult Network

TECHNICAL STAFF

Technical Liaison	Rosetta Smith, R.N., M.P.H.
Systems Analyst	Michael Blair
Programmer Analyst	Helen Lieberman

GLOSSARY

CHAP: Child Health Assurance Program

Client: A child (0 - 20 years) who has been screened by EPSDT. If capitalized, it refers to a member of the sample Panel.

Effectiveness: Ratio of system process to system output.

Efficiency: Ratio of system output to system input. Can be expressed as the product of Effectiveness and Level of Implementation

Eligible: A child under 21 in a family with a Medical Assistance number registered with the Commonwealth of Pennsylvania, or a child with its own Medical Assistance number. If capitalized, it refers to a member of the sample Panel.

EPSDT: Early and Periodic Screening, Diagnosis, and Treatment

Five-county area: The southeastern portion of Pennsylvania in which PHMC administers the EPSDT program, including Bucks, Chester, Delaware, Montgomery, and Philadelphia Counties.

Implementation: Ratio of system process to system input. Can be expressed as the product of Workload and Productivity.

MA: Medical Assistance

Outreach: The process of identification, contact, and assistance of eligibles to use the EPSDT program.

Panel: All the children who became newly eligible for EPSDT as of January 1, 1979, and were either screened or remained eligible for three months. This group was tracked through the program for the purposes of this study.

Phase I: Outreach

Phase II: Administration of the EPSDT screening program.

Phase III: Actual screening by provider for EPSDT.

Phase IV: Reporting of results and compilation of reports for EPSDT.

Phase V: Follow-up and treatment for EPSDT.

PHMC: Philadelphia Health Management Corporation

Productivity: System process per unit resource.

Provider: A site certified by the Commonwealth of Pennsylvania to administer screens. If capitalized, it refers to a site which has screened a member of the Panel.

Workload: System input per unit resource.

TABLE OF CONTENTS

Technical Evaluation Committee		i
Technical Staff		i
Glossary		ii
Table of Exhibits		v
Selected Summary		vi
1.	Introduction	1
1.1	Performance Measures	3
1.2	Selecting Program Performance Measures	5
1.3	Models	6
2.	Conceptual Modeling EPSDT	7
2.1	EPSDT Subsystems	8
2.2	EPSDT Performance Measures	9
2.3	Input-Process-Output Schematic	9
3.	Outreach Performance-Measure Model	12
3.1	"Eligibility" - Definitional Digression	15
3.2	"Contact"	17
3.3	Efficiency of Implementation	17
3.4	Outreach Phase Model	18
4.	Analytic Paradigm	22
5.	Risk Identification Phase Model	23
6.	Risk Reduction Phase Model	28
	Appendices	35

TABLE OF EXHIBITS

Exhibit 1.	Comparison of Planning & System Models	6
2.	Schematic of EPSDT Events	10
3.	Outreach Model	19
4.	Program Data for Outreach Phase	20
5.	Performance Measures for Outreach Phase	20
6.	Analytic Paradigm	23
7.	Risk Identification Model	25
8.	Program Data for Risk Identification Phase	26
9.	Performance Measures for Risk Identification Phase	27
10.	Risk Reduction Model	30
11.	Program Data for Risk Reduction Phase (Sample Data)	31
12.	Performance Measures for Risk Reduction Phase (Sample Data)	32
13.	Program Data Reporting Format	38
14.	Performance Measures Reporting Format	39
15.	Program Data for Outreach Phase (Hypothetical Data)	42
16.	Performance Measures for Outreach Phase (Hypothetical Data)	43
17.	Standardized Effectiveness Rates for Outreach Workers	48

SELECTED SUMMARY

- A set of program performance measures should include the minimum amount of information regulatory and funding agencies need in order to monitor, assess and compare the performance of service programs (see Section 1.1). A conceptual model of the EPSDT program, specifying the relationships among administrative components of the program and based on an Input-Process-Output schematic, can be used to generate program performance measures (see Section 1.2).

- As a first step in specifying the conceptual model of EPSDT, three sequential phases or program subsystems are identified. Although these subsystems (casefinding; risk identification; and risk reduction) are independent, overall program quality requires a concerted, carefully coordinated administration of all three areas (see Section 2.1).

- For each phase of the program, performance measures are constructed by taking the ratio of a subsequent step in the schematic to a preceding one, and the output of one program subsystem becomes the input for the next phase (see Section 4).

- When the model is applied to recent Philadelphia data, the efficiency level of Outreach is found to be .09 (see Exhibit 5). This measure - the ratio of screens to "eligibles" - is the product of effectiveness (screens/contacts) and implementation (contacts/eligibles).

- Based on data for all five southeastern Pennsylvania counties, performance measures are calculated for the Risk Identification phase. The effectiveness level (the ratio of newly discovered problems to all problems) is substantially higher in Philadelphia than in the suburban counties, and the level of implementation (problem screens/all screens) is lowest in Chester and Delaware counties (see Exhibit 9).

- Performance measures for Risk Reduction are calculated on a sample of children who were screened initially during a three-month period in 1977 and who were still eligible for EPSDT when due for rescreening two years later (see Exhibit 12).

Second Interim Report:

Conceptual Modeling

This is the second interim report of a two-year project funded by the Health Care Financing Administration (HCFA) to study data collected by the early and periodic screening, diagnosis, and treatment (EPSDT) program in southeastern Pennsylvania. The first interim report dealt primarily with the description of a large secondary data set; this report presents a conceptual model of the program. We note that the performance-measure model developed here is certainly not the only model possible, neither in kind of model nor in degree of complexity. But it is the nature of models to be simplifications of reality. Loss of detail is the price paid for enhanced understanding of complex systems and, in the particular model offered here, for improved program control and quality.

1. Introduction

Conceptual modeling of the EPSDT program concerns specification of (a) the components of the administration of EPSDT apart from their empirical referents (i.e., the concepts) and (b) the interrelationships of these components (i.e., the model). This level of abstraction permits us to discuss, for instance, the concept of a screen and its relationship to an outreach contact without having to operationally define "screen" or "contact." A conceptual definition of a term identifies the relationships that that term has with other, presumably more familiar concepts: the term being defined is placed, so to speak, in a conceptual matrix of related concepts much as a dictionary defines one word with others that are more generally familiar. Conceptual modeling is similar to such a definition except that the focus is broadened to encompass several related terms simultaneously. A conceptual model emerges when the relationship among concepts is made explicit.

In spite of the avowed theoretical intent of this modeling exercise, we are mindful that simply understanding the way a system works neither enables the program's administrators to control it, regulatory and funding agencies to evaluate it, nor policy-makers to improve it. Hence, we attempt to express the model's components in terms of empirical performance measures. More accurately, the model itself generates those measures when it is formulated in terms of a general analytic paradigm based on the definition of efficiency as system output per unit input. The resulting conceptual model may not only have immediate use in EPSDT administration, but the analytic paradigm upon which it rests may have application for evaluation in other similar service programs such as CHAP.

The report that follows presents the models that have been developed in this second stage of the project. A note on the organization of this report seems warranted since the flow of presentation may seem to be interrupted by the explication of the underlying analytic paradigm: after an overview of conceptual modeling of EPSDT (Section 2), the discussion of the outreach phase model (Section 3) details the rationale for the model as it is applied to that phase. Rather than passing directly to the other phases, we then pause to consider the paradigm itself (Section 4) but continuing to illustrate each component with applications from outreach. We then return to the remaining two phases in applying the model and presenting results. Finally, the appendix provides three additional, somewhat more technical applications.

1.1. Performance Measures

There has been an increasing demand at federal, state, and local levels for accountability of health and social service programs funded by public monies. Responses to this demand have taken many forms: establishment of professional review organizations (PSROs), increased involvement of citizens and consumers in evaluating programs, legal mandates which require specific percentages of program budgets to be set aside for evaluation (e.g., Community Mental Health Centers

Act), and certification of facilities and programs (e.g., Joint Commission on Accreditation of Hospitals). Program performance measures are another manifestation of this concern for accountability.

Program performance measures, in this sense, refer to the minimum amount of information regulatory and funding agencies need in order to monitor, assess, and compare the performance of service programs. This definition does not mean that these measures have no applicability to the management and planning of programs at the local level. On the contrary, as will be demonstrated in this report, these measures, if based on a conceptual model of program operation, probably represent information necessary for the successful administration of any program. However, most managers will likely require additional information, some of which is specific to the target population, location, and organization of the program.

The above definition of program performance measures deserves further clarification: specifying the minimum amount of information needed implies an effort to limit the burden and expense of data collection incurred by service programs while collecting a sufficient amount of information for optimal decision-making; monitoring program performance indicates an interest in examining performance over time, i.e., longitudinal analyses, and not simply conducting one-time evaluations or cross-sectional analyses; assessing program performance implies examining the ultimate impact of service programs on clients; and finally, comparing program performance requires that these measures be applicable across programs, i.e., they should be valid, reliable, and readily available across service programs.

1.2. Selecting Program Performance Measures

The first step in selecting program performance measures is to indicate how those measures are to be used. We are seeking the minimum amount of information needed for monitoring, assessment, and comparison of service programs' performance as discussed above. The next step is to posit some conceptual model of program operation that specifies at least the input to the program, its output and process. (See Exhibit 1.) Ideally, a service program begins with the identification of some need; this identification of need often leads to the articulation of program objectives that, in turn, must identify a target or "at risk" population to be served - the program inputs. Attempts to meet these needs lead to the creation of a set of program activities (processes) which are delivered to the target population. The result of these activities are program outputs that minimally should meet program objectives and hopefully meet the identified need as well.

Exhibit 1. Comparison of Planning and System Models

Planning Model		System Model
Identified Need		
↓		
Program Objectives		
↓		
TARGET POPULATION	. . . .	INPUTS
↓		↓
PROGRAM ACTIVITIES	. . . . . .	PROCESSES
↓		↓
PROGRAM RESULTS	. . . .	OUTPUTS

It seems obvious that clearly specifying a model of program operation is a prerequisite to the development and selection of performance measures which will allow valid and reliable monitoring, assessment, and comparison of service programs. Additionally, such a model should help identify the minimum amount of information needed to account for the operation of the program. Unfortunately, such an approach has not often been taken, and performance measures are used in EPSDT and elsewhere that are at best of limited utility and at worst misleading.

1.3. Models

It is useful at this point to distinguish among various types of models. Ackoff (Scientific Method, Wiley, 1962, Chapter 4) characterizes models as one of three types: iconic, analogue, or symbolic. They vary in their levels of abstraction and, especially, in ease of manipulation. Iconic models represent reality usually by a change in scale, such as an orrery or a "model" train. Analogues characterize one feature of reality with another, such as on a road map where color can represent topography and a line can stand for a highway. Symbolic models caste those phenomena we wish to model into abstract symbols such as mathematical equations, e.g., $e = mc^2$.

2. Conceptual Modeling EPSDT

Ideally, we want to develop a symbolic model of program operation, since this will be the easiest to test and manipulate and hence should assist us in our effort to control and improve the operation of service programs. The remainder of this report will describe the development and application of such a conceptual model of program operation.

The need to formally model EPSDT is closely related to the need to monitor, evaluate, and control that program. Numerous flow diagrams of EPSDT are available that are themselves analogue models, but their use is more didactic than instrumental since they are oriented more to teach or demonstrate how the system works than to control the system or even to make comparisons across systems. For this latter purpose, we shall attempt to formalize the model beyond a simple analogue or flow diagram of events by specifying mathematically the relationships among performance measures.

In their abstract, symbolic form, models can be more readily manipulated for testing and are more susceptible to critique and improvement than are iconic or analogue models. As will be noted in the following exercise, due to the complexity of the overall system it becomes increasingly difficult to maintain a clear verbal picture of it as we move from the original input of eligibles to the final output of reduced risk. The use of symbols and mathematical relationships provides us with a handy shorthand for manipulation of complex systems but with a loss of the intuitive appeal that is the forte of less abstract models.

2.1. EPSDT Subsystems

For the purposes of constructing a model of EPSDT, the program may be divided into three sequential phases: (a) casefinding/outreach, (b) risk identification/health assessment, and (c) risk reduction/health assurance. This conceptual convenience should not be confused with a causal chain: an increase or improvement in outreach does not *per se* lead to an increase or improvement in health assessment, nor does improvement in the latter necessarily lead to improvement in health assurance. Due to this independence, overall program quality would be difficult to assess with only one measure or index such as penetration rate or screening volume; several indices would seem more appropriate. The independence of phases should not be overstated, however, since program quality is surely a result of a concerted, carefully coordinated administration of all three areas.

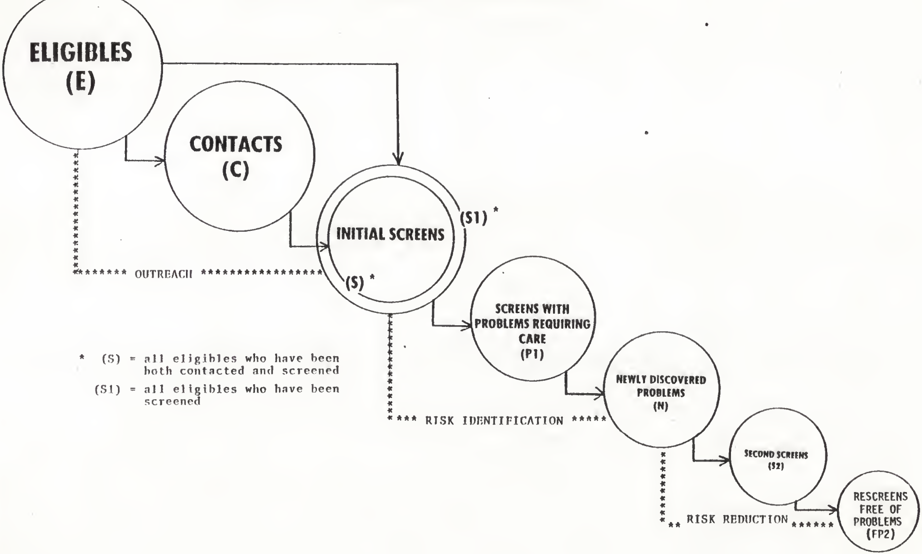
2.2. EPSDT Performance Measures

There are numerous overlapping and sometimes inconsistent performance measures that are or can be used to monitor EPSDT programs. Some conceptual scheme or model is needed to guide us in selecting measures that are valid and reliable and, when possible, that are currently available in administering programs. Ideally, we seek the smallest set of measures that will provide a consistent and complete picture (model!) of the program; however, our concern in this modeling is less to approximate a theory of how EPSDT works than it is to improve our ability to monitor and control the program. While these tasks need not be mutually exclusive, a conceptual model will be needed that is more instrumental than theoretical in focus.

2.3. Input-Process-Output Schematic

The instrumental approach that is used here posits three related, partial models of EPSDT, presented graphically in Exhibit 2 as a schematic flow of events. (See Exhibit 2.) Ideally, in the outreach phase, EPSDT

Exhibit 2. Schematic of EPSDT Events



eligibles are contacted and informed of the program; contacted eligibles are then screened. Clearly, not all eligibles can be contacted nor do all those contacted end up being screened. Thus, while it is not shown in this ideal schematic, at every stage a child can exit the system.

The "output" of outreach, screens, then becomes an "input" for the <u>risk identification</u> phase where children with problems are identified. (The other source for this phase is screens of uncontacted eligibles who entered the system through other means.) Some of the problems discovered are <u>newly</u> found as a result of the EPSDT screen; others are problems already known or under care. The discovery of these new problems found on initial screen is the output of risk identification and the input for the <u>risk reduction</u> phase of the schematic.

Since data on treatment follow-up are seldom available, the only estimation of risk reduction has to be based on those children who are rescreened. In one sense, this limitation is only apparent: EPSDT should, in fact, go beyond episodic care to bring the medically indigent into some regular health care system by establishing a continuing client-provider relationship. Rescreening itself may thus be indicative of the risk reduction process if the notion of "risk" is broadened to include potential and probable, as well as actual, problems. Returning to the schematic, those with new problems found on the initial screen are rescreened and problems can again be identified.

The opportunities for measuring performance based on this schematic may seem obvious, and, in fact, some measures are in current use, e.g., certain penetration rates such as the number screened over either the number of eligibles or number contacted. Less obvious are the ways in which these performance measures are related, their systemic interpretation, and even which measures of many are needed to model the program sufficiently to control it.

There follows first an example of how this model can be built (using first the outreach phase as an example) then an attempt to generalize the procedure, and finally an application to the risk identification and reduction phases.

3. Outreach Performance-Measure Model

We begin by introducing the notion of efficiency, defined as a ratio of system output to system input. The same definition was used in a similar context by Deniston, et al. ("Evaluation of Program Efficiency," Public Health Reports, LXXXIII, 7, July 1968, pp. 603-10):

> A definition of efficiency in public health programs may be formulated by referring to the classical definition of physical efficiency - the

> ratio between the energy output of a machine and the energy input supplied to it. In public health programs, efficiency may be defined as the ratio between an output [of the program] . . . and an input [of the program]. (p. 604)

Based on the schematic (Exhibit 2), for outreach we would use the proportion of eligibles (input) who are screened (output), or the ratio of the number of contacted children who were screened (S) to the number of children who are eligible (E): S/E. [This ratio, expressed as a percentage, has been (mis)used as a "penetration rate" for the entire EPSDT system. (See First Interim Report, Chapter 2.)]

Next, we consider effectiveness, a component of efficiency, which we define as a ratio of system process to system output. (For a somewhat different approach to effectiveness, see the companion article to that cited above, "Evaluation of Program Effectivenss," April 1968, pp. 323-35.) If we view outreach as essentially "selling" the program to potential clients, we can then measure outreach effectiveness by the ratio of screens (S) to contacts (C), i.e., the rate at which those "contacted" (however defined) actually get screened. The closer the ratio S/C comes to unity, as S approaches C, the more effective outreach becomes.

While effectiveness may be a component of efficiency, it is at least conceivable that the most effective program or system will be inefficient if it fails to reach the individuals it intends to serve. Outreach workers' ability to get individuals they contact into screening (the effectiveness of the contact) <u>interacts</u> with the rate at which outreach workers are able to contact eligibles. We may label this latter the level of <u>implementation</u>. Implementation level from the schematic is indicated by the proportion of eligibles who are contacted or C/E. Outreach efficiency can thus be expressed as the <u>product</u> of implementation and effectiveness:

$$\underline{\text{IMPLEMENTATION}} \times \underline{\text{EFFECTIVENESS}} = \underline{\text{EFFICIENCY}}$$

$$C/E \times S/C = S/E.$$

In this form, the interdependency of these components of the outreach phase is emphasized and can be interpreted. For example, S/C is interpretable as the <u>probability</u> that a contacted person will be screened (as well, in some studies, as a "penetration rate" - see "Final Evaluation Report - Phase 4, February 1976 - June 1978, EPSDT Demonstration in an Urban Setting - Dallas, Texas," Health Services Research Institute, The University of Texas Health Science Center at San Antonio, September 15, 1978). If this probability is constant <u>and</u> the number of contacts increases, then efficiency (S/E) must increase.

According to this model, for a fixed number of eligibles, if outreach implementation were increased by increasing the number of eligibles who are contacted (C), we might expect an increase in efficiency, the proportion of eligibles who receive screening (S/E). Clearly, without an even greater increase in the number screened, the effect will be cancelled out.

A slight but justifiable digression needs to be made at this point due to the importance in this model of the number of those considered "eligible" (E), which depends in turn on the definition of eligibility. We shall also address the operationalization used in this report for "contact."

3.1. "Eligibility" - Definitional Digression

If the model (or the ratio) is truly to reflect program performance, in the sense of penetration rate, it would seem reasonable to specify the target population. Familiarity with large data systems, such as that used by PHMC to administer EPSDT, usually reveals that a list of current eligibles is obsolete almost as soon as it is produced: the target not only moves, it changes its composition. In the case of PHMC's administrative area, as in other large-scale programs, the reason is "data system lag." An eligible individual is identified at one point in time by a state-run agency; the child enters PHMC's data system sometime later (perhaps months) and is reported, after that, to outreach workers as a "target."

There are several opportunities for an individual to exit the system (i.e., lose eligibility) between registering for Medical Assistance and the point that a screen or contact takes place. Most of these "fast exits" are not <u>at risk</u> for EPSDT long enough to affect or be affected by the program, but their influence on performance measures can be substantial: within 90 days of being posted as eligible, nearly 20% of newly eligible children were deleted from eligibility by the state agency before being screened (see <u>First Interim Report</u>, Section 1.3.2).

Unfortunately, it is sometimes only after efforts have been extended by outreach workers to contact an individual that it is discovered the family has moved, or never lived at a given address. In some cases, the family refuses to participate after being informed of the program. Whether such families - ones that have moved, broken up, never existed, or that refuse to participate for whatever reason - are truly at risk is problematic.

We may consider as a general rule for determining at risk that when outreach resources are used in trying to contact a family, even when the result is no response or a refusal to participate, outreach efficiency is affected. On the other hand, when eligibility is terminated prior to outreach activity, <u>these individuals are not at risk and should not be included in the number of eligibles</u> subject to program action. [If a "penetration rate" were required for later phases of the program and not just outreach, then eligibility should probably be further restricted to families known to be willing to participate.] The adjustment to the number of eligibles (E) to include only those at risk can be substantial, as indicated above, when the time "window" for eligibility of one group of children is confined to three months.

In monitoring outreach performance in this study, eligibility has been restricted to a sample panel of individuals listed for 90 continuous days as eligible plus those who were contacted and/or screened. (See First Interim Report, Chapter 2.)

3.2. "Contact"

Another definitional problem concerns what constitutes a "contact." The relevant dimension underlying effectiveness seems to run from informing eligibles personally to informing them through other, less aggressive means such as mail notification. For the purposes of the model, we are concerned only that "contact" be uniformly defined. In this study, a contact involves at least a personal, face-to-face encounter attempt with the outreach worker.

3.3. Efficiency of Implementation

If modality of contact is consistently applied across programs, we can then elaborate the outreach model a step further to analyze the efficiency of implementation as a function of outreach resources available. By including a resource variable in the model (e.g., number of outreach workers), we can express implementation itself as an efficiency index dependent on another level of implementation and effectiveness:

$$O/E \times C/O = C/E,$$

where O = number of outreach workers. Effectiveness, C/O, is simply the number of contacts per outreach worker; implementation, O/E, is the number of outreach workers per eligible. For our purposes, this latter index can also be expressed more conventionally as its reciprocal: the number of eligibles per outreach worker. We would then rearrange the equation as follows:

$$(C/O)/(E/O) = C/E.$$

In this form, we can interpret effectiveness (contacts per worker, C/O) as productivity and the level of implementation (eligibles per worker, E/O) as workload. According to the model, an increase in workload (E/O) will have no effect on overall implementation efficiency (C/E) without an even greater improvement in productivity.

3.4. Outreach Phase Model

We now restate the elaborated model of outreacn in

Exhibit 3.

Exhibit 3. Outreach Model

$$[(C/O) / (E/O)] \times S/C = S/E$$

Exhibit 4 presents the data called for by the model for outreach in Philadelphia County only since PHMC does not administer this phase in the suburban counties. These data are for the Panel of eligibles who were new to the data system as of January 1, 1979, and cover the first quarter of that year. (See First Interim Report, Section 1.3.2, for a review of the panel sample approach.)

Interpretation of the performance measures in Exhibit 5 would be improved if they could be compared to those of other outreach programs. Indeed, a great deal of caution should be exercised in interpreting these measures absolutely: their primary importance, at least at this stage of the model's development, seems to be comparative. It might be useful, with data on outreach restricted to just one program (Philadelphia County), to apply the same model to individual outreach worker's performance. (See Appendix, Example 3, for such an application.) Eventually, we could also compare efficiencies across phases within the same program.

That the outreach phase is "only" 9% efficient should probably not be interpreted as "low" given both

Exhibit 4. Program Data for Outreach Phase

(Panel data for first quarter, 1979)[a]

	Input	Process	Output	Resource
	Eligibles (E)	Contacts (C)	Screens (S)	Outreach Workers (O)
Philadelphia	7225	2981	637	66

Exhibit 5. Performance Measures for Outreach Phase

(Panel data for first quarter, 1979)[1]

	Input/Resource	Process/Resource	Process/Input	Output/Process	Output/Input
	WORKLOAD E/O	PRODUCTIVITY C/O	IMPLEMENTATION C/E	EFFECTIVENESS S/C	EFFICIENCY S/E
Philadelphia	109.47	45.17	0.41	0.21	0.09

[a] The Panel consists of all new Eligibles entering the data system for the first time as of January, 1979. (See First Interim Report, Section 1.3.1.)

the lack of comparable data as well as the nature of outreach. An active intervention program such as the outreach phase of EPSDT may well be inherently inefficient but still the only practical means of program recruitment and entry, more so for certain segments of the eligible population than others (e.g., the so-called "hard to reach" individuals).

In terms of the model, concern over the level of efficiency would lead to questions about the relatively low effectiveness of getting large numbers of children into screening with face-to-face outreach contact. Again, effectiveness is only low compared to the rate at which eligible children are personally contacted (level of implementation). Clearly, outreach may still be the most, if not the only, effective way to get the most needy of the medically indigent (e.g., those families on general welfare) into EPSDT. [Due to the recent termination by the state of this aggressive kind of outreach in Philadelphia, it will be possible to test this differential impact hypothesis by monitoring not only the changes in screening volume but also the proportions of families in various categories of assistance.]

Other interpretations can be made of this model by holding all variables on the left side of the equation constant ("ceteris paribus") except one and observing the effects on efficiency. It is more useful at this point, however, for us to attempt to generalize the model to the analytic paradigm that underlies it. After that we shall return to the remaining phases.

4. Analytic Paradigm

The initial step in (instrumental) modeling of a system is to identify the input and output of the system and the process whereby output for this system is produced. In the preceding example, the "system" was actually a subsystem of EPSDT: the outreach/casefinding phase. Once these factors are identified, we can place them into the schematic form of sequence of events,

INPUT -> PROCESS -> OUTPUT.

Efficiency was defined as output per unit input: the number of screens per number of eligibles. We define effectiveness in terms of output per unit process: number of screens per number contacted. Process (measured in some convenient units) per unit input is then defined as level of implementation: of the number of people who could receive the service, the proportion of those who did. In the outreach example, the proportion was the number contacted per number eligible. The analytic paradigm becomes,

IMPLEMENTATION X EFFECTIVENESS = EFFICIENCY

(Process/Input) X (Output/Process) = (Output/Input)

The performance measures are constructed by taking the ratio of a subsequent step in the schematic to a preceding step.

Further disaggregation of the level of implementation is accomplished by conceptualizing implementation itself as a measure of efficiency. This "nesting" approach to model elaboration seems appropriate here since it is logical to speak of the efficiency of implementation as well as overall system efficiency. Now, process becomes an output since it too can be seen as a function of some other controllable event: the number contacted is a function of, among other things, the number of outreach workers in a program.

In the outreach example, the model is elaborated by expressing efficiency in terms of measures of effectiveness, workload and productivity:

$$[(C/O) / (E/O)] \times (S/C) = (S/E),$$

or generally,

Exhibit 6. Analytic Paradigm

[PRODUCTIVITY / WORKLOAD] X EFFECTIVENESS = EFFICIENCY

We now apply the analytic paradigm to the risk identification and risk reduction phases.

5. Risk Identification Phase Model

From the schematic diagram (Exhibit 2), efficiency of risk identification can be indicated by the ratio

N/S1, or the number of children with newly identified problems divided by the number of initial screens. (S1 refers to all initial screens rather than to screens of only individuals who were contacted, S.) Interpretation of this and other indices is helped by considering the ratio at its extremes of zero and unity.

For example, under what circumstances could risk identification be completely inefficient? According to the index, N/S1, zero efficiency occurs when no new problems are found on initial screen. A perfectly efficient screening procedure (N/S1=1), where some new problem was found on every screen, seems both unrealistically pessimistic about the health status of those screened and unrealistically optimistic about the sensitivity of the screen. Still, it is at least conceivable, and has been argued by some physicians, that everyone has some "abnormality" that can be detected and treated; most otherwise "healthy" individuals do not consider such abnormalities worth the effort to correct. Suffice it to say that we expect neither extreme, but both are conceivably possible.

Effectiveness of risk identification is indicated by dividing the number of screens with problems newly found, as a result of the EPSDT screen, by the number of all screens with problems indentified on initial screening, N/P1. In certain environments, where medical services are readily available (e.g., in the middle and upper socio-economic strata), EPSDT risk identification would be inefficient because, in such an environment, it would be ineffective in locating problems not already under care. A similar argument may be made for environments replete with equivalent care resources that are accessible by EPSDT eligibles.

Implementation, the ratio of screens with problems to all screens (P1/S1), is analyzed in terms of <u>providers</u> - the resource variable that is analogous to the number of outreach workers as used in the analysis of outreach. Productivity in this phase is indicated by the ratio of problem screens to the number of provider sites performing these screens; workload is simply the number of screens per provider. With D1 = number of provider sites performing the initial screen (S1) the model for the risk identification phase is restated in Exhibit 7.

Exhibit 7. <u>Risk Identification Model</u>

$$[(P1/D1) / (S1/D1)] \times (N/P1) = (N/S1)$$

Exhibit 8 presents the Panel risk identification data for the entire program administered by PHMC as well as, for comparison, a breakdown by counties. In Exhibit 9, the performance measures are calculated to show an overall efficiency in this phase of 0.25, with Philadelphia County the most efficient (0.28). This higher efficiency is due to substantially greater effectiveness of risk identification (0.63) in Philadel

Exhibit 8. Program Data for Risk Identification Phase

(Panel data for first quarter, 1979)[a]

	Input	Process	Output	Resource
	Initial Screens (S1)	Screens with Problems Requiring Care (P1)	Screens with New Problems (N)	Providers Performing Screens (D1)
Total	1657	712	408	152
Philadelphia	1146	514	322	85
Suburban	511	198	86	67
Bucks	244	99	44	19
Chester	71	20	5	11
Delaware	117	25	12	15
Montgomery	99	54	25	22

[a] The Panel consists of all new Eligibles entering the data system for the first time as of January, 1979. (See First Interim Report, Section 1.3.1.)

Exhibit 9. Performance Measures for Risk Identification Phase

(Panel data for first quarter, 1979) [a]

	Input/Resource	Process/Resource	Process/Input	Output/Process	Output/Input
	WORKLOAD S1/D1	PRODUCTIVITY P1/D1	IMPLEMENTATION P1/S1	EFFECTIVENESS N/P1	EFFICIENCY N/S1
Total	10.90	4.68	0.43	0.57	0.25
Philadelphia	13.48	6.05	0.45	0.63	0.28
Suburban	7.63	2.96	0.39	0.43	0.17
Bucks	12.84	5.21	0.41	0.44	0.18
Chester	6.45	1.82	0.28	0.25	0.07
Delaware	7.80	1.67	0.21	0.48	0.10
Montgomery	4.50	2.45	0.54	0.46	0.25

[a] The Panel consists of all new Eligibles entering the data system for the first time as of January, 1979. (See First Interim Report, Section 1.3.1.)

phia County. In other words, of the screens with problems, 63% of the screens reported problems that were newly found as a result of EPSDT. Philadelphia's level of implementation is not very different from two of the other counties, but Delaware and Chester Counties both seem to have relatively low numbers of problems per provider compared to the number of screens per provider. Whether the cause is lack of screening rigor or of actual problems in the particular group of children screened is a question that can be answered with more intense analysis of screening results from these counties such as that conducted on the abnormality rates of all (Panel) providers in the First Interim Report. In this regard, the model seems quite useful in spotting actual or potential problems in this phase; program administrators might now wish to concentrate on finding provider sites in these counties with high screen volume but lower than normal abnormality rates.

6. Risk Reduction Phase Model

The overall efficiency of the risk reduction phase can be assessed (for those children who are rescreened) by the ratio of the number of screens that are free of problems on rescreen to the number of initial screens that had new problems, FP2/N. If risk has been reduced by the time of the second assessment, we would expect there to be more problem-free screens at rescreen than at initial screen.

This efficiency ratio raises once again the issue of eligibility: how many children are truly at risk for rescreening. To be consistent in this study, we should consider only those who were eligible at the time their rescreens were due. By not limiting N to this at-risk subset, the number of eligibles would be inflated causing efficiency of this phase to appear lower than it should.

Efficiency would be zero if the health status of the rescreened children did not change, and it would be one if all rescreens were problem-free. There is probably a theoretical upper limit to efficiency of risk reduction since some children will develop problems between screens. Still, if the provider-client relationship has been established, we would hope that some of these interim problems will have been treated and that aggregate health status, in terms of the number of problem-free screens, will have been improved.

Effectiveness of risk reduction is indicated by dividing the number of rescreens without problems by the number of rescreens of those children who had (new) problems on initial screen. [Witness the difficulty of expressing FP2/S2 verbally!] Keeping in mind that all rescreens in the schematic are of screens that had had (new) problems on a previous screen, effectiveness of this phase is appropriately indicated by FP2/S2.

As with risk reduction, implementation is analyzed in terms of providers (in this instance, providers of these rescreens, D2). Productivity is the ratio of rescreens performed to the number of providers doing the rescreening (S2/D2); workload is the total number of such children who are due for rescreen per provider (N/D2).

Exhibit 10. Risk Reduction Model

$$[(S2/D2) / (N/D2)] \times (FP2/S2) = FP2/N.$$

The data and rates presented in Exhibits 11 and 12 should be interpreted even more carefully than previous PHMC data because of the diminished sample size. Various data constraints required selecting another somewhat expanded panel of individuals. This panel consists of children screened initially during July, August, and September of 1977. (July was selected as a start for this "window" since a new screening data form was initiated at that time. A three month window was required because of the attrition rate of children leaving the system; a one-month sample would not provide a sufficiently large sample of rescreens.)

Since the notion of "at risk" is still as valid for rescreening eligibility as it was for initial screening eligibility, several criteria had to be met by the children to be included in this sample: (1) at screening, the child had to be between 18 months and 19 years of age; (2) a new problem was found as a result of that initial screen; (3) the child was eligible for

Exhibit 11. Program Data for Risk Reduction Phase

(Sample data)[a]

	Input	Process	Output	Resource
	Screens with Newly Discovered Problems (N)	Second Screens (S2)	Rescreens Free of Problems (FP2)	Providers Performing Second Screens (D2)
Total	2612	252	115	78
Philadelphia	2308	214	100	60
Suburban	304	38	15	18
Bucks	65	19	7	9
Chester	55	1	0	1
Delaware	137	14	6	6
Montgomery	47	4	2	2

[a] The sample is a panel of all individuals who were screened during a three-month period in 1977 and who were due for rescreening two years later.

Exhibit 12. Performance Measures for Risk Reduction Phase

(Sample data)[a]

	Input/Resource	Process/Resource	Process/Input	Output/Process	Output/Input
	WORKLOAD N/D2	PRODUCTIVITY S2/D2	IMPLEMENTATION S2/N	EFFECTIVENESS FP2/S2	EFFICIENCY FP2/N
Total	33.49	3.23	0.10	0.46	0.04
Philadelphia	38.47	3.57	0.09	0.47	0.04
Suburban	16.89	2.11	0.12	0.39	0.05
Bucks[b]	7.22	2.11	0.29	0.37	0.11
Chester[b]	55.00	1.00	0.02	0.00	0.00
Delaware[b]	22.83	2.33	0.10	0.43	0.04
Montgomery[b]	23.50	2.00	0.08	0.50	0.04

[a] The sample is a panel of all individuals who were screened during a three-month period in 1977 and who were due for rescreening two years later.

[b] All rates for this county have been calculated on a small sample size and should therefore be interpreted with caution.

rescreen at least 18 months after the initial screen date.

Criterion 1 reflects first our immediate concern with children on the two-year rescreen period rather than with infants (children up through 18 months) who can be rescreened more frequently. Secondly, we exclude those over 18 years since they will be ineligible for EPSDT screening after two years and thus not at risk.

The second criterion is in accord with the model (see Exhibits 2 and 10), viz., that the input to the risk reduction phase is N, the number of initial screens with new problems.

Finally, the third criterion assures us that the sample children were members of the target population when they became due for rescreen. (Note that this criterion does not imply that any child was continuously at risk until due for rescreen.) Of the 4151 children (3629 in Philadelphia and 522 in the suburban counties) whose 1977 screens met the first two criteria for inclusion in the sample, 62.9% (2308 in Philadelphia and 304 in the suburban counties) were eligible for rescreening in May 1979; 24.5% were ineligible as of that date. Since the eligibility status of the remaining children could not be determined, the values for N in Exhibits 11 and 12 may underestimate the number of children who were at risk for rescreening.

Performance measures for this phase indicate an efficiency level (0.04) below that of all other phases but an effectiveness level of 0.46: about 46% of children with newly found problems requiring care on initial screen had no problem - old or new - on rescreen. Clearly there has been improvement in health status. In fact, this statistic probably underestimates the health status of the rescreened group (S2) since a child would not be included in the problem-free group (FP2) if there were <u>any</u> problem still evident, even those that are chronic and already under care, or if all but one problem for multiple-problem children had been alleviated.

The efficiency of risk reduction is being attenuated by the level of implementation: S2/N = 0.10 represents a <u>restricted</u> penetration rate for rescreening of the target group (N = 2612). (Note, however, that N as a target is restricted as described above by requiring a child to meet specific inclusion criteria.) Possibly the relatively low implementation level is due to late rescreens (i.e., occurring later than 25 months after first screen) or late receipt of completed screening forms from the rescreening providers. [If needed, revisions of S2 can be made at a later date to allow additions to S2, but a significant increase is not anticipated.]

APPENDICES:

Further Examples of a General Evaluative Model

In the previous sections, a general evaluative model of EPSDT was presented which specifies (1) the EPSDT subsystems (casefinding, risk identification, and risk reduction), (2) the components of the model (workload, productivity, effectiveness, and (3) the mathematical relation among these components. Such a model has broad application in the attempts to monitor, control, and ultimately improve the program.

Our original research objective was, according to the Proposal, to provide "a model of reality which can be tested"; however, in addition to providing a framework for further EPSDT research, the evaluative model that has been developed has more general application to program management and design of performance measures.

This section will provide examples of potential and actual applications in EPSDT of the model from three perspectives: across programs, within a particular program and for one component of one phase. The first example describes how a regulatory or funding agency could design and apply program performance measures. The second example illustrates how a particular manager or planner could improve the design or operation of an EPSDT program through the application of this model. The final example uses the model to focus on one aspect of program management: evaluating individual outreach worker effectiveness.

Example 1: <u>Development and Application of Program Performance Measures</u>

The following example describes how a regulatory agency might develop and apply program performance measures using the model. The approach suggested here involves a series of actions.

(1) Establish minimum standards for the performance of EPSDT Programs in terms of measures of workload, productivity, effectiveness, and efficiency for each phase or subsystem of the program.

Once the ultimate goals of the program are established (e.g., 60% of all eligibles will be screened by 1986), minimum performance levels could be established and annually incremented until programs eventually reach these goals. Establishing performance measures and revising standards could be accomplished using (estimates of) national baselines derived from generally available program data. These data could help policy makers and regulatory agencies set and revise interim goals based on estimates of current national performance, perhaps adjusted for relevant local conditions (e.g., urban versus rural settings).

(2) Calculate specific values for each

performance measure for all programs being monitored.

Each EPSDT program would submit performance data to the regulatory agency periodically. Given the three program phase models in Exhibits 3, 7, and 10, program administrators would need to submit as few as four data items (excluding cost data) for each phase.

Exhibit 13 is an example of a reporting form that programs could submit periodically (e.g., semi-annually) to regional offices. These data could then be used to derive program performance measures for each program. Exhibit 14 provides an example of a program performance report that could be used for monitoring state and local programs.

(3) Determine whether individual performance measures are within an acceptable range.

Examination of summary measures that are comparable across programs will reveal which programs or phases deviate significantly from expected results. (A statistical technique for determining exceptional performance was demonstrated in the First Interim Report, Section 4.2.2, and applied to identify providers finding an abnormally high or low number of screens with problems. See also Example 3 below.) Administrators of programs might then be required to document and explain exceptional deviations; programs found consistently in violation would be eligible for on-site eval-

Exhibit 13. Program Data Reporting Format

PHASE I: Outreach

(INPUT)	Number of Eligibles (E)	(____)
(PROCESS)	Number of Contacts (C)	(____)
(OUTPUT)	Number of Screens Resulting from Contacts (S)	(____)
(RESOURCES)	Number of Outreach Workers (O)	(____)

PHASE II: Risk Identification

(INPUT)	Number of Screens (S1)	(____)
(PROCESS)	Screens with Problems Requiring Care (P1)	(____)
(OUTPUT)	Newly Discovered Problems (N)	(____)
(RESOURCES)	Number of Providers (D1)	(____)

PHASE III: Risk Reduction

(INPUT)	Number of Screens with Newly Discovered Problems (N)	(____)
(PROCESS)	Number of Screens with New Problems which are Rescreened (S2)	(____)
(OUTPUT)	Screens with Problems Requiring Care at Rescreen (P2)	(____)
(RESOURCES)	Number of Providers Doing Rescreens (D2)	(____)

Exhibit 14. Performance Measures Reporting Format

PHASE I: Outreach

WORKLOAD	(Input/Resource)	E/O	(____)
PRODUCTIVITY	(Process/Resource)	C/O	(____)
IMPLEMENTATION	(Process/Input)	C/E	(____)
EFFECTIVENESS	(Output/Process)	S/C	(____)
EFFICIENCY	(Output/Input)	S/E	(____)

PHASE II: Risk Identification

WORKLOAD	(Input/Resource)	S1/D1	(____)
PRODUCTIVITY	(Process/Resource)	P1/D1	(____)
IMPLEMENTATION	(Process/Input)	P1/S1	(____)
EFFECTIVENESS	(Output/Process)	N/P1	(____)
EFFICIENCY	(Output/Input)	N/S1	(____)

PHASE III: Risk Reduction

WORKLOAD	(Input/Resource)	N/D2	(____)
PRODUCTIVITY	(Process/Resource)	S2/D2	(____)
IMPLEMENTATION	(Process/Input)	S2/N	(____)
EFFECTIVENESS	(Output/Process)	FP2/S2	(____)
EFFICIENCY	(Output/Input)	FP2/N	(____)

uations and legislated penalties. Conversely, exemplary programs could also be noted and possibly provided with incentives. Since programs will be providing the information on which evaluative and monetary decisions are made, random audits may be necesary to insure the validity of the data being reported to regulatory agencies.

(4) Summarize and publish regional and national results.

Once a substantial data base has been established, semi-annual and annual reports could be produced regionally and nationally. These reports would provide local and state administrators with feedback concerning their programs' performance, and demonstrte that decision-making and policy development on the federal level are empirically based.

Example 2: Application of a General Evaluative Model to Program Management

This example is based on a hypothetical EPSDT program that has been in operation for six months. It illustrates how a manager could collect, organize, and interpret information based on this model, and eventually develop and implement recommendations to improve the operation of the program. The data presented in Exhibit 15 represent the results for the first months from the outreach phase of the program. Outreach workers have been organized into three teams. Each team is responsible for eligibles residing in a specific geographic area of the county.

As indicated by Exhibit 16, in Melrose Center, the outreach approach has been very effective (S/C = 0.60), but the level of efficiency is low (S/E = 0.10) because workload (E/O) is very high. In suburban Mayfair Township, on the other hand, the opposite is true: many contacts were made, but the outreach technique was relatively ineffective. Mayfair Township has the most productive outreach team, Melrose Center has the most effective approach, and Llanerch Township has the most efficient operation.

Using these facts, program managers can assess performance (e.g., the Mayfair Township team will work on raising its level of effectiveness to 0.45).

Exhibit 15. Program Data for Outreach Phase

(Hypothetical data)

	Input Eligibles (E)	Process Contacts (C)	Output Screens (S)	Resource Outreach Workers (O)
Total	4165	1254	548	24
Melrose Center	2500	440	264	8
Mayfair Township	1040	550	165	8
Llanerch Township	625	264	119	8

Exhibit 16. Performance Measures for Outreach Phase

(Hypothetical data)

	Input/Resource	Process/Resource	Process/Input	Output/Process	Output/Input
	WORKLOAD E/O	PRODUCTIVITY C/O	IMPLEMENTATION C/E	EFFECTIVENESS S/C	EFFICIENCY S/E
Total	173.54	52.25	0.30	0.44	0.13
Melrose Center	312.50	55.00	0.18	0.60	0.11
Mayfair Township	130.00	68.75	0.53	0.30	0.16
Llanerch Township	78.12	33.00	0.42	0.45	0.19

Example 3: Application of a General Evaluative Model to Monitoring Outreach Worker Performance

The proposed model provides a useful method for monitoring individual outreach worker activity. Using the approach outlined previously, an EPSDT Administrator could apply the concepts of efficiency, effectiveness, productivity, and workload to the individual outreach worker.

The first step in (instrumental) modeling is to identify the input and output of a system, and to specify the process whereby output for this system is produced. For the individual outreach worker, inputs are "assigned eligibles"; outputs are "screens"; and the process is the outreach "contact."

Efficiency of each worker is defined as screens (output) per assigned eligibles (input); effectiveness is defined as screens (output) per contacts (process); level of implementation is defined as contacts (process) per eligibles assigned (input). These performance measures can then be combined into the general model of IMPLEMENTATION X EFFECTIVENESS = EFFICIENCY thus: (Contacts/Assigned Eligibles) X (Screens/Contacts) = (Screens/Assigned Eligibles).

Level of implementation of outreach for each worker can be viewed as a function of two factors: workload and productivity. Workload has been defined as input per unit resource, and productivity was defined as process per unit resource. For outreach workers, the most basic resource is the amount of time available for outreach activities. Using time in this manner provides us with a means of standardizing our assessment of workers' activities to allow for variation in the number of hours devoted to outreach (e.g., sick days, holidays, part-time). Returning to our definitions, workload then becomes assigned cases per day of outreach and productivity is equal to contacts per day of outreach. Applying the general model for efficiency of implementation, PRODUCTIVITY / WORKLOAD = IMPLEMENTATION, results in (Contacts/Day) / (Assigned Cases/Day) = (Contacts/Assigned Cases).

A monthly outreach activity report which incorporates these performance measures could probably be produced by many existing EPSDT management information systems. The report profiles of each worker enable a program manager to identify (1) how well resources are being allocated (workload measures), (2) who are the more <u>productive</u> and <u>unproductive</u> workers, and (3) who are the more <u>effective</u> and <u>ineffective</u> workers. These last two characteristics may not necessarily be identical, that is, some workers may be productive (large number of contacts per day) and yet be ineffective (low number of screens per contact); however, others may be effective, but unproductive, while others may be both or neither. Classifying outreach workers according to these characteristics may aid administrators in developing a series of management strategies designed to maximize outreach worker productivity, effectiveness, and eventually, efficiency. For example, an administrator might use highly effective workers (high ratio of screens per contact) as in-house trainers of other or new workers; or a supervisor could establish quotas (e.g., minimum number of contacts per assigned eligibles) for workers with low productivity.

A relatively straightforward statistical monitoring system could be used to identify those outreach workers with exceptionally low or exceptionally high rates of productivity or effectiveness. The basic simplifying assumption underlying any attempt to identify exceptional rates is that the expected or baseline rates will be replicated, within statistical limits, by each outreach worker. Thus, since the sampling distribution of proportions is approximately normal (when corrected for continuity for workers with low numbers of contacts), each performance measure can then be expressed in terms of its distance (in standard deviation units) from the expected (i.e., population) proportion, and a decision rule can be adopted which defines the "normal" range for the performance measures.

In the example presented here, we are evaluating the effectiveness rates of a group of outreach workers, and we have adopted a relatively conservative definition of "normal." A rate (p) will be considered normal if, when standardized, it lies within one standard score of the population rate, P. Standardization of the difference into a z-score is based both on the distribution of sample proportions for sample size n (the number of contacts made by the particular worker) and population proportion P (the overall effectiveness rate for all such workers): thus $z = (p - P \pm .5/n)/SDp$, where $SDp = \sqrt{PQ/n}$. The correction for continuity, $\pm .5/n$, adjusts proportions based on small sample sizes to more closely approximate a normal distribution. (See Hays, Statistics, 2nd ed., 1973, p. 308.)

In our example, standardized effectiveness rates have been calculated for the 66 outreach workers who contacted those Philadelphia families who became newly eligible for EPSDT in January 1979. Each worker was classified as being either exceptionally high ($p > P + SDp$), exceptionally low ($p < P - SDp$) or normal ($P - SDp \leq p \leq P + SDp$). The overall rate for all 66 workers considered was 0.30.

Using these criteria, most outreach workers (61%) were classified as "normal," 24% were categorized as "highly effective" and 15% were classified as ineffective. Exhibit 17 summarizes these results.

Exhibit 17.

Standardized Effectiveness Rates for Outreach Workers