Heesuk Yun

Korea Development Institute

Performance Management in the Health Sector :

Performance Management in the Health Sector

PRESENTERS

Performance Management in the Health Sector

by
Heesuk Yun, Korea Development Institute

1. Need for Performance Evaluation in the Health Field

Most public interventions in the healthcare field tend to increase access to medical care or provide more appropriate forms of treatment, but usually result in increased expenditures for health services. In Korea, the Ministry of Health and Welfare's budget accounted for 7.38 percent of the gross national product in 2003 versus 3.82 percent in 1995. Also, total expenditure for public health shows a steeper increase over the same period. The budget size for public health and medical care, and medical security increased from 1.3 trillion to 5.2 trillion, an increase of nearly 300 percent (Ministry of Health and Welfare 2003).

In the case of the United States, unprecedented growth in social and health programs for the disadvantaged began during the Lyndon Johnson presidency in 1964, when Medicare, Medicaid, and other health care programs were launched, and Congress mandated and funded performance evaluations. Considering this, it is natural that there is increasing awareness for cost containment and cost-effectiveness in the health field in Korea nowadays.

Further, Korea is becoming a more advanced country in terms of public healthcare. Last year the secretary of Ministry of Health and Welfare declared the 'War against Cancer,' which suggests that the targets of public health actions have now expanded beyond infectious diseases to include chronic diseases. This is a common trend indicating a certain degree of development in public health. As shown in the experiences of more advanced countries, the area of public health actions will keep expanding into violence, emerging pathogens, and the social contexts that influence health disparities. With the continuing rise in healthcare costs expected to coincide with this trend, attempts to investigate new proposals for various public interventions will be needed while examining ongoing interventions carefully for more effective management of public health care.

In practice, in advanced countries, large expenditures generally increase the importance of accountability of public and private resources invested in health programs and medical services, and Congress is concerned about holding programs accountable for the funds. Also, citizens, employers, the government, and other purchasers of health services are concerned with clinical and fiscal accountability, or evidence that healthcare actions deliver effective and quality services in an efficient manner. Pressure of cost containment is being strengthened in Korea as well. The Ministry of Health and Welfare, faced with higher costs, is already considering abandoning the traditional Fee For Service (FFS) medical coverage and expanding Diagnosis Related Groups (DRG) to manage healthcare bills.

Furthermore, increasing emphasis on prevention is another factor stimulating interest in performance evaluation in the health field. Many people and health

professionals believe that prevention may be a better alternative to a cure. As a society enters into a more developed stage, more preventive programs and technologies tend to emerge to maintain or improve the nation's health. Immunizations to prevent disease, mammography screening to detect breast cancer, and campaigns promoting smoking cessation are such. However, since the potential efficacy of preventive care is hard to prove with certainty, the need for rigorous evaluation will increase in the near future in Korea.

Against this background, the paper attempts to summarize the essential elements and special characteristics of performance evaluation in the health field in relation to a general evaluation framework. Also, emphasizing the need to build an infrastructure for evaluating health interventions, the paper introduces the U.S. experience of devising indicators for measuring health performance. In doing so, standards for effective evaluation of health interventions will be briefly reviewed, and implications for establishing a performance evaluation in the health field in Korea will be extracted.

2. Types of Performance Evaluation in the Health Field

Any organized public health action can be subject to a performance evaluation. Public health actions include direct service interventions, community mobilization efforts, research initiatives, surveillance systems, policy development activities, outbreak investigations, laboratory diagnostics, communication campaigns, infrastructure-building projects, training and educational services, and administrative systems. For convenience, these can be divided into three categories; the evaluation of health programs, the evaluation of the health care system, the evaluation of health services (Grembowski 2001).

The evaluation of health programs: this type targets programs created to reduce or eliminate a health problem or achieve a specific objective. Comprehensive inventory of the categories of health programs that can be implemented to achieve specific health objectives is the subject of evaluation. For instance, exercise programs created to increase physical activity or prenatal care programs to prevent medical complications associated with low birth weight are such.

The evaluation of the health care system: this evaluates the performance of the health care system. A health care system has a structure defined by laws and regulations, the availability of personnel, facilities, the organization and type of financing. The characteristics of the population that the system serves, the physical, social, and economic environment where they live are included in the structure components. The structure of the system influences the process or delivery of health services, which in turn produces outcomes. The system performance typically examines the influence of the structure and structure components on the outcomes. For example, if whether the patents' payment method to primary care physicians has any influence on patients' utilization of health services.

The evaluation of health services: this type of evaluation examines the efficiency or effectiveness of specific health services provided to patients in the system. Evaluation is focused on measurement of the benefits, of health outcomes, of a specific health service associated with particular medical technology compared to the costs of producing them. Technologies that produce large benefits at a low cost have greater worth than technologies that offer few benefits at a high cost.

3. Objectives of Public Health Intervention

Most interventions in the health field have multiple objectives, and the objectives consist of a mix of different dimensions such as time, place, method and generality. This multiplicity of objectives is often a source of unproductive disagreement among evaluators. Confusion about objectives could be eliminated by recognizing that these objectives can be classified in a number of ways, and the most common way is ordering objectives by generality (Suchman 1967).

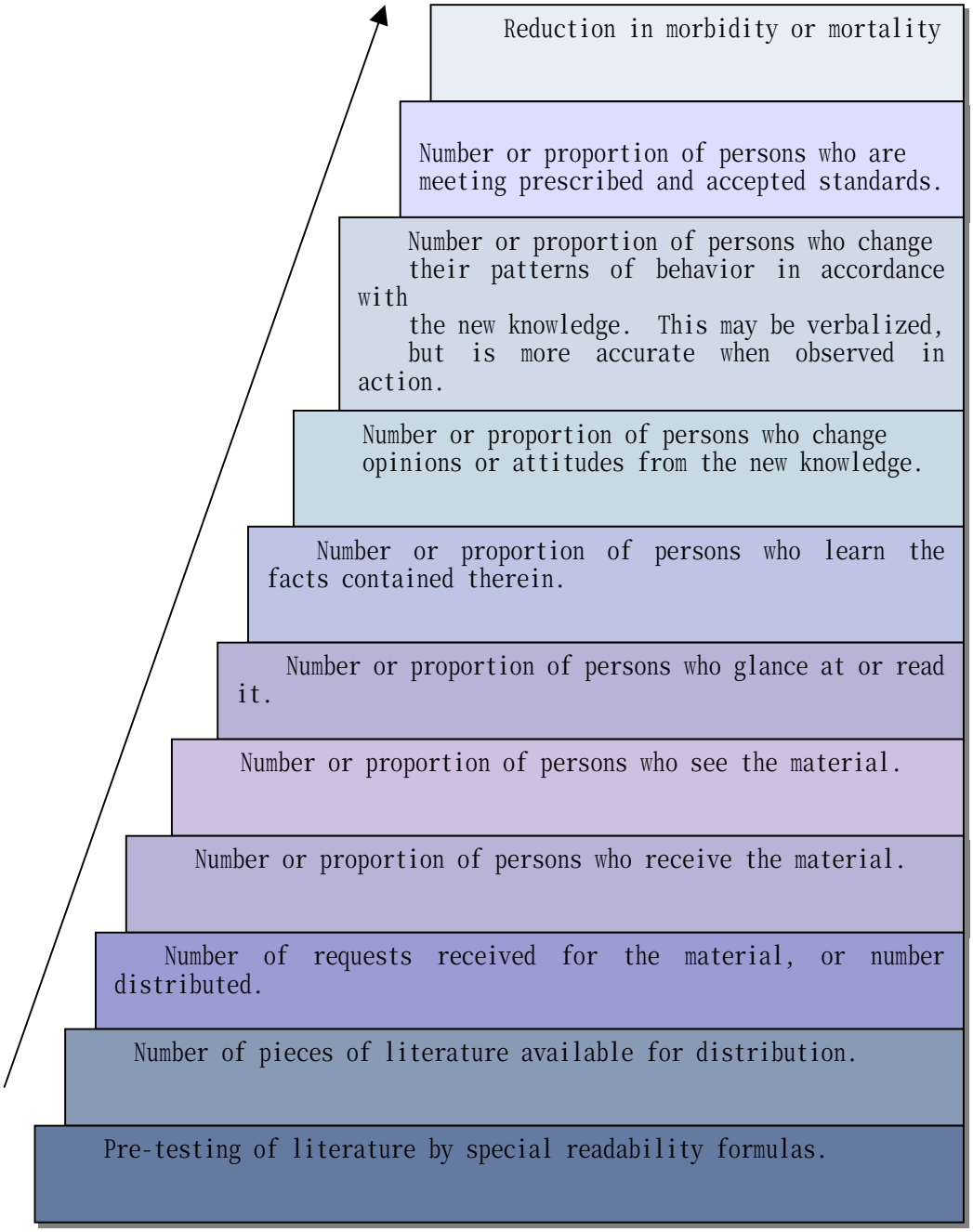
In this case, objectives may range from the most general such as reducing mortality, or to the very specific such as reading a health pamphlet. Furthermore, objectives and sub-objectives can be hypothesized corresponding to the various steps or actions that make up a program. While these are a continuous series of events, dividing them into a hierarchy of sub-goals is essential for evaluation purposes, each of which may be the result of successfully achieving the preceding goal, and a precondition to the next higher goal. <Figure 1> illustrates a cumulative chain of objectives progressing from the most immediate practical objectives toward the ultimate ideal goal in a step-wise chart showing different level of evaluation for health educational literature.

After determining the degree in which an objective has been achieved, this finding becomes a step toward achieving the next higher objectives.

This logic is based on the assumption that mortality will eventually be affected if a chain of prior accomplishments is fulfilled. In terms of the ultimate objective, this kind of evaluation assumes that the literature will reach a large proportion of the population for whom it was intended for and be read. Which in turn will have some effect in motivating the reader to go through the recommended procedures, and hence, reducing mortality as a result of it.

An assumption must be made whenever one moves from a higher-order objective to a lower one. Thus, every lower-level objective must assume all of the assumptions made for all of the objectives above it. Any program based on a false set of assumptions cannot be justified, even if sound evaluations are completed for each of the other objectives individually. In order to perform a successful evaluation, one needs to prove the intervening assumptions through research.

Figure 1. Chain of Objectives in the Health Field



Source : Suchman ((1967), (1967))

4. Special Characteristics of Performance Evaluation in the Health Field

Performance evaluation in every field follows general steps. First, questions calling for the evaluations, which are often too general and vague, need to be transformed into questions which are precise enough to be measured and evaluated. Followed by evaluation designing, analysis with credible evidence and presenting conclusions, and then lastly, sharing lessons and ensuring use.

Keeping in mind this commonality, however, particularity of evaluation in the health field should also be emphasized, which brings special features in each step of the evaluation process.

A. **Scientific Rigor Required**

Since the beginning of modern medicine, rigorous scientific inquiry has been the most important factor that has characterized the adoption of new procedures and practices in the health field. And this in turn characterizes performance evaluation in the area of health in a way that all causal links, explicit or implicit, in the intervention logic should be certified with scientific rigor.

Intervention logic is a hypothesis on the sequence of events for bringing about change by synthesizing the main program elements into a picture of how the program is supposed to work. As a conceptual link from a program's inputs, or resources devoted to it, to its output, and to the achievement of results and outcome, intervention logic summarizes the program's overall mechanism of change by linking processes to eventual effects (CDC 1999). Elements in a logic generally include inputs (e.g., trained staff), activities (e.g., identification of cases), outputs (e.g., persons completing treatment), and results ranging from immediate (e.g., curing affected persons) to intermediate (e.g., reduction in tuberculosis rate) to long-term effects (e.g., improvement of population health status).

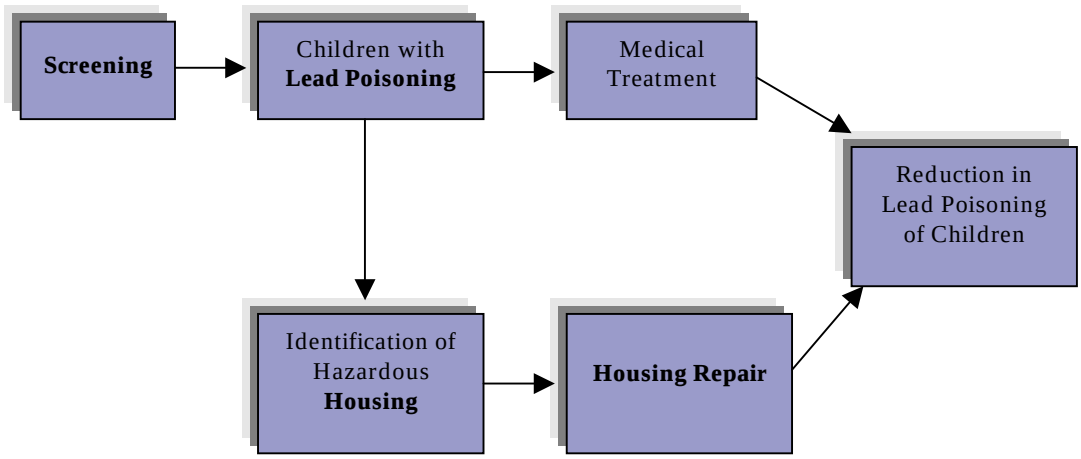
A main purpose of evaluation is to assess effects, and intervention logic, an explanation of what the public action is supposed to achieve and how it is supposed to achieve it, is the basis of the assessment. Since there are usually causal assumptions hidden beneath the surface relating to how the program is supposed to generate its supposed effects, and the state of the external environment such as other policy interventions and other external factors, the logic needs to be investigated carefully. In other words, the main job of the evaluator is to examine the relationship between activities of interest and observed consequences. This can be accomplished by identifying and critically assesing these hidden assumptions concerning conditions for the intervention to be effective. In the health field, intervention logic often involves scientific theory in laboratory diagnosis of disease or pharmaceutical knowledge, which makes it more complicating to identify the hidden assumptions and to investigate the their uncertainties. Inevitably, scientific specialty is heavily involved.

A noticeable example is the growing interest in preventive care. While advocates of preventive care assert that diagnosing and treating illness at an early stage can save a significant amount of money and successfully restore the patient to good health, the efficacy of preventive care is difficult to prove. It requires conceptualization of those specific conditions and socioeconomic characteristics of patients for whom preventive care might have some impact. For example, little is known about the circumstances under which various screening programs, such as Pap smears to detect cervical cancer, are effective.

<Figure 2> shows an example of a community-based program to reduce lead poisoning among children living in low-income, dilapidated housing areas in New York City in 1970. The lead poisoning problem mostly affected young children who ate paint chips based with lead that peeled from house walls. In order to reduce children's exposure to lead-based paint, the city started a program to screen 120,000 children for lead poisoning. If the lead level in a child's blood was high, the child's apartment was inspected, and if evidence of lead exposure was detected, wallboards were erected to prevent further exposure. The intervention logic illustrated in <Figure 2> assumes the housing repairs would have direct influence in reducing lead poisoning among children.

The result was that the incidence of lead poisoning decreased by 80% in 3 years. However, it was not the installation of wallboards that caused the decrease of incidences. Instead, due to the wide-scale screening and publicity about the program, parents monitored their children's behavior more diligently, preventing them from eating paint chips. This example shows that a causal assumption in the intervention logic, which is that children's exposure to lead paint would be reduced by house repairs, had missed important links. In this case, the program worked as intended but not because of the reasons specified in its intervention logic. It might have had a better result if the resources allocated for housing repairs had been invested differently such as expanding the number of houses being screened and campaign on children being exposed to lead poisoning.

Figure 2. Lead Poisoning Program



Source: Grembowski (2001)

<Figure 3> shows an example of a smoking cessation program to reduce low birth weight among low-income, pregnant women. Sheehan (1998) examined the association between stress, addictive behaviors, and low birth weight among 5,295 low-income, inner-city women in Hartford, Connecticut. The arrows in <Figure 3> indicate direct and indirect relationships between various factors and low birth weight. Addictive behaviors such as smoking and alcohol consumption have direct effect on low birth weight in the causal relationships. If a pregnant woman smokes or drinks alcohol more, the child's birth weight would be lowered. The standardized regression coefficients indicate the size of the change in low birth weight produced by a standardized change in various factors. Economic stress has direct effects on family stress and social support, and family stress

and social support, in turn, have direct effects on addictive behavior. Hence, the three stress variables have indirect effects on low birth weight through their direct effects on smoking and drinking behavior.

This causal relationship has some implications about the program's expected outcome. First, the program alone will likely reduce but not eliminate the low birth weight problem. In the cause and effect structure, 90% of the variation in low birth weight is not explained by the variables in the structure. Also, the figure indicates that both medical risks and addictive behavior have direct effects on low birth weight. Thus, because the unexplained variance is large and other risk factors exist beyond the program's control, the program will at most reduce the low birth weight problem.

The size of reduction can also be projected. It is suggested that the program's indirect effect on low birth weight will be smaller than 0.28. If the program reduces smoking by 0.25, then the program's indirect effect on low birth weight would be $0.25 \times 0.28 = 0.07$.

Second, the causal relationship in the figure also suggests that the program's performance would be increased if the program was modified. As shown, because stress has a direct effect on addiction and indirect effects on low birth weight, the program would perform better if it also includes a part for stress reduction.

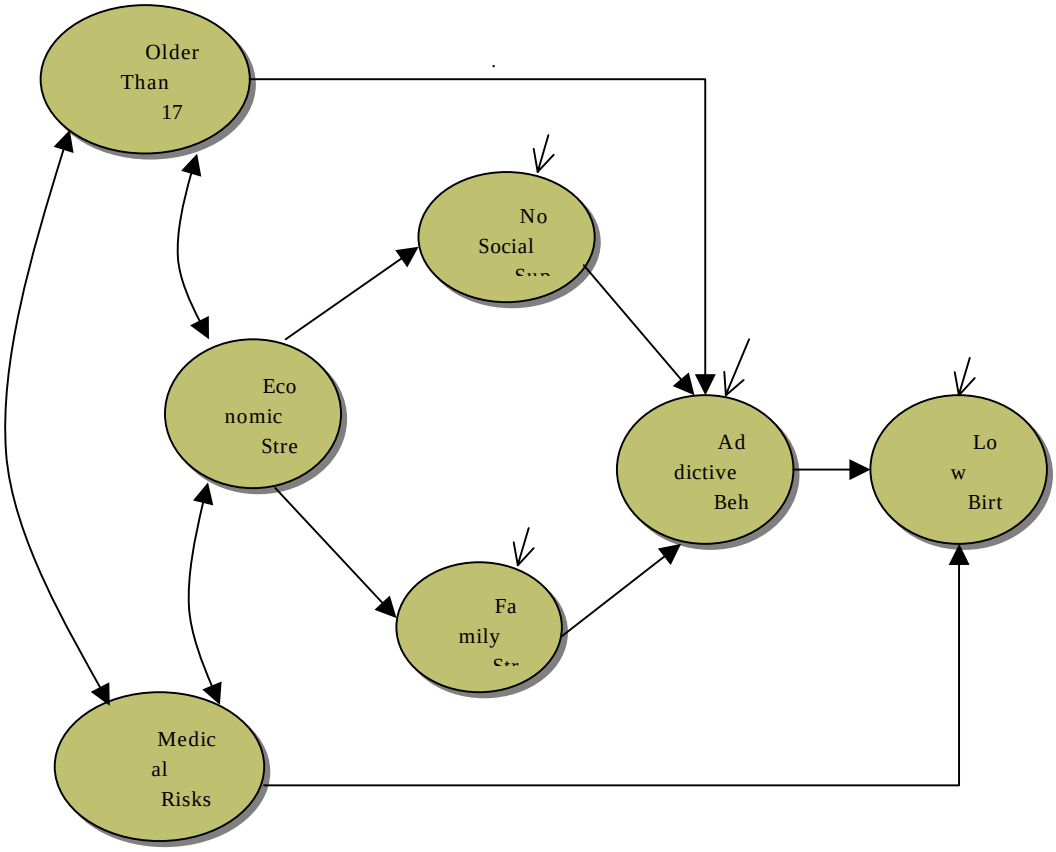
This is an example showing that scientific and statistic reasoning offers an understanding of how the program is supposed to work, and how efficient the program will be in achieving its objectives. Evaluation research in the health field is normally based on theory from laboratory experiments, economic reasoning and statistical projection, and it should be constantly adjusted with scientific rigor whenever new information about the evaluation logic is added to the information set or the contextual situation is changed.

B. Contextual Complexity

In most cases, health programs lie in complicated contexts including historical, geographical settings, political, social and economic conditions, and influences of related or competing organizations. The most important setting is the socio-economic conditions. The evaluator should consider the needs of the target population and the particular problems which the program is designed to address. It is difficult to identify what effects are genuinely caused by a program and to separate these effects from other influences on the socio-economic problems. Hence, the evaluator needs to understand these intertwined factors in order to design a context-sensitive evaluation, to interpret findings accurately and to assess the generalizability of the findings.

It should also be noted that because health programs are often created to reduce health disparities across socio-economic groups, evaluations also must consider political aspects among various income groups and politicians, and interest groups surrounding a health care program (Shortell 1978). Since politics is how we attach values to policy issues, politics and values are inseparable from the evaluation of health programs in nature. Consequently, evaluations are always conducted in a political context where a variety of interest groups compete in their own interests.

Figure 3. Smoking and Low Birth Weight



Source: Sheehan "Stress an Low Birth Weight" reproduced from Grembowski (2001)

5. Methods in Performance Evaluation for Health Care Intervention

A. Estimating Counterfactual Situation

When we say that certain effects were produced or caused by an intervention, this means that if the intervention had not taken place or had taken place in a different degree, those effects would not have occurred, or would nor have occurred in the same degree. Simply put, evaluations are carried out to determine what would have happened without the intervention. This is called the counterfactual situation (European Commission 1997), and ideally, deriving the counterfactual situation with certainty is desired.

Evaluating would be easy when comparing 2 groups that are identical in every respect except that one group is exposed to the intervention while the other group is not. In the real world, however, this is not the case since we can never find 2 groups that are identical in every respect except for being exposed to a program. The non-equivalence of the 2 groups means that the counterfactual situation has to be estimated rather than derived, and this emphasizes the evaluation design's scientific accuracy and data's credibility.

The methods for evaluation design are drawn from social, behavioral, statistical and health sciences. Design types include experimental, quasi-experimental, and observational designs, and according to the specific evaluation case, the best fitting design should be selected to provide the appropriate information to address the case specific questions (Fink 1993). Experimental designs use random assignment to compare the effect of an intervention with otherwise equivalent groups. Quasi-experimental methods compare nonequivalent groups (e.g., health program participants with those on a waiting list) or use multiple waves of data to set up a comparison. Observational methods use comparisons within a group to explain unique features of its members (CDC 1999). The choice of design has implications for what will count as evidence, how that evidence will be gathered, and what kind of claims can be made. In particular, health problems are usually entangled with other living conditions or other health problems where the origin of the cause is hard to trace, hence, special attention must be given when selecting a design. At the same time, data availability has enormous importance in performing evaluation in all types of design.

B. Customizing Tools for Evaluation

Though it is true that more rigorous methodological tools and designs focusing on the impact or outcome of evaluations are essential parts to ensuring the quality of healthcare evaluation research. What is more pressing is the need to understand specific program components and the process by which they are implemented since every health program has its special features due to scientific hypotheses or the contextual particularity in which it lies in. The interdependency of the health program with other political, environmental factors stated above should also be considered carefully.

For this reason, the observed golden rule is that there is no single evaluation methodology which is universally applicable for the entire health field. Instead, the choice of method should be determined by the particular evaluation problems at hand. Actually, arbitrary choice of method at the beginning without careful examination of the case often results in a poor evaluation. The evaluation methodology should be carefully chosen after examining the targeted health evaluation case, and more importantly, should be customized and modified to reflect the particularity of the case.

Also, if a program's activities are aligned with those of other programs operating in the same setting, or with change of social, economic conditions, certain effects cannot be attributed solely to one program or another. In this case, the goal of the evaluation is to gather credible evidence that describes each program's contribution in the combined change effort.

6. Prerequisites for Start-uping Performance Evaluation in the Health Field: Evaluation Infrastructure

A. Construction of Representative Frameworks for Evaluating Public Healthcare Intervention

Although it was previously emphasized that each evaluation case uses a different and customized methodology for the targeted program. It should also be highlighted that there exists a strong need for a recommended framework for health field evaluation that is both a synthesis of existing evaluation practices and a standard for further improvement.

The existence of a standardized framework supports a practical approach to

evaluation that is based on steps and standards applicable in public health settings. Because the framework is purposefully general, it provides a guide for designing and conducting specific evaluation projects across many different areas of public health intervention. In addition, the framework can be used as a template to create or enhance program-specific evaluation guidelines that further operationalize the steps and standards in ways that are appropriate for each program (CDC 1999).

By integrating the principles of the general framework into all program operations, it is possible to stimulate innovation toward the improvement of outcomes and better detection of program effects. Consequently, detecting these effects in a more efficient and timely manner will enhance the evaluator's ability to translate findings into practice. Furthermore, guided by the steps and standards in the framework, a basic approach to program planning will also evolve, and integrated information systems will support a more systematic measurement. Also, lessons learned from evaluations can be used more effectively to guide changes in public health strategies.

It should be noted that the standard framework should be considered as a practical, non-prescriptive tool for individual evaluation work; in that it only acts as a guide for public health professionals in their use of program evaluation. However, designed to summarize and organize essential elements of program evaluation, the framework comprises steps in program evaluation practices and standards for effective program evaluation. Adhering to the steps and standards of this framework will allow a better understanding of each program's context and will improve the way program evaluations are conceived and conducted.

It is also worth noting that the general framework can encourage evaluation research to be integrated with routine program operations. Because performance evaluation is in nature a case specific analysis, it is not easy to practice it consistently across program areas, or sufficiently integrate it into the day-to-day management of most public health interventions. Hence, there needs to be a standard framework that offers operating principles for guiding public health activities.

These principles of evaluation work as guidelines to improve how public health activities are planned and managed. It would be possible to ensure that new and existing programs follow these principles by requiring each program to conduct routine, practical evaluations that provide information for management and improve program effectiveness. In that way, public health professionals will recognize that the basic steps of the evaluation's framework are parts of their routine work. In day-to-day public health practice, program goals are defined; core questions are stated; data are collected, analyzed, and interpreted; judgments are formed; and lessons are shared.

B. Constructing Indicators for Major Healthcare Evaluation Types

Evaluation is about revealing the value of a program. This involves making value judgements on the degree to which a program's performance has been good or bad, so predetermined and transparent indicators are needed to ensure that value judgements do not become arbitrary. For this purpose, a prerequisite for evaluation research is to set up a system of criteria by which to measure the observed effects of a program and the benchmarks on the proper function or success of a program. However, setting the system of indicators may be difficult for a number of reasons. For one, objectives can sometimes be expressed in very vague terms, and a single program may have multiple objectives, either in terms of results of outcomes, some of which may carry relatively more weight, or even be incompatible with others. Also, objectives may also evolve over time, as the program's environment evolves. Furthermore, indicators should ideally allow evaluators to compare the performance with other policy instruments in the same field of action

because even if a program falls short of achieving its objectives, its performance may not necessarily be unsatisfactory.

In the health field, where intervention logic is grounded on scientific knowledge and all the political, social, and economic interests, indicators should function as the guiding rules that pertain to the evaluation's focus and questions. The role of indicators should translate general concepts regarding the health program, its context, and its expected effects into specific measures that can be interpreted, while providing a basis for collecting evidence that is valid and reliable for the evaluation's intended uses. For this reason, setting up a system of indicators in the area of health is more complicated than in other fields, because indicators need to present even the puzzling aspects of the laboratory hypothesis in a meaningful form for monitoring.

Indicators include measures of the program's capacity to deliver services, the participation rate, levels of client satisfaction, the efficiency of resource use, and the amount of intervention exposure. Also included are outcome measures such as changes in participant behavior, community norms, policies or practices, health status, quality of life, and the settings or environment around the program. Furthermore, multiple indicators are needed for tracking the implementation and effects of a program.

Developing multiple indicators should be based on intervention logic, and in this sense, intervention logic can be used as a template to define indicators leading from program activities to expected effects. For each step of the model, qualitative and quantitative indicators could be developed to suit the concept in question, linking assumption, the information available, and the planned usage of data. Relating indicators to intervention logic also allows detection of small changes in performance faster than when relying on a single outcome as the only performance measure. In addition, lines of responsibility and accountability are clarified in this approach as the measures are aligned with each step of the program strategy. Consequently, combining intervention logic and indicators results in a set of measures that detect the consequences of intermediate effects on health outcomes of the program.

Even during evaluation, indicators might be modified or new indicators can be adopted. According to changes in the environment or context, indicators need to adjust. It should also be noted that measuring performance by tracking indicators is only part of an evaluation and must not be the only basis for judging if whether an intervention is a success or failure. Serious problems can result from using performance indicators as a substitute for the entire evaluation process and reaching final conclusions. For example, an indicator such as a rising rate of disease might be assumed to reflect a failing program when, in reality, the indicator is influenced by changing conditions that are beyond the program's control.

C. Preparing Data Sources

When an indicator is proposed for use in performance monitoring, appropriate data must be available to support its use. Ideally, data would be collected from the specific population of interest, within the relevant time frame, using valid, reliable, and responsive measures. However, collecting data is expensive work, and realistically, efforts to expand data collection can reduce the resources available for programs. As a result, performance evaluations often have to rely on data collected for another purpose, so evaluators must understand limitations on the applicability of the data. Whether or not specific populations of interest are included in samples from which data are drawn and whether data are collected sufficiently often, or are made available soon enough, need to be considered in the monitoring process.

Another problem arises when only data sources that provide information on a particular group are available, and comparisons need to be made among groups. In this case, understanding the limits of available data has critical importance in drawing appropriate inferences (Drummond et al. 1997).

7. Evaluating Indicators for Healthcare Evaluation: Experience of U.S.

At the request of the U.S. Department of Health and Human Services (DHHS), the Panel on Performance Measures and Data for Public Health Performance Partnership Grants was established. The panel was given the task of examining the state of performance measurement for public health and to recommend performance measures in ten areas including: chronic disease; sexually transmitted disease (STD), human immunodeficiency virus (HIV) infection, and tuberculosis; mental health; immunization; substance abuse; and three areas of prevention of special interest to DHHS- sexual assault, disabilities, and emergency medical services. The panel focused on measures the government would be able to use over the next 3-5 years to negotiate agreements and monitor performance in these areas.

More than 3200 measures were proposed to the panel through various outreach efforts. The panel used four guidelines for assessing them, and measures that scored the highest were recommended for use in performance monitoring. The objective of recommending performance measures is to provide technically sound methods for assessing progress in meeting public health objectives and to provide the government practical and useful tools to advance their public health objectives (National Research Council 1997).

A. The Panel's Framework for Assessing Suggested Indicators

Although health outcome measures are widely used and have intuitive appeal, they are insufficient by themselves for monitoring the efforts of a given program to reduce complex public health problems. Many measures of health outcomes are affected by various factors that are not under the health intervention's control, subsequently, compromising the validity of measures of program effect. Hence, changes in outcomes cannot be attributed only to specific program effectiveness. Another constraint is that many important public health objectives, such as lowering the incidence of cancer and HIV infection, cannot be achieved over short periods of time to derive an outcome measure.

For example, if the policy intervention's goal is to reduce the mortality rate from breast cancer, it can reduce the risk of such adverse health outcomes by increasing the number of mammograms it provides to women aged 50 and over. However, the effectiveness of the intervention cannot be shown in a short period of time. Furthermore, there are a series of process activities (health education programs, surgical and nonsurgical treatment, and postoperative follow-up care) and capacity indicators (number of trained staff and facilities offering mammography screening) that are believed to be related to the level of mortality from breast cancer.

In this case, the evaluator needs to consider whether desired health outcomes are achieved, whether specific agency commitments are carried out, and whether the agency has the capacity to conduct all the necessary processes. Outcomes are fundamental, and any process or capacity measure used to assess performance should be widely accepted as being closely related to them.

For this reason, the panel recommends using "intermediate" outcome measures,

such as risk status, for which there is general consensus that the result being measured is related to the health status outcome, and that performance monitoring must make use of process and capacity measures to complement available measures of outcomes. Each process and capacity measure should be accompanied by reference to published clinical guidelines or other professional standards that describe the relationship between the process measure or capacity measure and the desired health outcome.

> Health outcome: change in the health of a defined population related to an intervention.
>
> Risk status (intermediate outcome): change in the risk demonstrated or assumed to be associated with health status.
>
> Process: what is done, for, with, or be defined individuals or groups as part of the delivery of services, such as performing a test or procedure or offering an educational service.
>
> Capacity: the ability to provide specific services, such as clinical screening and disease surveillance, made possible by the maintenance of the basic infrastructure of the public health system, as well as specific program resources.

To sum up, the effectiveness of a state program in using resources can most appropriately be evaluated by assessing the degree to which desired changes in health outcomes are achieved, together with a judgement of the degree to which those changes can be attributed to a program. When a firm causal link between the resources and processes used and the health outcome sought has not been established, as is often the case, or when the program resources are a small part of all the resources that contribute to the outcome, the panel believes that performance assessment must necessarily depend on a combination of health outcome, process and program capacity measures.

And performance measures should be understood and adopted as the product of an evolutionary process, to be revised as additional empirical evidence is obtained and better methods of data collection are implemented.

B. Guidelines for the Assessment of Proposed Indicators

1. Measures should be aimed at a specific objective and be result oriented. Measures must clearly specify a desired public health result, including identifying the population affected and the time frame involved. Process and capacity measures should clearly specify the health outcome, or long-term objective, to which they are thought to be related.
2. Indicators should be meaningful and understandable. Performance measures must be seen as important to both the general public and policy makers at all levels of government and they should be stated in non-technical terms.
3. Data should be adequate to support the indicator. Adequate data on the populations of interest must be available for the use of indicators and have the following characteristics:

   Data to track any objective must meet reasonable statistical standards for accuracy and completeness.

   Data to track any objective must be available in a timely fashion, at appropriate periodicity, and at reasonable cost; and

   Data applied to a specific indicator must be collected using similar methods and with a common definition throughout the population of interest. Comparisons of an indicator

across regions are valid only if the definition and collection methodology are consistent across regions.

4. Indicators should be valid, reliable, and responsive.
   Indicators should as much as possible, capture the essence of what they purport to measure, be reproducible (reliable), and be able to detect movement toward a desired objective (responsive).

C. Example: CHRONIC DISEALSE

Prevention of chronic disease is the primary goal of many health programs, but chronic disease incidence and mortality data are not useful because the expected time period between most prevention activities and the effect of those activities on disease incidence or mortality tends to exceed several years, which also exceeds the time that health departments are generally willing to wait to assess the intervention's effectiveness. Therefore, potential chronic disease measures are focused on risk reduction and screening, and process measures.

Potential Risk Status Measures Risk status measures represent intermediate health outcomes (see fn. 1).	
Tobacco	
Individual adult	Percentage of (a) persons aged 18-24 and (b) persons aged 25 and older currently smoking tobacco
Individual youth	Percentage of persons aged 14-17 (grades 9-12) currently smoking tobacco
Individual pregnant woman	Percentage of women who gave birth in the past year and reported smoking tobacco during pregnancy
Individual working adult	Percentage of employed adults whose workplace has an official policy that bans smoking
Nutrition	
Content	Percentage of persons aged 18 and older who eat five or more servings of fruits and vegetables per day
Content	Percentage of persons aged 14-17 (grades 9-12) who eat five or more servings of fruits and vegetables per day
Total calories	Percentage of persons aged 18 and older who are 20 percent or more above optimal body mass index
Exercise	
Individual adult	Percentage of persons aged 18 and older who do not engage in physical activity or exercise
Individual youth	Percentage of persons aged 14-17 (grades 9-12) who do not engage in physical activity or exercise
Screenings and Tests	
Hypertension	Percentage of persons aged 18 and older who had their blood pressure checked within past 2 years
Cholesterol	Percentage of women aged 45 and older and men aged 35 and older who had their cholesterol checked within past 5 years
Breast Cancer	Percentage of women aged 50 and older who received a mammogram within past 2 years
Colon Cancer	Percentage of adults aged 50 and older who had a fecal occult blood test within past 12 months or flexible sigmoidoscopy within past 5 years
Cervical Cancer	Percentage of women aged 18 and older who received a Pap smear within past 3 years
Diabetes	
HbA1C	Percentage of persons with diabetes who had HbA1C checked within past 12 months

Foot exam	Percentage of persons with diabetes who had a health professional examine their feet at least once within past 12 months
Eye exam	Percentage of persons with diabetes who received a dilated eye exam within past 12 months

Examples of Process Measures	
Nutrition Program Strategy	Enable children to learn healthy dietary habits
Process Measure	Percentage of schools with menus that meet dietary guidelines for fat content and five or more servings of fruits and vegetable daily
Physical Activity Program Strategy	Increase opportunities for sedentary working adult to exercise
Process Measure	Percentage of worksites with worksite wellness programs that include physical exercise
Smoking Program Strategy	See Table 3-1
Screening Program Strategy	Educating patients regarding need for and appropriate timing of screening tests
Process Measure	Percentage of persons with diabetes receiving diabetes health education
Screening Program Strategy	Improving access to screening services
Process Measure	Percentage of managed care organizations in which patients can schedule mammograms at convenient times for them
Screening Program Strategy	Implementing tracking and recall systems
Process Measure	Proportion of providers with chart-based or other real-time system for identifying women in need of mammography
Examples of Capacity Measures	
Resources	Number of full-time health department employees for chronic disease prevention
	Number of public service messages prepared by state agency shown annually for chronic disease prevention
Proficiencies	Number of key surveillance systems and data sets (i.e., death certificates, cancer registry data, birth certificates, Behavioral Risk Factor Surveillance System (BRFSS), Youth Risk Behavior Surveillance System (YRBSS), hospital discharge data, Medicare and Medicare encounter information and other relevant local data sets) that are establish and maintained
	Percentage of local health departments receiving technical assistance and training
	Percentage of labs that meet quality standards
Planning	Percentage of population served by systematic community planning process, with leadership provided by the official health agency and participation of all relevant groups (e.g., consumers, providers, advocators)
	Percentage of population covered by written comprehensive chronic disease prevention plan(s) containing priorities and objectives based on needs, resources, and local demands
Community Involvement	Percentage of health care providers working under agreements established with public health departments to provide population-based prevention programming to reduce major risk factors for premature morbidity and mortality
	Proportion of health department programs that operate within the framework of a community coalition or have a community advisory group

D. Example: STDS, HIV, AND TUBERCULOSIS

HIV infection, tuberculosis, and many of the STDs resemble noninfectious chronic disease. The interval between the acquisition of infection and the development of serious consequences may be years long. Monitoring the long-term consequences of these

infections does not provide a useful short-term indication of the performance of prevention efforts, so measuring in this public health area is more complex than for others. Three general strategies for preventing these diseases are: reducing the risk of exposure, reducing the probability that an exposed person becomes infected, and reducing the duration of the infectious state among persons who became infected. Indicators are measure best suited to these purposes.

Potential Health Status Outcome Measures	
Incidence rates of selected STDs	Rate of reported gonococcal urethritis in men
	Rate of reported chlamydial urethritis in men
	Rate of reported cases of primary and secondary syphilis
	Rate of reported cases of congenital syphilis
Incidence rates of HIV infection	Rate of reported newly diagnosed cases of HIV infection
	Rate of perinatally acquired HIV infection of infants
Prevalence rates of selected STDs	Prevalence rate of gonococcal infection in women in defined populations
	Prevalence rate of chlamydial infection in defined populations
	Prevalence rate of syphilis in defined risk groups, e.g., pregnant women
	Prevalence rate of rectal gonococcal infection in men
Prevalence rate of HIV infection	Seroprevalence of HIV infection in defined populations at high risk of the infection, e.g., pregnant women who abuse drugs
Potential Consumer Satisfaction Outcome Measure	
Rates of consumer satisfaction with STD, HIV, and tuberculosis treatment programs	
Potential Risk Status Measures	
Risk status measures represent intermediate health outcomes (see fn 1).	
Rates of sexual activity among adolescents aged 14-17	
Rates of sexual activity with multiple sex partners among people aged 18 and older	
Rates of condom use during last episode of sexual intercourse among sexually active adolescents aged 14-17	
Rates of condom use by persons aged 18 and older with multiple sex partners during last episode of sexual intercourse	
Rates of condom use during last episode of sexual intercourse among men having sex with men	
Rates of injection drug use among adolescents and adults	
Completion rates of treatment for STDs, HIV infection, and tuberculosis	
Example of process Measures	
Program Strategy	Reduce barriers to receiving treatment from specific providers
Process Measure	Percentage of patients with insurance coverage for specific treatments
Process Measure	Percentage of patients reporting no transportation barriers to obtain necessary services
Process Measure	Percentage of physicians and other care providers receiving cultural competency training
Program Strategy	Improve quality of services provided
Process Measure	Percentage of cases followed up after most recent contact
Process Measure	Percentage of known intravenous drug users with access to needle exchange program

Examples of Capacity Measures	
Resources	Percentage of high-risk communities with nearby testing and screening services from multiple types of health providers and public health organizations
Planning	Percentage of the state's population who reside in communities that are engaged in formal community processes for assessing and planning for HIV / STD / TB prevention and treatment services

E. Example: MENTAL HEALTH

Although there is little agreement on linkages between specific process and mental health outcomes, there is some agreement on the dimension that are important in evaluating mental health services; quality assurance, access to services and utilization of services, consumer satisfaction with services, and psychological and social outcomes.

Potential Health Status Outcome Measure
Percentage of persons aged 18 and older receiving mental health services who experience reduced psychological distress
Potential Social Functioning Outcome Measure
Percentage of persons aged 18 and older receiving mental health services who experience increased level of functioning
Percentage of persons aged 18 and older receiving mental health services who report increased employment (including volunteer time)
Percentage of persons aged 18 and older with serious and persistent mental illness receiving mental health services who live in integrated, independent living situations or with family members
Percentage of children aged 17 and younger with serious emotional disorders receiving mental health services who live in noncustodial living situations
Percentage of persons aged 18 and older with serious mental illness who are in prisons and jails
Percentage of children aged 17 and younger with serious emotional disorders who are in juvenile justice facilities
Percentage of homeless persons aged 18 and older who have a serious mental illness
Potential Consumer Satisfaction Outcome Measures
Percentage of adolescents aged 14-17 or family members of children and adolescents or both who are satisfied with: (a) access to services, (b) appropriateness of services, and (c) perceptions of gain in personal outcomes
Percentage of persons (aged 18 and older) or their family members or both who are satisfied with: (a) access to mental health services, (b) appropriateness of services, and (c) perceptions of gain in personal outcomes

Examples of Process and Capacity Measures	
Program Strategy	Improve access to services / utilization of services
Process Measure	Percentage of adults with serious and persistent mental illness who use health services
Process Measure	Percentage of youth with serious emotional disorders who use mental health services
Process Measure	Percentage of those who use services that are voluntary
Process Measure	Percentage of people requesting services who begin receiving services within 2 week of the initial request
Capacity Measure	Percentage of primary care providers who receive supplemental traning in mental health services

Program strategy	Improve quality assurance
Process Measure	Percentage of service plans that include input from consumers and family members
Process Measure	Percentage of children less than 5 years old who are screened and assessed for mental health intervention
Capacity Measure	Percentage of primary care providers who use standardized screening tools for assessing the mental health status of primary care clients

F. Example: IMMUNIZATION

The vaccine preventable diseases re well understood. Cases are easily identified, and health outcomes are well defined. A strong, causal relationship exists between immunization and disease prevention. Clearly defined process (vaccination rate) and outcome (disease incidence rate) measures are preferred, and they should be responsive to effective interventions within a short period of time.

Potential Health Outcome Measures

Reported incidence rate of representative vaccine-preventable diseases

Potential Health Outcome Measures

Risk status measures represent intermediate health outcomes (see fn. 1).
Age-appropriate vaccination rates for target age groups for each major vaccine group:
- Children aged 2 years; children entering school at approximately 5 years of age
  - Mumps, measles, and rubella
  - Diphtheria-tetanus-pertussis
  - Polio
  - hemophilus influenza B
  - varicella
- Adults aged 65 years and older
  - diphtheria-tetanus
  - hepatitis B
  - influenza
  - pneumococcal pneumonia

Examples of Process Measures

Program Strategy	Improve access to immunization services
Process Measure	Percentage of population who do not cite financial resources as a barrier to immunization
Program Strategy	Increase parent education and awareness
Process Measure	Percentage of parents with children under 18 who believe that the benefits of immunization outweight the risks
Process Measure	Percentage of parents with children under 18 who report receiving immunization reminders from their immunization providers

8. Summary

As stated in the beginning, performance evaluation in the health field is an increasingly important area of study for future government officials, health professionals, and services providers. The essence of the evaluation project is to isolate causes of particular events or outcomes. If a particular intervention appears to be associated with a beneficial effect, one must find out whether the effects can really be attributable to the intervention or whether they might be the result of some other factors. With this commonality, what distinguishes evaluation of public health action from other fields is the use of rigorous scientific methods involving laboratory frontier knowledge and contextual complexity requiring socio-economic and political awareness.

In order to set up an evaluation system in the field, building infrastructure is required more than anything. The framework for evaluating public health intervention needs to be devised, indicators need to be collected and assessed for future use. Also, to utilize indicators, appropriate data sources for each indicator should be gathered and organized. Furthermore, since indicators are designed to detect changes in a limited time period, following up data collection should be made within a reasonable time. Although it will take a tremendous amount of resources to accomplish these prerequisite tasks, now is the time for Korea to tackle them in order to remain accountable and to manage surging healthcare costs.

Example Indicators for Chronic Disease; STD, HIV, and Tuberculosis; Mental Health; Immunization

Reference

Centers for Disease Control and Prevention, “Framework for Program Evaluation in Public Health,” 1999.

Drummond Michael F., Bernie J. O’Brien, Greg L. Stoddart, George W. Torrance, “Methods, for the Economics Evaluation of Health Care Programmes,” Oxford University Press, 1997.

European Commission, “Evaluating EU Expenditure Programmes: A Guide,” 1997.

Fink Arlene, “Evaluation Fundamentals: Guiding Health Programs, Research, and Policy,” SAGE Publictions, 1993.

Grembowski, David, “The Practice of health Program Evaluation,” SAGE Publications, 2002.

Ministry of Health and Welfare, “Yearbook of Health and Welfare Statistics," 2003.

National Research Council, “Assessment of Performance Measures for Public Health, Substance Abuse, and Mental Health,” Edward B. Perrin and Jeffrey J. Koshel eds, 1997.

Shortell Stephen M. and William C. Richardson, “Health Program Evaluation,” The C. V. Mosby Company, 1978.

Suchman, Edward A., “Evaluative Research: Principles and Practice in Public Service & Social Action Programs,” Russel Sage Foundation, 1967.