



Technical

Discussions



Washington, D.C.
September 1957

CD10/DT/1 (Eng.)
6 September 1957
ORIGINAL: SPANISH

*Spanish printed in
Boletín O.S.P. #3(6): 540-57, Dec. 1957*

INTRODUCTORY STATEMENT ON THE TOPIC

"BASES AND METHODS FOR THE EVALUATION OF HEALTH PROGRAMS"

By Guillermo Arbona, M.D., M.P.H.
Secretary of Health
Department of Health of the Commonwealth of Puerto Rico

CONTENTS

	<u>Page</u>
Introduction	1
Definition	1
The Evaluation Process	2
Evaluation Methods and Techniques	5
Evaluation and Scientific Investigation	10
Responsibility for Evaluation of Programs	11
Purposes of Evaluation	12
References	13

BASES AND METHODS FOR THE EVALUATION OF HEALTH PROGRAMS

Introduction

The objective evaluation of public health work is a subject that has commanded increasing interest in all parts of the world. There are many reasons for this trend, two of them being of special significance. First, the resources available for public health work are usually quite limited and must therefore be invested carefully to produce maximum returns; through the periodic evaluation of programs, one can see whether those resources are being invested as profitably as possible. Second, it is very difficult to find a yardstick capable of measuring the effect of a large part of the public health activities, and efforts to convince others of the merits of programs are therefore often frustrated unless objective evidence is put forth.

Human beings have a natural tendency to evaluate; every decision they make is preceded by a process of weighing or measuring. If someone thinks of buying a pair of shoes, he first considers how many pairs he already has, their color, style and condition, the amount of money he has to spend, etc. On buying a new pair, he takes into account the price, the quality of material, the shape and color, and so forth. Certain features of the shoes are objective, tangible, and measurable, such as size and cost; others, such as looks of the shoe, form or color, are of an intangible and subjective character.

The same is true in the evaluation of health programs: some of the activities can be measured objectively, while others can be appraised only in subjective terms. We are not, however, in agreement with methods of measurement that depend too much on subjective appraisal either because objective methods are not known or because those that are known are difficult and costly to apply.

Definition

In the public health field the term evaluation covers various processes of measurement. Chiefly, it is used to indicate the processes for assessing the state of individual and collective health and the extent to which a program is effective in achieving the objectives for which it was planned.

The World Health Organization (1) defines the term as follows: "Generally speaking, evaluation is the process of determining qualitatively or quantitatively, and by appropriate means, the worth of a thing

or event. In WHO, however, the term is used to describe the administrative and technical procedures by which the value of planned activities is measured for the benefit of those responsible for their control. In practice and as applied to a WHO project, evaluation consists of the factual reporting and assessment of the progress made towards attaining the objectives set out in the plan of operation. It is therefore essential that the process of evaluation should start at the inception of a project, and that it be recognized as an integral and continuous part of the responsibility for running a project."

The Committee on Evaluation of the Adult Education Association of the USA (2) defines evaluation as "the process of assessing the degree to which one is achieving his objectives. It is looking at one's present position in regard to one's goal. It is a comparison of the actual with the ideal."

The Diabetes Program Guide, published by the U. S. Department of Health, Education, and Welfare, (3) gives this definition: "Evaluation is objective appraisal of a program (or some element of a program) in relation to a goal, by comparing effort expended with results obtained."

Otis Anderson (4) refers to evaluation as follows: The purpose of evaluation is to determine the value of work performed. It is done by measuring accomplishment or progress against pre-established goals. It is also related to determining the justification and validity of the goals.

Evaluating is a process of measuring and this, in turn, is a process of comparing. In evaluating, we compare a present situation with a prior one, or a situation in one place with a similar one in another place, or an actual situation with a potential one; that is, one in which the program we wish to evaluate has already been applied, with another in which it has not been applied.

The Evaluation Process

Public health programs are so widely varied that it is impossible in a study of this type to describe all the methods used for their evaluation. However, the principles underlying practically all the methods are the same; only the manner of applying them varies.

From the definitions given above we can deduce that the evaluation process is used to determine: (a) the extent to which a program is attaining its objectives; (b) the effectiveness with which it is being conducted; and (c) the extent to which the program itself is justified.

During the technical discussions last year on methods for the preparation of national public health programs, (5) we referred to the following steps in the planning process:

- a. Study and analysis of the problems.
- b. Study and analysis of the resources.
- c. Definition of objectives, long-term and short-term.
- d. Formulation of plans of operation.
- e. Evaluation.

Through this process we arrive at a basis for planning specific and general programs in a given area, taking into account both the needs and the resources available. Stress should be laid on the importance of making provision for the evaluation of the program. No plan can be considered complete unless it includes measures for determining the extent to which established objectives are being achieved, the efficiency with which the plan is being carried out and, finally, the extent to which the plan is justified.

The most commonly used evaluation measures are based on the experimental method. A program is planned with the aim of changing a given situation into another that is considered better. Three factors are taken into account in evaluating the program:

1. The situation existing at the start of the program.
2. The program as such.
3. The situation existing at the time of evaluation.

Theoretically, the difference between the situation at the start of the program and that at the time of evaluation indicates the results of the program. In practice, this is not usually the case, for the program is only one of the many factors that have a bearing on the situation. Thus, the evaluation process should indicate not only the changes in the situation that we wish to improve but also the extent to which the changes result from the program as such. To make this possible, it is essential to define clearly the situation that is to be improved, as we pointed out last year in discussing the definition of existing needs. We emphasized also that the description of the situation we wish to improve should be as objective as possible, and we referred to the importance of statistics on mortality and morbidity and data on environmental conditions such as number of houses supplied with potable water, proper measures for sewage disposal, etc.

As a second step, we must define our goals in the same objective way: that is, the situation we expect to achieve in each stage of the program. Thus, if the objective is to reduce mortality or morbidity due to one or more diseases, it can be expressed in terms of the reduction we expect to bring about. If we wish to improve sewage disposal, we can express this goal in terms of the number of houses in which it is planned to install adequate disposal facilities.

Thirdly, the program must be defined in clear and specific terms. If it is proposed to reduce disease mortality and morbidity, the plan should describe how this is to be done; if the basic measure is immunization of susceptible individuals, we must indicate what immunizing agent is to be used, to what groups it will be administered, when and by whom it will be applied, etc. Or if it is a question of improving sewage disposal, we must show by what means houses are to be furnished with the lacking facilities.

Often, the program objectives are very general and long-range in scope, in which case evaluation may be quite difficult. For example, the program's goal may be to raise the level of child health through activities directed toward educating mothers in child growth and development. This is a long-range objective and a number of years must elapse before we can determine whether it is being achieved. Also, other programs may at the same time be contributing toward the improvement of child health.

To overcome this type of difficulty, other methods must be brought into play in order to appraise, not the extent to which a program is achieving its objectives, but what probability there is of reaching those goals. Certain standards have been established for this purpose. Some of them have been put to experimental tests, whereas others are hypothetical and have been adopted on the strength of authoritative opinion. Experience has shown, for example, that if a given percentage of children in a community (the proportion varying according to conditions) is kept immunized against diphtheria, the prevalence of that disease will drop significantly. Some years ago, the American Tuberculosis Association set up a standard according to which, in order to reduce tuberculosis prevalence in a community, that community's hospitals should have a ratio of 2.5 tuberculosis beds for each tuberculosis death occurring annually. With this type of standard, the progress of the program can be measured indirectly, in the assurance that if the standard is met the program will in all probability be achieving its goal. Each country, of course, would adjust such standards in accordance with prevailing local conditions.

Evaluation Methods and Techniques

We mentioned earlier that the methods which have been used for the evaluation of programs are usually modifications of the experimental method.

One recent example was the evaluation of the Salk poliomyelitis vaccine. The situation that was to be changed was the incidence of polio in the population. The program consisted in the application of that vaccine. The situation sought was a lower incidence of the disease. In this case, the goal was to prove the efficacy of Salk vaccine. It was necessary to select two population groups that were similar in all but one respect: one group received the polio vaccine, while the other did not. The two groups thus had to be identical as to age, sex, race, clinical history, exposure to the disease and to climatic, geographic, and many other conditions. One of the groups was injected with the vaccine and the other with a placebo. The results of these experiments are well known to everyone.

In countries where polio vaccine is now being used, it is not submitted to such a rigorous evaluation; this is unnecessary because we already know that the vaccine is efficacious. Thus, we now depend on the establishment of standards for evaluating the applied program of vaccination.

Let us suppose that the incidence of poliomyelitis in one country is virtually limited to children under 5 years of age. The objective is set of vaccinating all children in that age group and a plan is devised for reaching that goal in a period of two years and at a given cost. In this case we would evaluate the program, from the short-term view, not on the basis of a reduction in polio incidence -- since we know that the incidence varies from year to year and only after long-term observation can we determine whether it has decreased -- but in terms of the number of children under 5 who receive the vaccination. In other words, we evaluate a program of this type on a short-term basis in terms of the efficiency with which the operations are carried out. From the long-term view, of course, we would evaluate it on the basis of the incidence of poliomyelitis in the population.

I had the occasion, recently, of talking with Dr. Luis Vargas, of Mexico, who is participating in the evaluation of the malaria eradication program in that country. The long-term objective of the program is to eradicate the disease in five years, and it is expected to reach this goal by conducting, among other activities, biannual spraying of houses with insecticides, intensive search for malaria cases, and treatment of the disease.

The long-range evaluation will depend on what reports are received of malaria cases. If after five years of operations cases of the disease continue to be reported, then the program will be deemed not to have

achieved its objective. But such an evaluation cannot be made until the end of five years. Therefore, other measurements, in addition to morbidity rates, must be made to evaluate the program's effectiveness.

The evaluation of the Mexico program is based on the assessment of the efficiency with which the program is being developed and the methods are being applied. One goal of the program is to spray periodically all houses in the malarious area of the country. The need is obvious, therefore, to count those houses and to count also the number sprayed during the year. The goal will have been reached if all the houses are sprayed.

It must also be determined whether the house-spraying is being done in accordance with the accepted and tested standards or criteria governing such factors as the concentration of the insecticide used, the amount remaining on the walls of the house, etc.

Programs for the control of specific diseases, such as poliomyelitis and malaria, are usually easy to evaluate, for over the years we have been developing the necessary methods for this purpose and gaining experience in their application. The same cannot be said of health promotion programs or, up to a certain point, of health restoration programs. Actually, the methodology we have developed for the control of communicable diseases is relatively simple, particularly if the active participation of the members of the community is reduced to a minimum. In health promotion programs the active participation of the individual is, of course, essential.

Let us take the case of activities directed toward improving child nutrition in a child care clinic, as part of a child health program. The intent is to educate the mother to feed her child according to the best known methods. To evaluate these activities, we would have to take into account the mother's knowledge and reactions at the time the child is admitted to the clinic, the educational service she receives there, and then the knowledge or attitudes she acquires or the changes that occur in her outlook. These changes in the mother should result in a better fed and healthier child. We know, however, that the mother's knowledge and attitude as regards the child's nutrition are influenced by a number of other factors, such as her conversations with neighbors and relatives, advertisements of food products in newspapers and magazines, on radio and television, etc.

To be able to arrive at definite conclusions, we would have to conduct an experiment similar to that made in evaluating the poliomyelitis vaccine, taking two groups of mothers who are similar in all respects except that one group was exposed in the child care clinic to education on the feeding of the child, while the other group was not. Since such an experiment would have to be of long duration, it would be difficult to select two groups of mothers who would differ only in that one respect. The cost of this type of evaluation is generally beyond the

means of the local health services. It is therefore frequently recommended that countries set up pilot projects for testing methods and techniques and developing suitable standards for this activity. Thereafter, indirect evaluation can be carried out in the local health services by means of supervision, so that the accepted methodology and techniques will be properly applied.

Thus, the objective of a program for the promotion of child health could consist in furnishing child hygiene services to the greatest possible number of children and in accordance with given standards. The standards would govern such factors as visits of the child to the physician, visits of the nurse to the home, age at which the child should be vaccinated against specific diseases, etc. The program is evaluated on the basis of the proportion of children receiving the service in accordance with the established standards. To make certain that the services are being carried out according to these standards, an adequate system of supervision or control is required.

Supervision of public health personnel forms part of the process of evaluation. In programs whose efficiency is difficult to measure in terms of final results, we depend on the measurement of work performed from both the quantitative and qualitative viewpoints. The number of visits a nurse makes to children in their first year of life tells us the number of mothers influenced by the service, but it does not tell us how effective the visits were, how long they lasted, or how successful the nurse was in grasping the mother's and the family's problems, in imparting information, and in instilling favorable attitudes in the mother. It is through supervision that we obtain information of this type. Quantitative information on the work performed and qualitative information on the efficiency with which it was done are used in evaluating programs whose final objectives are achieved on a long-term basis and are difficult to ascribe to the program being evaluated, since other programs, activities, and factors outside our control may contribute toward those same objectives.

In summary, evaluation methods may be classified into two groups. First, there is the direct evaluation of programs whose efficiency we measure by the experimental method or a modification of that method. In such cases we measure the extent to which a program is achieving the proposed goals.

Secondly, we use indirect evaluation methods by which we measure the quantity and quality of the effort expended. The quantity and quality of this effort are compared with previously established standards which are known from experience (derived from the use of direct evaluation or the experimental methods in pilot projects) to produce the desired results. We recognize that, at times, it is necessary to establish such standards arbitrarily.

Usually, a combination of these methods is applied in evaluating public health programs.

The American Public Health Association, through its Committee on Administrative Practices, has for many years promoted the evaluation of public health services in local communities and prepared forms for this purpose. At first, the forms gave prominence to work done and effort expended. They were used to stimulate the local public health services and prizes were given to communities that had the most complete programs and did the largest amount of work. The use of the method for these purposes degenerated into a race to win prizes. To correct the situation, the Association changed the method, discontinued the prizes, and redrafted the form to give more importance to the results of the work than to the work itself.

The Association recommends to local health services in the United States the periodic use of this revised form as a means of assessing health conditions in the community they serve and the progress made in the health work. This method evaluates not only the official health programs but all other programs of the community, whether official or not, for it is well known that in the United States many health programs are carried out on the initiative of private and civic groups.

In some states of the country, the form has been modified for use by community leaders, such as those serving on health boards. Reports from communities where this method is used are most encouraging. Self-evaluation appears to be a valuable means of education through which the community recognizes its own health problems and stimulates the organization of new programs, voluntary or official, or the improvement of existing ones to meet the community's needs.

When the form is used by local health services, it is filled out by the staff of the service itself and thus draws the attention of that staff to the work that needs to be done. In many of the local services, the form is used for the in-service training of personnel.

The work reports of a service can also be used as instruments for evaluation. The World Health Organization depends to a great extent on monthly and annual reports for evaluating the programs it sponsors. Health departments in all parts of the world use such reports for similar purposes.

Unfortunately, the forms and methods used for reporting on work done are not always as complete as could be desired. In the many countries we have visited we have always noticed dissatisfaction with the methods employed. Generally, the reports measure the effort made rather than the results achieved. We do not as yet have a system that covers both aspects and makes allowance for the fact that under certain circumstances objectives can be achieved with a minimum of effort while in others even the maximum of effort is not sufficient.

There is an obvious need for devising better reporting forms and models that will take into account the many factors which in the long run determine the success or failure of the program and which will reflect both the effort made and the results obtained.

It is even more difficult to determine the justification for a specific program than to measure its effectiveness on the basis of the results achieved and efficiency with which it is carried out. Yerushalmy (6) in referring to the evaluation of the needs and requirements of services, states:

"The fact is that no machinery is available for the routine recording and reporting of data which bear directly on needs for health services when this need is interpreted in its broadest sense. There may be reporting systems by which it is possible to evaluate certain demands for services, but it must be emphasized that there is a distinction between needs and demands for services. As a rule the demand for a specific service represents only a part of the need, that part which either the individual or the community is conscious of.

It is almost hopeless in many instances to attempt an evaluation of the total need in a community for a specific service. Since the nature of the activities in public health are in prevention and early detection, it follows that the needs for health services must often be evaluated at a stage before these needs have manifested themselves to the individual in the form of severe disease."

In countries such as the United States where the public health services have responsibility for health promotion and disease prevention and where resources come nearer to filling the needs, it is easier to justify programs than in countries where such services have the added responsibility of providing medical care and where resources are limited because the public demand is usually for medical care rather than for preventive or promotion services.

We are reminded of a decision taken by Dr. Eduardo Garrido Morales when he was Health Commissioner of Puerto Rico. There was public demand, voiced in the press and by the political and civic leaders and the medical profession, for the Health Department to take over the medical care services, which were then the responsibility of the municipalities. At that time -- about 1938 -- the Department was concentrating its efforts on two programs, one against tuberculosis and the other against malaria. His approach was as follows: he calculated the cost of improving public medical-care services, and the cost of continuing the tuberculosis and malaria programs; he also made an estimate of Puerto Rico's immediate and potential future resources for those programs. It became evident that only by sacrificing the tuberculosis and malaria programs could the Department

assume responsibility for medical care, and even thus the improvement in medical care services would be very small. He deemed it preferable to continue the preventive programs against tuberculosis and malaria so as in this way to reduce general morbidity in the population. He had the ability to convince public opinion. Years later, when the tuberculosis mortality and morbidity rates were effectively reduced and malaria was eradicated, Puerto Rico's Health Department decided to undertake a program aimed at improving medical care.

We mention this incident because we believe that, in the field of public health, in countries with limited resources and obvious need for services, what is most difficult is not to justify the execution of a program but to justify not executing it: to decide that a service that is necessary and for which there is public demand will not be provided because the available resources can be used more advantageously in another service, the need for which is not as obvious to the community, the political leaders, and the press of the country.

Evaluation and Scientific Investigation

There are two types of investigation: one that provides fundamental knowledge, in the laboratory, the hospital, or the field; and the other that provides methods for application of the new knowledge, generally in the field. Some refer to the two types as basic and applied, respectively. One is just as scientific as the other; the technique used is the same, namely, the experimental method we mentioned before.

The need for more applied investigation is evident. Applied investigation is the only effective means of sharing the experience acquired in the field of administration of public health programs, and the way to make this type of investigation is to evaluate, on a strictly scientific basis, modern methods and the modification of those methods that are used in the administration of health programs. Up to the present, in our health services there has been too much reliance on good judgment and improvisation in adopting methods, and frequently we make mistakes, because what is good for one country or for one region of a country is not necessarily good for another country or another region of the same country.

As we have said before, it is neither possible nor necessary to make an extensive evaluation of all the methods used in a public health service; moreover, it would often not be practical. However, we feel that an attempt should be made to undertake a broad evaluation of changes made in methods and of the adoption of new methods and procedures. For this purpose, we suggest that the national public health services organize programs for investigation and demonstration in cases where existing methods are changed or new methods adopted. The Canadian Public Health Ministry has a division of investigation for this purpose. The Ministry of Public Health of Mexico has a department of experimental

studies; the Central American countries and Panama have the INCAP, where, taking into account of course differences in national needs, investigation and evaluation are carried on for the preparation and application of new methods and procedures.

Responsibility for Evaluation of Programs

In our paper on methods for the preparation of national public health plans, we stressed the fact that planning, responsibility for which lies with the executive head of a service, can and should be delegated to the staff at all levels. We are of the same mind with respect to evaluation. The planning of the program for a health center serving a small community should be initiated by its staff, and if possible with the cooperation of community leaders; we believe that the evaluation should also be delegated to this personnel. This must hold true if we accept the premise that evaluation begins with planning and should be an integral part of the program.

This procedure has many advantages. Let us take the case of a rural sanitary inspector who is responsible for improving sewage disposal in an area. He determines the need for his services in that area and draws up his plan of work to cover the objectives from year to year. In other words, he sets a specific goal. Of course his immediate supervisor may or may not accept the goal set by the inspector, but they will of necessity reach an agreement. They will then establish the criteria for evaluating the program or the work to be done. Thus, the inspector will know what his work will be, why it is to be done, etc., and, in addition, how it is to be evaluated and when it can be considered a job well done or not. In the public administration field this procedure is effective in improving the employee's performance and in maintaining good personnel relations.

In this manner, program planning, which includes procedures for evaluating programs, is delegated down the line to the various organizational levels. The personnel of a local unit of health services plan their own program; the chief of the local unit is responsible for the program of that unit and, through this system, shares this responsibility with his personnel. Similarly, the local chiefs are responsible for their respective programs to the regional chief (state or provincial), and the latter, in turn, to the national chief.

It is obvious that if this down-the-line delegation of responsibility is clarified through the approval of criteria and procedures for the evaluation of programs, the most frequent causes of friction in an organization of this type will be considerably reduced.

Purposes of Evaluation

We have referred to the advantages to be gained through the evaluation of programs. We would conclude this paper by summarizing the purposes of evaluation.

1. Evaluation is essential for the proper planning of programs. Through the evaluation of health conditions in a community or country, we discover the health needs and resources of the community or the country and arrive at a basis for knowing the type of program required to maintain or improve the health of the population.

2. Evaluation is essential for the direction of programs. It aids us in determining the efficiency and effectiveness with which the programs are being developed and in deciding on the changes needed to improve them.

3. Evaluation helps us to determine how sound our decisions on the programs were, in both the planning and the execution phases. It enables us to profit from experience and is therefore an educational tool.

4. Evaluation helps us in justifying the programs to ourselves, to the public, and to the authorities.

5. The evaluation of programs is a source of satisfaction for those responsible for their execution. If the program fails, evaluation shows us the reasons, and while this is not a cause for satisfaction it lessens the frustration.

6. Evaluation helps the higher public health officials in appraising the ability and efficiency of their staff.

REFERENCES

1. World Health Organization. Administrative Manual, Section on Program Evaluation, X.4, pp. 1-8, 28 September 1956.
2. Adult Education Association of the United States of America, Committee on Evaluation. Program Evaluation in Adult Education. Adult Education Association, Washington, D. C., 1952, p. 6.
3. U. S. Public Health Service, Division of Special Health Services. Diabetes Program Guide. U. S. Govt. Printing Office, Washington, D. C., 1956, p.42.
4. Michigan, University, School of Public Health. First National Conference on Evaluation in Public Health. Ann Arbor, 1955.
5. Arbona, Guillermo. "Métodos para la Formulación de Planes Nacionales de Salud Pública." Boletín de la Oficina Sanitaria Panamericana. Vol. XLII, No. 1, Jan. 1957, pp. 6-15.
6. California, University, Institute of the Evaluation of Public Health Practices. Proceedings. Berkeley, 1948.

Note: A bibliography on evaluation prepared by Dr. Vlado Getting appears in the American Journal of Public Health, Vol. 47, No. 4, April 1957, pp. 409-413.