

Performance Measurement and Program Monitoring

**Prairie Research Associates Inc.
Winnipeg, Manitoba**

Presented
to
Program Standards, Evaluation and Monitoring
Manitoba Health

October 21, 1997

Outline of the workshop

The Program Standards, Evaluation, and Monitoring Branch is charged with the responsibility of creating standards to ensure quality of health services and accountability.

The central mandate for the new branch is to determine whether *“the resources put into the health system are making a positive difference to people’s health.”* Other key questions are:

- Were needs accurately defined?
- Were the goals appropriate?
- Were the cause and effect relationships valid?
- Is there new knowledge we can apply to meet needs and improve health?¹

These questions are essential for measuring performance.

¹ Adapted from, “A Planning Framework to Promote, Preserve and Protect the Health of Manitobans.”

The workshop presents the following themes:

- (A) Implementing performance measurement and results based activity
- (B) Creating standards and using benchmarks
- (C) Defining/developing performance measures
- (D) Auditing/measuring performance
- (E) Assessing and managing risk in health services delivery
- (F) Working to ensure acceptance, support, and internal promotion of standards.

Preliminaries 1

Performance measurements and standards are firmly based on the business model of operation.

- performance measures as profits, sales, market share or whatever objective management believes serves the objectives of the firm.
- as objectives change, so will performance measures.

Example 1: A profit maximizing firm may not seek to expand market share, being content to return the highest possible dividends to the share holder. A firm that wishes to expand market share, perhaps as a defensive strategy against a potential competitor, may sacrifice profit. These two objectives will have quite different standards and quite different measures of performance.

Preliminaries 2

Standards need to be referenced to an external measure that is credible and acceptable to all stakeholders.

Standards can be moveable as part of a total and continuous quality process.

Standards must be consistent with objectives and performance measures.

Deviations of performance from standards must be accepted as “real” by stakeholders and used as the basis for action.

Preliminaries 3

Laws of performance²

- | | |
|-------|--|
| Law 1 | If you try to be the best at everything, you'll be the best at nothing. |
| Law 2 | People are more important than the process, but good process is important to people. |
| Law 3 | If you can't describe it you can't improve it. |
| Law 4 | If you don't have a goal you can't score. |
| Law 5 | Measuring the activity usually improves the activity, but not the result. |
| Law 6 | If you know the score, you should be able to predict the result. |

² From Performance Based Management, Department of Energy, United States (<http://www.llnl.gov/PBM/handbook>).

Example 2 Infections resulting from intravenous injection of chemotherapy

<i>The case</i>	<i>The Intervention</i>	<i>The results</i>
<i>Some patients get infections from a permanent catheter fitted to a blood vessel. The usual response is to remove the catheter, treat the infection, and re-insert. The option tried was to retain the catheter and use stronger antibiotics. However, the reinfection rate climbed from 10% to 45% and these patients had to have their catheter removed and be treated with increased antibiotics.</i>	<i>The initial cost saving of leaving the line in was \$2000 by saving two procedures (removal and reinsertion). The increased costs associated with treating reinfected patients (with catheter removal and increased antibiotics) was \$15,000 each.</i>	<i>A 72% reduction in re-admissions saved \$530,000 per year.</i>
<i>Source: Quality Measures Hit Prime Time, Modern Health, May 5, 1997.</i>		

Example 3 Testing for a marker enzyme after a heart attack

<i>The case</i>	<i>The Intervention</i>	<i>The results</i>
After a heart attack, a routine scan for an enzyme secreted by the damaged heart muscle was used to confirm diagnosis.	Analysis found that the test was bundled as part of the standard test battery. The results did not affect treatment or outcomes.	Dropping the test (86% reduction) for an annual savings of \$300,000.
Source: Quality Measures Hit Prime Time, <i>Modern Health</i> , May 5, 1997.		

Example 4 Maryland Quality Indicator Project

This set of clinical indicators seeks to summarize the quality of care through a set of 12 composite indicators. Some indicators comprise a set of sub-indicators.

Inpatient indicators

- hospital acquired infections
- surgical site infections
- inpatient mortality
- neonatal mortality
- inpatient preoperative mortality
- Caesarian sections
- unscheduled re-admissions for the same or related conditions
- unscheduled admissions following an ambulatory procedure
- unscheduled returns to a special care unit
- unscheduled returns to the operating room

Ambulatory indicators

- unscheduled returns to emergency for the same or a related condition
- registered patients' time in the emergency department
- emergency cases where a discrepancy between initial and later x-rays changed the treatment
- patients leaving before treatment is completed
- cancellation of ambulatory procedure on day of procedure.

(A) Implementing performance measurement and results-based activity³

Several core steps support a departmental/agency move toward outcome (results) based planning.

Step 1 Define the goals for the department/agency/unit

Step 2 Measure performance

Step 3 Use performance data to identify program changes

Step 4 Reinforcement

³

This section is based on Effectively Implementing the Government Performance and Results Act, *Comptroller General of the United States*, GAO/GGD - 96-118.

Step 1 Define the goals for the department/agency/unit

As the Branch creates the standards and measures process, it must ensure that a number of key conditions are fulfilled:

- **Involve stakeholders** including government, senior policy makers and managers, service providers, and consumers. Techniques such as public consultations, in-depth interviews, conventional market research, and feedback processes all may form the basis for creating and communicating standards.
- **Complete an internal and external scan** to identify factors that may both reinforce and counter the performance measurement process. Collective agreements and consumer interest groups are a common factor in using performance indicators.
- **Align activities and core processes with resources.** A key barrier for most departments is their limited ability to adjust resources from activities that exceed standards towards those that do not. Often political and institutional constraints frustrate this alignment. Create a production model for the department/agency/unit.

Step 2 Measure performance

- **Produce measures at each activity level.** Indicators of performance need to capture the outcomes of an activity unit and lead managers to a solution.
- **Collect data.** Data collection needs to be woven into the culture of the organization.

Step 3 Use performance data to identify program changes

- **Identify gaps between performance and expectations/standards.** Gaps between performance and standards create the “tension” that motivates change and reallocates resources to high “needs” units.
- **Report on gaps.** Regular reporting (monitoring) is mandatory and should be an automatic outcome of the measurement process. Monitoring reports need to be rolled-up to management summaries.
- **Use information to change operations/delivery**

Step 4 Reinforcement

- **Devolve and decentralize to meet performance targets.** Empowering units creates the climate for change.
- **Incentives (and sometimes penalties) motivate change** at the individual unit level.

Example 5: The Department of Veterans Affairs in the US offers an award to recognize performance and service to managers who successfully promote and implement quality change.

- **Develop expertise** within the department/agency to act as a resource to managers. Staff training is essential to diffuse the performance monitoring technology and to ensure that all parts of the organization are “on side.”

(B) Creating standards and using benchmarks

The entire quality control process is based on the objective for the organization.

Each unit within the organization that contributes to the overall results needs to have goals.

A unit that has no goals, is by definition, a non-contributor to the results for the organization and could be eliminated.

By working backwards from the results of a process (e.g., delivering public opinion results or servicing furnaces, or reducing health effects of smoking) one may construct the *essential* steps that are needed to produce that service.

Methods for creating standards

Selecting the norm against which performance will be assessed is a central strategic decision. Several basic goals shape what form of standard is to be prepared.

- **Legislative standards** are designed to ensure service that meets equity needs (e.g., “*all clients of mental health in Manitoba deserve access to the same care regardless of where they live*”).
- **Developmental goals/standards** are set low to encourage new groups to meet/exceed targets.
- **Quality improvement standards** often start low and are adjusted to encourage continuous improvements.

Sources of standards

Standards may be developed from a number of sources:

- **Expert panels** can define what is technically possible based on the mechanics of operation. At times, experts can overstate the actual performance capacity of a system.
- **Stakeholder wishes** often define a standard. This approach may result in lower standards, but ones that are achievable.
- **Client expectations/satisfaction scores** are a common standard.
- **Historical norms** can be comfortable, but may reflect the output levels of different organizations' structures. New technologies may have increased performance — downsizing may have lowered performance.
- **Benchmarking** is common to health activities, especially hospital services where procedures are widely similar. These forms of standards are useful in comparing services in jurisdictions of similar population densities and fiscal capacities.

Benchmarking

Example 6 *HClA - Mercer's Top 100 Hospitals*

This study uses common indicators to benchmark hospitals as a basis for setting improvement goals. The indicators used for the rating comprise:

- *clinical practices*
 - *risk adjusted mortality index*
 - *risk adjusted complications index*
 - *severity adjusted length of stay*
- *operations*
 - *index of outpatient activity*
 - *expense per adjusted discharge*
- *financial management*
 - *profitability*
 - *long-term growth in equity*
 - *productivity (total asset turnover)*

(C) Defining/developing performance measures

Frameworks for Performance Measurement

Many frameworks exist for defining a system of performance measures:

- **Supply chain model**⁴ moves from inputs through to customer outcome.
 - *efficiency* (cost per unit input, unit cost of activity)
 - *effectiveness* (outcomes - satisfaction of customers, customer needs met)
 - *productivity* (output/unit input)
 - *profitability*
 - *quality* (rejects, rework, re-admissions)
 - *innovation*
 - *quality of worklife*

⁴

Sink. D. Scott and Tuttle, Thomas C., *Planning and Measurement in Your Organization of the Future*, Industrial Engineering and Management Press, Norcross Georgia, 1989.

Frameworks for Performance Measurement (Continued)

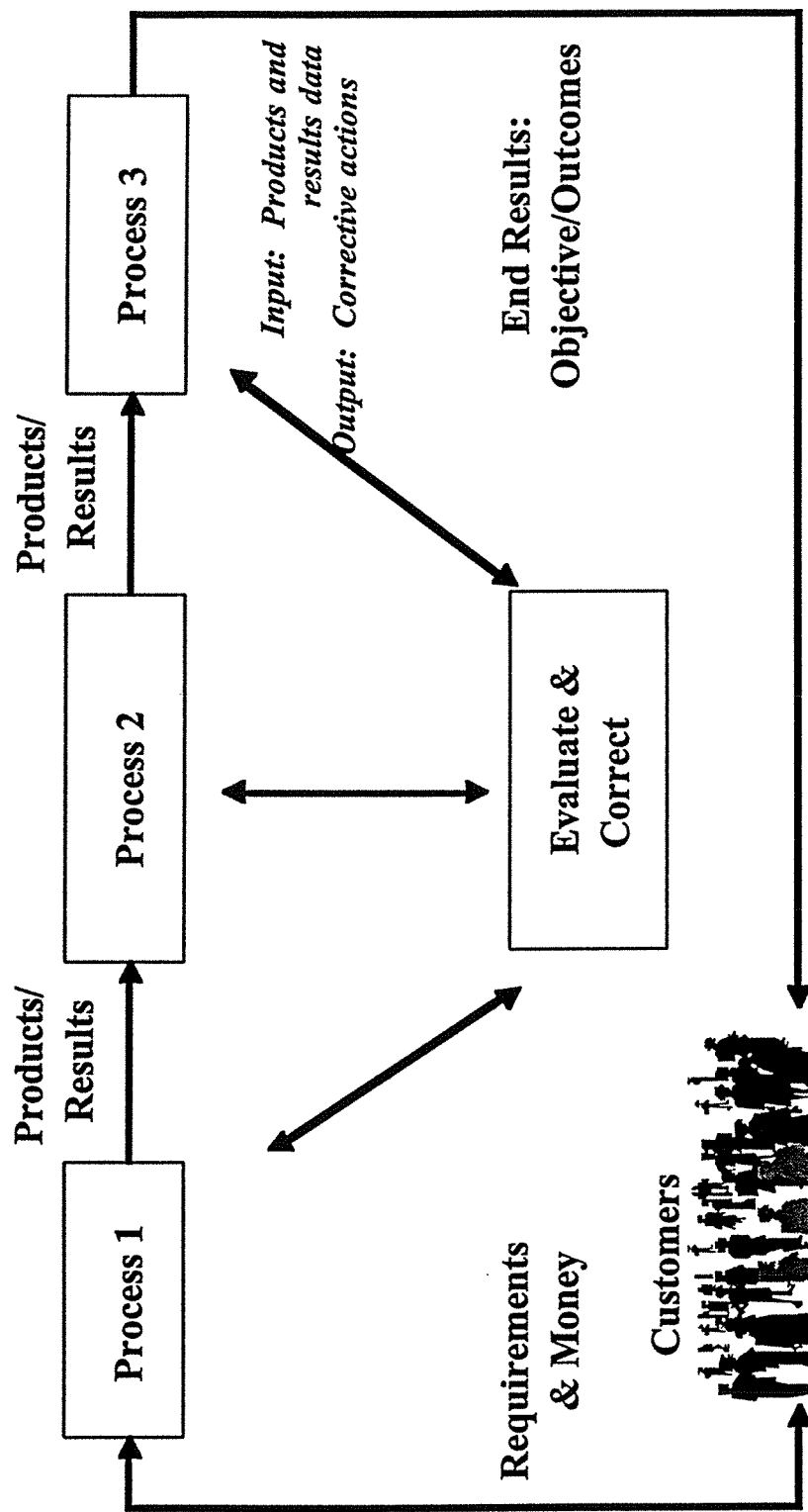
- **Performance Improvement Measurement Methodology (PIMM)**⁵ is oriented toward a Research & Development environment and measures objectives, achievement, cost performance, and risk reduction.
- **Balanced Scorecard**⁶ is designed for senior managers to maintain a summary assessment of the entire organization.
- **Program Logic Model**⁷ is developed collaboratively with managers and staff who create a picture of the production process. This picture leads directly to a performance matrix that moves from goals through to measurements. This perspective is closely aligned with the evaluation model used to assess the contribution of individual programs.

⁵ METC Performance Improvement Measurement Methodology, U.S. Department of Energy, (www.metc.doe.gov/tqm.html).

⁶ Kaplan, R.S. and Norton, David P. The Balanced Scorecard, Measures that Drive Performance, *Harvard Business Review*, January-February 1993.

⁷ Rush, Brian and Ogborne, Alan, Program Logic Models: Expanding Their Role and Structure in Program Planning Evaluation, *Canadian Journal of Program Evaluation*, 6:2, 1991.

Logic modeling for the production process



Generic system diagram⁸

⁸ From Department of Energy, *ibid.*

Example 7 - Firewood⁹

Consider a logic model for producing firewood. The first step identifies the linkages between the products/results and the elements/processes. Customers offer their needs or demands in the form of an order. These trigger requirements that the firm must fulfill and produce an expected income. The firm must then organize permits for cutting, secure interim financing to pay workers and create job orders. Labour and equipment are mobilized along with control documentation (time sheets). Firewood is cut, packaged, and delivered and these in turn generate profits and customer satisfaction.

Element/Process	Products/Results
Customers	Income, Requirements
Manage	Funding, Permits, Orders
Mobilize	Labour, Equipment, Documentation
Produce & Deliver Firewood	Cut/Package and Deliver Firewood
Evaluate & Correct	Profits, Customer Satisfaction

⁹ Department of Energy, Ibid.

Example 7 (continued)

The next "layer" consists of the goals for each process/product combination. The goals are development by management.

Process	Products	Goals
Customers (Output)	Income Requirements	25% of the regional market Twice last year's volume
Manage	Funding Permits Orders	+10% - 0% of budget request 100% timely availability 98% error-free
Mobilize	Equipment Documentation	97% uptime 99% error-free
Produce & Deliver Firewood	Delivered Firewood	Customer cost \$90 per cord 50 cords per day
Evaluate & Correct	Corrective Actions Surveys Costs Customer Satisfaction	100% on-time completion Positive perception 80% +0%, -10% of budget See Customer output goals

Example 7 (continued)

Each goal is measured conceptually in the fourth layer.

Process	Products	Goals	Measures
Customers	Income	25% of the regional market	Income compared with regional markets
	Requirements	Twice last year's volume	Volume delivered compared with last year
Manage	Funding	+10% - 0% of budget request	Available funds compared with budget request
	Permits	100% timely availability	Permits available when needed
	Orders	98% error-free	Proportion of error-free orders
Mobilize	Equipment	97% uptime	Proportion of available equipment time
	Documentation	99% error-free	Proportion of error-free documents
Produce & Deliver Firewood	Delivered Firewood	Selling price \$90 per cord 50 cords per day	Average selling price Average daily volume produced
	Corrective Actions	100% on-time completion	Proportion of corrective actions completed on time
Evaluate & Correct	Surveys	Positive perception 80%	Perception scores
	Costs	+0%, -10% of budget (90% to 100% of budget)	Spending

Example 7 (continued)

In the fifth layer the specific indicator or metric is added that will measure performance at each stage.

Process	Products	Goals	Measures	Metrics *
Customers	Income	25% of the regional market	Income compared with regional markets	<u>\$ income</u> \$ regional
	Requirements	Twice last year's volume	Volume delivered compared with last year	# cords this year # cords last year
	Funding	+10% - 0% of budget request	Available funds compared with budget request	<u>\$ funding</u> \$ requested
Manage	Permits	100% timely availability	Permits available when needed	# permits available # permits needed
	Orders	98% error-free	Proportion of error-free orders	# error-free orders # orders issued
Mobilize	Equipment	97% uptime	Proportion of available equipment time	<u>Equipment hrs. avail.</u> Equipment hrs. needed
	Documentation	99% error-free	Proportion of error-free documents	# error-free documents # documents issued
Produce & Deliver Firewood	Delivered Firewood	Selling price £ \$90 cord	Average selling price	<u>\$ income</u> # cords sold
		50 cords per day	Average daily volume produced	# cords stocked
Evaluate & Correct	Corrective Actions	100% on-time completion	Proportion of corrective actions completed on time	# completed on time # scheduled
	Surveys	Positive perception 80%	Perception scores	<u>Total actual score</u> Total possible score
	Costs	+0%, -10% of budget	Spending	<u>\$ spent</u> \$ budgeted

Properties of Good Performance Metrics	
Good Metrics are	Comment
Measurable	Performance metrics that are not based on data that are countable will not be credible. A bias exists toward administrative data that can be quantified and are a by-product of an operational process. <i>Example 8 A metric of the number of surgical site infections is preferred to the opinions of operating room personnel.</i>
Normalized	Normalized data have a "base." <i>Example 9 Recently MPI announced the rate at which various car models have been stolen in 1996/97. This is a useless measure without "normalizing" the data by the numbers of each brand registered in the province.</i> <i>Example 10 Disease rates are normalized by population (cases per 100,000).</i>
Reliable	The metric should not be affected by who does the measuring. <i>Example 11 Client satisfaction surveys should only show variation because the level of satisfaction has changed not because the company doing the survey has changed.</i>
Valid	Valid metrics capture the essence of a phenomenon.
Easy to produce	Metrics that require little effort to produce are more readily accepted by staff and management.
Independent	Metrics that are impervious to manipulation by staff are clearly preferable.
Direct	Metrics that directly capture inputs, outputs and outcomes are preferred.
	<i>Example 12 Number of flu shots captures inputs to an immunization while number of cases per 100,000 is the outcome. Complications, (cases of illness resulting from the shot per 100,000) is an unintended outcome.</i>

Breakout session 1

Develop logic model -- assessment and delivery process for home care

Home care delivery can be visualized as two main activities:

- assessment of clients that are referred by professionals and care-givers (friends and family)
- delivery of services to qualified clients

For each process you need to determine:

The Production Flow

What is the result?

What is the process used to produce the result?

What evaluations and measures may be attached to the individual steps in the process?

(D) Auditing/measuring performance

Service delivery relies on standards and systems to measure and audit operations.

Private sector measurement/auditing

Private sector performance stresses financial measures that are an outcome of the process of accepting payments for goods delivered and making payments for resources used.

Balance sheets and income statements evolve to meet the needs to raise capital from third parties and to service loans.

Various measures such as profit, leverage, and working capital are used to diagnose the health of a firm.

Standards are defined in relation to the alternative use of capital.

Investors will not place their capital in areas with low rates of return. Owners of firms cease operation if the profit rate is lower than alternatives.

Financial measures are incomplete and are supplemented by market research based indicators (client satisfaction, brand equity, market share...)

Example 13 Increasing post elective surgery readmission rates (a routine indicator) will usually require a purposeful data collection process (interviewing staff, observing procedures, interviewing patients).

Example 14 Two recent trends in health promotion are:

- *increasing teen smoking in the face of a preponderance of evidence about the effect of second hand smoke and health promotion.*
- *reduced condom use among sexually active teens.*

Purposeful diagnostic research is needed to understand these trends as well as the limits to conventional health promotion.

Public sector measurement and auditing

Public sector activities typically do not generate indicators such as revenues, profit, cost, or productivity as a natural outcome of operations.

Therefore, these measures must be developed and monitored consciously as part of any public sector operation.

Two general classes of measurement activity include:

- **Routine operational information** is a by-product of existing technical and financial operations. Although these data require the development of recording forms and regular reporting, measurement is routine and may be accomplished without disturbing operations.
- **Purposeful diagnostic information** requires occasional data gathering and analysis. If done regularly, such as a client satisfaction survey, these may become routine. However, no measurement and auditing system will ever be able to rely solely on routine indicators.

(E) Assessing and managing risk in health services delivery

Risk assessment is central to the creation of standards of care or performance.

- Setting a standard and using monitoring, one always runs the risk of not meeting the standard.
- Standards in public sector activity create an implicit contract with the public. The government is offering a warranty that patients can expect/demand a specific level of care.
- Risk varies with almost every socio-demographic variable, and thereby influences the standard of care.
- Health services affect risk directly by altering the incidence of disease through service delivery, and altering behavior through health promotion. Effective health promotion relies on raising the perception of risk in a target population.

Example 15 Living in a remote rural area increases response times and raises the risks associated with medical emergencies.

Example 16 Fiscal decisions, such as investment in diagnostics (MRI, CAT equipment), allocate service across time. Delays in procedures may increase risks associated with progressive diseases.

Any standard such as eight weeks wait time for an MRI scan is clearly related to both actual and perceived risks. Assessing the risk is a determinant of the standard, along with the risks associated with other elements of the health care system, and fiscal capacity to make these investments.

Risk assessment/management and health care

Health care is supplied to alleviate pain, prolong life and maintain the quality of life. Associated with each illness and injury are risks specific to each individual.

The growth of scientific medical science and health care is the expectation that these risks can be both assessed and managed. In Canada the public provider as well as the individual is seen as the agent that is responsible for assessing and managing these risks.

The interplay between private and public responsibility is complex.

Example 17 A smoker who contracts lung cancer, an ex-smoker who quit a decade ago and contracts lung cancer, and someone who never smoked are seen as having differing individual responsibilities for their condition.

The public provider may be seen by some as having equal responsibility for delivering care, although many may argue that the smoker “deserves” a lower standard of care.

Public perceptions of risk are a critical factor in the creation of acceptable health and safety standards.

What is risk?

Two polar views of risk frame the response to this question:

- Positivists assert that risk is an objective number based on measuring rates and using other scientific information. The scientific process is able to assess risk quantitatively.
- Relativists argue that risk is purely subjective and framed by perception and experience arising from personal and social circumstances.

Risk can be expressed as incidence for many health outcomes.

Example 18 The risk of heart attack can be computed for a class of individual based on sex, age, and specific indicators (blood pressure, cholesterol, weight, height, level of activity and family history).

The risk for a specific individual is harder to compute.

Risk and standards

Risk is often used as a synonym for probability. Where events are sufficiently frequent, these incidence levels can be used to develop standards. However, as standards and quality control are implemented the incidence may fall or rise.

- Low standards may encourage increased incidence of a disease and risk will rise.
- High standards may lower incidence and risk.
- The implementation of a measurement process may produce incidence figures that were previously non-existent. These incidence levels could be seen as either acceptable or unacceptably high.
- Incidence data, especially those related to child deaths have the potential to increase the perceived risks.

Risk perception and communication

Example 19 *Vaccination against a number of diseases is mandated in most developed countries. Recently, a counter movement has emerged to challenge the belief/standard that all children should be vaccinated. Immunization rates are falling and increasing numbers of physicians accept parents' decisions not to vaccinate.¹⁰*

Reduce immunization stems from several factors:

- cases of adverse reaction to vaccine
- increasing advocacy against vaccination
- decreased advocacy by health practitioners

Advocacy against vaccination arises from a number of points:

- assertions that diseases have disappeared without the use of vaccines
- arguments that periodic outbreaks of disease in vaccinated populations demonstrates non-efficacy
- omission bias (the belief that actions are riskier than inactions)
- rare but extreme adverse effects
- more common mild adverse effects
- the near eradication of some diseases reduces their perceived risks.

¹⁰

Bostrom, Ann, Vaccine risk communication: lessons from risk perception, decision making and environmental risk communication research, *Risk*, vol 8, p.173.

The basis for risk perception

Research in risk perception has confirmed the following:

- Risk (fear) is asymmetrical - it is easier to prevent fear that reassure those who are afraid.
- Knowledge of true risk tends to be higher for males, younger adults and those who are better educated.
- Perceived risk is dependent on:
 - the benefit that the activity conveys (e.g. “smoking relaxes me”)
 - the extent to which individual action can mediate the incidence
 - the number of deaths in an average setting
 - the “seriousness” of each death (a lingering painful death is dreaded more than a swift death)
 - the extent to which the incidence and effects are known.

Risk assessment and risk management

Standards and performance measurement are the most common risk management processes in health (as well as environmental safety, law enforcement, air transport, ...)

Risk assessment and risk management need to be connected.

Example 20 *Most consumers want their physician to both assess and manage the risks of their condition.*

A danger exists if standards are created in isolation from the management of health care. Those who are involved in identifying the risk, need to be involved in setting the standards.

At the same time, risk assessment stresses the gathering of information to increase the precision of the estimates of harm.

- Doctors may order many tests to reduce the risk from a wrong diagnosis and to increase income.
- At the same time the public health provider must develop a set of standards about diagnostic procedures.

(F) Working to ensure acceptance, support, and internal promotion of standards

1. Collaboration is important

- Risk assessment and standards setting require close consultation with program managers, senior decision-makers, health practitioners and the public.
- Be selective and strategic in the selection of stakeholder consultants.
- Benchmarking from other similar jurisdictions is a useful starting point in any standards exercise.
- A logic model of the production and organizational process will serve as a visual metaphor and allow all stakeholders to participate.

2. Design the measures for each activity area

- Identify information needs from strategic and operational plans.
- Select a few balanced measures that have well documented reliability and validity properties.
- Use measures that have information content (i.e., avoid yes/no, milestone, and threshold type measures since they do not motivate performance).

3. Collect data

- Determine data collection methodology (outside expertise will often prove beneficial).
- Accept that data measures will be redefined through the collection process.
- Allow a trial process where staff and managers can comment on the measures.
- Recognize that information collection is intrusive and respect the limits of staff in preparing data.
- Collect only what will be used.

4. Use the data